12-15-2016 12:00 AM

Producing a Subunit Vaccine for Porcine Epidemic Diarrhea Virus

Zayn Khamis, *The University of Western Ontario*

Supervisor: Dr. Rima Menassa, *The University of Western Ontario*
Joint Supervisor: Dr. Susanne Kohalmi, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biology

Abstract

Porcine epidemic diarrhea virus (PEDv) causes disease and mortality to piglets worldwide. Most vaccines used to combat the disease have been ineffective live attenuated virus vaccines. The goal of this project was to produce a plant-made subunit vaccine based off the membrane protein of the virus. This is the first time this protein has been produced in plants. An elastin-like polypeptide fusion membrane protein accumulated up to 0.8 mg/g of fresh leaf weight when transiently expressed in *Nicotiana benthamiana*. Virus-like particles were also produced for the first time for PEDv, and were able to form with just the membrane protein, or from co-expression of the membrane and envelope protein. This adds to the limited body evidence that the membrane protein is the only necessary component to make coronavirus-like particles, and represents the first time coronavirus-like particles have been made in plants.

Keywords

Porcine epidemic diarrhea virus, recombinant protein, subunit vaccine, membrane protein, envelope protein, coronavirus, virus-like particle, plant biotechnology, *Nicotiana benthamiana*

Acknowledgments

Thank you for teaching me…

First, my supervisor, Dr. Rima Menassa: …that good things can come when you believe in people. It was a genuine pleasure being your student.

My co-supervisor, Dr. Susanne Kohalmi: …that guidance paired with a smile goes a long way.

My advisors, Dr. Aiming Wang and Dr. Rob Cumming: …to be critical always, and hold myself to the highest standards.

Dr. Justin Renaud: …mass spectrometry, and for his assistance with mass spectrometry analysis.

Karen Nygard and Dr. Richard Gardiner: …how to perform TEM analysis, and use of the Biotron Experimental Climate Change Research Centre facilities.

Dr. Norm Hüner and Beth Szyszka: …how to use an ultracentrifuge, and allowing me to use theirs.

Alex Molnar: …how to make a beautiful figure, and for his assistance in figure preparation.

Angelo and Hong: …for everything I know about labwork. Truly.

Reza: …that hard work can be paired with a little goofiness.

The Menassa lab – Adam, Coby, Igor, Israel, Jacqueline, Jenny, Kira, Livia, Sean: …that good friends make work a pleasure.

Emily: …that some people are crazy enough to support me at all hours of the day.

Everybody in the department, and all my friends at AgCan: …the power of a community.

And most of all, thank you to my family and close friends for teaching me that good things take time, patience, and little bit of love and support.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

3CL	3C-like proteinase
a.a.	Amino acid
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride
AGC	Automatic gain control
ANOVA	Analysis of variance
APC	Antigen presenting cell
Apo	Apoplast
ATP	Adenosine triphosphate
Bp	Base pair
BSA	Bovine serum albumin
CaMV	Cauliflower mosaic virus
CBD	Cellulose binding domain
CCoV	Canine coronavirus
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CHO	Chinese hamster ovary
COE	CO-26K equivalent
CoVLP	Coronavirus-like particle
CTAB	Cetrimonium bromide
DIVA	Distinguish infected from vaccinated animals

DMSO	Dimethyl sulfoxide
DSB	Detergent screen buffer
DTT	Dithiothreitol
E	Envelope protein
ECL	Enhanced chemoluminescence
EDTA	Ethylenediaminetetraccetic acid
ELP	Elastin-like polypeptide
ER	Endoplasmic reticulum
ERGIC	Endoplasmic reticulum intermediate compartment
FCoV	Feline coronavirus
FEB	Final extraction buffer
GOI	Gene of interest
GST	Glutathione S-transferase
HCoV	Human coronavirus
HEK293	Human embryonic kidney 293 cells
HFBI	Hydrophobin
His	Histidine
HRP	Horseradish peroxidase
IBV	Avian infectious bronchitis virus
IFN-α	Alpha interferon

IgA	Immunoglobulin A
IgG	Immunoglobulin G
IT	Injection time
Kb	Kilobase
kDa	Kilodalton
KDEL	ER retrieval signal
L	Linker
LB	Lysogeny broth
M	Membrane protein
MES	2-(N-morpholino)ethanesulfonic acid
MHC	Major histocompatibility complex
MHV	Mouse hepatitis virus
N	Nucleocapsid protein
Nos	Nopaline synthase
OD	Optical density
PAR	Photosynthetically active radiation
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEB	Protein extraction buffer
PED	Porcine epidemic diarrhea

PEDv	Porcine epidemic diarrhea virus
PEG	Poly(ethyleneglycol)
pI	Isoelectric point
PLP	Papain-like proteinase
PMSF	Phenylmethanesulfonylfluoride
Pr1b	Pathogenesis-related signal peptide 1b of tobacco
PTGS	Post-translational gene silencing
PTM	Posttranslational modification
PVDF	Polyvinylidene difluoride
PVPP	Polyvinylpolypyrrolidone
S	Spike protein
SARS-CoV	Severe acute respiratory syndrome-associated coronavirus
SDS	Sodium dodecyl sulfate
TBS	Tris-buffer saline
TBS-T	Tris-buffer saline-Tween-20
TCoV	Turkey coronavirus
tCUP	Tobacco cryptic upstream promoter
TEM	Transmission electron microscopy
TEV	Tobacco etch virus
TGEV	Transmissible gastroenteritis coronavirus

TMV	Tobacco mosaic virus
TSP	Total soluble protein
UPR	Unfolded protein response
UTR	Untranslated region
VLP	Virus-like particle
v/w	volume per weight
w/v	weight per volume
w/w	weight by weight

1 Introduction

Porcine epidemic diarrhea virus (PEDv) is a coronavirus that causes porcine epidemic diarrhea (PED), a disease which affects pigs, and in particular, newly born piglets. PEDv has struck around the world, with the first known outbreaks occurring in Europe in the 1970s. PEDv later spread to Asia, where it was first detected in Japan in 1982, and then in South Korea, China, and Thailand (Chen et al., 2014). Since October 2010 China has seen a severe outbreak of PEDv, resulting in high porcine mortality rates and economic losses (Sun et al., 2012). PEDv was first detected in the United States (U.S.) in May of 2013 (Stevenson et al., 2013) and in Canada in January of 2014 (Ojkic et al., 2015). The serotypes identified in North America are most closely related to a recently emerged Chinese serotype (Huang et al., 2013).

PEDv causes the destruction of villus enterocytes and atrophy of intestinal villi. The disease is 95% fatal for neonatal piglets in naïve unvaccinated herds (Stevenson et al., 2013). The death of millions of suckling piglets and diarrhea-derived weight loss in fattening pigs caused severe economic losses in the U.S. and Canada (Chen et al., 2014). An effective vaccine is needed against PEDv, but no vaccines are available in Canada, and only two vaccines with unknown efficacy have conditional approval in the U.S. (see section 1.1.3.2 below).

1.1 Vaccine Design

Vaccines can be live attenuated viruses, inactivated viruses, or subunit vaccines. While live attenuated viruses have been used commercially, they potentially can mutate and become pathogenic again. They allow for genome segment re-assortment between vaccine and field strains, resulting in potentially dangerous new strains (reviewed in Calvo-Pinilla et al., 2014). One Chinese field strain of PEDv is thought to have potentially evolved from a live attenuated vaccine (Chen et al., 2010). Attenuated strains may not replicate enough to induce protective immunity at viral infection site (Saif, 1993). Inactivated viruses raise concerns over reliability of inactivation, and have a high cost of production (Calvo-Pinilla et al., 2014). With either live attenuated viruses or

inactivated viruses, serological testing to “distinguish infected from vaccinated animals” (DIVA) assays are not reliable. Subunit vaccines present a stable and safe alternative, as they are synthesized proteins that are not directly derived from live viruses. As isolated molecules they also lack genetic material avoiding any possibility for virulence or mutation. Subunit vaccines for infectious bronchitis virus, a gammacoronavirus, were shown to induce higher immune responses than inactivated virus vaccines when both were injected (Liu et al., 2013).

1.1.1 Virus-Like Particles

Subunit vaccines, while safer than inactivated or live attenuated vaccines, do not always have high immunogenicity. Virus-like particles (VLPs) represent advancements in subunit vaccine development, showing higher immunogenicity. VLPs are structures made up of assembled viral proteins that resemble the morphology of their respective pathogen (reviewed by Kushnir et al., 2012). Like other subunit vaccines, they benefit from being non-replicative and non-infectious due to not having any genetic material. In addition, VLPs display viral protein epitopes in correct conformations, and as exogenous, particulate antigens, are processed and presented on antigen presenting cells (APC) by MHC class I or II molecules. This means that they can effectively stimulate humoral and cellular immune responses as the native viruses would, and do not require the use of adjuvants (reviewed by Grgacic and Anderson, 2006). It is also important that VLPs can target dendritic cells, APCs involved in innate and adaptive immunity. Dendritic cell stimulation for cytokine production requires an intact virion. Here VLPs have an advantage over live attenuated and inactivated viruses, as both are shown to interfere with dendritic cell activation (reviewed by Grgacic and Anderson, 2006).

One vaccine success story involves hepatitis B virus-like particles. In the 1980s, existing products on the market were expensive, and limited in supply. VLPs based on hepatitis B small surface antigen were produced in yeast at a cheaper cost, which continued to decrease, allowing 110 countries to routinely immunize infants for hepatitis B today (reviewed by Rybicki, 2014). Almost all recombinant vaccines available for humans are VLP based, including GlaxoSmithKline’s Cervarix® (against human papilloma virus)

(GlaxoSmithKline, 2014), and Merck and Co., Inc's Recombivax HB® (against hepatitis B virus) (Merck, 2011b) and Gardasil® (against human papilloma virus) (Merck, 2011a).

1.1.2 Platforms for Recombinant Protein Expression

To produce subunit vaccines, a suitable platform must be chosen for recombinant protein production. While bacteria are commonly used to produce recombinant proteins, use of bacteria is not preferred due to differences with eukaryotes in their protein-folding machinery (Jacob et al., 2007), absence of a mammalian-like posttranslational modifications (PTM) such as glycosylation, and risk of bacterial endotoxin contamination of the final product (reviewed by Kushnir et al., 2012). Bacteria are used to produce non-enveloped VLPs, but not enveloped VLPs (reviewed by Kushnir et al., 2012). Mammalian cell lines such as Chinese hamster ovary (CHO) and human embryonic kidney 293 (HEK293) cells are currently the gold standard for biopharmaceutical production, but are expensive and bear the risk of harboring mammalian pathogens (Fischer et al., 2012). Research has shown that plants are also an efficient platform for the production of recombinant proteins (reviewed by Rybicki, 2014).

1.1.2.1 Plants as Bioreactors

Plants have numerous advantages as a platform for recombinant protein production. This platform is easily scalable, and safe from mammalian and bacterial pathogens (Menassa et al., 2012). Plants have the capability to fold and glycosylate complex proteins and can be grown in greenhouses using current farming techniques (Fischer et al., 2012; Menassa et al., 2012). Plants allow for an easy delivery method for vaccines, as plants can be fed without processing or extracting the protein. It is believed that plant components, including the plant cell matrix, can act as adjuvants, stimulating antigen-specific and non-specific immune responses (Bae et al., 2003). The plant cell wall also protects the antigen from degradation in the gastro-intestinal tract. Proteins are then released in the gut lumen, and can cross the intestinal epithelium with the use of tags that utilize endocytic subpathways (reviewed by Kwon and Daniell, 2016). This property is useful for a PEDv vaccine, as the epithelium is the major area of disease. In addition, recombinant proteins are stable in lyophilized plant tissue stored at room temperature, thus shrinking the

storage capacity needed, and simultaneously killing most microbes (Lakshmi et al., 2013). The ability to orally feed lyophilized plant tissue means reduced costs of purification, refrigeration and transport, allowing increased access in developing countries hit by PEDv (Lakshmi et al., 2013).

One plant bioreactor that has been widely studied is *Nicotiana benthamiana*, as it is amenable to transient expression using *Agrobacterium tumefaciens* (Conley et al., 2011). Transient expression does not require the transgene to integrate into the nuclear genome, and allows for quick production of recombinant proteins, with no generation of genetically modified plants. This is an advantage that regulatory bodies look favourably upon, due to reduced risk for transgene dissemination via pollen or seed (reviewed by Rymerson et al., 2002). *N. benthamiana* is also a good platform for transient recombinant protein production due to its small size and quick growth (Conley et al., 2011). *N. benthamiana* is also advantageous from a regulatory standpoint as it is not a food or feed plant, reducing the risk of contaminating the food chain (reviewed by Rymerson et al., 2002). One highly publicized story of this platform has been with the production of monoclonal antibodies against Ebola. Three Ebola-specific mouse-human chimeric monoclonal antibodies were transiently produced in *N. benthamiana* using the magnICON tobacco mosaic virus (TMV)-based viral vectors (Gleba et al., 2005). Plant-produced antibodies were three times as potent as those produced in CHO cells. These antibodies were used in a cocktail of anti-Ebola monoclonal antibodies called ZMapp, which were shown to rescue 100% of rhesus macaques when administered five days post live virus challenge (Qiu et al., 2014). ZMapp was an important player in the recent Ebola outbreak, used to save the lives of two U.S. healthcare workers that had contracted Ebola (reviewed by Rybicki, 2014).

1.1.2.2 Plant-Made VLPs

Enveloped VLPs acquire lipid envelopes, characteristically with embedded immunogenic glycoproteins, when they bud off. Therefore, they are only produced in eukaryotic systems (Lua et al., 2014). Plants are capable of producing both enveloped and non-enveloped VLPs. Enveloped plasma membrane-derived VLPs for influenza were produced through transient expression in *N. benthamiana*, and these VLPs were able to

confer complete protection to mice against a lethal challenge, even though mice were challenged with a genetically dissimilar strain. It was also found that immunization with these VLPs induced antibody responses four- to six-fold higher than when compared to immunization with fifty times more of flu antigen not in VLPs (D'Aoust et al., 2008).

Complex non-enveloped VLPs have also been produced, including for bluetongue disease, which causes lameness and mortality in ruminants. VLPs for bluetongue virus (BTV) required simultaneous expression of four different proteins in varying amounts. *N. benthamiana* was used for transient expression using the pEAQ expression vector system, and the VLPs were able to elicit antibody response, and provide protective immunity (Thuenemann et al., 2013). Medicago Inc., a Canadian company, is a success story of plant-made VLPs, using a fully automated greenhouse to efficiently agroinfiltrate large numbers of *N. benthamiana* plants. Medicago Inc. is on track to produce ten million doses of VLP pandemic influenza vaccine per month (Medicago, 2016).

VLP purification can be very challenging, and despite advances geared towards purifying large scale amounts, technical challenges persist in ensuring no contaminating DNA or proteins are trapped in the VLPs and that VLPs are homogenous (reviewed by Lua et al., 2014). Additionally, although VLPs are self-adjuvanting, currently all licensed VLP-based vaccines are formulated with aluminum salts, an adjuvant (reviewed by Lua et al., 2014). As plant components are believed to have adjuvanting ability (Kolotilin et al., 2014), plant cell-encapsulated VLPs may not need additional adjuvants. VLP structure and immunogenicity is preserved when plant tissue is lyophilized, and shown to be stable for at least one year (Czyz et al., 2014). Lyophilized tissue as a delivery system then may be cost-effective, foregoing both VLP purification, and potentially the use of additional adjuvants, as well as requiring less storage capacity as previously discussed (1.1.2.1).

1.1.3 PEDv Vaccine Design

With the economic losses incurred around the world due to PEDv, the importance of an effective vaccine to prevent PED has gained importance. When piglets are first born, their immune system is not fully developed and they cannot be vaccinated. However, when sows are vaccinated IgA immunocytes secrete IgA antibodies into the colostrum

and milk, passing on passive lactogenic immunity to suckling pigs (Bae et al., 2003). It has been suggested that aside from the fecal-oral route of transmission, piglets may also be infected through vertical transmission of PEDv from sow milk (Sun et al., 2012). Vaccinating sows would both prevent vertical transmission of PEDv, and also provide passive immunity. These characteristics need to be considered when designing a PEDv vaccine.

1.1.3.1 PEDv Structure

To produce an effective subunit vaccine for PEDv, the coronavirus structure must be understood. The Coronavirinae subfamily consists of three genera: alphacoronavirus, betacoronavirus, and gammacoronavirus. PEDv is an enveloped alphacoronavirus encoded by a 28 kilobase single-stranded, positive-sense RNA genome (Song and Park, 2012). Coronaviruses have the largest known RNA genomes of all viruses (King, 2011). The genome has a 5' cap, a 3' polyadenylated tail and seven open reading frames (ORFs), which code for three non-structural proteins (ORF 1a, ORF1b, and ORF3), and four structural proteins (spike (S), envelope (E), membrane (M) and nucleocapsid (N)) (Figures 1, 2, Figure 2 designed based on reports of protein abundance below) (Song and Park, 2012). The two overlapping open reading frames, ORF1a and ORF1b code for two polyproteins. These are processed by three virus-encoded proteases, a 3C-like proteinase (3CLpro) and two papain-like proteinases (PLP) which results in 16 non-structural proteins required for genome replication and mRNA transcription (John et al., 2016; reviewed by Prentice et al., 2004). The accessory protein ORF3 is a potassium ion channel, but its role is not well defined (Wang et al., 2012). Of the four structural proteins, S and M are the proteins considered most important for antigenicity. Reports on transmissible gastroenteritis coronavirus (TGEV), another alphacoronavirus, indicate that S and E are only present in the virion in small quantities, with E estimated to occur 20 times in a virion (Godet et al., 1992), while N and M occur in higher numbers, at a ratio of 1N:3M (King, 2011). A recent study examining mouse hepatitis virus (MHV), a betacoronavirus, severe acute respiratory syndrome-coronavirus (SARS-CoV), a betacoronavirus, and feline coronavirus (FCoV), an alphacoronavirus, has determined

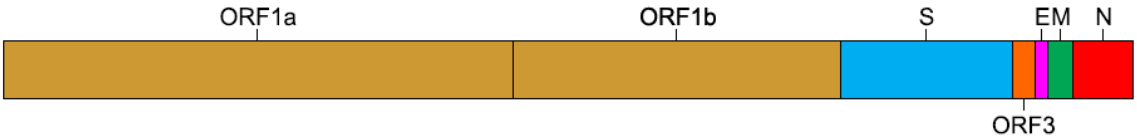


Figure 1. A schematic of the genome of PEDv.

The first two ORFs from the 5’ end of the genome cover two thirds of the genome. The first codes for polyprotein 1a, and the second for polyprotein 1b. Towards the 3’ end, the S protein is then coded for, followed by ORF3, E, M, and finally the N protein.

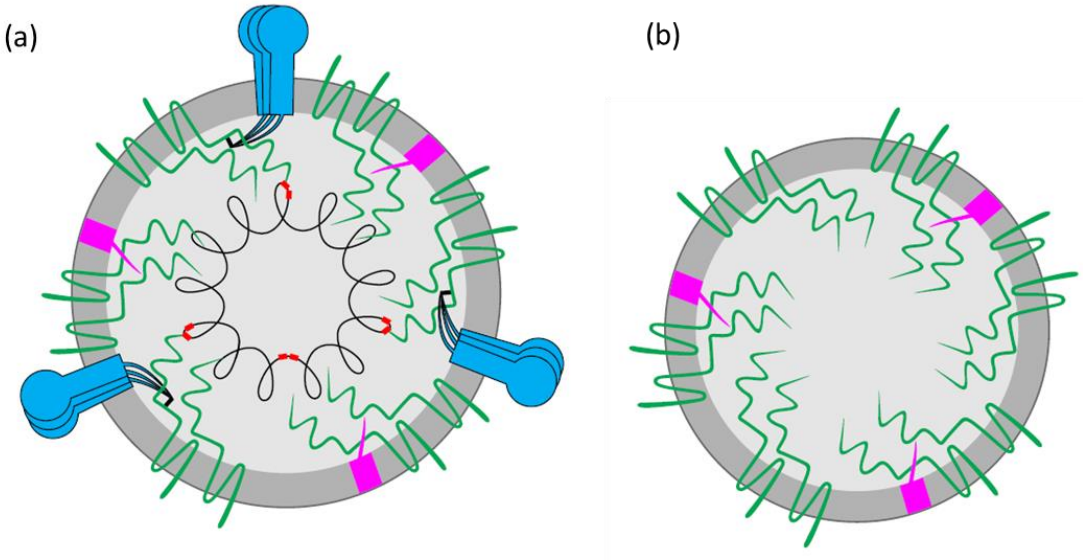


Figure 2. A schematic of an assembled PEDv virus, and a PEDv VLP.

(a) The nucleocapsid protein (N, red squares) in the centre forms a ribonucleoprotein complex with viral RNA (black line). The envelope protein (E, pink) is embedded in the membrane (darker grey) as is the membrane protein (M, green), which is the most abundant envelope component, encompassing an amino-terminal domain outside the virus, three transmembrane segments, and a longer carboxy-terminal domain inside the virus. The spike protein (S, blue) also embeds in the membrane, and forms surface projections, or ‘spikes’. (b) A theoretical PEDv VLP is smaller in diameter, and forms with the M and E proteins.

coronaviruses have approximately 1100 M dimers, 90 S trimers, and N proteins in a ratio from 3M:1N to 1M:1N (Neuman et al., 2011). Thus, M is the most abundant structural protein displayed at the viral surface.

1.1.3.2 Existing PEDv Commercial Vaccines

Much of the research done in developing PEDv vaccines has occurred in Asia, where outbreaks have been most severe, but none of the produced vaccines are completely effective on Asian PEDv strains (Song and Park, 2012). Available Asian vaccines are based on strains that are genetically different from those sequenced in the U.S. (Huang et al., 2013).

Two PEDv vaccines were given conditional licenses in the U.S. by the United States Department of Agriculture (USDA). The first was a vaccine originally produced by Harrisvaccines, Inc. This vaccine is based on their SirraVax^{SM} RNA platform. Using this platform, part of the RNA genome of a Venezuelan equine encephalitis alphavirus is replaced with a gene for PEDv S protein. The resulting RNA particle looks like the alphavirus, but carries the S gene. After injection, the pig's dendritic cells produce the S protein and an immune response is launched against the produced protein (Harrisvaccines, 2015)**.** Merck Animal Health acquired Harrisvaccines in 2015 (Merck Animal Health, 2015), but the product is still sold under a conditional license in the U.S., where safety and field trials are ongoing. The second is an inactivated virus particle vaccine produced by Zoetis, Inc. (Zoetis, 2014). Efficacy and potency studies are still in progress for the Zoetis vaccine, and duration of immunity has not been evaluated. It must be refrigerated, and used all at once when opened (Zoetis, 2016). Neither vaccine is sold in Canada – veterinarians have to apply for an import permit from the Canadian Food Inspection Agency to bring batches of the drugs into the country (Merck Animal Health, Personal Communication 2016; Zoetis, Personal Communication, 2016). There is one Canadian research organization that is testing a prototype PEDv vaccine (Brusky, 2016)

None of the PEDv vaccines used commercially, at least in the U.S. and Canada, are made *in planta*. The ability to vaccinate orally through feeding plant tissue is important as Song et al. (2007) demonstrated that oral vaccination was more effective than injection for their

PEDv vaccine. When comparing oral to intramuscular administration of their attenuated virus vaccine, Song et al. found that more IgA's were produced by orally vaccinated pigs, and that the mortality rate for this group was 13% in comparison to 60% for the intramuscular group.

1.1.3.3 S Protein

To date, the S protein has been the primary focus of subunit vaccine design. This is due to its antigenicity and the role it plays in viral entry, as it regulates interactions with host cell receptor proteins (reviewed by Song and Park, 2012). An epitope capable of inducing virus-neutralizing antibodies, called CO-26K equivalent (COE), was identified through sequence homology with transmissible gastroenteritis coronavirus (TGEV) (Chang et al., 2002) at amino acids (a.a.) 1495-1914. Another motif (1368-1374 a.a.; Cruz et al., 2008) , as well as the epitope region (636-789 a.a.) have been found to induce the production of neutralizing antibodies (Sun et al., 2006).

However, the use of S is problematic, as it was shown to be prone to mutations through serial passages (Sato et al., 2011), and has high genetic variability among different PEDv strains. In fact, this variability is utilized to study genetic relatedness of different PEDv strains (Chen et al., 2014). This genetic variation, particularly apparent in attenuated strains, and importantly, in the epitope, may impact the efficacy of live attenuated virus vaccines (Sun et al., 2014). Strains have even been found where 582 nucleotides were deleted from the S gene (Masuda et al., 2015). The failure of vaccines on the market may be attributed to differences in the sequence of S between the delivered vaccine and infectious strain (Sun et al., 2012).

The majority of recombinant production of the spike protein has been performed in plants. Recombinant PEDv protein production in plants has only focused on producing the S-COE epitope, and is summarized in Table 1. Plant made S-COE is able to elicit antibodies. Bae et al. (2003) produced transgenic tobacco plants expressing the S-COE at levels of 10 mg/kg of wet weight. They demonstrated that feeding ground lyophilized transgenic tobacco leaves with S antigen suspended in phosphate buffered saline (PBS)

Table 1. Production of S-COE in plants[1].

Plant Host	Transient or Transgenic	Fusions	Yield	Promoter and/or Enhancer Used	Reference
Nicotiana tabacum	Transgenic	-	10 mg/kg of protein per fresh weight[2]	2x35S, TOL	(Bae et al., 2003)[3]
N. tabacum	Transient	-	5% TSP	TMV RNA	(Kang et al., 2004)
No-nicotine *N. tabacum*	Transgenic	-	2.1% TSP	2x35S, TOL	(Kang et al., 2005a)
N. tabacum	Transgenic	-	0.1% TSP	2x35S, TOL	(Kang et al., 2005b)
Solanum tuberosum	Transgenic	-	0.1% TSP	2x35S, TOL	(Kim et al., 2005)
Lemna minor	Transgenic	-	Not reported	35S	(Ko et al., 2011)
Ipomoea batatas	Transgenic	-	Not reported	35S	(Yang et al., 2005)
Zea mays seed	Transgenic	-	0.122% TSP	2x35S, maize intron Hsp70	(Kun et al., 2014)*
Daucus carota	Transgenic	-	Not reported	2x35S, TOL	(Kim et al., 2003)
Lactuca sativa	Transgenic	LTB	0.048% TSP	Ubiquitin promoter	(Huy et al., 2009)
Oryza sativa endosperm	Transgenic	LTB	1.3% TSP	HMW-Bx17-p, Act1-i	(Oszvald et al., 2007)
Oryza sativa endosperm	Transgenic	LTB	1.9% TSP	HMW-Bx17-p, Act1-i	(Tamás, 2010)
N. tabacum	Transgenic	LTB	1.6% TSP	Ubiquitin promoter	(Kang et al., 2006)
Oryza sativa calli	Transgenic	Co1	0.083% TSP	RAmy3D	(Huy et al., 2012)[3]

[1]2x, duplicated; 35S, cauliflower mosaic virus (CaMV) 35S promoter; Act1-I, rice actin first intron; Co1, M cell-targeting ligand; HMW-Bx17-p, wheat high molecular weight glutenin subunit Bx17 endosperm-specific promoter; LTB, heat-labile enterotoxin B subunit of *Escherichia. coli*; RAmy3D, rice α-amylase 3D promoter; TOL, TMV Omega-prime leader, containing transcriptional and translational enhancer from the coat protein gene of TMV; TSP, total soluble protein; all yield values are highest levels reported.

[2]Total soluble protein levels were not reported in this study.

[3]Study also showed antibody production against protein.

was effective in inducing systemic and mucosal immune responses in mice. Serum from immunized mice inhibited PEDv plaque formation by 49.7% in comparison to controls.

1.1.3.4 M Protein

1.1.3.4.1 Structure and Conservation

M is an N-glycosylated transmembrane protein, and is the most abundant component of the viral envelope (Neuman et al., 2011; Utiger et al., 1995). M is predicted to have three transmembrane segments, with two flanking domains one short, and one long (The UniProt Consortium, 2015). It was experimentally shown that the shorter domain lies outside the virion on the amino terminus, and a longer carboxyl tail is found inside the virion (Utiger et al., 1995).

In contrast to S, M is more conserved (Chen et al., 2014; Sato et al., 2011). PEDv was propagated through serial passages in Vero cells, and gene sequences were analyzed at the 34^{th}, 61^{st}, and 100^{th} passage. Vero cells are a cell line derived from kidney epithelial cells from African green monkey of the *Chlorocebus* genus, and are commonly used for viral propogation. The *S* gene had 6, 10, and 18 nucleotide changes at each respective passage, causing 5, 9, and 13 amino acid substitutions respectively, in comparison to the original sequence. The amino acid sequence of the *M* gene was unaffected until the 61^{st} passage, after which there was one mutation that changed the amino acid sequence (Sato et al., 2011). To ensure that a vaccine candidate is effective against multiple strains, it is better to use a protein that is not prone to variation or adaption. A comparison of M and S protein sequences of PEDv strains from 16 different pig farms in China, showed that M proteins shared 97.8%-100% amino acid identity, while in comparison S proteins showed 94.6%-100% amino acid identity, and importantly showed amino acid substitutions in the neutralizing epitope region (Sun et al., 2014).

1.1.3.4.2 Antigenicity of M Protein

In addition to showing strong conservation over serial passages, the coronavirus M protein has been shown to induce both humoral and cellular immune responses. The M protein of SARS-CoV, a betacoronavirus, has two cytotoxic T-cell epitopes in the second

and third transmembrane domain (Liu et al., 2010). The M protein of mouse hepatitis virus (MHV), also a betacoronavirus, has CD4+ T-cell epitopes on its carboxyl tail (Xue et al., 1995).

The M protein of SARS-CoV has B-cell epitopes both on the N and C-termini (He et al., 2005). The C-terminus of avian infectious bronchitis coronavirus (IBV), a gammacoronavirus, M protein is also found to have a B-cell epitope (Xing et al., 2009). Further, monoclonal antibodies against M of TGEV have virus-neutralizing activity in the presence of complement, the part of the immune system that enhances antibody ability to clear pathogens (Woods et al., 1988). Recombinant M protein of SARS-CoV also induces neutralizing antibodies in the presence of complement. The neutralizing capacity of M protein from SARS-CoV was higher or the same when compared to the neutralizing capacity of 8 individual S protein fragments (Pang et al., 2004).

Coronavirus M protein also is able to stimulate interferon-α production, which in turn increases major histocompatibility complex (MHC) class I expression, and thus antigen presentation in cells (Laude et al., 1992).

The B-cell epitope of IBV M protein was experimentally determined to be amino acids ^{199}FATFVYAK^{206} on the C-terminal tail (Xing et al., 2009). The homologous region on the PEDv M protein C-terminus contains the B-cell epitope, amino acids ^{195}WAFYVR^{200}, and this epitope is conserved in ten different PEDv strains analyzed. Nine different coronaviruses (PEDv, TGEV, canine coronavirus (CCoV), feline coronavirus (FCoV), human coronavirus 229E (HCoV 229E), HCoV OC43, MHV, SARS-CoV, turkey coronavirus (TCoV), IBV) were compared and three of the five amino acids A^{196}, Y^{198} and V^{199} are conserved among all nine coronaviruses (Zhang et al., 2012). HCoV 229E M protein is of interest, as it is the most closely related to PEDv M, with 57% sequence identity (Duarte et al., 1994). This further shows the conserved nature of M as useful for vaccine design. Future studies may find that PEDv M also contains the B-cell and T-cell epitopes found on M of other coronaviruses, especially as monoclonal antibodies were shown to bind to the N terminus of PEDv M (Utiger et al., 1995).

1.1.3.4.3 Role in Viral Assembly

Apart from its antigenicity, M is also a useful candidate for vaccine design due to its key role in viral assembly, allowing it to be used to produce VLPs. Coronavirus assembly and budding takes places in the endoplasmic reticulum-Golgi intermediate compartment (ERGIC). M-M interactions drive the formation of the envelope, and M-M associations are thought to take place in pre-Golgi compartments (de Haan et al., 2000), with M having been shown to be localized to the *trans* Golgi, *cis*-Golgi, or trans Golgi network, depending on the coronavirus (reviewed by Ujike and Taguchi, 2015). ER-retained M is able to recycle from the Golgi complex to early compartments, potentially to allow M proteins that did not assemble into the virion another chance to do so through retrograde transport (de Haan et al., 2000). It is thought that M forms a lattice-like circular matrix within the viral envelope with open positions, which S and E proteins can fill through specific interactions with M when in the ER and ERGIC membranes (de Haan et al., 2000; Lim and Liu, 2001). In this process M efficiently excludes foreign proteins (Nguyen and Hogue, 1997). The N protein then attaches to the M matrix (de Haan et al., 2000). M will only bind to N when N is complexed with RNA (Narayanan et al., 2000). It is hypothesized that E may be involved in membrane scission to have the virion bud off and impact the rearrangement of the organelles of the ER-Golgi complex for virion trafficking, but this requires further research, particularly as M-only VLPs have been produced (reviewed by Ruch and Machamer, 2012).

Changes in any part of the M protein, including the amino terminus, transmembrane spanning regions, amphipathic domain and carboxyl tail have effects on M assembling into enveloped particles (de Haan et al., 1998). Mutations of the M protein are better tolerated when M is in the presence of all structural proteins, rather than when M and E VLPs are being produced. This is likely because of the stabilization provided by interaction with all proteins, particularly the nucleocapsid (de Haan et al., 1998). However, the carboxyl terminus is critical for envelope assembly, with truncation of just the single terminal residue preventing virion or VLP formation (de Haan et al., 1998). The carboxyl tail is involved in Golgi retention, and contains a conserved domain important for viral envelope formation, and in mediating M-M interactions during viral

assembly, though other domains, especially the transmembrane domains, are also important (Arndt et al., 2010; de Haan et al., 2000) . The carboxyl tail may also be important in M-N interactions, allowing N to potentially stabilize M complexes (Arndt et al., 2010).

1.1.3.4.4 M is Crucial for VLP Formation

The minimum requirements for coronavirus-like particle formation have been controversial, and vary depending on the virus. It was previously thought that M and E were the minimal requirements to form VLPs, as shown with MHV, where it was also observed that glycosylation of M makes no difference on viral assembly (de Haan et al., 1998). However, conflicting studies have emerged for SARS-CoV. An initial study found M and E were sufficient for VLP formation (Ho et al., 2004), aligning with previous coronavirus research, but a subsequent study instead found that M and N were both sufficient and necessary to form VLPs (Huang et al., 2004). Another study has shown that SARS CoVLPs can be produced with only M, or M and N (Tseng et al., 2010). Further, VLPs for IBV were produced by co-expressing M and S (Liu et al., 2013). The requirements to produce PEDv VLPs are not known, as they have never been reported in the literature.

1.1.3.4.5 Previous Recombinant Expression

As a membrane protein, M can be difficult to express. Some groups have attempted to express full length PEDv M in *E. coli* only to resort to expressing different fragments of the protein (Zhang et al., 2012), or just expressing a truncated version of the protein (Shenyang et al., 2007). Recombinant expression of M has mostly utilized prokaryotic platforms, and M is almost always expressed as a fusion protein. Two groups have produced M in eukaryotic platforms (Ren et al., 2012; Utiger et al., 1995). While some of the produced M proteins have induced antibodies, none of these antibodies have been tested for virus neutralizing activity. Table 2 lists all known studies that have produced recombinant PEDv M protein.

Table 2. List of all work on recombinant expression of PEDv M protein.

Full Length?	Modifications	Host Species[1]	Induces antibodies?	Reference
Carboxyl tail	N-terminal Glutathione s-transferase (GST) tag, C-terminal His_6 tag	*E. coli*	Yes	(Zhang et al., 2012)
N terminus	N-terminal GST	*E. coli*	Yes	(Shenyang et al., 2007)
Signal peptide deletions	-	*E. coli*	Yes	(Ren et al., 2011)
Full length	N-terminal GST	*E. coli*	N/A	(Fan et al., 2015)
Full length	N-terminal maltose binding protein (MBP), C-terminal cellulose-binding domain (CBD)	*E. coli*	N/A	(Zhang et al., 2011)
Full length	None, or fusion with N	*Salmonella choleraesuis*, Vero cells	Yes: *S. choleraesuis* as a DNA vaccine, with M-N fusion	(Ren et al., 2012)
Full length	-	Sf9 cells[a], Vero cells[b]	Yes	(Utiger et al., 1995)

[1]Sf9 cells, insect cells from the ovarian tissue of *Spodoptera frugiperda*; Vero cells, kidney epithelial cells from an African green monkey of the *Chlorocebus* genus.

No reports to date show expression of PEDv M, or any full length PEDv protein, in plants. Three reasons may be behind this: first, M can be difficult to express, second, membrane proteins in general present difficulties (discussed below), and third, that S COE has been the focus of vaccine design, with the B-cell epitope of M only having been identified in 2012 (Zhang et al., 2012).

1.2 Membrane Protein Expression and Extraction

Recombinant production of eukaryotic membrane proteins in heterologous expression systems can be a difficult task (Grisshammer, 2006). Membrane protein synthesis most commonly requires the polypeptide to be targeted to a protein-conducting channel (translocon) in the ER, to then be released into the ER membrane where it folds into its correct conformation. Only proteins that are correctly folded pass into the Golgi, otherwise proteins are degraded (reviewed by Grisshammer, 2006). The detection of incorrectly folded proteins triggers the unfolded protein response (UPR), which either targets these proteins for degradation, or causes cell apoptosis, in both mammalian and plant cells (Iwata and Koizumi, 2005). In addition, the use of strong promoters resulting in a large number of mRNAs which are translated on ribosomes and inserted into the ER may overwhelm the ER protein folding machinery, leading to increased misfolding and UPR. This may result in reduced accumulation of membrane proteins. To confound matters further, not all membrane proteins elicit UPR (Grisshammer, 2006). Why some membrane proteins insert and fold better than others is not understood. The ER translocon recognizes transmembrane helices, inserting them into the membrane rather than allowing the helices to pass into the aqueous compartment, and arranging the orientation of the newly synthesized protein. This topology outcome is driven by synthesis rate as well as the length of the hydrophobic segment, and the positioning of key charged residues in the translocon, which can inform orientation of the protein. However, these residues are not conserved in all translocons, leading to difficulties and trial and error in heterologous membrane protein expression (reviewed by Bowie, 2005). As some proteins may require post-translational modifications that only some eukaryotic cell lines have the machinery for, it can be necessary to test various expression systems for each protein (reviewed by Junge et al., 2008).

Aside from expression, membrane proteins are difficult to study due to their hydrophobic surfaces and unstable nature. This is problematic as 20-30% of most proteomes (Krogh et al., 2001), and 60% of drug targets (Overington et al., 2006) are membrane proteins (Carpenter et al., 2008). Despite this, of the 121, 654 proteins whose structure has been identified and deposited in the Protein Data Bank, only 2, 829 are identified as membrane proteins (Kozma et al., 2013). A difficult component of studying membrane proteins is their extraction. Membrane protein extraction is reliant on the use of detergents, where often a series of detergents needs to be tested to see which extracts the highest amount of a particular protein (Carpenter et al., 2008).

1.2.1 Importance of Detergents

1.2.1.1 Detergent Properties

Membrane proteins are by definition, at least partially embedded in membranes. Biological membranes are typically amphiphilic lipid bilayers (Arnold and Linke, 2008). Detergents have properties to mimic the lipid environment which helps to solubilize membranes, allowing for stable extraction of membrane proteins, or membrane protein complexes. Detergents are amphiphilic, with a hydrophilic headgroup and a hydrophobic section, normally made up of an extended hydrocarbon chain (Arnold and Linke, 2008). Detergents allow membrane protein extraction by partitioning into lipid bilayers, where cooperative detergent-detergent interactions destabilize the bilayer to give mixed lipid-detergent fragments. More detergent then leads to the dissolution of the bilayer and protein solubilisation (reviewed by le Maire et al., 2000). The final result is that the transmembrane domains of the proteins are covered by a torus of detergent molecules (reviewed by Garavito and Ferguson-Miller, 2001).

1.2.1.2 Detergent Types

Four classifications of detergents exist, based on the electrical charge of the hydrophilic group: non-ionic, zwitterionic, anionic, and cationic. The properties of the detergents within these categories depends on the structure and size of the polar headgroup, and length of the attached hydrocarbon chain (Arnold and Linke, 2008). Detergents have the ability to denature proteins. Ionic detergents in particular are harsh on proteins, as they

have the ability to disrupt both inter- and intra-molecular protein interactions. Thus ionic detergents are not typically used for procedures requiring protein stability. Non-ionic detergents are better suited for keeping proteins active and not affecting structural features (le Maire et al., 2000). Some membrane proteins are not active in pure detergents, therefore it can be undesirable to wash off most of the membrane lipids with too high of a detergent concentration. Solubilisation, and the concentration at which delipidation occurs, depends on detergent choice, the protein, and medium conditions in the buffer (le Maire et al., 2000).

There are multiple kinds of non-ionic detergents, which as a group are characterized by having an uncharged hydrophilic headgroup. One family is the *tert*-Octylphenol poly(ethyleneglycolether)$_n$ detergents, which includes the Triton, Nonidet, and Igepal detergents. These have a poly(ethyleneglycol) (PEG) headgroup, and a *tert*-octyl chain. Detergents in this family are mild, and rarely denature membrane proteins (Arnold and Linke, 2008). Detergents in this family typically have slight differences in the average size *n* and size distribution of their headgroups. Triton X-114 has been frequently used for phase separation due to its cloud point of 22ºC. Triton X-100 can also be used for phase separation, as its cloud point can be brought down with the addition of PEG, dextran, glycerol, NaCl or $(NH_4)_2SO_4$ (Arnold and Linke, 2008).

There are other groups of non-ionic detergents. One is polyoxyethylene sorbitan esters of fatty acids, commonly known as Tween. These are also mild detergents. Tween detergents are mixtures of complex polymers (Ayorinde et al., 2000). Another is polyoxyethylene glycol monoether detergents. These are named C_xE_y for their alkyl chain length and number of polyoxyethylene glycol units in the headgroup respectively (Arnold and Linke, 2008). These typically go by trade names, for example $C_{12}E_{23}$ goes by Brij 35.

Anionic detergents are those such as sodium dodecyl sulfate (SDS). These are harsh detergents with small headgroups that denature most proteins (Arnold and Linke, 2008). This detergent is primarily used for SDS-PAGE. This technique relies on full denaturation of proteins using SDS and heat followed by electrophoretic separation of proteins in a matrix such as polyacrylamide to assess molecular weight. Some membrane

proteins are resistant to SDS denaturation, and also have heat modifiability – where apparent molecular weight changes depending on whether the samples with SDS in them are heated before separation by SDS-PAGE. SDS is also known to precipitate at low temperatures (Arnold and Linke, 2008). Cationic detergents also typically denature proteins, though some proteins remain fully functional (Arnold and Linke, 2008). CTAB is an example of a cationic detergent.

1.3 Hypothesis and Objectives

I hypothesize that using *N. benthamiana* as a bioreactor, the full length M protein of PEDv will be transiently produced and extracted, and VLPs for PEDv can be formed either with M alone, or in conjunction with E, S, or N.

The specific objectives of this study were:

1. Design and identify which gene constructs will lead to the highest transient expression levels of PEDv M in *N. benthamiana*.

2. Transiently express gene constructs and determine which detergents are most efficient for extracting PEDv M.

3. Investigate the capacity of plants to produce virus-like particles with the M protein, or with M in combination with E, S, or N.

2 Materials and Methods

2.1 Gene Synthesis

Peptide sequences for the membrane, spike, nucleocapsid and envelope proteins were obtained from the National Center for Biotechnology Information (NCBI) database (accession number KF650373). This sequence is from a PEDv strain isolated in the U.S. that showed ≥99.7% nucleotide identity with nine other U.S. strains that have been sequenced (Chen et al., 2014). The *M*, *N*, and *E* nucleotide sequences were synthesized by BioBasic Inc. (Markham, Ontario), optimized for nuclear expression in *N. tabacum* with a C-terminal fusion to a *Strep*II tag (Schmidt et al., 1996) for protein purification. The gene sequence is flanked with *attL* sites for Gateway® recombination (Hartley et al., 2000) into transient expression vectors, or *attB* sites for Gateway® recombination into entry vectors. The *S* nucleotide sequence was synthesized by GenScript (Piscataway, U.S.A), optimized for nuclear expression in *N. tabacum*, flanked by *attL* sites and contained a C-terminal fusion to a *Strep*II tag.

2.2 Transient Assays

2.2.1 Composition of Cloning and Infiltration Reagents

The composition of the buffers used for cloning and agroinfiltration is detailed in Table 3.

2.2.2 Gateway Cloning®

Gateway® cloning technology (Invitrogen, Thermo Fisher Scientific, Waltham, U.S.A.) was used to recombine *M*, *N*, and *E* gene cassettes into pCaMGate expression vectors (Pereira et al., 2014). Gene constructs that were synthesized with flanking *attB* sites underwent a BP reaction to be cloned into the pDONR vector. This was followed by an LR reaction to introduce the inserts into various expression vectors. For the BP

Table 3. Recipes for cloning and infiltration reagents

Reagent	Recipe
Lysogeny broth (LB), Miller	10 g tryptone, 5 g yeast extract, 10 g NaCl in 1 L Milli-Q water
LB agarose	LB + 1.5% agar
LB/LB agarose with kanamycin	LB/LB agarose with50 μg/μl kanamycin
LB agarose with kanamycin and rifampicin	LB/LB agarose with50 μg/μl kanamycin and 10 μg/μl rifampicin
Infiltration medium	10 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) at pH 5.6, 100 μM acetosyringone, 50 μg/ml kanamycin, and 10 μg/ml rifampicin in LB
Gamborg's solution	3.2 g/L Gamborg's B5 with vitamins, 20 g/L sucrose, 10 mM MES (pH 5.6), 200 μM acetosyringone

reaction, 100 ng of the pUC57 vector containing the different genes of interest (GOI) was recombined with 300 ng of pDONR using 1 µl of BP Clonase® enzyme mix (Invitrogen) overnight at room temperature. After transformation into *E. coli*, (described in 2.2.4), and plasmid extraction (described in 2.2.4), the recombined plasmid underwent the LR reaction. 100 ng of pDONR with GOI were recombined into 300 ng of expression vector using 1 µl of LR Clonase® enzyme mix (Invitrogen) overnight at room temperature.

2.2.3 Golden Gate Cloning

Golden Gate cloning (Engler et al., 2008) was used to recombine the *S* gene construct into the elastin-like polypeptide-endoplasmic reticulum (ELP-ER) pCamGate expression vector (Pereira et al., 2014). In a microcentrifuge tube 200 ng of expression vector and pUC57 plasmid containing the GOI, 2 µl of CutSmart buffer (New England Biolabs, Whitby, Canada) supplemented with 10 mM adenosine triphosphate (ATP), 18 µl of Milli-Q water, and 1 µl of both *Bsa*I (New England Biolabs) and T4 DNA Ligase (New England Biolabs) were mixed. The reaction was incubated for 1 hour at 37 °C.

2.2.4 *E. coli* Transformation

pCaMgate expression vectors with GOIs were transformed in to *E. coli* XL1-Blue, using the Gene Pulser II system (Bio-Rad Laboratories Inc, Hercules, U.S.A.) 1.5 µl of BP or LR reaction, or 2 µl of Golden Gate reaction was added to 40 µl of *E. coli* which were then electroporated. Cells were then immediately diluted with 960 ml of LB, incubated for 1 hour at 37°C at 250 rpm (Innova® 42 Incubator, Eppendorf, Hamburg, Germany), and 20 and 200 µl aliquots were plated on LB agarose plates with kanamycin. Plates were incubated overnight at 37°C.

Colonies were screened using polymerase chain reaction (PCR) with gene-specific primers to select for colonies with the GOI. Colonies that were found to be positive were used to inoculate 5 mL of LB with kanamycin and grown at 37°C at 250 rpm overnight. Plasmid DNA was extracted using the QIAprep® Spin Miniprep Kit (Qiagen, Venlo, Netherlands) following manufacturer's protocols. Extracted plasmid DNA was screened using PCR.

2.2.5 *Agrobacterium tumefaciens* Transformation

Electro-competent *A. tumefaciens* EHA 105 cells were transformed like *E. coli*, except incubated at 28°C at 250 rpm for one hour, and plated on LB agarose plates with kanamycin and rifampicin, then incubated for two days at 28°C.

2.2.6 Transient Expression in *Nicotiana benthamiana*

Transformed *A. tumefaciens* colonies, with expression constructs with either GOI or p19, were used to inoculate 3 ml of LB with kanamycin and rifampicin. Cultures were incubated from 1PM to the following morning at 28°C at 250 rpm. These cultures were used to inoculate 50 ml of infiltration medium. A 1/4000 dilution of *A. tumefaciens* that contained the GOI expression vectors, and 1/2000 dilution of *A. tumefaciens* containing p19 was used to inoculate infiltration medium. The inoculated infiltration culture was incubated overnight at 28°C at 250 rpm until an optical density at 600 nm (OD_{600}) of 0.5-1.0 was reached, measured with a Nanodrop 2000c spectrophotometer (Thermo Fisher Scientific). Cultures were then centrifuged at 6000 x *g* for 10 minutes, and resuspended in Gamborg's solution to an OD_{600} of 1.0. Cultures were then incubated at room temperature with gentle agitation for at least 1 hour. *A. tumefaciens* cultures containing GOI expression constructs were mixed with an equal volume of *A. tumefaciens* cultures containing the p19 expression vector. This was sometimes then diluted with an equal volume of Gamborg's solution, giving an *A. tumefaciens* OD_{600} of either 0.5 or 0.34 for each culture. If multiple GOI were being co-expressed, they were all mixed in equal volumes without Gamborg's solution, such that the final OD_{600} of each culture was 0.34, 0.25, or 0.2, for co-expression of 2, 3, or 4 cultures with GOI expression constructs, respectively.

These cultures were used to infiltrate 5-7 week old *N. benthamiana* plants that were grown in a growth room with the following conditions: 16 hours of light/8 hours of dark, 21-22°C, 55% humidity, and receiving approximately 100 μmol/photons $m^{-2}s^{-1}$ of photosynthetically active radiation (PAR). If syringe infiltrations were performed, a 3 ml syringe was used to infiltrate the *A. tumefaciens* suspensions into the underside of the leaves, occasionally with the help of a needle to make a cut. Three different biological

replicates were used, with constructs having equal representation on the lower, middle, and upper leaves of the plants. After infiltration, plants were returned to the growth chamber until tissue was collected two to eight days later.

If vacuum infiltrations were performed, the soil of the pot was covered with a plastic film, and the plant was inverted into a 1L beaker containing a mixture of Gamborg's solution and *A. tumefaciens* containing expression constructs for GOI and p19. The pot was supported by two metal slats sitting on top of the beaker. The beaker was placed in a sealed chamber. The connected VPD3 pump (VIOT, Champagne, U.S.) was turned on until a vacuum of 25" mercury was reached. The valve connecting the pump and chamber was then closed, and the chamber was held at 25" mercury for 1 minute to allow air in the interstitial spaces of *N. benthamiana* leaves to escape. The vacuum was then released over 30 seconds to allow *A. tumefaciens* culture to enter the interstitial spaces.

2.2.7 Small Scale Tissue Collection and Protein Extraction

To collect tissue, a 7.1mm diameter cork borer was used to collect three leaf discs from each infiltrated leaf. For pooled samples, one disc was taken from the infiltrated leaf of each biological replicate. Tissue was collected in 2 ml microcentrifuge tubes containing three 2.3 mm ceramic beads. Tissue mass was measured using a Sartorius (Göttingen, Germany) B 120S fine balance, and the tissue was then flash frozen in liquid nitrogen, and subsequently stored at -80°C until extraction.

To extract protein, microcentrifuge tubes containing tissue were placed in homogenization blocks that had been pre-cooled to -80°C. Tissue was homogenized twice for 1 minute at 30 hertz in a TissueLyser (Qiagen). Blocks were then centrifuged for 1 minute at 2254 x *g*.

Various extraction buffers were used, and details are given in Table 4.

Table 4. Composition of various buffers used for protein extraction[1,2].

Extraction Buffer	Detergent	Protease Inhibitors	Buffer
Plant Extraction Buffer (PEB)	0.1% (v/v) Tween-20	2% (PVPP) (w/v), 1 mM EDTA, 1 mM PMSF, 1 µg/ml leupeptin, 100 mM sodium L-abscorbate	1xPBS, pH 7.8
Detergent Screen Buffer (DSB)	1% (v/v) sodium dodecyl sulfate (SDS), 1% (v/v) Brij L23, 1% (v/v) Triton X-100, 1% (v/v) Tween-20, 7 M urea/2 M thiourea/4% (w/v) CHAPS, or no detergent	2% PVPP, 1 mM EDTA, 1 mM PMSF, 1 mg/ml leupeptin, 100 mM sodium L-abscorbate	100 mM Tris pH 7.5-8
ProteoExtract® Native Membrane Protein Extraction Kit[3]	0.2% digitonin (Buffer 1) 0.5% Triton X-100 (Buffer 2)	Protease Inhibitor Cocktail (Calbiochem 539134; 100 mM AEBSF, 80 µM aprotinin, 5 mM bestatin, 1.5 mM E-64, 2 mM leupeptin, 1 mM pepstatin A, in DMSO)	300 mM Saccharose 15 mM NaCl 10 mM Pipes pH 7.2
Final Lab Extraction Buffer (FEB)	1.5% (v/v) Triton X-100	2% PVPP, 1 mM EDTA, 1 mM PMSF, 1 mg/ml leupeptin, 100 mM sodium L-abscorbate	100 mM Tris pH 8, 300 mM sucrose, 15 mM NaCl
VLP Extraction Buffer (VEB)	0.1% Triton X-100	2% PVPP, 1 mM EDTA, 100 mM sodium L-abscorbate	50 mM Tris-HCl pH 7.5 140 mM NaCl

[1]Initial extractions were performed with Plant Extraction Buffer (PEB), extraction optimization was performed with the Detergent Screen Buffers, as well as with the ProteoExtract® Native Membrane Protein Extraction Kit, the optimized extraction buffer was Final Extraction Buffer, and the VLP Extraction Buffer was used for VLP analysis.

[2] AEBSF, 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride ; CHAPS, 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraccetic acid; PMSF, phenylmethanesulfonylfluoride.

[3]The ProteExtract® Native Membrane Protein Extraction Kit contained two buffers for sequential extraction of soluble proteins (Buffer 1), then membrane proteins (Buffer 2).

Two extraction methods were used – one for lab made buffers, and one for the commercial kit. When using lab made buffers, after homogenization blocks were centrifuged, 200 µl of buffer was added to each microcentrifuge tube. Microcentrifuge tubes were mixed for three seconds on a vortex mixer, inverted and mixed for another three seconds. Microcentrifuge tubes were centrifuged at 4ºC for 15 minutes at 17,949 x *g*. The supernatant from each microcentrifuge tube was collected into a 1.5 ml microcentrifuge tube and then centrifuged again for 15 minutes at 4ºC. Extract was mixed in a 4:1 ratio with a sample buffer containing 0.3 M Tris-HCl pH 8.0, 5% (w/v) SDS, 10% (v/v) glycerol, 100 mM dithiothreitol (DTT) and 0.05% (w/v) phenol Red. Extracts were then run on SDS-PAGE gels or stored at -80ºC.

If the ProteoExtract® Native Membrane Protein Extraction Kit was used, the manufacturer's protocol was followed. After centrifugation of homogenization blocks, 5 µl of the kit's Protease Inhibitor Cocktail and 0.5 ml of Buffer 1 were added to each microcentrifuge tube. Microcentrifuge tubes were mixed by two 3 second pulses on a vortex mixer. Microcentrifuge tubes were then incubated for ten minutes at 4ºC with gentle agitation using a rotary shaker. All microcentrifuge tubes were then centrifuged at 16,000 x *g* at 4ºC for 15 minutes. The resulting supernatant from each microcentrifuge tube was mixed with sample buffer as described for lab-made buffer protein extract, and labelled the soluble fraction. The pellet was resuspended in 2.5 µl Protease Inhibitor Cocktail and 0.5 ml Buffer 2 using a pipette. The suspension was incubated for 30 minutes at 4ºC with gentle agitation using a rotary shaker, then centrifuged at 16,000 x *g* at 4ºC for 15 minutes. The supernatant was mixed with sample buffer as described for lab-made buffer protein extract.

2.2.8 Large Scale Tissue Collection and Protein Extraction and Concentration for VLP Detection

Infiltrated leaves were collected and placed in plastic bags, in 8 gram portions. The tissue was then immediately stored at -80ºC until it was extracted.

To extract protein, leaf tissue was placed in a large mortar, mixed with liquid nitrogen and ground into a fine powder with a pestle. Three parts (v/w) extraction buffer were

added to the mortar, and the mixture was ground in to a paste. Using a serological pipette the solution was then transferred to a conical tube. The solution was centrifuged twice for ten minutes at 4,200 x *g*. To concentrate any produced VLPs, 22 ml of supernatant was applied to a discontinuous sucrose gradient ranging from 30% to 60% in 10% 3 ml steps. The gradient was then ultracentrifuged using an Optima™ L-100 (Beckman Coulter) and a SW28 rotor at 86,329 x *g* at 4ºC. Ultracentrifuge ubes had a hole pierced in the bottom through which each fraction was collected and stored on ice at 4ºC, or stored at -80ºC until analyzed.

2.2.9 Western Blotting Procedure

Protein extracts were analyzed by immunoblot to determine protein accumulation, and the reagents used are detailed in Table 5. Other than where noted, samples were not boiled. 20 to 40 µl of sample were loaded onto Bio-Rad Mini-Protean® TGX™ Precast 4-20% (w/v) polyacrylamide gradient gels, except where noted. Gels were run at 115 V to start, then voltage was increased up to 140 V. Gels were run until dye reached the bottom of the gels. Proteins were transferred from gel to polyvinylidene difluoride (PVDF) membranes (Bio-Rad Laboratories, Inc.) using a Bio-Rad Trans-Blot® Semi-Dry Transfer Cell apparatus at 25 V and the membranes were blocked overnight in blocking solution.

To verify the detected protein was not a plant peroxidase, membranes were washed with TBS-T for 10 minutes, ECL Prime Western Blotting Detection Reagent (GE Healthcare Life Sciences, Little Chalfont, UK) was incubated with the membrane for 1 minute, and proteins were then visualized in a MicroChemi 4.2 (DNR Bio-Imaging Systems, Jerusalem, Israel). Membranes were then washed twice for ten minutes in TBS-T.

Membranes were then hybridized with primary antibody solution for one hour, under gentle agitation. Membranes were washed 3 x 15 minutes in TBS-T, then hybridized for one hour with secondary antibody solution. Three more TBS-T washes were performed under gentle agitation for 10 minutes each. Membranes were then rinsed in 1xTBS, after which ECL Prime Western Blotting Detection Reagent (GE Healthcare Life Sciences,

Table 5. Reagents used for Western blotting procedures.

Reagent	Recipe
10x Tris-buffered saline (TBS)	20 mM Tris, 300 mM NaCl, pH 7.5
1x TBS	10x TBS diluted to 1x with Milli-Q water
TBS-T	1x TBS with 0.1% Tween-20
Blocking solution	5% (w/v) skim milk powder in TBS-T
Primary antibody solution	1:4000 dilution of monoclonal mouse anti-*c-Myc* primary antibody (GenScript) in 0.5% (w/v) skim milk powder in TBS-T
Secondary antibody solution	1:4000 polyclonal goat anti-mouse secondary antibody (Bio-Rad Laboratories, Inc.) conjugated to horseradish peroxidase in 0.5% (w/v) skim milk powder in TBS-T

Little Chalfont, UK) was incubated with the membrane for 1 minute. Proteins were then visualized in a MicroChemi 4.2 (DNR Bio-Imaging Systems, Jerusalem, Israel).

2.3 Protein Quantification

Total recombinant protein was quantified through the use of dot blots. Sample extracts and negative controls from the same biological replicates were spotted on to nitrocellulose membranes (Bio-Rad Laboratories, Inc.) in a dilution series. Known amounts of a cellulose binding domain (CBD) synthetic protein standard (GenScript) were also spotted in two sets of dilutions for densitometry analysis. Membranes were allowed to dry for forty five minutes, and then blocked overnight in blocking solution. Membranes were processed as described in 2.2.9 for the Western blotting procedure.

Recombinant protein was quantified using Totallab TL100 software (Nonlinear Dynamics, Durham, USA). Two sets of CBD dilutions were used to develop a standard curve which was used to determine the amount of protein in samples. The amount of protein in mg per g was then calculated using the tissue weight of each sample. Final accumulation amount was determined by subtracting any detected protein in negative control dilutions from the amount of protein detected in sample dilutions. A balanced analysis of variance (ANOVA), followed by Tukey Pairwise Comparisons was performed with Minitab (State College, Pennsylvania, USA) to analyze differences in mean accumulation levels. Statistical significance level for all the tests was defined as 0.05 or lower.

2.4 VLP Analysis

2.4.1 Transmission Electron Microscopy

To assess whether VLPs were formed, large scale infiltrations were performed, and protein extracted and concentrated, as described above, using the VLP extraction buffer (Table 4). Analysis of sucrose gradient fractions for VLP formation was performed using transmission electron microscopy (TEM). A droplet of extract was placed on carbon grids (Electron Microscopy Sciences, Hartfield, U.S.), and allowed to sit for two minutes. Liquid was then drawn off, and the grid was washed in three consecutive drops of water

for two minutes each. Finally, a drop of 2% uranyl acetate was placed on the grid, and allowed to sit for a minute before being drawn off. The negatively stained grids were examined with a CM-10 transmission electron microscope (Philips, Amsterdam, Netherlands) equipped with a digital camera (Advanced Microscopy Techniques, MA) at 80 kV.

2.4.2 TEM Immunogold Labeling

A droplet of protein extract from the 40% sucrose fraction of either M-ELP, co-expressed M-ELP and E-ELP, or wild type tissue was placed on a carbon grid (Electron Microscopy Sciences), and allowed to sit for two minutes. Liquid was then drawn off. Grids were blocked by placing them specimen-side down in a drop of goat normal serum (25596; Aurion, Netherlands) for fifty minutes. Grids were washed in two consecutive drops of dilution buffer (1% bovine serum albumin (BSA), 0.2% BSA-c™ (Aurion), 0.05% Tween-20 in PBS pH 7.35) for two minutes each. One set of grids was incubated specimen-side down in a droplet of mouse anti-*c-Myc* primary antibody diluted 1:10 with dilution buffer , while a second set of grids was simultaneously incubated with just dilution buffer as a negative control. All grids were then washed in three consecutive drops of dilution buffer for two minutes. All grids were incubated for one hour in a drop of secondary antibody diluted 1:10 in dilution buffer. Secondary antibody was goat anti-mouse IgG conjugated to 10 nm gold particles (Aurion). Grids were washed in three consecutive drops of dilution buffer for ten minutes each, and then in four consecutive drops of Milli-Q water for three minutes each. Once dry, grids were stained with 2% uranyl acetate and examined under TEM as in 2.4.1.

2.5 Mass Spectrometry

The reagents used for mass spectrometry are described in Table 6. Protein extracts were run on an SDS-PAGE gradient gel as described in 2.2.9. The gel was incubated in gel fixing solution for fifteen minutes. The gel was subsequently washed with Milli-Q water 3 x 5 minutes each. 20 ml of GelCode™ Blue Stain Reagent (Thermo Fisher Scientific) was applied to the gel, which was then incubated for two hours. The gel was then

Table 6. Reagents used for mass spectrometry methods.

Reagents	Recipe
Gel fixing solution	50% methanol, 7% acetic acid in Milli-Q water
Gel destaining solution	50% methanol, 1% acetic acid in Milli-Q water
Destaining solution	80 mg of ammonium bicarbonate, 20 ml of acetonitrile, 20 ml of Milli-Q water
Digestion buffer	10 mg of ammonium bicarbonate with 5 ml of Milli-Q water
Reducing buffer	3.3 µl *tris*(2-carboxyethyl)phosphine with 30 µl digestion buffer
Alkylation buffer	50 mM iodoacetamide in digestion buffer

incubated in gel destaining solution 2 x 5 minutes before rinsing in Milli-Q water overnight. Bands of interest were then cut out of the gel and stored at -80ºC.

To digest proteins in the gel, gel pieces were thawed, 200 µl of destaining solution was added to each gel piece, and samples were incubated at 37ºC for 30 minutes, after which the destaining solution was discarded. This was repeated twice. Samples were then reduced by incubating in reducing buffer at 60ºC for ten minutes. Samples were allowed to cool, and reducing buffer was discarded. Samples were incubated in 30 µl of alkylation buffer in the dark for one hour. Alyklation buffer was then discarded, and samples were washed by incubating them in 200 µl of destaining buffer for fifteen minutes at 37ºC with shaking. Destaining buffer was discarded, and samples were washed again. Gel pieces were shrunk through incubation with 50 µl of acetonitrile for fifteen minutes. Gel pieces were air-dried for ten minutes, and incubated with 10 µl of activated trypsin for fifteen minutes. 25 µl of digestion buffer was added, and samples were incubated at 30ºC overnight with shaking. Digestion mixtures with peptides were then placed in clean microcentrifuge tubes.

These peptide digests were then analyzed in the chemistry facility of the London Research and Development Centre of Agriculture and Agri-Food Canada by Dr. Justin Renaud using the Easy-nLC 100 nano system at 75 µm x 15 cm Acclaim C18 PepMap™ column (Thermo Fisher Scientific) coupled to a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific) and the following protocol. The flow rate used was 300 nL min^{-1} and 10 µl of the protein digest was injected. Equilibration of the C18 was performed with 98% mobile phase A (water 0.1% formic acid) and 2% mobile phase B (acetonitrile 0.1% formic acid) and elution with a linear gradient from 2-30% B over 18 minutes followed by 30-98% over 2 minutes and maintained for 10 minutes. Nanospray voltage was set at 1.95 kV, capillary temperature 275°C, and S-lens RF level 60. Q-Exactive was operated in top 10 data-dependent acquisition mode with a full scan mass range of 300-1500 m/z at 70,000 resolution, automatic gain control (AGC) of $1x\ 10^6$ and maximum injection time (IT) of 250 ms. The MS/MS scans were acquired at 17,500 resolution, AGC of $1x\ 10^6$, maximum ITof 110 ms, intensity threshold of $8 x\ 10^4$, normalized collision energy of 27 and isolation window of 1.2 m/z. Unassigned, singly

and >4 charged peptides were not selected for MS/MS and a 20 s dynamic exclusion was used. The Thermo .raw files were converted to .mgf using Proteowizard v2 (Kessner et al., 2008).

MS/MS scans were then searched against the target/reverse proteome from Sol Genomics Network (Bombarely et al., 2012) and the amino acid sequences of the four recombinant PEDv proteins using X! Tandem (Craig and Beavis, 2004) search algorithm operated from the SearchGUI v2.8.6 (Vaudel et al., 2011) interface and processed in PeptideShaker v1.12.2 (Vaudel et al., 2015). An 8 ppm precursor ion mass error and a 0.05 Da product ion error were used along with carbamidomethylation as a constant modification and oxidation of methionine as a variable modification. A 1% false discovery rate was used at the protein, peptide, and peptide spectrum match level.

2.6 *In Silico* Sequence Analysis

2.6.1 Protein Structure Analysis

M and E peptide sequences obtained from NCBI, and the peptide sequences for the M-ELP and E-ELP fusion proteins were submitted to the I-TASSER server for protein structure prediction (Zhang, 2008). The sequence-based predictions of secondary structure given by PSSpred algorithm were compared to determine structural similarity between proteins.

2.6.2 VLP Structure Analysis

The pixel count of TEM scale bars were measured to deduce the nm:pixel ratio for each image. The pixel count of VLP envelope thickness was measured in three areas for each VLP, and the measurements for all VLPs were averaged. This average was multiplied by the nm:pixel ratio to obtain an estimate of VLP envelope thickness.

3 Results

3.1 Transient Expression of M

3.1.1 Optimal Construct Choice

The first objective of this study was to determine which *M* constructs would lead to the highest expression levels when transiently infiltrated into leaves of *N. benthamiana*. The PEDv gene sequence used was from strain ISU13-22038 isolated in 2013 (Chen et al., 2014). Sequence optimization has led to significantly change levels of PEDv protein accumulation in plants (Kang et al., 2005a), so the *M* gene sequence was optimized for expression in the nuclear genome of *N. tabacum*. The gene was cloned into four constructs (Figure 3) distinguished by the subcellular targeting signals and fusion peptides they contained. The recombinant protein was targeted to either the ER or apoplast (Apo). The ER targeted protein was fused either to an ELP or hydrophobin (HFBI) tag to potentially increase protein accumulation (Conley et al., 2009a; Joensuu et al., 2010), and had a C-terminal KDEL peptide to retrieve it to the ER. *A. tumefaciens* cultures carrying GOI expression constructs were co-infiltrated into *N. benthamiana* leaves with *A. tumefaciens* carrying a p19 expression construct, a suppressor of gene silencing (Silhavy et al., 2002).

Samples were collected from plant leaves four days post infiltration (DPI). Total soluble protein (TSP) was then extracted with PEB (Table 4), and separated on 12% SDS-PAGE gels, and analyzed by Western blotting to detect the recombinant protein. Samples were collected from three different biological replicates consisting of different plants. Figure 4 shows that only the ELP-ER cassette resulted in accumulation of M protein.

Previous literature indicated that the M protein of SARS-CoV showed thermal aggregation when heated in Laemmli buffer, resulting in protein being trapped at the stacking/separating gel interface (Lee et al., 2005). Thus both the stacking and separating gel were transferred to PVDF membrane. No noteworthy amount of M was detected at the gel interface (Figure 4). Instead, it is possible that the protein is trapped in the

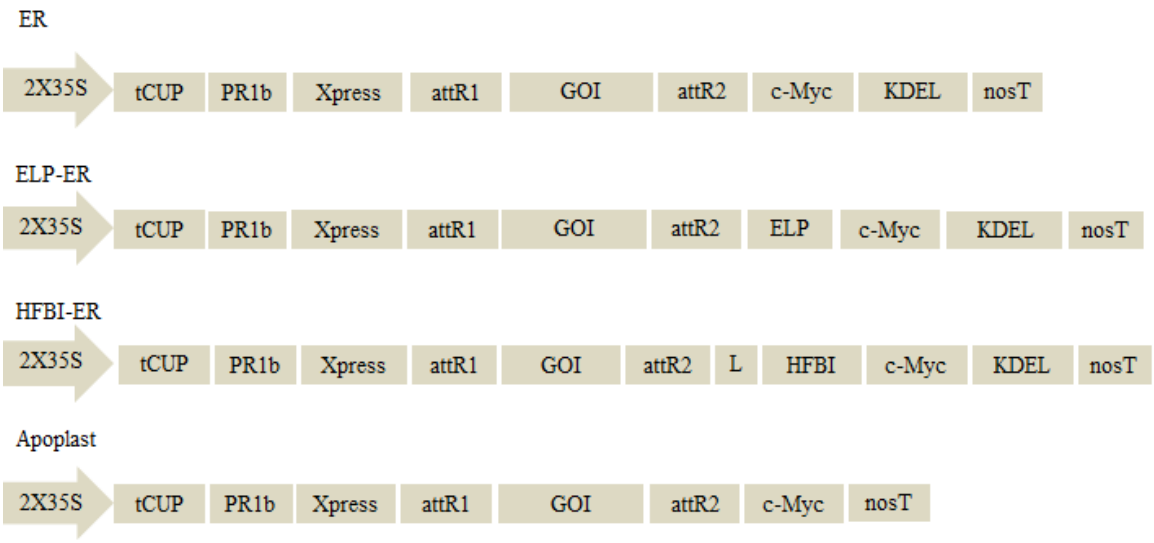


Figure 3. Schematics of the gene constructs used for *N. benthamiana* expression.

Constructs contained several elements for expression. The four constructs are for the M-ER, M-ELP, M-HFBI, and M-Apo protein products. The constructs included a double enhanced cauliflower mosaic virus promoter (2x35S) (Covey et al., 1981), and a translational enhancer from tobacco (tCUP) (Wu et al., 2001). Two of the constructs targeting the ER also contained sequences encoding an elastin-like polypeptide (ELP) or a hydrophobin (HFBI) tag with a linker (L). Two tags were included for immunodetection, N-terminal *Xpress* and C-terminal *c-Myc*. For subcellular targeting, tobacco pathogenesis-related-1b signal peptide (PR1b) (Huub and Van Loon, 1991) and the KDEL ER retrieval tetrapeptide were used. The schematic is not to scale.

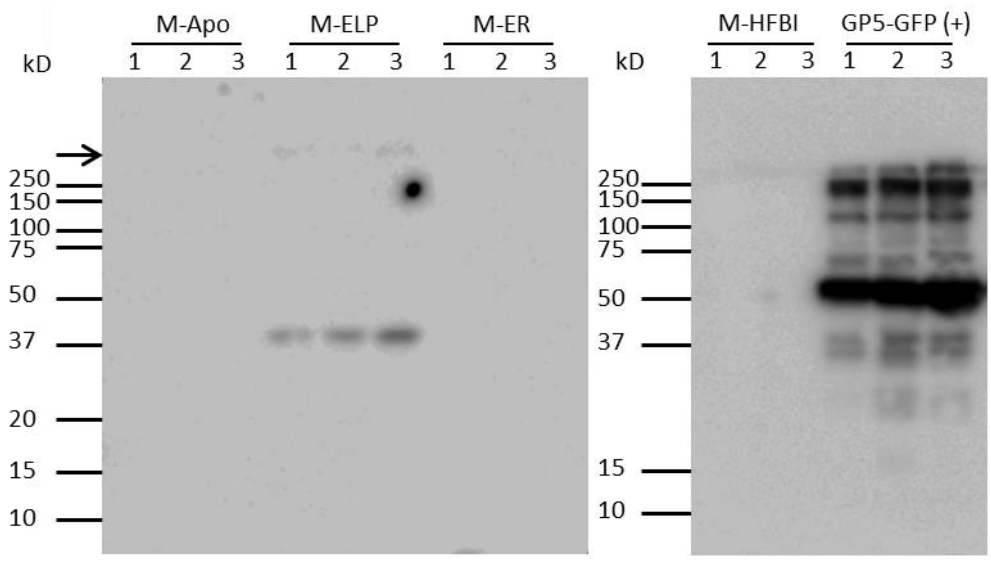


Figure 4. Western blot of M expressed with various expression constructs.

A band of the expected size, 43 kDa is observed for M-ELP-ER, but not for the other three constructs. GP5-GFP, a recombinant protein successfully produced previously, was used as a positive control. The arrow points to the barrier between the stacking and separating gel. Protein was extracted with PEB from three biological replicates, labelled as 1, 2, and 3, and run on 4% stacking, 12% separating polyacrylamide gels. This and all subsequent membranes were probed with anti-*c-Myc* primary antibody and goat anti-mouse secondary antibody

membranes and lost during soluble protein extraction.

3.1.2 Extraction and Detection Choice

As the M protein has three transmembrane domains and probably embeds in plant cell membranes, alternative extraction methods were explored. Bünger et al. (2009) compared five different commercial extraction kits for their ability to isolate and purify membrane-bound proteins. The study found the Qproteome® Cell Compartment Kit by Qiagen and the ProteoExtract® Native Membrane Protein Extraction Kit by EMD Millipore most effective. EMD Millipore's kit has been successfully tested on plant tissue, while Qiagen's has not been tested on plant tissue (Qian et al., 2008). Thus, EMD Millipore's kit was used to see if more M protein could be detected. The kit gives two protein extract fractions, one with soluble proteins, and one with membrane-bound proteins. As before, no protein accumulation was observed for any construct other than ELP-ER (Figure 5). More M-ELP was detected than with the previous extraction buffer. M-ELP was observed at the correct size of approximately 43 kDa in the soluble fraction, but potential dimers were also detected. In the membrane fraction, a large amount of aggregates of various molecular weights were present. Because the ELP-ER cassette was the only one that resulted in M protein accumulation, this construct was used exclusively for the rest of this study.

Lee et al. (2005) documented that the SARS-CoV M protein undergoes aggregation when heated in denaturing buffer, therefore it was hypothesized that heating may be the cause of the aggregates observed in the membrane-bound fraction. To test this hypothesis, protein extracts were separated by gradient PAGE, omitting the heating step of the procedure (Figure 6). This resulted in far fewer aggregates in the membrane-bound fraction, and more protein instead being detected at the correct molecular weight. Hence protein was not boiled any longer before loading onto SDS-PAGE gels. When extracts were not boiled, potential dimers and trimers became evident in place of thermal aggregates. Gradient gels were also used henceforth for clear separation of multimers.

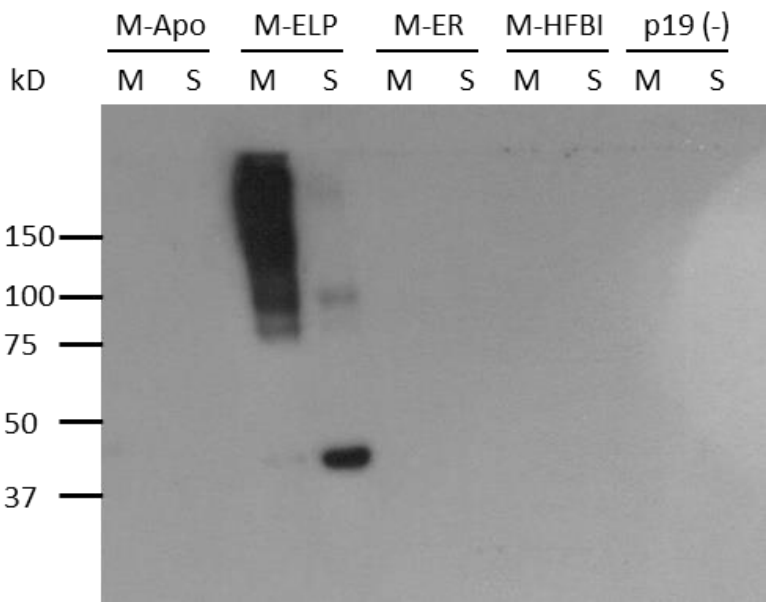


Figure 5. Extraction of infiltrated constructs with commercial membrane protein extraction kit.

N. benthamiana tissue was infiltrated with four different constructs and proteins were extracted with EMD Millipore's ProteoExtract® Native Membrane Protein Extraction Kit. Each extraction resulted in one fraction containing membrane-bound proteins (M), and one fraction containing soluble proteins (S). The only construct that showed protein accumulation is the M-ELP-ER construct. Using this kit, more protein was visualized than previously with PEB (Figure 4). Protein was separated on 4-20% gradient polyacrylamide gels.

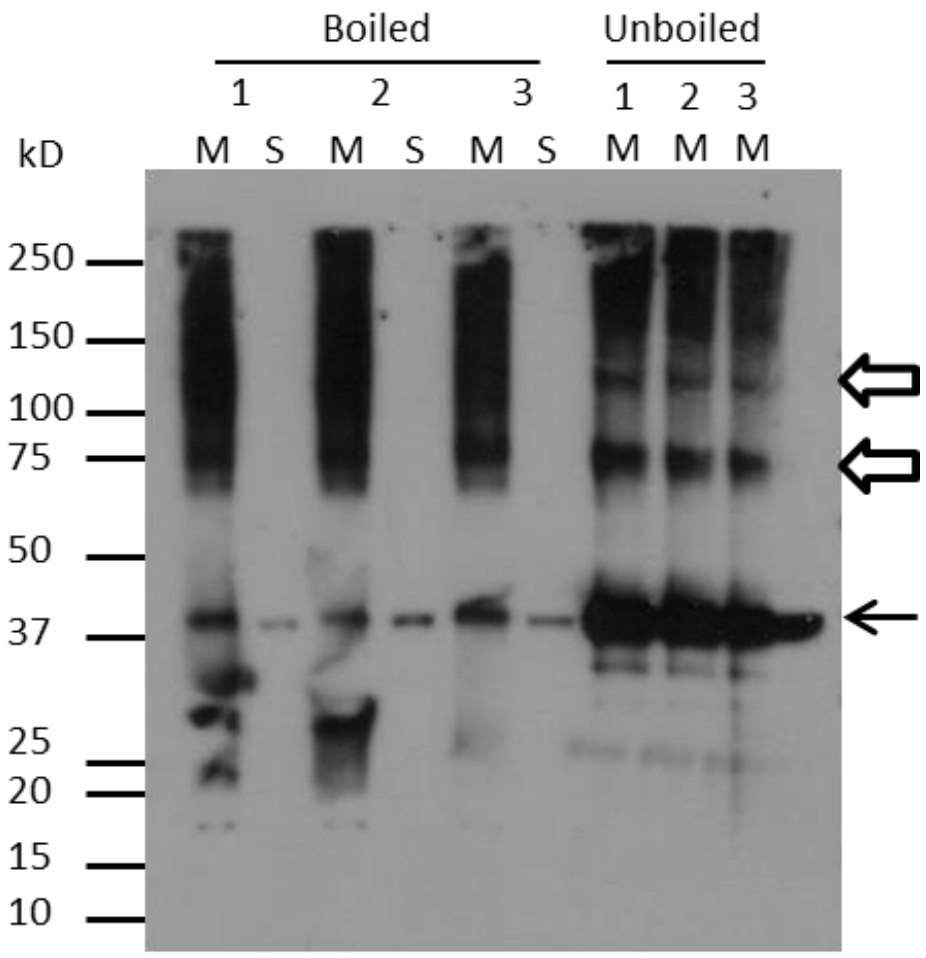


Figure 6. M-ELP extracted with commercial kit, and both boiled and unboiled.

Three biological replicates were extracted, and the membrane-bound fraction (M) and soluble fraction (S) of each were separated on gradient PAGE. The membrane-bound fraction of each was also not boiled to compare effects on protein aggregation. Higher molecular weight aggregates were observed in the membrane fraction, but more protein ran at the expected size when the extract was not boiled prior to running the gel (solid arrow). When extracts were not boiled, potential dimers and trimers became evident (empty arrowheads).

As extracts were not boiled, to ensure that the stronger 43 kDa protein band observed is M-ELP and not a plant peroxidase that would catalyze the chemiluminescent reaction used for detection, with a new set of samples the soluble and membrane-bound protein extract fractions were dotted directly on a nitrocellulose membrane (Figure 7). Enhanced chemiluminescence (ECL) substrate was applied, and it became evident that peroxidase was present when protein was not boiled. More peroxidase was detected in the soluble fractions than in the membrane-bound protein fractions.

The previous experiment was then repeated, and the membrane was incubated with ECL substrate before probing with primary antibody. No peroxidase was observed on the membrane exposed to ECL substrate before probing in either boiled or unboiled lanes (Figure 8). As before, the protein is observed after the blot is hybridized with the appropriate antibodies and detected with ECL substrate. Thus, the increased amount of protein detected at 43 kDa without boiling is in fact M-ELP. Perhaps particularly with dot blots, presence of plant peroxidase should be monitored.

To develop an extraction buffer and not rely on a commercial buffer with unknown composition, several buffer components for enhanced membrane protein extraction were investigated. Buffer choice is important to maintain the stability of proteins at the chosen pH and ionic strength. Proteins are denatured easily once no longer in the cell, and can be sensitive to pH changes (Cseke et al., 2011). Buffering capacity is important when acidic plant cell vacuoles are burst during the extraction procedure, to ensure buffer pH does not drop from the ideal range. pH 8 was chosen to avoid protein precipitation due to M-ELP's isoelectric point of 6.65. Tris was chosen as a buffer due to its buffering range of pH 7.5-9, and because it increases membrane permeability, which can help in the solubilisation of membrane proteins (Irvin et al., 1981). Detergent choice is a key component when optimizing an extraction buffer, as detergents disrupt membranes, and thus are necessary to solubilize a membrane protein (Arnold and Linke, 2008). Protein was extracted in a Tris-based buffer that contained either non-ionic detergents (1% Triton X-100, 1% Brij 35, 1% Tween-20), a harsh denaturant (7 M urea/2 M thiourea/4% CHAPS), or no detergent/disruptant.

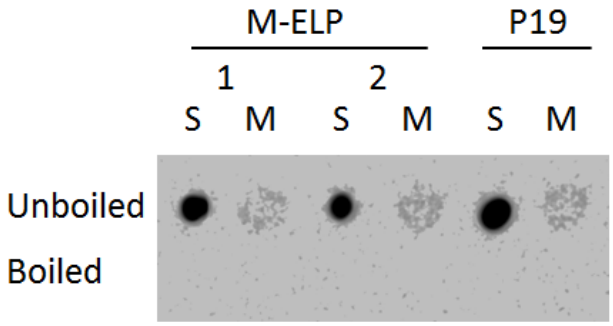


Figure 7. Dot blot showing presence of plant peroxidase in unboiled protein extracts.

Most of the peroxidase is in the soluble protein extract. Plant peroxidase is not increased with viral membrane protein expression, it is found with both M-ELP and p19 infiltration, and just infiltration of p19 (negative control). S, soluble protein fraction, M, membrane-bound protein fraction. Protein was extracted with EMD Millipore's ProteoExtract® Native Membrane Protein Extraction.

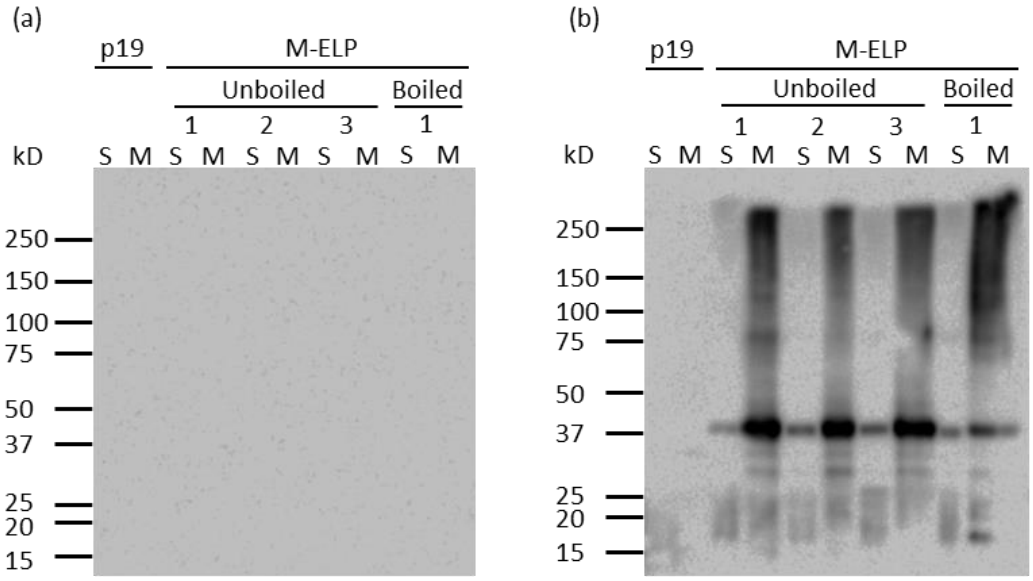


Figure 8. Western blot detected before and after probing to look for peroxidase activity.

(a) ECL substrate was applied after blocking the membrane but before probing. Neither the negative control (p19) nor M-ELP lanes have any peroxidase detected in the soluble (S) or membrane-bound (M) protein extraction fractions. (b) As before, more protein is observed in the membrane-bound fraction (Figure 6). When protein is not boiled prior to loading gel, less aggregation is observed. p19 protein extracts were not boiled. Protein was extracted with EMD Millipore's ProteoExtract® Native Membrane Protein Extraction Kit from three biological replicates, (1, 2, and 3).

Using dot blots, the amount of protein in each extract was quantified as mg of protein per g of fresh leaf weight. Tween-20 and Triton X-100 extracted the highest levels of protein, at 0.80 ± 0.43 mg/g and 0.75 ± 0.39 mg/g of fresh leaf weight, respectively (Table 7). The difference between protein extracts was visible, with no detergent and Tween-20 having the palest extracts, and SDS and the urea extract appearing a darker green, indicating lysis of the chloroplast membrane (data not shown).

To determine if the effects of the detergents on M-ELP extraction were significantly different among treatments, a balanced ANOVA was performed, which takes in to account variability between replicates. For this, the log value of the protein accumulations was used to satisfy the assumptions of homoscedasticity. A balanced ANOVA showed that accumulation levels between extraction methods were significantly different ($p < 0.001$). Tukey Pairwise Comparisons were performed to determine which treatments were significantly different from each other. Buffers with detergents extracted significantly more protein than when no detergent was used, but the difference between detergents was not significant (Figure 9). Use of detergent resulted in an over 350x increase in the amount of protein extracted, and choice of detergent resulted in a 4-fold increase in the amount of protein extracted (Table 7).

Tween-20 and Triton X-100 became prime candidates for future experiments. After this experiment was performed, the composition of EMD Millipore's protein extraction kit was released, and reported to use Triton X-100 as its detergent. Due to Triton X-100's success in extracting M protein, its commercial use, its utility in phase separation, and to stay consistent with VLP extraction protocols, which use Triton X-100, it was the detergent used for further experiments.

Research has shown that sucrose can stabilize proteins by increasing the activation energy of protein unfolding. Sucrose is preferentially excluded from protein domains, thus increasing the free energy of solutions, and making the unfolded state of proteins less favourable (Lee and Timasheff, 1981). Another set of extractions was performed to compare Tris-based buffer with 1% Triton X-100, with and without sucrose, to EMD

Table 7. Protein extraction levels with different detergents/disruptants.

Detergent/Disruptant Used	Accumulation of M-ELP in mg/g of fresh weight leaf tissue[1]
7M urea, 2M thiourea, 4% CHAPS	0.620 ± 0.30
1% Triton X-100	0.748 ± 0.388
1% Tween-20	0.800 ± 0.433
1% Brij35	0.202 ± 0.012
No detergent	0.00214 ± 0.00142

[1]Mean value of three biological replicates, ±, standard error of the mean value of the three biological replicates.

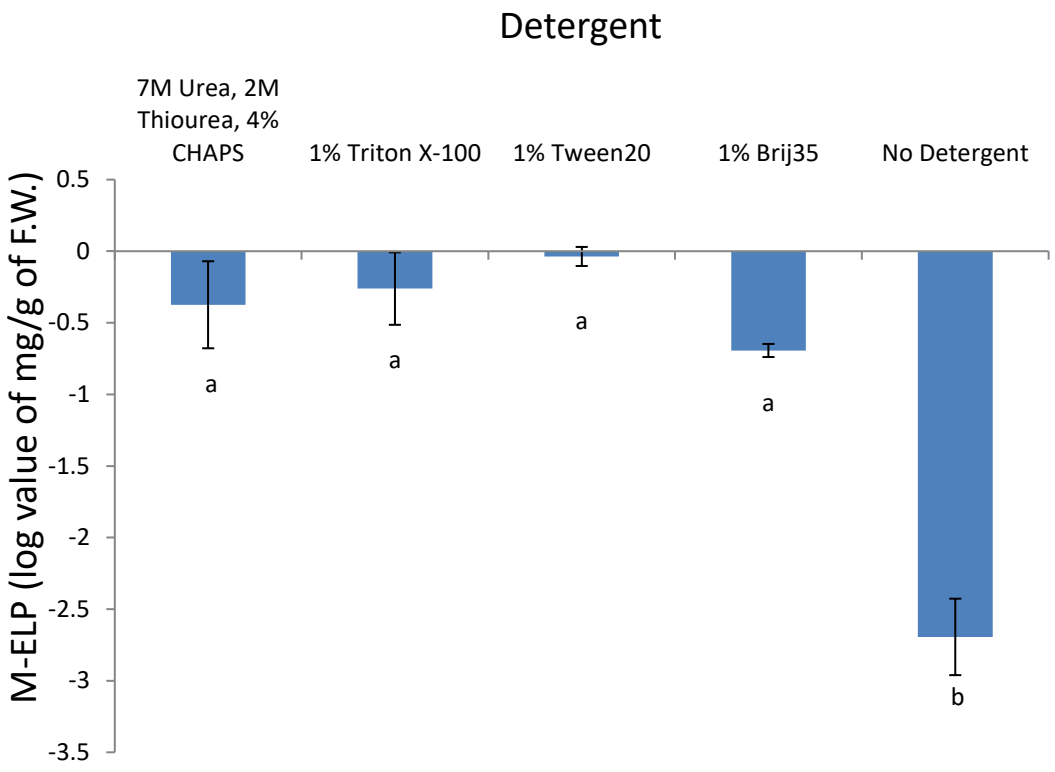


Figure 9. Comparison of protein extraction levels with different detergents/disruptants.

Triton X-100 and Tween-20 extracted significantly more protein than the other extraction methods, and all detergents extracted significantly more protein than extraction without detergent. Each column represents the mean value of three biological replicates and error bars indicate the standard error of the mean. Means that do not share a letter are significantly different ($F_5 = 21.16$, $p < 0.001$).

Millipore's kit. Protein extraction values for the soluble and membrane-bound fraction were pooled and compared to the single extraction values of the lab buffer protein extracts (Table 8).

The Tris-based buffers extracted more protein than the commercial buffers despite resulting in a paler extract (data not shown). The addition of sucrose had a positive impact, but the differences between the extractions were not significantly different [F_2 = 1.21, p >0.39]. Due to the observed trend of increased extraction with sucrose, and its inclusion in the commercial kit, sucrose was added to the composition of FEB.

Lastly, the effect of increasing the detergent concentration in the lab buffer was studied. Increasing the concentration of the detergent marginally increased the amount of protein detected, evidenced by comparing the detection of degraded protein products (Figure 10). Due to any potential increase in the amount of protein extracted Triton X-100 concentration was increased to 1.5%.

3.2 VLP Formation

3.2.1 *In silico* Analysis of M and M-ELP

Having determined that M-ELP is produced in *N. benthamiana* at above 0.7 mg/g of fresh leaf weight, the next objective to examine was whether M was assembling into VLPs. While M from SARS-CoV and TGEV was shown to independently assemble into VLPs (Tseng et al., 2010; Zhenhui et al., 2015), it is unknown if a fusion of PEDv M with ELP would assemble into VLPs. Therefore, the structure of the fusion protein was compared to the native protein structure by *in silico* protein structure prediction software.

The I-TASSER 3D protein prediction software (Zhang, 2008) using the PSSpred algorithm predicted that the beginning of the native protein was a short coil (1-7), followed by three helices, representing the transmembrane segments (8-34, 44-65, 76-105) with coils in between each, and a long C-terminal coil-strand mix (Figure 11a). This gave further insight into what had been previously reported about coronavirus M structure. Previous predictions have suggested that the N-terminal outside the virus

Table 8. Protein extraction levels with lab and commercial buffers.

Extraction Buffer Used	Accumulation in mg/g of fresh weight leaf tissue of M-ELP[1]
1% Triton X-100	0.57 ± 0.17
1% Triton X-100 (with sucrose)	0.73 ± 0.17
Combined Soluble and Membrane fractions of ProteoExtract® Native Membrane Protein Extraction Kit	0.49 ± 0.056

[1]Mean value of three biological replicates, ± standard error of the mean value of the three biological replicates.

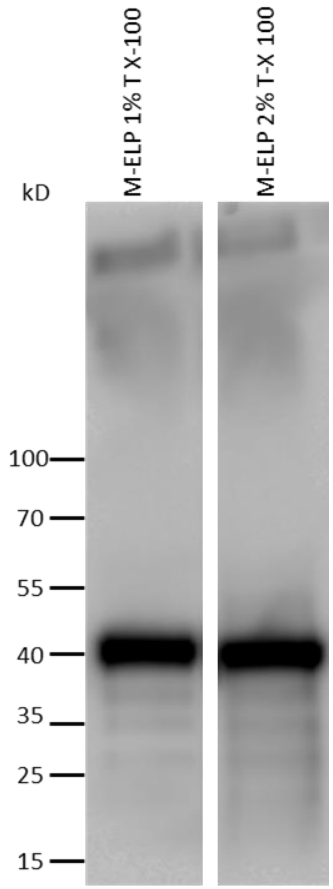
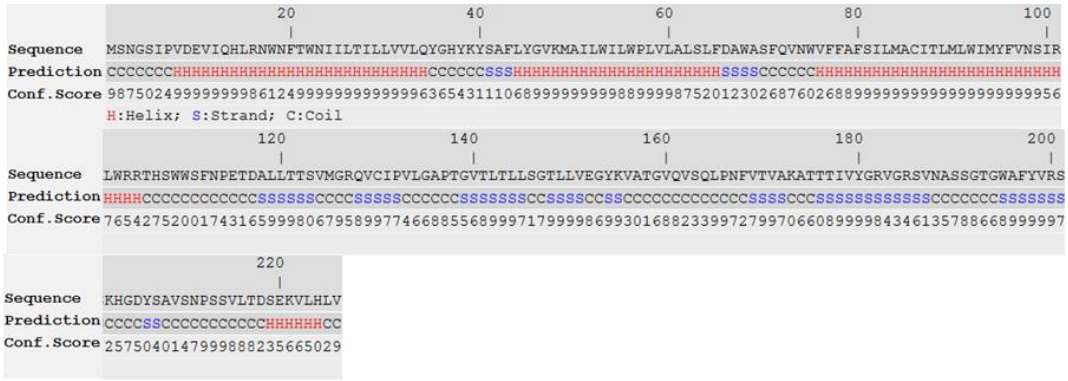


Figure 10. Visual comparison of differences in Triton X-100 based buffers.

It is observed that increasing the concentration of Triton X-100 potentially increases the amount of protein detected on a Western, particularly when comparing degraded protein products.

(a)



(b)

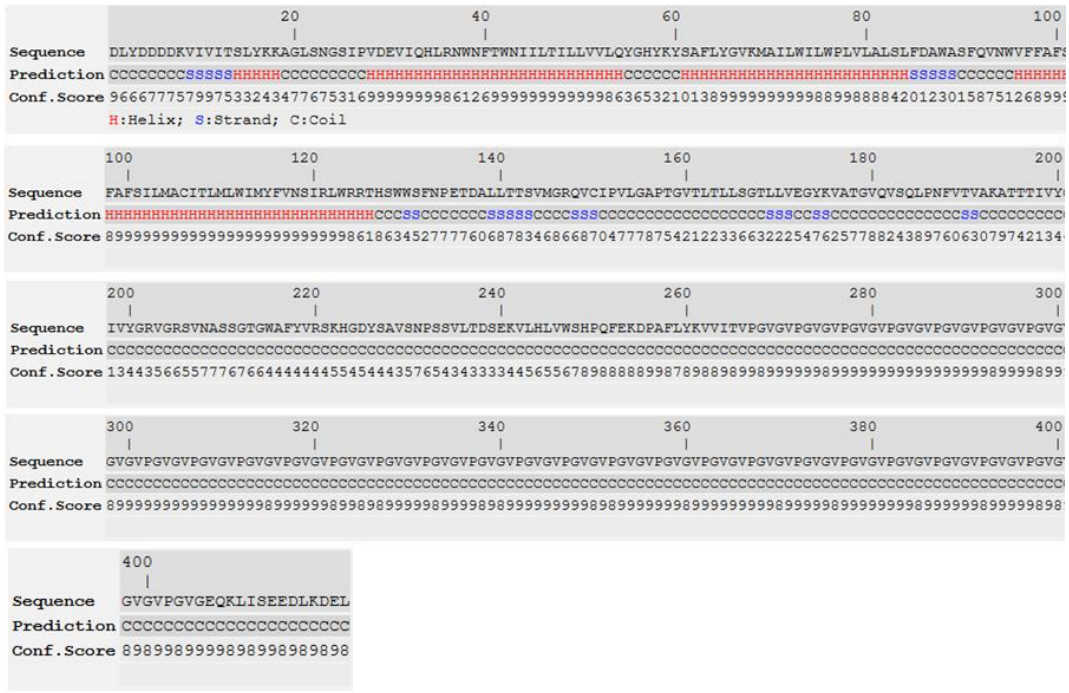


Figure 11. Predicted protein structures of M and M-ELP.

(a) Native M amino acid sequence and PSSpred predicted protein structure, showing the characteristic three helices that represent the transmembrane segments.(b) M-ELP amino acid sequence and PSSpred predicted protein structure, showing the three helices, as well as extended N-terminal and C-terminal fusions. Numbers represent the predicted confidence score from 1-10. C, coil; H, helix; S, strand.

stretched from amino acid positions 1-17. This analysis instead suggested that this domain was made up of only amino acids 1-7. This led to a difference in the length of the transmembrane domain that had previously been reported. Previously, transmembrane segments were predicted to be amino acid sequences 18-38, 42-62, and 76-96 (The UniProt Consortium, 2015). With no coronavirus membrane proteins submitted to the Protein Data Bank database, neither estimate is definitive.

Using the same software to compare the structure of M-ELP, the N-terminal coil and the triple transmembrane segments separated by coils were preserved. The N-terminal additions (att sites and Xpress tag) are predicted to result in an additional coil at the N-terminus, and the addition of the ELP polypeptide results in an extension of the C-terminal coil (Figure 11b). Whether there is an impact on the typical coronavirus M protein amphipathic domain inside the virus is not clear, as this was not predicted for the native or fusion protein. With the fundamental structure of the protein predicted to be unchanged, the potential for M-ELP to form VLPs as native coronavirus M protein was explored.

3.2.2 Expression of S, N, E, and Co-Expression

To determine whether VLPs can be produced through co-expression of several proteins, it was first necessary to determine which proteins could be produced in *N. benthamiana*. The genes for the three remaining structural proteins, N, S, and E were cloned into ELP-ER construct since the ELP fusion was the only gene construct that expressed with M. These gene constructs were expressed transiently by agroinfiltration in *N. benthamiana,* and protein was extracted six days post-infiltration. Recombinant proteins were separated by SDS-PAGE and analyzed by immunoblotting.

The M-ELP lane showed the expected 43 kDa band, but additionally showed a smaller band above 25 kDa (Figure 12). While degradation products had previously been observed, for example in Figure 10, a clear single smaller band had not been detected. The observed band is likely a result of proteolytic cleavage (see 3.2.3.2).

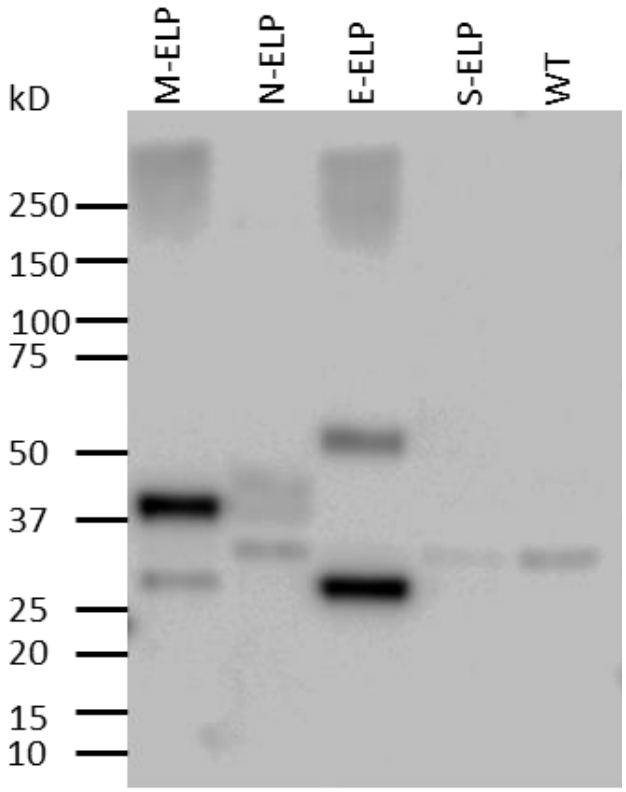


Figure 12. Expression of all four PEDv structural genes.

M-ELP and E-ELP are clearly detected. A smaller than expected band is observed alongside the monomer and a putative peroxidase band for M-ELP, and E-ELP is found to accumulate as a monomer and dimer. There is putative plant peroxidase observed in all lanes just under 37 kDa in size. Bands larger than the potential plant peroxidase are observed in the N-ELP lane, but these bands are smaller than the expected size of 66.6 kDa. S-ELP is predicted to be 169.4 kDa, and neither full size S nor a breakdown product were observed. Protein was extracted with FEB from pooled leaf tissue from three plants with FEB.

The E-ELP construct showed a major band of the predicted size as well as a putative dimer (26 kDa and slightly larger than 50 kDa, respectively). N-ELP showed two diffuse bands at ca. 37 and 45 kDa, instead of the expected 66.6 kDa, and S-ELP did not produce any protein (Figure 12). All lanes showed a band smaller than 37 kDa thought to be plant peroxidase, due to its presence in the wild type tissue extract.

Previous coronavirus studies have indicated that M and E, and sometimes N, are critical to VLP formation (Ho et al., 2004; Huang et al., 2004). Since M-ELP, E-ELP and possible N-ELP were produced in *N. benthamiana*, co-expression studies were conducted. Aside from expression of each construct alone, three combinations of gene constructs were co-expressed: M-ELP with E-ELP, M-ELP with E-ELP and N-ELP, and M-ELP with E-ELP, N-ELP, and S-ELP. While it was clear that S-ELP was not accumulating on its own (Figure 12), all four structural proteins were co-expressed to assess whether this could assist in S-ELP accumulation.

To confirm whether the consistent 35 kDa band observed in Figure 12 was plant peroxidase, and whether the bands around 37 kDa in the N-ELP extract were breakdown products of N-ELP or plant peroxidase aggregates, the protein extracts were run on another PAGE and the membrane was treated with the ECL substrate prior to antibody binding. Knowing that E-ELP was also able to express, co-expression studies also began. The proposed 35 kDa peroxidase band in Figure 12 was found in every sample in Figure 13a, confirming that it is a plant peroxidase catalyzing the chemiluminescent reaction. The larger bands in the N-ELP extract were not observed, indicating that they correspond to cleavage products of N-ELP. The band in the M-ELP extract just larger than 25 kDa is also not present, indicating it is likely a hypothesized cleavage product of M-ELP. It has previously been reported that the M protein of TGEV undergoes N-terminal proteolytic cleavage (Laude et al., 1987). While this is not expected for PEDv M due to the lack of predicted signal sequence, incubation with Triton X-100 is suggested to make a protease cleavage site accessible, leading to the presence of a fragment (Utiger et al., 1995). Based on previous reports, cleavage of M-ELP should result in an N-terminal fragment of 23.61 kDa, and a C-terminal fragment of 19.39 kDa (Utiger et al., 1995). The fragment detected most strongly is instead observed to be slightly larger than 25 kDa (Figure 13a).

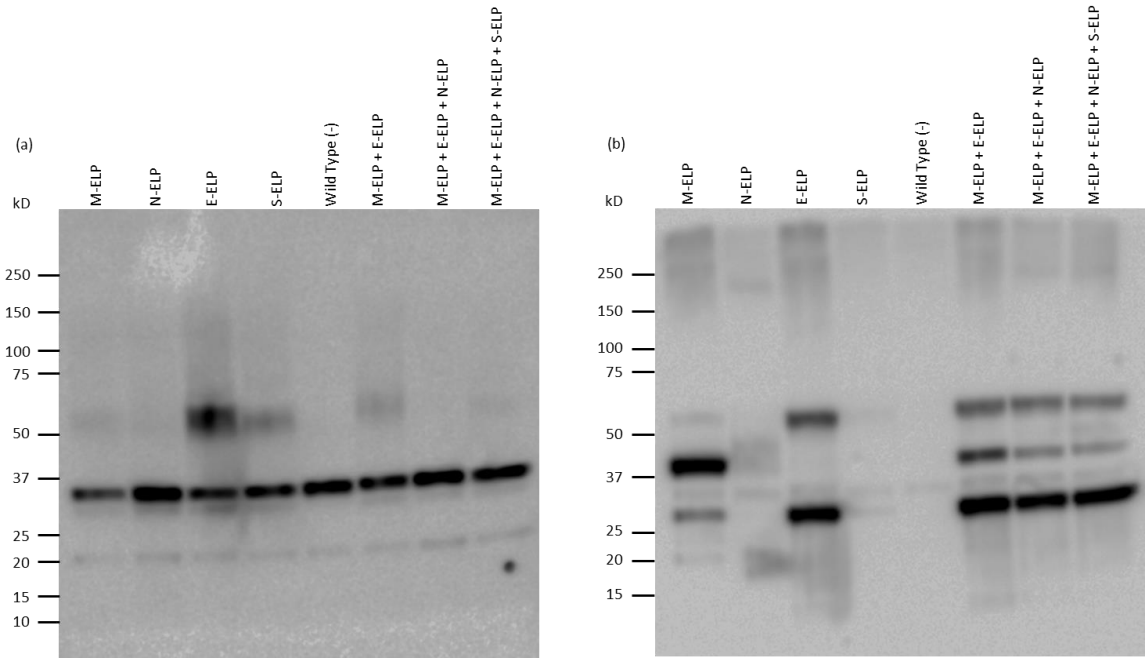


Figure 13. All four structural genes, individually and co-expressed, with ECL substrate before and after antibody hybridization.

(a) The PVDF membrane was treated with ECL substrate before antibodies were hybridized to proteins. Bands show presence of plant peroxidase, catalyzing the ECL reaction. A consistent band slightly smaller than 37 kDa is observed in all lanes, this same band is observed in Figure 12. A band just above 20 kDa is also present in all lanes, and a band between 50 and 75 kDa is present in all lanes except wild type (non-infiltrated) tissue, and potentially M-ELP+E-ELP+N-ELP. (b) The M-ELP lane shows smaller degradation products, as well as the putative cleavage product larger than 25 kDa, and monomer. The 35 kDa and 50+ kDa peroxidase bands are observed more clearly than expected. The N-ELP, E-ELP, S-ELP and wild type (non-infiltrated) tissue lanes show the same banding patterns as in Figure 12. When proteins were co-expressed the banding patterns of the respective proteins were all found, with the N-ELP fragments detecting most strongly when all four proteins were co-expressed. In contrast, M-ELP and E-ELP bands are weaker with the co-expression of N-ELP or N-ELP and S-ELP. Protein was extracted from pooled leaf tissue from three plants with FEB.

In order to verify the identity of the detected protein, mass spectrometry analysis was performed. Analysis of the putative fragment resulted in peptide spectra matching to three areas of M-ELP, the N-terminal of the protein (a.a. 19-37), across the middle of the protein, in M-ELP's intravirion tail (a.a. 176-,194, 194-203, 206-221, 223-242), and to the C-terminal c-Myc tag and ER retention sequence (a.a. 409-420). Finding that the >25 kDa band contains peptides from both the N- and C-terminal of M-ELP may indicate that fragments of the protein are self-associating to give the observed >25 kDa band.

In addition, a band slightly larger than 50 kDa is found in several lanes of the blot. While it is barely visible in the M-ELP and N-ELP extracts, it is pronounced in the E-ELP extract, and to a lesser extent, in the S-ELP lane, and visible when protein extracts involve either protein. This band may represent a different peroxidase that is induced by specific viral protein expression (G. Lomonossoff, personal communication).

The membrane was then washed, probed with primary and secondary antibodies and imaged by chemiluminescence. The 35 kDa putative peroxidase band is far fainter than before probing. Though in Figure 13b the membrane was imaged for a sixth of the time, the amount of light captured was over five times higher. This indicates that recombinant protein bands are stronger, with the faint 20 kDa band of Figure 13a no longer showing in all lanes. The truncated M-ELP product here is distinctly smaller than the peroxidase band. Other bands are also observed in the M-ELP lane which may represent the smaller and larger peroxidase bands in Figure 13a.

The putative E-ELP dimer migrates at a similar size to the peroxidase in Figure 13a; however it is a much tighter band. It is clear that this band is in fact E-ELP, as the 50+ kDa putative peroxidase band is very faint in co-expressed protein lanes in Figure 13a, but is clearly visible in all three co-expressed lanes in Figure 13b. Mass spectrometry analysis of the putative E-ELP dimer band confirmed the presence of E-ELP, confirming that E-ELP is forming a dimer (Appendix 2). No peroxidase of the correct molecular weight was detected in the protein composition of the band, likely because the peroxidase observed in Figure 13a is present at levels below the detection limits of the mass spectrometry equipment. Thus the tight >50 kDa band observed in Figure 13b is

primarily composed of the E-ELP dimer. Co-expression of M-ELP and E-ELP resulted in detection of the E-ELP monomer and dimer, M-ELP monomer, and the 35 kDa peroxidase band. Co-expression of N-ELP, or N-ELP and S-ELP with M-ELP and E-ELP results in the detection of N-ELP fragments. These fragments detect most strongly when all four proteins are co-expressed. When all four proteins are co-expressed M-ELP and E-ELP bands detect less strongly, likely as the *A. tumefaciens* OD_{600} of each construct is decreased.

It became apparent through these experiments that leaves that had been infiltrated with E-ELP were much paler post-infiltration than leaves that had been infiltrated with M-ELP. This phenomenona persisted when M-ELP and E-ELP were co-expressed (Figure 14abc). This potentially is due to the upregulation of the >50 kDa peroxidase.

When proteins were extracted with the commercial kit, as opposed to FEB (Table 4) like in Figure 13, a laddering effect was observed (Figure 15b). When M-ELP and E-ELP are co-expressed the laddering bands are stronger than those in the M-ELP/N-ELP/S-ELP/E-ELP extract, similar to what was observed in Figure 13a. With no S-ELP or N-ELP accumulation observed, infiltrating all four genes together meant that the number of *Agrobacterium* bacteria carrying each construct was reduced, thus leading to a reduction in protein accumulation of M-ELP and E-ELP. This provides evidence that laddering is an effect of M-ELP and E-ELP co-expression. A time course was also performed, showing the highest amount of protein accumulation six days post infiltration for M-ELP+E-ELP lanes, and four days post infiltration for co-expressed M-ELP+E-ELP+N-ELP+S-ELP lanes. M-ELP only lanes did not have discernibly different accumulation between days 4 and 6, but both days show higher accumulation than day 2.

In the M-ELP+E-ELP lanes, bands the size of the E monomer and dimer were observed (one and two squares, Figure 15b), but a band was not detected at the size of the M monomer , rather a band at ca. 80 kDa was observed that might correspond to a dimer of M (one and two diamonds, Figure 15b). This dimer was also the only recombinant protein observed in the M-ELP only lanes. This was the only time M-ELP was only

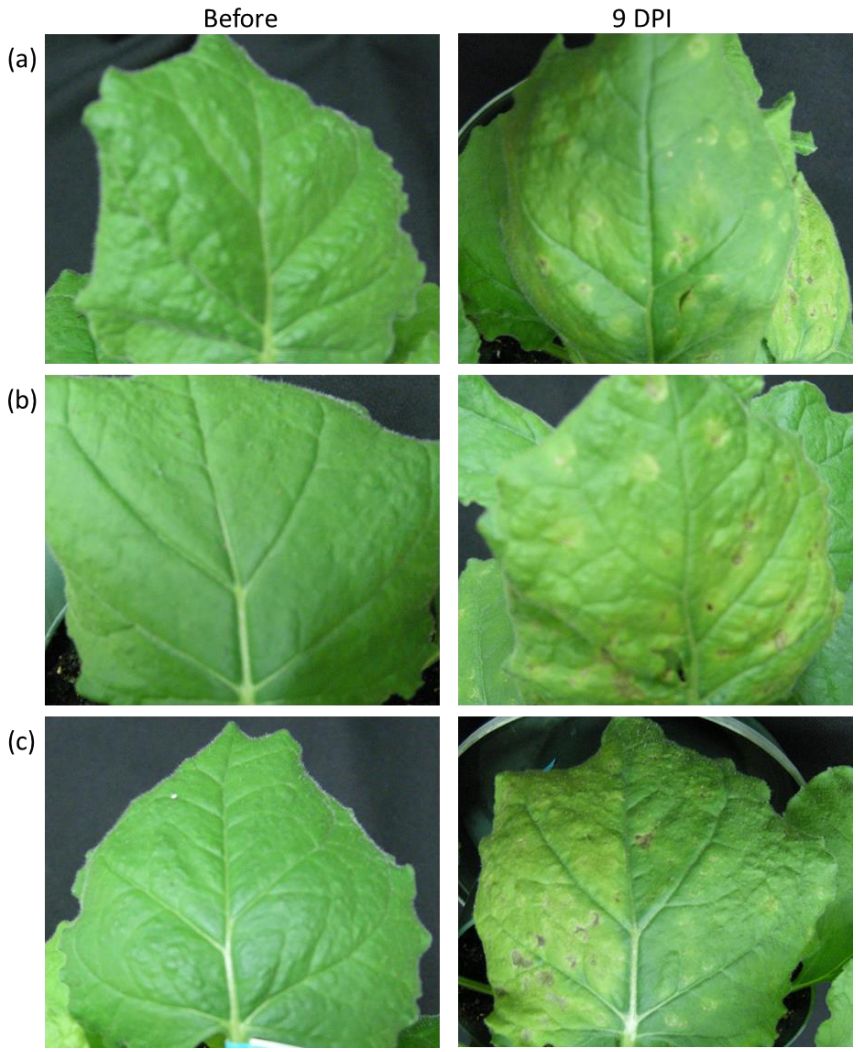


Figure 14. *N. benthamiana* leaf infiltrated with M-ELP, E-ELP, or M-ELP and E-ELP.

(a) M-ELP infiltrated leaf. The leaf is lighter post-infiltration, but not to a large degree. (b) E-ELP infiltrated tissue. The leaf is much lighter than before infiltration. (c) M-ELP and E-ELP infiltrated leaf tissue (left side of the leaf) and p19 (right side of the leaf, negative control). The yellowed leaf phenotype seen with E-ELP infiltrated tissue is maintained when both constructs are infiltrated together, and is clearly not due to the infiltration of p19.

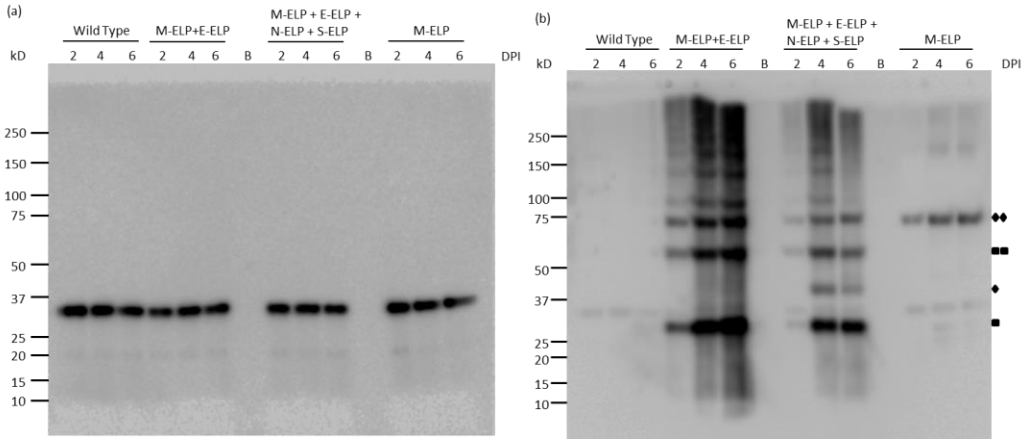


Figure 15. Co-expression of PEDv proteins, and time course.

Extraction of co-expressed proteins with the commercial kit showed a laddering effect was possible. (a) Membrane before probing. ~35 kDa peroxidase band is observed, and the 20 kDa band is observed faintly. (b) M-ELP and E-ELP co-expression showed the highest accumulation of protein on day six. Bands are lighter when all four proteins are co-expressed, due to lower OD_{600} of each construct being infiltrated. Protein accumulation was highest on day four when all four proteins were co-expressed. E-ELP monomer and dimer size bands are observed (one and two squares). When M is expressed alone or together with E-ELP, the M-ELP monomer is not observed (one diamond), but a dimer is (two squares). M-ELP only tissue shows the highest amount of protein four or six days post infiltration (DPI). Wild type (non-infiltrated) tissue is the negative control; B, blank lane; DPI, days post infiltration. Protein was extracted with ProteoExtract® Native Membrane Protein Extraction Kit, membrane extract for one of three plants is shown.

observed as a dimer, and may correlate to the observed laddering. The higher molecular weight bands observed on this blot likely indicate multimerization and VLP formation. The lanes with all four constructs showed a band a band the size of the M-ELP monomer, potentially the result of reduced protein accumulation and interaction.

To determine if these samples showed any difference in peroxidase expression, the membrane was imaged before and after antibody probing. The pre-probed membrane in Figure 15a shows the ~35 kDa and ~20 kDa peroxidase bands observed previously, but the 50+ kDa band was notably absent from the M-ELP + E-ELP extracts. This is different than what was observed in Figure 13, where E-ELP seemed to cause production of a peroxidase with a molecular weight above 50 kDa. Despite the laddering detected on the probed blot, peroxidase size is not affected.

3.2.3 VLP Formation

3.2.3.1 Sucrose Gradient Analysis

Knowing that both M-ELP and E-ELP are produced in *N. benthamiana* leaves, and that a laddering effect was observed which could mean that the proteins are assembling into higher order structures, the possibility that VLPs were being assembled in plant cells was explored. For this, leaf tissue was vacuum-infiltrated with either M-ELP, or M-ELP and E-ELP, and 8 g of infiltrated and wild type tissue was harvested four days post-infiltration. Protein was extracted in VLP Extraction Buffer (Table 4), and the protein extract was spun in an ultracentrifuge through a 30-60% discontinous sucrose gradient to concentrate any VLPs (Figure 16).

The resulting gradient was fractionated into 1 ml aliquots, and samples were analyzed by immunoblot. The full-length M-ELP protein was detected strongly in all fractions, with the highest amount of protein in the 40% fraction. The 25 kDa cleavage product previously observed was detected less in the heavier sucrose gradient fractions, which would contain sedimented VLPs (Figure 17a). Truncated M-ELP not being detected in heavier sucrose fractions may indicate that the protein is not being incorporated into VLPs. Previous literature has noted that the M fragment does not incorporate into virions (Utiger et al., 1995). As peptide spectra coverage did not match to the transmembrane

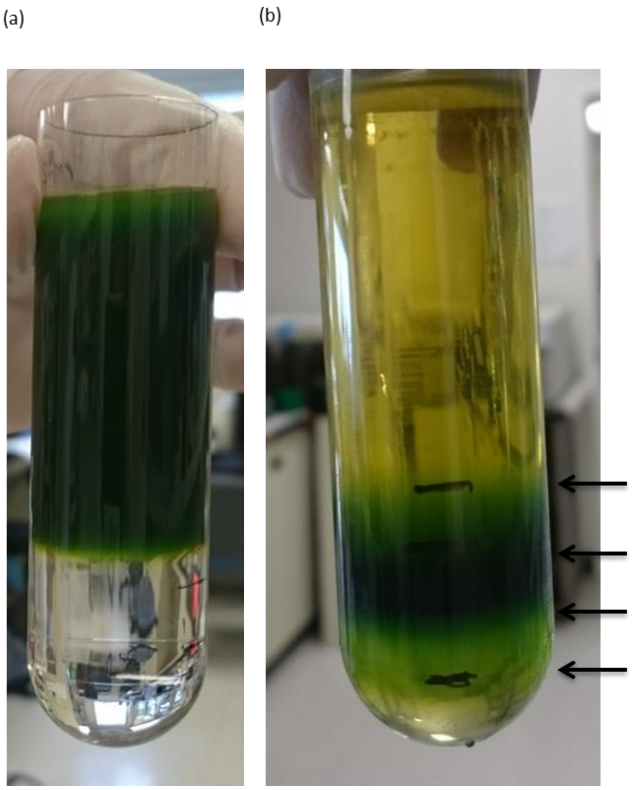


Figure 16. Protein extract ultracentrifuged through a sucrose gradient.

(a) 24 ml of protein extract loaded onto 12 ml of discontinuous sucrose gradient. (b) After the spin, each gradient is clearly shown, as well as the supernatant. From the bottom up, arrows indicate the interphase between the 60%, 50%, 40%, 30% sucrose and supernatant fractions. VLPs were expected to be found in the 40% fraction. Protein was extracted from leaf tissue with the VLP Extraction Buffer.

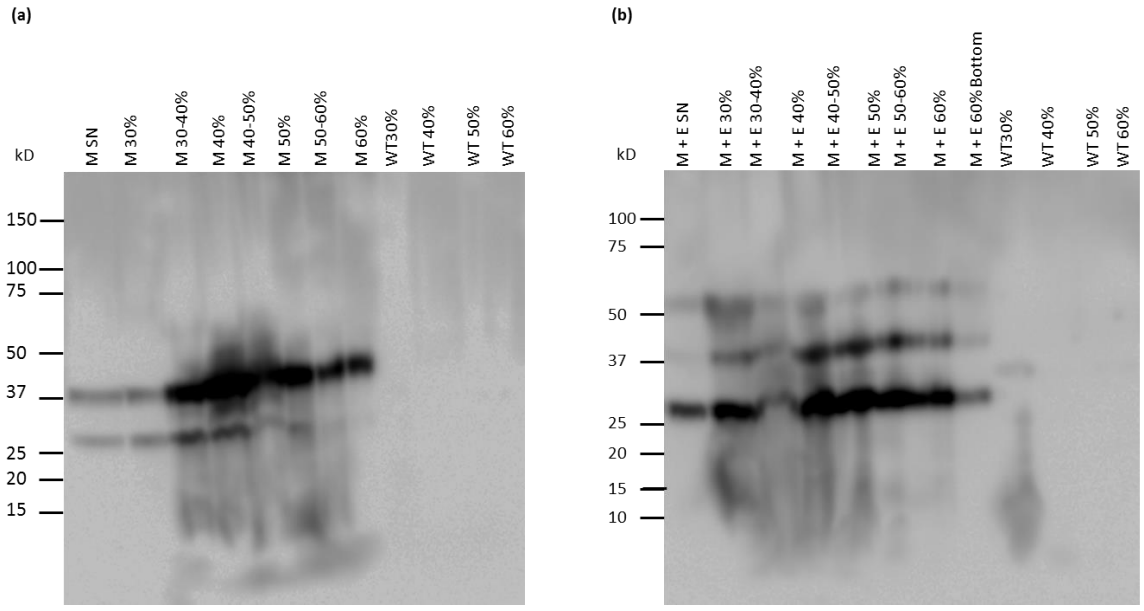


Figure 17. Immunoblot analysis of sucrose gradient fractions of M-ELP and M-ELP+E-ELP protein extracts.

(a) M-ELP is observed throughout all the fractions and in the supernatant. The M-ELP band in the 40% fraction appears heaviest. The expected band of 43 kDa is observed, as well as the cleavage product band just above 25 kDa. (b) Bands of the correct size are observed for M-ELP and E-ELP monomers, and the E-ELP dimer. The 40% fraction is observed to have the most protein overall.

domains, PEDv M may require these domains for inter-M interactions, as has been reported for MHV (de Haan et al., 1998)

Analyzing the sucrose gradient fractions of co-expressed M-ELP and E-ELP showed the combined banding pattern of the M-ELP monomer and the E-ELP monomer and dimer. The highest amount of total detected protein in this protein extract is also in the 40% fraction; the E-ELP and M-ELP monomers are heaviest at 40% sucrose, and the putative E-ELP dimer is heaviest at 30% sucrose (Figure 17b). The smallest band of the M-ELP/E-ELP sucrose gradient was analyzed, and found to contain both E-ELP, and the truncated M-ELP Spectral counts were used to determine relative protein amounts. In the 30% fraction, where the most truncated M-ELP would be expected, it was found that there was twice as much E-ELP than M-ELP present. This ratio likely increases in the heavier sucrose gradients based on the pattern observed in Figure 17a. The wild type extract showed no protein in most of the fractions, with the exception of the 30% fraction where, the 35 kDa peroxidase band was observed. As both M-ELP and M-ELP+E-ELP extracts showed the highest amount of protein at 40% sucrose, the same density where PEDv virions are expected to be found (Hofmann and Wyler, 1989), these fractions were the primary choice for TEM analysis.

3.2.3.2 VLP Analysis

The final question that needed to be addressed was whether M-ELP or the combination of M-ELP and E-ELP would assemble into VLPs. For this, the 40% fraction from the sucrose gradients of both treatments and the wild type were analyzed by TEM. The 40% fraction of M-ELP extract was found to contain circular particles (Figure 18a). While native virions, excluding spike projections are just larger than 100 nm, M VLPs for SARS were observed to be 50 nm in diameter (Tseng et al., 2010). The particles observed in Figure 18a were also about 50 nm in diameter. A distinct feature of coronaviruses is that M forms a lattice in the viral envelope, resulting in a membrane twice the thickness of typical biological membranes (4 nm) (Bárcena et al., 2009). The particles observed have an envelope thickness of approximately 9.3 nm. The size and envelope thickness of the particles indicates that these are VLPs formed with only the M protein of PEDv.

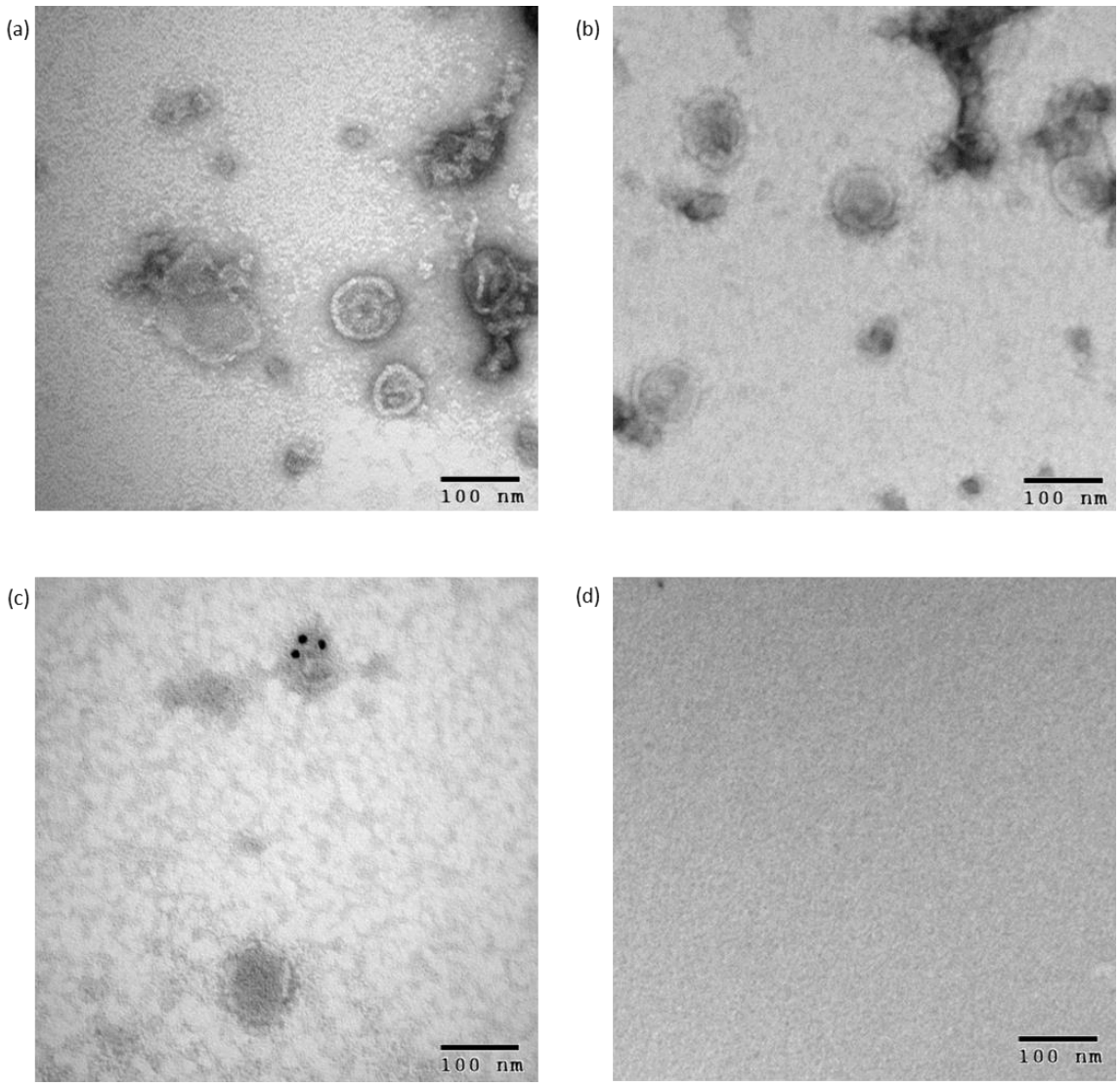


Figure 18. TEM and Immungold TEM analysis of 40% sucrose gradient fractions.

(a) M-ELP protein extract. VLPs observed are approximately 50 nm. (b) M-ELP+E-ELP protein extract. VLPs are observed to be approximately 80-85 nm in size. (c) Immunogold labelled M-ELP protein extract, using anti-c-Myc primary antibody. Gold particle bound secondary antibodies are able to bind to burst VLPs (evidenced by partial membrane structure), as the C-terminal of M-ELP is exposed. No antibody is able to bind to the intact VLP at the bottom of the picture. (d) Wild type tissue protein extract. No VLP structures observed.

Extracts of co-expressed M-ELP and E-ELP showed similar particles. Figure 18b shows that when the two proteins were co-expressed, VLPs were larger than with M-ELP alone. The larger size correlates to what has been presented in the literature, with MHV M+E VLPs having diameters of approximately 75 nm (Neuman et al., 2011). The VLPs observed in Figure 18b have a thick membrane of 10.1 nm, and are approximately 80 nm in diameter, suggesting they are VLPs. To provide further proof that the observed structures are VLPs containing M-ELP, immunogold TEM was done with the same M-ELP extract as above. Extracts had been frozen since previous grid preparation, and this reduced the total number of VLPs. The primary antibody used for detecting M-ELP is a commercial c-Myc tag antibody. This was not ideal, as the c-Myc tag is at the C-terminus, which folds inside the VLP. Previous reports have shown that using a C-terminal antibody on M results in no binding when virions are intact, but that binding can be observed on burst virions (Utiger et al., 1995). Similarly, it was observed that c-Myc antibody was able to bind to burst but not intact structures, giving further evidence that the structures are VLPs, and that protein is folding in the correct conformation (Figure 18c). Thus, VLPs are thought to have been produced for PEDv with expression of M and E, or with just M alone as ELP fusion proteins. These results may add to the evidence that coronavirus VLPs can be produced with only the M protein.

4 Discussion

4.1 PEDv vaccine design

The goal of this study was to produce a plant-made vaccine candidate to immunize pigs against PEDv. Previous work has focused on the S protein. As M is less prone to mutations (Sato et al., 2011), has an epitope (Zhang et al., 2012), and is involved in viral assembly (de Haan et al., 2000), M, and M based VLPs were the focus for this study's vaccine design. Future work can focus on optimizing N and S expression to then produce VLPs with all four proteins. The results show that high accumulation levels of M are attainable using plants as bioreactors with the use of a fusion tag, and that M-ELP is able to form VLPs independently, as well as together with E-ELP.

4.2 Transient production of M

One of the findings of this study was that the addition of the ELP tag was necessary to detect any accumulation of M transiently produced in *N. benthamiana*. ELP tags are not susceptible to proteolysis, and are hydrophilic and soluble (Raucher and Chilkoti, 2001). As such, ELP tags allow the accumulation of higher levels of M by either protecting the fused protein from proteolysis, or increasing solubility of the whole protein, potentially allowing for increased stability.

While the hydrophobin tag has worked in increasing recombinant protein accumulation previously (Joensuu et al., 2010), no protein was detected when M was fused with a hydrophobin tag, indicating increased accumulation is protein specific, as previously found (Pereira et al., 2014).

ELP has a positive effect on accumulation levels for the ER-targeted protein, but not for protein targeted to the cytoplasm, apoplast, or chloroplast (Conley et al., 2009b). As such, ELP was only fused to M in the construct targeting the ER. As no recombinant protein was detected when M was directed to the apoplast or ER without ELP, it is suggested that ELP is the main source of increased protein accumulation.

Having recombinant protein accumulate in the ER is beneficial. Targeting recombinant protein to the ER rather than the extracellular space in tobacco has shown to result in increased accumulation levels from 10 to 20 fold (Fiedler et al., 1997). It has been suggested that incorrectly folded proteins in the ER are degraded. This indicates that ER retention can promote correct protein folding, leading to less degradation and higher stability. Further, ER retention allows stable storage of recombinant proteins that retain their activity after lyophilisation (Fiedler et al., 1997). The ERGIC is where PEDv virions naturally bud off. Targeting M-ELP to the ER may assist in the interaction of M-ELP proteins, or M-ELP and E-ELP proteins, resulting in VLP formation.

M-ELP was found in both soluble and membrane-bound protein extracts (Figure 6). Protein found in the membrane-bound fraction is likely M-ELP that is retained in the ER, or VLPs that have not been secreted extracellularly, while soluble M-ELP may be detected due to the formation of soluble secreted VLPs.

Full length M has been difficult to express, and this is the first time it has been reported to be expressed in plants. Given the high levels of accumulation and the expanding knowledge of coronavirus replication, this may be an ideal protein to study membrane protein expression in plant cells, to better understand heterologous transmembrane segment identification by translocons, and to reach a better understanding of why some membrane proteins express better than others in eukaryotic systems.

4.3 Extraction of M-ELP

M-ELP was observed to form aggregates when boiled. When protein was not boiled before separation by PAGE, fewer high molecular weight aggregates and more protein at the monomer size was observed, as well as distinct dimer, trimer and tetramer size bands. Thermal aggregates have been observed for M of another coronavirus, SARS-CoV. Lee et al. (2005) found that boiling M caused aggregation in the stacking gel. The hydrophobic regions of SARS-CoV M were necessary for this thermal aggregation to occur, and the addition of reducing agent had no effect on aggregation. That the transmembrane domains of SARS-CoV M were needed for thermal aggregation may be related to roles the transmembrane domains have in protein interaction. However the

transmembrane domains also play a critical role in protein interaction for MHV, although this virus does not form thermal aggregates (de Haan et al., 2000; Lee et al., 2005). Further research may detail the mechanisms behind the observed aggregation.

Higher molecular weight bands were only observed in M-ELP extracts when EMD Millipore's ProteoExtract® Native Membrane Protein Extraction Kit was used. Formation of discrete M oligomers has also been shown for HCoV 229E, where reduction-sensitive complexes with molecular weights corresponding to two, four, and eight copies of M were reported (Arpin and Talbot, 1990). Given the importance of M-M interactions on virion formation, the formation of these complexes is logical. The cysteine in M was hypothesized to play a role in forming intermolecular disulfide bridges. PEDv M protein has two cysteines, at positions 85 and 129 (Figure 11). The second cysteine is likely significant, as it is on the cytoplasmic tail, which was proven to be crucial in VLP formation (de Haan et al., 1998). PEB, DSB and FEB contain sodium L-ascorbate, which can reduce disulfide bridges (Giustarini et al., 2008), while the commercial kit does not have a similar reagent. The lack of a compound that can break disulfide bridges may allow the formation of M-ELP complexes. While all protein extracts were stored in sample buffer with DTT, a reducing agent, the reducing ability of DTT is limited at pH 7 or below. Both extraction buffers of the commercial kit have a pH of 7.2, at which point the activity of DTT may be suboptimal, explaining the detection of higher molecular weight complexes. This pH may also be too close to the isoelectric point (pI) of M-ELP, 6.65. When the pH of a solution is the same as a protein's pI, the protein is no longer charged, and will aggregate and precipitate (Xia, 2007). To avoid precipitation of recombinant protein, a final pH of 8 was used for FEB.

Extracting membrane proteins is difficult (Carpenter et al., 2008), and established protocols are not focused on extracting membrane proteins from plants (Arnold and Linke, 2008). Thus it was necessary to have an optimized extraction buffer to accurately quantitate M-ELP accumulating in *N. benthamiana* leaves. Detergents are known to be key components of the extraction buffer for membrane proteins (Arnold and Linke, 2008). This study found that the addition of detergent made a difference of 0.798 mg/g of protein detected per fresh weight of leaf, a 350x increase over an extraction buffer with

no detergent (Table 7). Non-ionic detergents often strike the middle ground between not denaturing proteins, and efficiency in covering the hydrophobic surface of membrane proteins. The commercial kit uses 0.5% Triton X-100 as its detergent. Triton X-100 is widely used as a non-denaturing detergent for membrane proteins, including previously for PEDv M (Shenyang et al., 2007), and has applications in phase separation. Triton X-100 is also used in VLP extraction protocols (Tseng et al., 2010). To keep detergent use consistent across experiments, and for all the above reasons, Triton X-100 was used in further experiments. Triton X-100 is able to solubilize protein complexes, increasingly so with incubation (Dencher and Heyn, 1978), and is known for its ability to solubilize the inner cell membrane of Gram-negative bacteria (Arnold and Linke, 2008). Both these properties may have helped in solubilizing ER-bound M-ELP.

In contrast to the commercial kit, FEB contains antioxidant/reducing agents. Antioxidants prevent the oxidation of both proteins and phenolics. Antioxidants reduce the activity of polyphenol oxidases. This prevents the oxidation of polyphenols, which can bind to proteins, forming complexes and changing their structure (Laing and Christeller, 2004; Ozdal et al., 2013). When protein was extracted with the commercial kit, polyphenols may have bound to and inactivated proteins, potentially making the *c-Myc* tag inaccessible. FEB contains sodium L-ascorbate and PVPP, which binds to polyphenols, to alleviate this problem.

Future work can explore the effect of adding DTT to the extraction buffer rather than in the sample buffer to prevent oxidation earlier, adding a larger range of protease inhibitors and the effects of sonicating protein extracts, using sound energy to break cell walls (Tang et al., 2002).

4.3.1 Peroxidase

Peroxidases are found in plants, animals, and microorganisms, and carry out a variety of roles in different physiological processes (Hiraga et al., 2001). One peroxidase that has been harnessed for its usefulness in research is horseradish peroxidase (HRP). HRP can catalyze the oxidation of luminol to 3-aminophthalate via intermediates. In the presence of substrate, HRP allows production of a detectable light signal. This application is useful

in Western blots. The primary antibody binds to a protein fixed to a specific spot on a membrane. When the secondary antibody is conjugated to HRP, it allows for the detection of the target protein upon incubation with a luminescent substrate (reviewed by Gazaryan et al., 1998). The anionic peroxidase (TOP) of *N. tabacum* was found to mimic HRP's activity in this chemiluminescent reaction (Gazaryan et al., 1998). A BLAST search of the draft *N. benthamiana* genome for predicted proteins (Bombarely et al., 2012) using the published sequence of *N. tabacum*'s anionic peroxidase (accession number: L02124) found one predicted protein, with a 93.73% sequence similarity to TOP, peroxidase 53. In comparison, TOP is 52% homologous to HRP, and retains function (Lagrimini et al., 1987). This indicates that peroxidase 53 may be the homologue to TOP in *N. benthamiana*. As peroxidases have conserved binding sites, peroxidase 53 is likely to retain function. Peroxidase 53 is predicted to have a molecular weight of 34.85 kDa – the same size of the detected peroxidase (Figure 4). In addition, more peroxidase 53 is observed in the soluble protein extract in comparison to the membrane protein extract (Figure 7). This correlates to data showing TOP is secreted extracellularly (Lagrimini et al., 1987).

Tobacco peroxidases are involved in lignin formation, as well as in pathogen response. Anionic isozymes are upregulated in leaves infected with tobacco mosaic virus, though not in leaves that are wounded (Lagrimini and Rothstein, 1987). Overexpression of two anionic peroxidases which differ in post translational modifications in *N. tabacum* was observed to cause severe wilting (Lagrimini et al., 1990). E-ELP lanes show more accumulation of a peroxidase that runs between 50 and 75 kDa than any of the other PEDv proteins (Figure 13), and these leaves show yellowing (Figure 14). This peroxidase may be upregulated in response to E-ELP expression, and thus causing yellowing of the leaf tissue. This peroxidase of molecular weight about 60 kDa may be either a dimer of peroxidase 53, or a different peroxidase – but nonetheless it is correlated with the appearance of yellowing leaves. E causes ER stress in swine epithelial cells (Xu et al., 2013); it may cause ER stress in plant leaves as well, and this ER stress may cause the upregulation of the higher molecular weight peroxidase. When M-ELP and E-ELP are co-expressed the amount of higher molecular weight peroxidase is reduced. E-ELP in VLPs may not cause as much ER stress. The infiltration process and viral components of

the expression vector also likely stimulate pathogen response, explaining why this higher molecular weight band was only seen in lanes of infiltrated leaves.

After probing, membranes showed a similar sized 50-75 kDa band for E-ELP extracts, thought to be the dimer of E-ELP. To clarify whether this band was the putative E dimer or the above-mentioned peroxidase, the intensity of the light emitted from the membrane before and after probing was examined. Comparing Figures 13a and b, peroxidase bands that showed heavily before probing were in comparison faint afterwards. In addition, the amount of light emitted after probing was five times higher, despite being imaged for a sixth of the time, demonstrating that plant peroxidase was not emitting as strong a signal as HRP. Finally, mass spectrometry was performed on the putative dimer band, and E-ELP was detected, indicating the dark band was due to E-ELP detection. Further credence is given knowing that TOP has a very low detection limit, does not need an enhancer substrate, and has a ten-fold higher signal intensity than HRP (Gazaryan et al., 1998). The weak band observed in Figure 13a is likely a result of low amounts of peroxidase, as at pH 9, the pH of the ECL substrate used, TOP gives a higher light intensity than HRP (Hushpulian et al., 2007). While it must be considered that the ECL substrate used is optimized for HRP activity, TOP shows higher light intensity at all luminol concentrations, and higher stability in high H_2O_2 conditions (Gazaryan et al., 1998), indicating weakened activity in comparison to HRP is unlikely because of buffer conditions.

4.4 PEDv S, N, and E, and Co-Expression

Accumulation of the S protein was not detected. This is not surprising, as others tried and failed to express it in plants (Zhu, Kaldis, Menassa, personal communication). S is a very large protein of calculated molecular weight of 169.4 kDa. All *in planta* work on expression of PEDv proteins has focussed on producing epitopes of the S protein (see 1.1.3.3), and full length S protein has not been produced in plants.

N-ELP has a predicted molecular weight of 66.6 kDa. While faint bands were observed in N-ELP samples, detected proteins were between 37 and 50 kDa. The presence of two distinct C-terminal fragments indicates that the fusion protein was likely produced, but

that there may be two protease cleavage sites. Testing a broader range of protease inhibitors in the extraction buffer may allow the detection of full length N-ELP in the future.

Recombinant expression of PEDv E was successful, and E-ELP was observed at the predicted molecular weight of 26 kDa, and as a dimer slightly above 50 kDa. This was ideal as E has been necessary for VLP formation for other coronaviruses (described in 4.5) (Vennema et al., 1996). Additionally, antibodies raised against MHV E neutralize viral infectivity in the presence of complement (Yu et al., 1994). The effect of PEDv E-GFP has been monitored on porcine intestinal epithelial cells. PEDv E-GFP was observed to cause upregulation of Bcl-2, an anti-apoptotic molecule associated with cell survival. E-GFP was also observed to cause ER stress in the porcine cells however, and this may relate to the yellowing phenomenon seen after infiltration (Figure 14) (Xu et al., 2013).

E-ELP was detected as both a monomer and dimer. Coronavirus E proteins demonstrate ion channel activity. However, with only one hydrophobic domain E would need to oligomerize to act as an ion channel (Ruch and Machamer, 2012). SARS-CoV E was predicted to form pentamers, and synthetic chemically produced SARS-CoV E was detected as a dimer, trimer, and pentamer (Torres et al., 2005). Oligomerization of biologically produced SARS-CoV E in the form of dimers and trimers was only observed under non-reducing conditions, indicating interaction through disulfide bonds (Liao et al., 2004). IBV E also can form oligomers, and these dissociate into monomers when exposed to detergent or reducing agent (Westerbeck and Machamer, 2015).

In this study E-ELP is observed as a monomer and dimer under detergent and reducing conditions in the extraction buffer and sample buffer. This gives further insight into the oligomerization of E. There is potential that the addition of ELP may have a stabilizing effecting on the E oligomer. If this is the case, it is unclear why the dimer is clearly visible instead of the trimers or pentamers observed in literature.

4.5 Virus-Like Particles

4.5.1 Formation Requirements

The minimum requirement for the formation of coronavirus-like particles has been the subject of ongoing research. With all reported VLPs requiring M, and most needing E as well, it was convenient that the two proteins that expressed well in this study were M-ELP and E-ELP. As PEDv VLPs have not been previously reported, the minimum requirements were unknown. This study's results show PEDv M is sufficient for VLP assembly. VLPs were also observed when the M-ELP and E-ELP were co-expressed. This is only the third time CoVLPs have been reported to be formed with just M.

Research on TGEV had shown that VLPs could be formed without the S or N protein, indicating that the M and E proteins were both necessary and sufficient for VLP formation (Baudoux et al., 1998). For MHV it was found that M protein was not released into the medium unless it was co-expressed with E, using a vaccinia viral vector and OST7-1 (Mouse L) cells (de Haan et al., 1998), but that N and S are dispensable. This was striking as E is not abundant in the viral envelope – estimated to be found only twenty times in TGEV virions (Godet et al., 1992). Thus, rather than E forming an important part of the viral structure, it was thought that E was necessary in specific parts of the envelope to induce curvature in the intracellular membranes, resulting in particle formation (Vennema et al., 1996).

However, research on SARS-CoV contradicted previous research on CoVLPs. This may be because SARS-CoV M is not as closely related to TGEV or MHV (Huang et al., 2004), or potentially because it has a unique method of coronavirus assembly. While a report found M and E sufficient for SARS-CoVLP assembly (Ho et al., 2004), corroborating previous coronavirus research, a conflicting report claimed co-expression of M and N was both sufficient and necessary for VLP formation. However, these VLPs were observed intracellularly, and did not bud off unless S was co-expressed (Huang et al., 2004). While the study's findings that M and N could form VLPs were an interesting advancement, the authors did not co-express M and E, making their claims about N being necessary unsatisfying. They did however express M alone, and found no VLPs were

formed. This potentially indicates an expression partner with M may be necessary for VLP assembly. That too was dispelled when SARS-CoVLPs were formed with the expression of only the M protein (Tseng et al., 2010). Recently it was found that expressing TGEV M alone also resulted in the formation of VLPs (Zhenhui et al., 2015), and that VLPs for IBV could be produced by co-expressing the M and S proteins (Liu et al., 2013) indicating that the formation requirements for CoVLPs are not as well understood as previously thought. Table 9 presents a brief non-exhaustive summary of CoVLP studies.

E has also been found to release into sedimentable particles. Using sedimentation analysis+ it was found that E is released into vesicles from MHV in infected cells and when independently expressed, and in both cases sedimented at the same sucrose density. This density was lighter than that of MHV. Using flotation analysis, it was determined that E was an integral membrane protein in vesicles, though these were not visualized (Maeda et al., 1999). E has been reported to be released into sedimentable particles for IBV as well (Corse and Machamer, 2000). This emphasizes that E may be needed for VLP formation, as it may play an important role in the budding process of coronaviruses. For both MHV and IBV the efficiency of release was very low, and sedimented particles were not visualized. As M forms the characteristic structure of coronaviruses, these vesicles will be distinct from CoVLPs (see 4.5.2). Further research must be done to distinguish the relationship between E released in vesicles and E oligomers (see 4.5.4).

Why other studies were only able to form VLPs when both M and E were co-expressed may be due to the system used for expression (i.e. animal cells vs plant cells). As research on SARS-CoV has conflicting reports on the need for E co-expression, differences in minimal requirements for VLP formation are likely not because of a difference in the viral assembly process, but more likely due to differences in experimental methods. Of the two studies that used HEK293T cells to produce SARS-CoVLPs, one reports the assembly of M VLPs the other does not (Table 9). The two studies utilized the same transfection methods, cells, and similar growth mediums, and the study that found M VLPs obtained its expression vector from the lab that did not. The major difference then was the amount of plasmid transfected – the study that found M

Table 9. Protein expression configurations resulting in CoVLP assembly.

Virus	Protein(s) (Co-)Expressed	VLPs Observed	Expression Platform	Reference
MHV	ME/MES	+	OST7-1[1]	(Vennema et al., 1996)
	M, S, MS	-		
TGEV	ME	+	RK13 cells[2]	(Baudoux et al., 1998)
SARS-CoV	ME/MES	+	Sf21[3]	(Ho et al., 2004)
	M	-		
SARS-CoV	MN (didn't bud off)/MNS/MNSE	+	HEK293T[4]	(Huang et al., 2004)
	N/S/M	-		
SARS-CoV	ME/MES	+	Sf9[5]	(Mortola and Roy, 2004)
	M/E/S/MS/MESN	-		
SARS-CoV	M/MN	+	HEK293T[4]	(Tseng et al., 2010)
IBV	MS	+	Sf9[5]	(Liu et al., 2013)

[1]OST7-1, mouse L cells (mouse embryo fibroblast cells transformed with Maloney Sarcoma Virus and T7 RNA polymerase gene); [2]Rk13 cells, rabbit kidney cells; [3]Sf21, fall armyworm ovary cells; [4]HEK293T, human embryonic kidney cells stably expressing SV40 large T antigen; [5]Sf9, fall armyworm ovary cells.

VLPs transfected six times more DNA. Other differences include trypsinized cells before transfection, and concentrated VLPs before TEM analysis (Huang et al., 2004; Tseng et al., 2010). From this comparison, the difference in VLP detection may be because of amount of transfected DNA or cell transfection and handling methods. This would impact either VLP formation, or detection.

The insertion of a histidine tag after the initiating methionine has previously been shown to "strongly impair" VLP formation for MHV, suggesting that only minor insertions were permissible at the N-terminus (de Haan et al., 1998). The M-ELP fusion protein in this study contained a cleaved signal peptide after the initiating methionine, followed by an Xpress immunodetection tag, a restriction enzyme site, and a Gateway® cloning site before the amino acid sequence of the M protein. VLPs were still able to form, indicating that for PEDv, insertion restrictions may not be as severe. It remains to be seen if removing all insertions at the N-terminus can increase the number of VLPs formed. Use of an N-terminal signal peptide to target M-ELP and E-ELP to the ER may have helped with VLP formation, as coronaviruses typically bud off from the ERGIC (de Haan et al., 2000).

4.5.2 CoVLP Structure

A distinct feature of coronaviruses is the thickness of the viral envelope, due to the matrix of M proteins in the envelope (Bárcena et al., 2009). The PEDv VLPs observed have the characteristic thick coronavirus envelope, averaging 9.3 nm and 10.1 nm, for M-ELP and M-ELP/E-ELP respectively (Figure 18), more than double the average thickness of biological membranes, approximately 4 nm. The envelope of MHV and TGEV are 7.8 ± 0.7 nm and 7.4 ± 0.6 nm respectively (Bárcena et al., 2009), and this characteristic is observed for coronavirus-like particles (Neuman et al., 2011), as demonstrated in this study.

4.5.3 CoVLP Size

The results of this study show that VLPs produced with M-ELP and E-ELP are approximately 80 nm and VLPs with just M-ELP are smaller, approximately 50 nm. This is the first report comparing the size of ME VLPs with M VLPs. Both are smaller than

the average PEDv virion size, which has been reported to have a diameter ranging from 60 nm (Ducatelle et al., 1982) to 190 nm, or just a 150 nm diameter (Hofmann and Wyler, 1989). The characteristic projections of the spike protein make up a large part of the coronavirus diameter, measuring 18 nm in length (Pensaert and De Bouck, 1978). While there is no existing data on PEDv VLPs, a look at the VLPs produced of other coronaviruses can provide insights on PEDv VLPs. SARS-CoV M VLPs were mostly found to be 40-50 nm in diameter, while MN VLPs were mostly found to have a diameter of 60-70 nm (Tseng et al., 2010). The absence of the nucleocapsid protein resulted in smaller VLPs. SARS-CoV virions typically have a diameter of 100 nm, indicating that MN VLPs may just be smaller by the size of the spikes on the virion surface.

A similar result was reported for MHV. MSE VLPs have a slightly smaller diameter than natural virions, indicating presence of N may result in a larger size particle. This group also found not expressing S with ME made no difference to the particle diameter, (excluding spikes) (Vennema et al., 1996). However, another study on MHV reported that ME VLPs were bigger than MEN VLPs, and that both were larger than MHV virions with the full complement of proteins. The authors suggested that the interactions between the proteins led to smaller virions (Neuman et al., 2011).

The lattice of M proteins is thought to have specific gaps for S and E proteins to integrate into the envelope. The absence of these interactions may result in a smaller diameter, as these gaps are closed due to lack of protein interaction. Similarly, lack of M-N interactions may also result in the smaller diameter of VLPs observed in this study. Why ME VLPs of the same coronavirus are observed to be either bigger or smaller than the corresponding virions may be correlated to why some groups have been able to produce VLPs with different minimum protein requirements. The potentially differing nature of M’s interactions may give rise to both differing particle sizes and what proteins are required to form VLPs.

4.5.4 CoVLP Sedimentation Analysis

Sucrose gradients are used for partial purification of VLPs, and for sedimentation analysis. PEDv M-ELP VLPs most heavily concentrated in the 40% sucrose fraction

(Figure 17a). PEDv virions have a density of 1.18 g/ml in sucrose (Hofmann and Wyler, 1989), equivalent to approximately 40% sucrose, and as such, PEDv M tightly packs into VLPs that are similar to whole virions. Previously, SARS-CoV M VLPs were analyzed and found to concentrate at a density of 1.13 g/ml, corresponding to just over 30% sucrose (all % sucrose values in w/v) (Tseng et al., 2010). Unlike what is often observed for other CoVLPs, sedimentation analysis in this study did not result in recombinant protein only being detected in a few sucrose gradient fractions. M-ELP was found to sediment into every fraction. This likely indicates that the VLPs produced are very pleomorphic, similar to natural PEDv virions (Hofmann and Wyler, 1989). An alternate explanation may be that the accumulation of M-ELP was too high to allow for all the protein to efficiently bud into VLPs. The presence of M-ELP in the supernatant in particular may indicate that not all protein was incorporated into particles, leading to protein that did not sediment from the supernatant in particles, or may indicate VLPs formed that were not dense enough to enter the 30% sucrose fraction.

When M-ELP and E-ELP were co-expressed, presence of M-ELP in the supernatant was reduced (Figure 17b). While the results show that M-ELP can independently form particles, the presence of E-ELP may allow for increased incorporation of the protein into VLPs. Fine tuning the ratios at which E-ELP and M-ELP are expressed may increase VLP assembly. Comparing levels of both M-ELP and the E-ELP monomer in the sucrose gradient fractions reveals that M-ELP/E-ELP VLPs concentrate most in the 40% fraction, like M-ELP VLPs and PEDv virions. TGEV ME VLPs have been compared to TGEV virions, and found to accumulate most highly in 27.8% and 33.9% sucrose respectively (Baudoux et al., 1998). Similarly MHV virions have been found to concentrate at the interface between 40% and 50% sucrose gradients, while ME VLPs concentrated at the 30% and 40% interface (Vennema et al., 1996). With the discontinuous sucrose gradient used in this study, if VLPs and virions concentrated closer to 40% or 50% sucrose respectively, they would still be found in the same gradient. A continuous sucrose gradient may be used to compare the densities between VLPs and virions in a future study.

While the E-ELP monomer in this study accumulates most highly at the same point as M-ELP, the E-ELP dimer peaks in the 30% sucrose fraction. E vesicles for MHV are observed at a density of 1.13 g/ml (Corse and Machamer, 2000), equivalent to 30% sucrose. This may indicate that E-ELP dimers bud off into vesicles with a smaller density, as was proposed for MHV. Alternatively, this may represent the sedimentation of higher molecular weight oligomers. Sedimentation analysis of IBV E on a 5 to 20% sucrose gradient showed an increase in IBV E oligomers at 20% sucrose (Westerbeck and Machamer, 2015). A wider gradient may have indicated E oligomers concentrating at a lighter gradient. The observation of E oligomers may be necessary for the release of E vesicles. IBV E oligomers were proposed to play a role in IBV MNE VLP formation, and IBV E monomers were not involved in this process (Westerbeck and Machamer, 2015). While in this study, E-ELP monomers sedimented at the same density as M-ELP, indicating the presence of E-ELP monomers in M-ELP/E-ELP VLPs, the presence of E-ELP dimers in a lighter sucrose fraction potentially indicates that the dimers are driving release of E-ELP only vesicles. TEM analysis of extracts with just E is needed to determine the true nature of sedimented E protein.

4.5.5 CoVLP Immunogenicity

The ability to produce PEDv VLPs has direct implications on immunogenicity. The importance of VLP formation was demonstrated in a comparison of TGEV VLPs to TGEV proteins not in VLPs on the ability to induce alpha interferons (IFN-α), Type I interferons. IFN-α are an important component of the immune system, inducing the differentiation of APCs, and enhancing T- and B-cell functions (reviewed by Rizza et al., 2011). Neither TGEV M in detergent-solubilized material nor TGEV M and S in protein-liposome mixtures were able to induce interferons. However, TGEV ME VLPs were able to consistently induce IFN-α. As neither M nor E have a sequence motif for interferon induction, it is predicted that induction is caused by a structural feature of the VLPs (Baudoux et al., 1998). Classical adjuvants, such as complete Freund's adjuvant, work in a process mediated by Type I interferons (reviewed by Rizza et al., 2011). The ability of CoVLPs to induce IFN-α may mean that these VLPs can be administered independently without the need of an adjuvant, and underline the importance of the specific VLP

structure. This could explain why even for other viruses the antibody response of VLPs is effective against a broader set of viral isolates, and results in a higher antibody titre than administering recombinant protein not in VLPs (Bright et al., 2007). CoVLPs have been demonstrated to effectively induce immune responses. IBV VLPs administered without adjuvants induce humoral immune responses comparable to inactivated IBV vaccine, and cellular immune responses that are significantly higher than those for inactivated IBV vaccine (Liu et al., 2013). The immunological importance of VLPs is clear, and animal trials are necessary to show if PEDv VLPs have similar immunogenicity. The successful production of PEDv VLPs gives hope for a low-cost effective PEDv vaccine that can be used to combat the disease worldwide.

5 Outlook

The production of the first reported VLPs for PEDv, and the first plant-produced coronavirus-like particles is an exciting step forward for both efforts to curb the spread of PED and research on molecular farming. The combination of the conserved nature of the M protein (Zhang et al., 2012) and the immunogenic potential of VLPs to be effective against multiple viral strains (D'Aoust et al., 2008) indicate that M-ELP+E-ELP VLPs may provide an ideal vaccine solution for multiple markets. The quick turnaround time of the plant-based molecular farming platform means that even when genetically dissimilar PEDv strains arise, custom PEDv VLPs can rapidly be produced.

To determine the viability of using the PEDv VLPs as vaccines, animal trials must be carried out. Future work may also include fine tuning the *A. tumefaciens* ratios of M-ELP and E-ELP to optimize VLP assembly, and troubleshooting the production of N-ELP and S-ELP in order to produce VLPs with all four structural proteins. Immunogold TEM using N-terminal antibodies for all three proteins embedded in the envelope can determine whether the proteins are being incorporated into VLPs.

References

Arndt, A.L., Larson, B.J. and Hogue, B.G. (2010) A conserved domain in the coronavirus membrane protein tail is important for virus assembly. *Journal of Virology* **84**, 11418-11428.

Arnold, T. and Linke, D. (2008) The use of detergents to purify membrane proteins. *Current Protocols in Protein Science* **Chapter 4**, Unit 4.8.

Arpin, N. and Talbot, P.J. (1990) Molecular characterization of the 229E strain of human coronavirus. In: *Coronaviruses and their Diseases* pp. 73-80. New York: Springer.

Ayorinde, F., Gelain, S.V., Johnson, J. and Wan, L.W. (2000) Analysis of some commercial polysorbate formulations using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **14**, 2116-2124.

Bae, J.-L., Lee, J.-G., Kang, T.-J., Jang, H.-S., Jang, Y.-S. and Yang, M.-S. (2003) Induction of antigen-specific systemic and mucosal immune responses by feeding animals transgenic plants expressing the antigen. *Vaccine* **21**, 4052-4058.

Bárcena, M., Oostergetel, G.T., Bartelink, W., Faas, F.G., Verkleij, A., Rottier, P.J., Koster, A.J. and Bosch, B.J. (2009) Cryo-electron tomography of mouse hepatitis virus: insights into the structure of the coronavirion. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 582-587.

Baudoux, P., Carrat, C., Besnardeau, L., Charley, B. and Laude, H. (1998) Coronavirus pseudoparticles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes. *Journal of Virology* **72**, 8636-8643.

Bombarely, A., Rosli, H.G., Vrebalov, J., Moffett, P., Mueller, L.A. and Martin, G.B. (2012) A draft genome sequence of *Nicotiana benthamiana* to enhance molecular plant-microbe biology research. *Molecular Plant-Microbe Interactions* **25**, 1523-1530.

Bowie, J.U. (2005) Solving the membrane protein folding problem. *Nature* **438**, 581-589.

Bright, R.A., Carter, D.M., Daniluk, S., Toapanta, F.R., Ahmad, A., Gavrilov, V., Massare, M., Pushko, P., Mytle, N. and Rowe, T. (2007) Influenza virus-like particles elicit broader immune responses than whole virion inactivated influenza virus or recombinant hemagglutinin. *Vaccine* **25**, 3871-3878.

Brusky, T. (2016) U of S VIDO-InterVac develops vaccine for devastating pig virus http://news.usask.ca/media-release-pages/2016/u-of-s-vido-intervac-develops-vaccine-for-devastating-pig-virus-.php. Accessed: September 1 2016

Bünger, S., Roblick, U.J. and Habermann, J.K. (2009) Comparison of five commercial extraction kits for subsequent membrane protein profiling. *Cytotechnology* **61**, 153-159.

Calvo-Pinilla, E., Castillo-Olivares, J., Jabbar, T., Ortego, J., de la Poza, F. and Marín-López, A. (2014) Recombinant vaccines against bluetongue virus. *Virus Research* **182**, 78-86.

Carpenter, E.P., Beis, K., Cameron, A.D. and Iwata, S. (2008) Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology* **18**, 581-586.

Chang, S.H., Bae, J.L., Kang, T.J., Kim, J., Chung, G.H., Lim, C.W., Laude, H., Yang, M.S. and Jang, Y.S. (2002) Identification of the epitope region capable of inducing neutralizing antibodies against the porcine epidemic diarrhea virus. *Molecules and Cells* **14**, 295-299.

Chen, J., Wang, C., Shi, H., Qiu, H., Liu, S., Chen, X., Zhang, Z. and Feng, L. (2010) Molecular epidemiology of porcine epidemic diarrhea virus in China. *Archives of Virology* **155**, 1471-1476.

Chen, Q., Li, G., Stasko, J., Thomas, J.T., Stensland, W.R., Pillatzki, A.E., Gauger, P.C., Schwartz, K.J., Madson, D., Yoon, K.J., Stevenson, G.W., Burrough, E.R., Harmon, K.M., Main, R.G. and Zhang, J. (2014) Isolation and characterization of porcine epidemic diarrhea viruses associated with the 2013 disease outbreak among swine in the United States. *Journal of Clinical Microbiology* **52**, 234-243.

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

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

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

Consortium, T.U. (2015) UniProt: a hub for protein information. *Nucleic Acids Research* **43**, D204-D212.

Corse, E. and Machamer, C.E. (2000) Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles. *Journal of Virology* **74**, 4319-4326.

Covey, S.N., Covey, S.N., Lomonossoff, G.P., Lomonossoff, G.P., Hull, R. and Hull, R. (1981) Characterisation of cauliflower mosaic virus DNA sequences which

encode major polyadenylated transcripts. *Nucleic Acids Research* **9**, 6735-6748;6747;.

Craig, R. and Beavis, R.C. (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **20**, 1466-1467.

Cruz, D.J., Kim, C.J. and Shin, H.J. (2008) The GPRLQPY motif located at the carboxy-terminal of the spike protein induces antibodies that neutralize porcine epidemic diarrhea virus. *Virus Research* **132**, 192-196.

Cseke, L.J., Kirakosyan, A., Kaufman, P.B. and Westfall, M.V. (2011) *Handbook of Molecular and Cellular Methods in Biology and Medicine, 3rd ed.* Boca Raton, Florida:CRC Press.

Czyz, M., Dembczynski, R., Marecik, R., Wojas-Turek, J., Milczarek, M., Pajtasz-Piasecka, E., Wietrzyk, J. and Pniewski, T. (2014) Freeze-drying of plant tissue containing HBV surface antigen for the oral vaccine against hepatitis B. *BioMed Research International* **2014**, 485689.

D'Aoust, M.A., Lavoie, P.O., Couture, M.M.J., Trépanier, S., Guay, J.M., Dargis, M., Mongrand, S., Landry, N., Ward, B.J. and Vézina, L.P. (2008) Influenza virus-like particles produced by transient expression in *Nicotiana benthamiana* induce a protective immune response against a lethal viral challenge in mice. *Plant Biotechnology Journal* **6**, 930-940.

de Haan, C.A.M., Kuo, L., Masters, P.S., Vennema, H. and Peter, J.M.R. (1998) Coronavirus Particle Assembly: Primary Structure Requirements of the Membrane Protein. *Journal of Virology* **72**, 6838-6850.

de Haan, C.A.M., Vennema, H. and Peter, J.M.R. (2000) Assembly of the coronavirus envelope: homotypic interactions between the M proteins. *Journal of Virology* **74**, 4967-4978.

Dencher, N. and Heyn, M. (1978) Formation and properties of bacteriorhodopsin monomers in the non-ionic detergents octyl-β-D-glucoside and Triton X-100. *FEBS Letters* **96**, 322-326.

Duarte, M., Tobler, K., Bridgen, A., Rasschaert, D., Ackermann, M. and Laude, H. (1994) Sequence analysis of the porcine epidemic diarrhea virus genome between the nucleocapsid and spike protein genes reveals a polymorphic ORF. *Virology* **198**, 466-476.

Ducatelle, R., Coussement, W., Debouck, P. and Hoorens, J. (1982) Pathology of experimental CV777 coronavirus enteritis in piglets. II. Electron microscopic study. *Veterinary Pathology* **19**, 57-66.

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

Fan, J.-H., Zuo, Y.-Z., Shen, X.-Q., Gu, W.-Y. and Di, J.-M. (2015) Development of an enzyme-linked immunosorbent assay for the monitoring and surveillance of antibodies to porcine epidemic diarrhea virus based on a recombinant membrane protein. *Journal of Virological Methods* **225**, 90-94.

Fiedler, U., Phillips, J., Artsaenko, O. and Conrad, U. (1997) Optimization of scFv antibody production in transgenic plants. *Immunotechnology* **3**, 205-216.

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

Garavito, R.M. and Ferguson-Miller, S. (2001) Detergents as tools in membrane biochemistry. *Journal of Biological Chemistry* **276**, 32403-32406.

Gazaryan, I.G., Rubtsova, M.Y., Kapeliuch, Y.L., Rodriguez-Lopez, J.N., Lagrimini, L.M. and Thorneley, R.N. (1998) Luminol Oxidation by Hydrogen Peroxide Catalyzed by Tobacco Anionic Peroxidase: Steady-State Luminescent and Transient Kinetic Studies. *Photochemistry and Photobiology* **67**, 106-110.

Giustarini, D., Dalle-Donne, I., Colombo, R., Milzani, A. and Rossi, R. (2008) Is ascorbate able to reduce disulfide bridges? A cautionary note. *Nitric Oxide* **19**, 252-258.

GlaxoSmithKline (2014) Cervarix® product monograph. http://ca.gsk.com/media/589880/cervarix.pdf. Accessed: September 1 2016

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

Godet, M., L'Haridon, R., Vautherot, J.-F. and Laude, H. (1992) TGEV corona virus ORF4 encodes a membrane protein that is incorporated into virions. *Virology* **188**, 666-675.

Grgacic, E.V.L. and Anderson, D.A. (2006) Virus-like particles: passport to immune recognition. *Methods* **40**, 60-65.

Grisshammer, R. (2006) Understanding recombinant expression of membrane proteins. *Current Opinion in Biotechnology* **17**, 337-340.

Harrisvaccines, I. (2015) SirraVax Platform. http://www.harrisvaccines.com/en/technology/sirravax_platform. Accessed: April 10 2015

Hartley, J.L., Temple, G.F. and Brasch, M.A. (2000) DNA cloning using *in vitro* site-specific recombination. *Genome Research* **10**, 1788-1795.

He, Y., Zhou, Y., Siddiqui, P., Niu, J. and Jiang, S. (2005) Identification of immunodominant epitopes on the membrane protein of the severe acute respiratory syndrome-associated coronavirus. *Journal of Clinical Microbiology* **43**, 3718-3726.

Hiraga, S., Sasaki, K., Ito, H., Ohashi, Y. and Matsui, H. (2001) A large family of class III plant peroxidases. *Plant and Cell Physiology* **42**, 462-468.

Ho, Y., Lin, P.-H., Liu, C.Y.Y., Lee, S.-P. and Chao, Y.-C. (2004) Assembly of human severe acute respiratory syndrome coronavirus-like particles. *Biochemical and Biophysical Research Communications* **318**, 833-838.

Hofmann, M. and Wyler, R. (1989) Quantitation, biological and physicochemical properties of cell culture-adapted porcine epidemic diarrhea coronavirus (PEDV). *Veterinary Microbiology* **20**, 131-142.

Huang, Y., Yang, Z.-Y., Kong, W.-P. and Nabel, G.J. (2004) Generation of synthetic severe acute respiratory syndrome coronavirus pseudoparticles: implications for assembly and vaccine production. *Journal of Virology* **78**, 12557-12565.

Huang, Y.W., Dickerman, A.W., Pineyro, P., Li, L., Fang, L., Kiehne, R., Opriessnig, T. and Meng, X.J. (2013) Origin, evolution, and genotyping of emergent porcine epidemic diarrhea virus strains in the United States. *mBIO* **4**, e00737.

Hushpulian, D.M., Poloznikov, A.A., Savitski, P.A., Rozhkova, A.M., Chubar, T.A., Fechina, V.A., Lagrimini, L.M., Tishkov, V.I. and Gazaryan, I.G. (2007) Biocatalytic properties of recombinant tobacco peroxidase in chemiluminescent reaction. *Biocatalysis and Biotransformation* **25**, 163-170.

Huub, J.M. and Van Loon, L.C. (1991) Pathogenesis-related proteins of plants. *Critical Reviews in Plant Sciences* **10**, 123-150.

Huy, N.-X., Kim, S.-H., Yang, M.-S. and Kim, T.-G. (2012) Immunogenicity of a neutralizing epitope from porcine epidemic diarrhea virus: M cell targeting ligand fusion protein expressed in transgenic rice calli. *Plant Cell Reports* **31**, 1933-1942.

Huy, N.-X., Kim, Y.-S., Jun, S.-C., Jin, Z., Park, S.-M., Yang, M.-S. and Kim, T.-G. (2009) Production of a heat-labile enterotoxin B subunit-porcine epidemic diarrhea virus-neutralizing epitope fusion protein in transgenic lettuce (*Lactuca sativa*). *Biotechnology and Bioprocess Engineering* **14**, 731-737.

Irvin, R., MacAlister, T. and Costerton, J. (1981) Tris (hydroxymethyl) aminomethane buffer modification of *Escherichia coli* outer membrane permeability. *Journal of Bacteriology* **145**, 1397-1403.

Iwata, Y. and Koizumi, N. (2005) Unfolded protein response followed by induction of cell death in cultured tobacco cells treated with tunicamycin. *Planta* **220**, 804-807.

Jacob, E., Horovitz, A. and Unger, R. (2007) Different mechanistic requirements for prokaryotic and eukaryotic chaperonins: a lattice study. *Bioinformatics* **23**, i240-248.

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

John, S.E.S., Anson, B.J. and Mesecar, A.D. (2016) X-Ray structure and inhibition of 3C-like protease from porcine epidemic diarrhea virus. *Scientific Reports* **6**, 25961.

Junge, F., Schneider, B., Reckel, S., Schwarz, D., Dotsch, V. and Bernhard, F. (2008) Large-scale production of functional membrane proteins. *Cellular and Molecular Life Sciences* **65**, 1729-1755.

Kang, K.-H., Kang, T.-J., Kim, J.-A., Kwon, T.-H., Jang, Y.-S. and Yang, M.-S. (2004) High-level expression of the neutralizing epitope of porcine epidemic diarrhea virus by a tobacco mosaic virus-based vector. *Protein Expression and Purification* **38**, 129-135.

Kang, T.-J., Han, S.-C., Yang, M.-S. and Jang, Y.-S. (2006) Expression of synthetic neutralizing epitope of porcine epidemic diarrhea virus fused with synthetic B subunit of *Escherichia coli* heat-labile enterotoxin in tobacco plants. *Protein Expression and Purification* **46**, 16-22.

Kang, T.-J., Kim, Y.-S., Jang, Y.-S. and Yang, M.-S. (2005a) Expression of the synthetic neutralizing epitope gene of porcine epidemic diarrhea virus in tobacco plants without nicotine. *Vaccine* **23**, 2294-2297.

Kang, T.J., Seo, J.E., Kim, D.H., Kim, T.G., Jang, Y.S. and Yang, M.S. (2005b) Cloning and sequence analysis of the Korean strain of spike gene of porcine epidemic diarrhea virus and expression of its neutralizing epitope in plants. *Protein Expression and Purification* **41**, 378-383.

Kessner, D., Chambers, M., Burke, R., Agus, D. and Mallick, P. (2008) ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534-2536.

Kim, Y.-S., Kang, T.-J., Jang, Y.-S. and Yang, M.-S. (2005) Expression of neutralizing epitope of porcine epidemic diarrhea virus in potato plants. *Plant Cell, Tissue and Organ Culture* **82**, 125-130.

Kim, Y.-S., Kwon, T.-H. and Yang, M.-S. (2003) Expression of porcine epidemic diarrhea virus spike gene in transgenic carrot plant. *Plant Resources* **6**, 108-113.

King, A.M. (2011) *Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses*. Amsterdam:Elsevier.

Ko, S.-M., Sun, H.-J., Oh, M.J., Song, I.-J., Kim, M.-J., Sin, H.-S., Goh, C.-H., Kim, Y.-W., Lim, P.-O., Lee, H.-Y. and Kim, S.W. (2011) Expression of the protective antigen for PEDV in transgenic duckweed, *Lemna minor*. *Horticulture, Environment, and Biotechnology* **52**, 511-515.

Kolotilin, I., Topp, E., Cox, E., Devriendt, B., Conrad, U., Joensuu, J., Stöger, E., Warzecha, H., McAllister, T., Potter, A., McLean, M.D., Hall, J.C. and Menassa, R. (2014) Plant-based solutions for veterinary immunotherapeutics and prophylactics. *Veterinary Research* **45**, 375.

Kozma, D., Simon, I. and Tusnady, G.E. (2013) PDBTM: Protein Data Bank of transmembrane proteins after 8 years. *Nucleic Acids Research* **41**, D524-529.

Krogh, A., Larsson, B., Von Heijne, G. and Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* **305**, 567-580.

Kun, M., Bing, Y., Jing, X., XiangJie, Q., HongZhuan, Z., FuZhou, X., ChengYang, X., FengPing, Y., LiQuan, Z. and XiangLong, L. (2014) Expression of core neutralizing epitope gene of porcine epidemic diarrhea virus in maize. *Journal of Agricultural Science and Technology (Beijing)* **16**, 28-35.

Kushnir, N., Streatfield, S.J. and Yusibov, V. (2012) Virus-like particles as a highly efficient vaccine platform: diversity of targets and production systems and advances in clinical development. *Vaccine* **31**, 58-83.

Kwon, K.C. and Daniell, H. (2016) Oral delivery of protein drugs bioencapsulated in plant cells. *Molecular Therapy*.

Lagrimini, L.M., Burkhart, W., Moyer, M. and Rothstein, S. (1987) Molecular cloning of complementary DNA encoding the lignin-forming peroxidase from tobacco: molecular analysis and tissue-specific expression. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7542-7546.

Lagrimini, L.M. and Rothstein, S. (1987) Tissue specificity of tobacco peroxidase isozymes and their induction by wounding and tobacco mosaic virus infection. *Plant Physiology* **84**, 438-442.

Lagrimini, M., Bradford, S. and Rothstein, S. (1990) Peroxidase-induced wilting in transgenic tobacco plants. *Plant Cell* **2**, 7-18.

Laing, W. and Christeller, J. (2004) Extraction of proteins from plant tissues. *Current Protocols in Protein Science* **38**, 4.7.1–4.7.7.

Lakshmi, P.S., Verma, D., Yang, X., Lloyd, B. and Daniell, H. (2013) Low cost tuberculosis vaccine antigens in capsules: expression in chloroplasts, bio-encapsulation, stability and functional evaluation *in vitro*. *PLoS One* **8**, e54708.

Laude, H., Gelfi, J., Lavenant, L. and Charley, B. (1992) Single amino acid changes in the viral glycoprotein M affect induction of alpha interferon by the coronavirus transmissible gastroenteritis virus. *Journal of Virology* **66**, 743-749.

Laude, H., Rasschaert, D. and Huet, J.-C. (1987) Sequence and N-terminal processing of the transmembrane protein E1 of the coronavirus transmissible gastroenteritis virus. *Journal of General Virology* **68**, 1687-1693.

le Maire, M., Champeil, P. and Møller, J.V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica et Biophysica Acta - Biomembranes* **1508**, 86-111.

Lee, J.C. and Timasheff, S.N. (1981) The stabilization of proteins by sucrose. *Journal of Biological Chemistry* **256**, 7193-7201.

Lee, Y.-N., Chen, L.-K., Ma, H.-C., Yang, H.-H., Li, H.-P. and Lo, S.-Y. (2005) Thermal aggregation of SARS-CoV membrane protein. *Journal of Virological Methods* **129**, 152-161.

Liao, Y., Lescar, J., Tam, J. and Liu, D. (2004) Expression of SARS-coronavirus envelope protein in *Escherichia coli* cells alters membrane permeability. *Biochemical and Biophysical Research Communications* **325**, 374-380.

Lim, K.P. and Liu, D.X. (2001) The missing link in coronavirus assembly: retention of the avian coronavirus infectious bronchitis virus envelope protein in the pre-golgi compartments and physical interaction between the envelope and membrane proteins. *Journal of Biological Chemistry* **276**, 17515-17523.

Liu, G., Lv, L., Yin, L., Li, X., Luo, D., Liu, K., Xue, C. and Cao, Y. (2013) Assembly and immunogenicity of coronavirus-like particles carrying infectious bronchitis virus M and S proteins. *Vaccine* **31**, 5524-5530.

Liu, J., Sun, Y., Qi, J., Chu, F., Wu, H., Gao, F., Li, T., Yan, J. and Gao, G.F. (2010) The membrane protein of severe acute respiratory syndrome coronavirus acts as a dominant immunogen revealed by a clustering region of novel functionally and structurally defined cytotoxic T-lymphocyte epitopes. *Journal of Infectious Diseases* **202**, 1171-1180.

Lua, L.H.L., Connors, N.K., Sainsbury, F., Chuan, Y.P., Wibowo, N. and Middelberg, A.P.J. (2014) Bioengineering virus-like particles as vaccines. *Biotechnology and Bioengineering* **111**, 425-440.

Maeda, J., Maeda, A. and Makino, S. (1999) Release of coronavirus E protein in membrane vesicles from virus-infected cells and E protein-expressing cells. *Virology* **263**, 265-272.

Masuda, T., Murakami, S., Takahashi, O., Miyazaki, A., Ohashi, S., Yamasato, H. and Suzuki, T. (2015) New porcine epidemic diarrhoea virus variant with a large deletion in the spike gene identified in domestic pigs. *Archives of Virology* **160**, 2565-2568.

Medicago (2016) Facility Overview. http://www.medicago.com/English/Manufacturing/Facility-Overview/default.aspx. Accessed: June 1 2016

Menassa, R., Ahmad, A. and Joensuu, J.J. (2012) Transient expression using agroinfiltration and its applications in molecular farming. In: *Molecular Farming in Plants: Recent Advances and Future Prospects* pp. 183-198. Dordrecht, Netherlands: Springer.

Merck (2011a) Gardasil®. Highlights of prescribing information. http://www.fda.gov/downloads/BiologicsBloodVaccines/Vaccines/ApprovedProducts/UCM111263.pdf. Accessed: September 1 2016

Merck (2011b) Recombivax HB®. Highlights of prescribing information. http://www.fda.gov/downloads/BiologicsBloodVaccines/Vaccines/ApprovedProducts/UCM110114.pdf. Accessed: September 1 2016

Merck Animal Health (2015) Merck Animal Health to acquire Harrisvaccines. http://www.merck-animal-health-usa.com/news/2015-11-12.aspx. Accessed: January 10 2016

Mortola, E. and Roy, P. (2004) Efficient assembly and release of SARS coronavirus-like particles by a heterologous expression system. *FEBS Letters* **576**, 174-178.

Narayanan, K., Maeda, A., Maeda, J. and Makino, S. (2000) Characterization of the coronavirus M protein and nucleocapsid interaction in infected cells. *Journal of Virology* **74**, 8127-8134.

Neuman, B.W., Kiss, G., Kunding, A.H., Bhella, D., Baksh, M.F., Connelly, S., Droese, B., Klaus, J.P., Makino, S., Sawicki, S.G., Siddell, S.G., Stamou, D.G., Wilson, I.A., Kuhn, P. and Buchmeier, M.J. (2011) A structural analysis of M protein in coronavirus assembly and morphology. *Journal of Structural Biology* **174**, 11-22.

Nguyen, V.P. and Hogue, B.G. (1997) Protein interactions during coronavirus assembly. *Journal of Virology* **71**, 9278-9284.

Ojkic, D., Hazlett, M., Fairles, J., Marom, A., Slavic, D., Maxie, G., Alexandersen, S., Pasick, J., Alsop, J. and Burlatschenko, S. (2015) The first case of porcine epidemic diarrhea in Canada. *Canadian Veterinary Journal* **56**, 149-152.

Oszvald, M., Kang, T.-J., Tomoskozi, S., Tamas, C., Tamas, L., Kim, T.-G. and Yang, M.-S. (2007) Expression of a synthetic neutralizing epitope of porcine epidemic diarrhea virus fused with synthetic B subunit of *Escherichia coli* heat labile enterotoxin in rice endosperm. *Molecular Biotechnology* **35**, 215-224.

Overington, J.P., Al-Lazikani, B. and Hopkins, A.L. (2006) How many drug targets are there? *Nature Reviews Drug Discovery* **5**, 993-996.

Ozdal, T., Capanoglu, E. and Altay, F. (2013) A review on protein–phenolic interactions and associated changes. *Food Research International* **51**, 954-970.

Pang, H., Liu, Y., Han, X., Xu, Y., Jiang, F., Wu, D., Kong, X., Bartlam, M. and Rao, Z. (2004) Protective humoral responses to severe acute respiratory syndrome-associated coronavirus: implications for the design of an effective protein-based vaccine. *Journal of General Virology* **85**, 3109-3113.

Pensaert, M. and De Bouck, P. (1978) A new coronavirus-like particle associated with diarrhea in swine. *Archives of Virology* **58**, 243-247.

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

Prentice, E., McAuliffe, J., Lu, X., Subbarao, K. and Denison, M.R. (2004) Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. *Journal of Virology* **78**, 9977-9986.

Qian, Q., Qi, J., Bu, Q., Li, X., Li, S., Li, C., Chen, M., Chen, Q., Chen, J., Sun, J., Liang, W., Zhou, Y., Chu, C., Ren, F., Palme, K. and Zhao, B. (2008) Mutation of the rice narrow *leaf1* gene, which encodes a novel protein, affects vein patterning and polar auxin transport. *Plant Physiology* **147**, 1947-1959.

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

Raucher, D. and Chilkoti, A. (2001) Enhanced uptake of a thermally responsive polypeptide by tumor cells in response to its hyperthermia-mediated phase transition. *Cancer Research* **61**, 7163-7170.

Ren, X., Suo, S. and Jang, Y.S. (2011) Development of a porcine epidemic diarrhea virus M protein-based ELISA for virus detection. *Biotechnology Letters* **33**, 215-220.

Ren, Y.P., Wang, X.Y., Yan, Q.G., Chen, D.S., Cheng, Y., Wan, L., Ma, L., Zhang, Q., Guo, L., Xie, Z.Y. and Guo, W.Z. (2012) Immunogenicity of recombinant attenuated *Salmonella choleraesuis* C500 strain co-expressing M and N protein of

porcine epidemic diarrhea virus (PEDV). *Journal of Animal and Veterinary Advances* **11**, 3234-3241.

Rizza, P., Capone, I., Moretti, F., Proietti, E. and Belardelli, F. (2011) IFN-α as a vaccine adjuvant: recent insights into the mechanisms and perspectives for its clinical use. *Expert Review of Vaccines* **10**, 487-498.

Ruch, T.R. and Machamer, C.E. (2012) The coronavirus E protein: assembly and beyond. *Viruses* **4**, 363-382.

Rybicki, E.P. (2014) Plant-based vaccines against viruses. *Virology Journal* **11**, 205.

Rymerson, R., Menassa, R. and Brandle, J. (2002) Tobacco, a platform for the production of recombinant proteins. In: *Molecular Farming of Plants and Animals for Human and Veterinary Medicine* pp. 1-31. Dordrecht, Netherlands: Springer.

Saif, L.J. (1993) Coronavirus immunogens. *Veterinary Microbiology* **37**, 285-297.

Sato, T., Takeyama, N., Katsumata, A., Tuchiya, K., Kodama, T. and Kusanagi, K. (2011) Mutations in the spike gene of porcine epidemic diarrhea virus associated with growth adaptation in vitro and attenuation of virulence in vivo. *Virus Genes* **43**, 72-78.

Schmidt, T.G.M., Koepke, J., Frank, R. and Skerra, A. (1996) Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin. *Journal of Molecular Biology* **255**, 753-766.

Shenyang, G., Enhui, Z., Baoxian, L., Xinyuan, Q., Lijie, T., Junwei, G. and Yijing, L. (2007) High-level prokaryotic expression of envelope exterior of membrane protein of porcine epidemic diarrhea virus. *Veterinary Microbiology* **123**, 187-193.

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

Song, D. and Park, B. (2012) Porcine epidemic diarrhoea virus: a comprehensive review of molecular epidemiology, diagnosis, and vaccines. *Virus Genes* **44**, 167-175.

Song, D.S., Oh, J.S., Kang, B.K., Yang, J.S., Moon, H.J., Yoo, H.S., Jang, Y.S. and Park, B.K. (2007) Oral efficacy of Vero cell attenuated porcine epidemic diarrhea virus DR13 strain. *Research in Veterinary Science* **82**, 134-140.

Stevenson, G.W., Hoang, H., Schwartz, K.J., Burrough, E.R., Sun, D., Madson, D., Cooper, V.L., Pillatzki, A., Gauger, P., Schmitt, B.J., Koster, L.G., Killian, M.L. and Yoon, K.J. (2013) Emergence of porcine epidemic diarrhea virus in the United States: clinical signs, lesions, and viral genomic sequences. *Journal of Veterinary Diagnostic Investigation* **25**, 649-654.

Sun, D., Feng, L., Shi, H., Chen, J., Liu, S., Chen, H. and Wang, Y. (2006) Spike protein region (aa 636789) of porcine epidemic diarrhea virus is essential for induction of neutralizing antibodies. *Acta Virologica* **51**, 149-156.

Sun, R.-Q., Cai, R.-J., Chen, Y.-Q., Liang, P.-S., Chen, D.-K. and Song, C.-X. (2012) Outbreak of porcine epidemic diarrhea in suckling piglets, China. *Emerging Infectious Diseases* **18**, 161-163.

Sun, R.Q., Leng, Z.M., Zhai, S.L., Chen, D.K. and Song, C.X. (2014) Genetic variability and phylogeny of current Chinese porcine epidemic diarrhea virus strains based on spike, ORF3, and membrane Genes. *Scientific World Journal* **2014**, 1-8.

Tamás, L. (2010) Molecular farming, using the cereal endosperm as bioreactor. *Acta Agronomica Hungarica* **58**, 55-64.

Tang, S., Hettiarachchy, N.S. and Shellhammer, T.H. (2002) Protein extraction from heat-stabilized defatted rice bran. 1. Physical processing and enzyme treatments. *Journal of Agricultural and Food Chemistry* **50**, 7444-7448.

Thuenemann, E.C., Meyers, A.E., Verwey, J., Rybicki, E.P. and Lomonossoff, G.P. (2013) A method for rapid production of heteromultimeric protein complexes in plants: assembly of protective bluetongue virus-like particles. *Plant Biotechnology Journal* **11**, 839-846.

Torres, J., Wang, J., Parthasarathy, K. and Liu, D.X. (2005) The transmembrane oligomers of coronavirus protein E. *Biophysical Journal* **88**, 1283-1290.

Tseng, Y.-T., Wang, S.-M., Huang, K.-J., I-Ru Lee, A., Chiang, C.-C. and Wang, C.-T. (2010) Self-assembly of severe acute respiratory syndrome coronavirus membrane protein. *Journal of Biological Chemistry* **285**, 12862-12872.

Ujike, M. and Taguchi, F. (2015) Incorporation of spike and membrane glycoproteins into coronavirus virions. *Viruses* **7**, 1700-1725.

Utiger, A., Tobler, K., Bridgen, A. and Ackermann, M. (1995) Identification of the membrane protein of porcine epidemic diarrhea virus. *Virus Genes* **10**, 137-148.

Vaudel, M., Barsnes, H., Berven, F.S., Sickmann, A. and Martens, L. (2011) SearchGUI: an open-source graphical user interface for simultaneous OMSSA and X! Tandem searches. *Proteomics* **11**, 996-999.

Vaudel, M., Burkhart, J.M., Zahedi, R.P., Oveland, E., Berven, F.S., Sickmann, A., Martens, L. and Barsnes, H. (2015) PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nature Biotechnology* **33**, 22-24.

Vennema, H., Godeke, G., Rossen, J., Voorhout, W., Horzinek, M., Opstelten, D. and Rottier, P. (1996) Nucleocapsid-independent assembly of coronavirus-like

particles by co-expression of viral envelope protein genes. *EMBO Journal* **15**, 2020.

Wang, K., Lu, W., Chen, J., Xie, S., Shi, H., Hsu, H., Yu, W., Xu, K., Bian, C., Fischer, W.B., Schwarz, W., Feng, L. and Sun, B. (2012) PEDV ORF3 encodes an ion channel protein and regulates virus production. *FEBS Letters* **586**, 384-391.

Westerbeck, J.W. and Machamer, C.E. (2015) A coronavirus E protein is present in two distinct pools with different effects on assembly and the secretory pathway. *Journal of Virology* **89**, 9313-9323.

Woods, R.D., Wesley, R.D. and Kapke, P.A. (1988) Neutralization of porcine transmissible gastroenteritis virus by complement-dependent monoclonal antibodies. *American Journal of Veterinary Research* **49**, 300-304.

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

Xia, X. (2007) Protein isoelectric point. *Bioinformatics and the Cell*, 207-219.

Xing, J., Liu, S., Han, Z., Shao, Y., Li, H. and Kong, X. (2009) Identification of a novel linear B-cell epitope in the M protein of avian infectious bronchitis coronaviruses. *Journal of Microbiology* **47**, 589-599.

Xu, X., Zhang, H., Zhang, Q., Dong, J., Liang, Y., Huang, Y., Liu, H.-J. and Tong, D. (2013) Porcine epidemic diarrhea virus E protein causes endoplasmic reticulum stress and up-regulates interleukin-8 expression. *Virology Journal* **10**, 26-26.

Xue, S., Jaszewski, A. and Perlman, S. (1995) Identification of a CD4+ T cell epitope within the M protein of a neurotropic coronavirus. *Virology* **208**, 173-179.

Yang, K.-S., Lim, S., Kwon, S.-Y., Kwak, S.-S., Kim, H.-S. and Lee, H.-S. (2005) Transgenic sweetpotato (*Ipomoea batatas*) expressing spike gene of porcine epidemic diarrhea virus. *Journal of Plant Biotechnology* **32**, 263-268.

Yu, X., Bi, W., Weiss, S.R. and Leibowitz, J.L. (1994) Mouse hepatitis virus gene 5b protein is a new virion envelope protein. *Virology* **202**, 1018-1023.

Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 1.

Zhang, Z.-B., Chen, J.-F., Shi, H.-Y., Chen, X.-J. and Feng, L. (2011) Expression of the whole M protein of porcine epidemic diarrhea virus CH/S strain in *Escherichia coli* and analysis of its reactogenicity. *Chinese Veterinary Science* **1**, 008.

Zhang, Z., Chen, J., Shi, H., Chen, X., Shi, D., Feng, L. and Yang, B. (2012) Identification of a conserved linear B-cell epitope in the M protein of porcine epidemic diarrhea virus. *Virology Journal* **9**, 225.

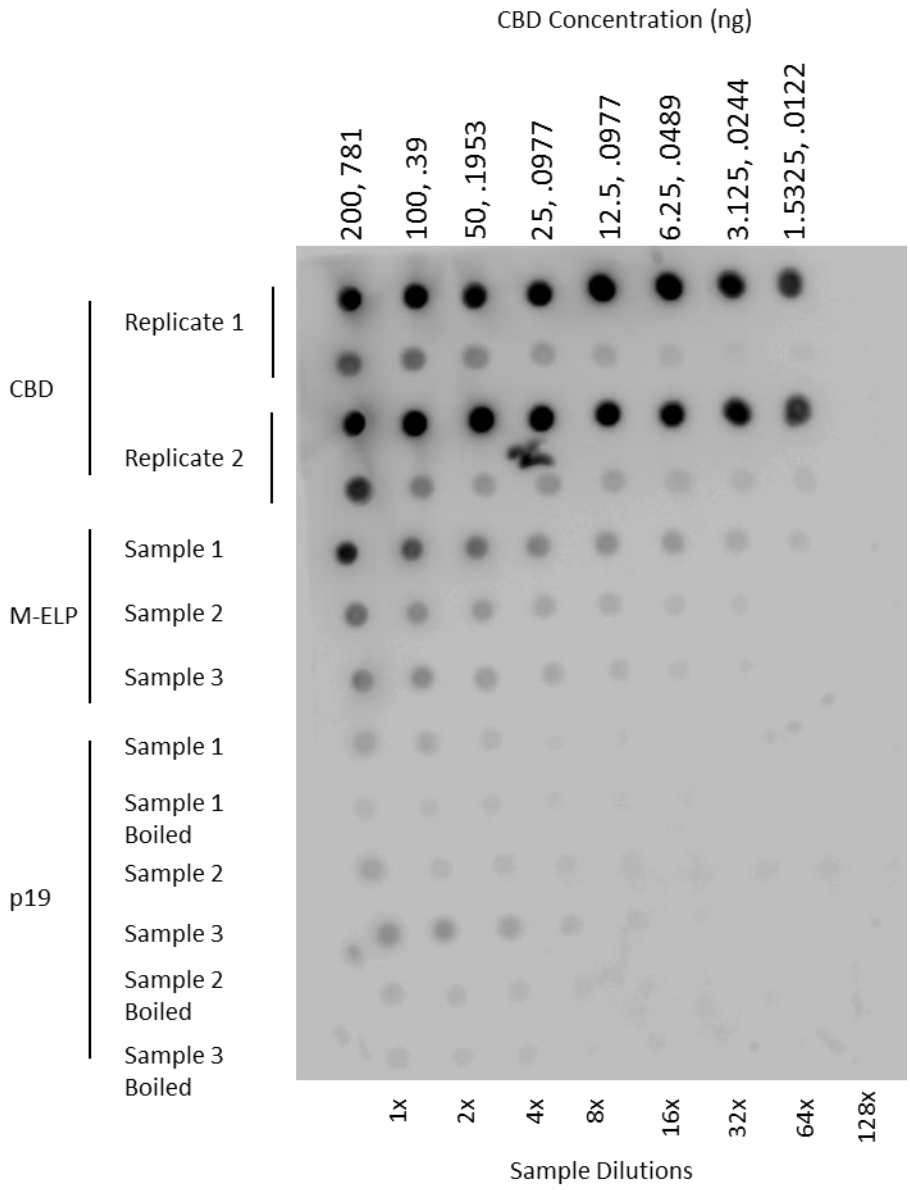
Zhenhui, S., Feng, H., Zhu, Z., Dai, X., Zhou, Y., Li, Y. and Cao, X. (2015) The assembly of virus-like particles of porcine transmissible gastroenteritisvirus *in vitro* by baculovirus expression system. *Turkish Journal of Veterinary and Animal Sciences* **39**, 302-307.

Zoetis (2014) Zoetis Granted Conditional License for Porcine Epidemic Diarrhea Vaccine*. http://news.zoetis.com/press-release/manufacturing/zoetis-granted-conditional-license-porcine-epidemic-diarrhea-vaccine. Accessed: January 10 2016

Zoetis (2016) Porcine Epidemic Diarrhea Vaccine*. http://www.zoetisus.com/products/pork/porcine-epidemic-diarrhea-vaccine.aspx. Accessed: July 17 2016

Appendices

Appendix A. Example of a dot blot used for quantification, with standards, M-ELP and p19 (negative control) tissue.



Appendix B. Mass spectrometry analysis of digested protein extracts.

Recombinant Protein	Amino Acid Sequences Detected	Percent of Possible Coverage Detected	Number of Peptide Spectrum Matches
Full length M-ELP	194-203; 255-262; 490-420	9.9%-21.96%	3; 1; 1
Truncated M-ELP	19-37; 176-194; 194-203; 206-221; 223-242; 409-420	24.67%-61.65%	3; 2; 2; 1; 1; 1
E-ELP (dimer band)	64-76	22.65%	2

Curriculum Vitae

Name:	Zayn Khamis
Post-secondary Education and Degrees:	The University of Western Ontario London, Ontario, Canada 2009-2013 B.Sc
	The University of Western Ontario London, Ontario, Canada 2009-2014 B.A.
	The University of Western Ontario London, Ontario, Canada 2014-2016 M.Sc
Honours and Awards:	Department of Biology Graduate Student Travel Award 2016 Graduate Student Travel Award
Related Work Experience	Teaching Assistant The University of Western Ontario 2014-2016
	Summer Student Agriculture & Agri-Food Canada 2014
Conference Presentations:	International Society for Plant Molecular Farming 2016 Oral Presentation
	Canadian Society for Plant Biotechnology/Canadian Society of Plant Biologists 2016 Oral Presentation