View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Helsingin yliopiston digitaalinen arkisto

Potato virus A as a heterologous protein expression tool in plants

Jani Kelloniemi

Department of Applied Biology
Faculty of Agriculture and Forestry
and
Viikki Graduate School in Molecular Biosciences
University of Helsinki

ACADEMIC DISSERTATION
in Plant Virology

To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public criticism in auditorium B3 at the Metsätieteiden talo (Latokartanonkaari 7) on May 30th 2008, at 12 o'clock.

Supervisors: Professor Jari PT Valkonen
Department of Applied Biology
University of Helsinki
Finland

Docent Kristiina Mäkinen
Department of Applied Chemistry and Microbiology
University of Helsinki
Finland

Reviewers: Professor Andres Merits
Institute of Technology
University of Tartu
Tartu
Estonia

Docent Kirsi Lehto
Department of Biology
University of Turku
Finland

Opponent: Doctor George P Lomonossoff
Department of Biological Chemistry
John Innes Centre
Norwich
Great Britain

ISSN 1795-7079
ISBN 978-952-10-4614-8 (paperback)
ISBN 978-952-10-4630-8 (PDF)
Yliopistopaino
Helsinki 2008

CONTENTS

ABBREVIATIONS **6**

ORIGINAL PUBLICATIONS **8**

ABSTRACT **9**

INTRODUCTION **11**

Vector viruses in plants 11

Vector viruses as research tools to understand viral functions in plants 11

Vector viruses for suppression of host gene expression in plants 13

Plant virus-based expression of heterologous proteins for industrial uses 15

Plants as heterologous protein production platforms *15*

Comparison of plant virus-based expression vectors and transgenic plants in heterologous protein production *17*

DNA viruses as overexpression vectors *18*

RNA viruses as overexpression vectors *20*

RNA viruses of the genus Potyvirus as overexpression vectors *27*

Epitope/peptide presentation vectors *32*

Vector viruses as stable transgenes *32*

Potato virus A 33

AIMS OF THE STUDY **35**

MATERIALS AND METHODS **36**

The PVA-based expression vectors 36

The cloning site within the P1 encoding region *36*

The cloning site between the P1 and HC-Pro encoding regions *37*

The cloning site between the NIb and CP encoding regions *37*

Optimization of the NIb/CP site *38*

Methods for virus inoculation and detection, and for analysis of expressed heterologous proteins 40

RESULTS AND DISCUSSION **41**

Infectivity of the PVA-based vectors in *N. benthamiana* 41

Disease symptoms and accumulation of the vector-viruses in *N. benthamiana* 42

Influence of inserts within the P1-encoding region in single-insert vectors 43

Influence of inserts at the P1/HC-Pro site in single-insert vectors 44

Influence of inserts at the NIb/CP site in single-insert vectors 45

Influence of multiple inserts 46

Double-insert vectors 46

The triple-insert vector 46

Optimizations of the NIb/CP cloning site 51

Variability in virus titers 52

Testing vector-viruses in *N. tabacum* cv Samsun nn 53

Testing vector-viruses in *S. tuberosum* 54

Stability of chimeric viruses during infection 54

Heterologous protein expression and accumulation in *N. benthamiana* 57

Expression of the jellyfish GFP 57

Expression of the seapansy luciferase 59

Yields of human S-COMT and bacterial GUS 59

Role of P1 in HC-Pro mediated suppression of RNA silencing 61

Sizes of virions of the PVA-based vectors 62

CONCLUSIONS **64**

ACKNOWLEDGEMENTS **66**

REFERENCES **67**

REPRINTS OF ORIGINAL PUBLICATIONS

ABBREVIATIONS

3′-UTR	3′-untranslated region
5′-UTR	5′-untranslated region
A_{405}	absorbance at 405 nm
BMV	*Brome mosaic virus*
BSMV	*Barley stripe mosaic virus*
CaMV	*Cauliflower mosaic virus*
cDNA	complementary DNA
CI	cylindrical inclusion protein
ClYVV	*Clover yellow vein virus*
CMV	*Cucumber mosaic virus*
CP	coat protein
CPMV	*Cowpea mosaic virus*
CS	cloning site
cv	cultivar
DAS-ELISA	double antibody sandwitch – enzyme-linked immunosorbent assay
DNA	dioxyribonucleic acid
dpi	days post-inoculation
dsDNA	double-stranded DNA
DsRed	*Discosoma* red fluorescent protein
ER	endoplasmic reticulum
FDMV	foot-and-mouth disease virus
GFP	green fluorescent protein
GUS	β-glucuronidase
HC-Pro	helper component proteinase
hGH	human growth hormone
IC-RT-PCR	immuno capture reverse transcription PCR
IgA	immunoglobulin type A
LMV	*Lettuce mosaic virus*
miRNA	micro RNA
MP	movement protein
mRNA	messenger RNA
NIa-Pro	nucleic inclusion protein a - proteinase
NIb	nuclear inclusion protein b
nt	nucleotides
ORF	open reading frame
P1	(potyviral) protein 1
PCR	polymerase chain reaction
PDS	phytoene desaturase
poly(A)	polyadenosine
PPV	*Plum pox virus*
PVA	*Potato virus A*
PVX	*Potato virus X*
RISC	RNA-induced silencing complex

RNA	ribonucleic acid
S-COMT	soluble catechol-*O*-methyltransferase
sgRNA	sub-genomic RNA
siRNA	small interfering RNA
ssDNA	single-stranded DNA
T-DNA	transfer DNA
TEV	*Tobacco etch virus*
TRV	*Tobacco rattle virus*
TMV	*Tobacco mosaic virus*
TuMV	*Turnip mosaic virus*
TVMV	*Tobacco vein mottling virus*
UV	ultraviolet
VIGS	virus induced gene silencing
VPg	viral genome-linked protein
wt	wild-type
ZYMV	*Zucchini yellow mosaic virus*

ORIGINAL PUBLICATIONS

I **Rajamäki ML, Kelloniemi J, Alminaite A, Kekarainen T, Rabenstein F & Valkonen JPT (2005).** A novel insertion site inside the potyvirus P1 cistron allows expression of heterologous proteins and suggests some P1 functions. *Virology* **342**, 88-101.

II **Kelloniemi J, Mäkinen K & Valkonen JPT (2006).** A potyvirus-based gene vector allows producing active human S-COMT and animal GFP, but not human sorcin, in vector-infected plants. *Biochimie* **88**, 505-513.

III **Kelloniemi J, Mäkinen K & Valkonen JPT (2008).** Three heterologous proteins simultaneously expressed from a chimeric potyvirus: infectivity, stability and the correlation of genome and virion lengths. *Virus Research*, doi: 10.1016/j.virusres.2008.04.006.

ABSTRACT

An infectious clone of *Potato virus A* (PVA) (genus *Potyvirus*, family *Potyviridae*) was engineered to be used as an expression vector for production of heterologous proteins. The PVA genome (9565 nt) is translated to a large polyprotein that is subsequently processed, yielding up to ten mature proteins. Hence, a foreign sequence inserted into an infectious cDNA clone of PVA will also be translated as part of the viral polyprotein in infected plants. Three sites in the genome of PVA were used for expression of heterologous protein encoding sequences in plants. Proteolytic cleavage sites for the viral proteinases were engineered and added for separation of the heterologous protein from the viral proteins.

A novel genomic location for foreign encoding sequence expression was tested by inserting the *Aequorea victoria gfp* sequence encoding the green fluorescent protein (GFP) into the P1 encoding region (genomic position 235/236). The vector-PVA expressing GFP (239 amino acids) accumulated to high titers in *Nicotiana benthamiana* and *N. tabacum* cv. Samsun nn cells. The vector-PVA continued to produce intact GFP in the systemically infected plants for 3 weeks post-inoculation. The vector construct caused much milder disease symptoms than the wild-type virus. The viral coat protein (CP) in these plants and in tobacco protoplasts accumulated to about 30-50% of the levels reached with the wt PVA. The role of P1 as an enhancer of RNA silencing suppression, which is mediated by the HC-Pro protein, was investigated by an overexpression system (agroinfiltration) assay in plant leaves. Up to 30-fold higher amounts of HC-Pro mRNA were produced when P1 was present, as compared to expression of HC-Pro alone.

The cloning site between the polymerase (NIb) and CP encoding regions of PVA was tested for expression of human proteins. Soluble catechol-*O*-methyltransferase (S-COMT), presumed to be involved in the development of Parkinson's disease, was produced from the vector PVA and constituted *ca.* 1% of the total soluble proteins in systemically infected *N. benthamiana* leaves.

However, another human protein, sorcin (a Ca^{2+}-binding protein), was not detected in infected plants although the sequence encoding it cloned into the NIb/CP site was stably present in the PVA genome for at least a month post-inoculation. Thus, sorcin was expressed in equal molar amounts with the viral proteins but it was apparently quickly degraded. The amount of PVA CP produced from the vectors carrying the S-COMT or sorcin encoding sequences was 60-100% of the wt virus levels. These data indicate that PVA can be used to produce at least some human proteins in plants, but further optimization may be needed with others.

The third cloning site used was between the P1 and HC-Pro encoding regions. The GFP and the *Renilla reniformis* luciferase encoding sequences were expressed from this site. In the systemically infected leaves, vector-virus titers were 40-70% of the wt virus levels.

Stability of inserts differed, depending of the cloning site and the heterologous sequence. Deletions within the *gfp* sequence located at the P1 cloning site were detected two to three weeks post-inoculation, whereas it was stable up to a month post-inoculation when inserted at the NIb/CP site, as was also the coding sequence of sorcin. On the other hand, deletions in the sequence encoding *Eschericia coli* β-glucuronidase (*UidA*) appeared in the NIb/CP site as soon as two weeks post-inoculation. The coding sequence for luciferase situated at the P1/HC-Pro site was intact in all plants tested up to a month post-inoculation.

All the three cloning sites were combined in the same vector-PVA clone to simultaneously produce three heterologous proteins: GFP from the P1 site, luciferase from the P1/HC-Pro site, and β-glucuronidase from the NIb/CP site. Vector-virus amounts in the systemically infected leaves of *N. benthamiana* were 15% of those of the wt virus. All three heterologous proteins were detected in the leaf sap in an active form. In conclusion, PVA can be used for simultaneous production of at least three proteins (together consisting of over 1000 amino acid residues) in plants, which possibly will be useful for some research purposes and for heterologous protein production.

INTRODUCTION

Vector-viruses in plants

The term 'vector-virus' denotes an infectious clone of a virus, into which a heterologous nucleotide sequence can be inserted without loss of viral infectivity and replication functions. The capacity for systemic movement within the host is also desired, but not always necessary. The vector-virus is used to infect a host organism, delivering the heterologous sequence into the host for expression. Other names for a vector-virus commonly used in literature are 'gene vector' and also 'virus vector' (for example Scholthof *et al.* 1996, Gleba *et al.* 2007); the latter can lead to confusion with vectors of the virus. In most cases, viruses are engineered to be used as heterologous protein expression vectors (overexpression vectors). A notable exception that has become increasingly common in use is the application of vector-viruses in virus-induced gene silencing (VIGS), where a fragment of a gene, inserted in a cDNA clone of the viral genome, is used to suppress the expression of the homologous endogene in infected host plants (see below).

Vector-viruses as research tools to understand viral functions in plants

Reporter proteins such as β-glucuronidase (GUS; from *Eschericia coli*), green fluorescent protein and its colour-shift variants (GFP; from jellyfish *Aequorea victoria*), red fluorescent protein (DsRed; from reef coral of genus *Discosoma*) and luciferases (from the firefly *Photinus pyralis* and the seapansy *Renilla reniformis*) can be visualized and/or quantified when expressed either as a free protein or as a fusion to a viral protein from vector-viruses in various tissues of plants. This has made them useful tools for following the course of virus infection in infected tissues and cells. For example, the use of a clone of *Tobacco etch virus* (TEV) (family *Potyviridae*) engineered to express GUS and histochemical GUS assay in time-course experiments allowed the visualization of virus activity in single, mechanically inoculated leaf epidermal cells, in neighboring epidermal and mesophyll cells, in phloem-

associated cells after long-distance transport, and in cells surrounding vascular tissues of organs above and below the site of inoculation (Dolja *et al.* 1992). Similarly, with a clone of *Potato virus X* (PVX) (family *Flexiviridae*) engineered to express either GUS or GFP, the cell-to-cell spread of the virus and its emergence in the upper non-inoculated leaves were observed (Chapman *et al.* 1992, Baulcombe *et al.* 1995). *Potato leaf roll virus* (family *Luteoviridae*) genome with a GFP-encoding sequence located after the P5 gene was encapsidated and aphid-transmissible, and enabled visualization of the early stages of infection after aphid transmission (Nurkiyanova *et al.* 2000). To elucidate the temporal expression pattern of four genes of *Beet yellows virus* (BYV, family *Closteroviridae*) in infected cells, the coding sequence for GUS was inserted between the first and the second codons of these four genes in different BYV clones (Hagiwara *et al.* 1999). The amounts of expressed GUS at various timepoints revealed the stages of infection at which the promoters for these four viral genes were active.

However, the functions of viruses with foreign sequences, and the functions of viral proteins fused to reporter proteins, may not fully reveal the functions of wild-type (wt) viruses and their proteins. For example, the movement protein (MP) of *Cauliflower mosaic virus* (CaMV) with GFP fused to either the N- or C-teminus was not observed to aggregate and form the tubules that are formed by the wt MP (Thomas & Maule 2000).

One application of the reporter proteins expressed from viruses has been characterization of mutant viruses. The TEV-GUS vector virus (Dolja *et al.* 1992) has been used to elucidate functions of the viral P1, HC-Pro and CP proteins by comparing the amounts of activity of the expressed GUS from the parent and mutant clones in infected protoplasts and host plants (Dolja *et al.* 1994, Verchot & Carrington 1995, Kasshau *et al.* 1997). TEV-GUS with two highly conserved charged residues in the CP substituted with alanine residues replicated equally as well as the parental vector-virus in protoplasts, but failed to move cell-to-cell. TEV-GUS lacking the whole P1 encoding sequence accumulated to 2-3% of the parental clone levels in protoplasts,

indicating that the P1 protein was involved but not absolutely required for replication. Up to 25 charged amino acids were substituted with alanine residues in HC-Pro of TEV-GUS and tested for amplification in tobacco protoplasts and systemic movement in tobacco plants. The results suggested that the central region of HC-Pro is necessary for efficient genome amplification and systemic movement. When the CP genes of PVX, *Brome mosaic virus* (BMV; family *Bromoviridae*) and *Cowpea mosaic virus* (CPMV; family *Comoviridae*) were replaced with the GFP encoding sequence, the chimeric viruses were restricted to single GFP expressing cells (Baulcombe *et al.* 1995, Schmitz & Rao 1996, Verver *et al.* 1998). Thus CP was essential for cell-to-cell movement of these viruses. Failure of cell-to-cell movement was also observed with GFP-expressing *Potato virus A* (PVA, family *Potyviridae*) clones with amino acid substitutions at putative phosphorylation sites within the CP (Ivanov *et al.* 2003). More examples of plant viruses expressing reporter proteins are shown in Tables 1 and 2.

Vector viruses for suppression of host gene expression in plants

Vector viruses are used for host gene characterization by exploiting the virus-induced gene silencing (VIGS) phenomenon (Kumagai *et al.* 1995). The vector-virus is engineered to carry a sequence of the host gene to be silenced. Viral ssRNA forms hairpin-like structures that are recognized by the host (Pantaleo *et al.* 2007). The structures are cut into small interfering RNA (siRNA) fragments of 21-24 nt by a cellular RNase (Dicer) and one strand of siRNA is incorporated into an RNA-induced silencing complex (RISC) (reviewed in Baulcombe 2005, Voinnet 2005). The siRNA is used as a 'probe' to recognize homologous ssRNA molecules, which are then degraded by the RISC. Due to a phenomenon called transitivity, siRNA are also formed from corresponding viral sequences outside the initially targeted regions (Vaistij *et al.* 2002), and hence also from the inserted heterologous sequence that targets the homologous host gene for silencing. Inserting the sequence in an antisense orientation in the vector-virus triggers silencing (Kumagai *et al.* 1995),

presumably by forming double-stranded RNA hybrids with the endogenous mRNA from the target gene. Most efficient induction of silencing is usually achieved when expressing the silencing-inducing sequence as an inverted repeat that forms a double-stranded hairpin structure recognised by the host silencing system (Waterhouse *et al.* 1998). However, an inverted repeat may not be well tolerated by a vector-virus, because the hairpin structure in the viral genome could interfere with viral funtions such as replication or translation.

In the seminal work on VIGS, a partial cDNA copy of tomato phytoene desaturase gene (PDS) (92% identical to the PDS encoding sequence of *N. benthamiana*) was inserted in antisense orientation into a TMV-based vector that was subsequently used to inoculate *N. benthamiana* plants (Kumagai *et al.* 1995). Approximately a week post-inoculation the systemically infected leaves showed a white 'photo bleaching' phenotype. A similar phenotype developed when *N. benthamiana* plants were sprayed with the herbicide norflurazon, which is an inhibitor of PDS (Kumagai *et al.* 1995).

Dozens of publications exploiting VIGS are available. The following exemplifies a few different vector viruses and their host species used in studies on VIGS. *Tobacco rattle virus* (TRV, genus *Tobravirus*), has been the most commonly used vector-virus for VIGS so far. One reason is the wide host range of TRV, including more than 400 species in more than 50 monocotyledonous and dicotyledonous families (Robinson & Harrison 1989). Liu *et al.* (2004) used a TRV-based vector to silence known defence-related genes in transgenic *N. benthamiana* plants carrying the TMV resistance gene *N*. Upper leaves of the plants were subsequently inoculated with TMV to identify genes, the downregulation of which suppressed resistance to TMV mediated by the the *N* gene. In a similar fashion, Anand *et al.* (2007) screened approximately 1000 plant genes to detect those involved in *Agrobacterium*-mediated plant transformation. Shoresh *et al.* (2006) characterized a putative mitogen-activated protein kinase of cucumber (*Cucumis sativus*) by inserting the corresponding coding sequence as an antisense copy into a vector based

on *Zucchini yellow mosaic virus* (ZYMV; family *Potyviridae*). Decreased levels of the corresponding host mRNA were correlated with increased sensitivity to pathogen attack. Similarly, VIGS was used to identify genes in barley (*Hordeum vulgare*) associated with fungal resistance by expressing them in a vector based on *Barley stripe mosaic virus* (BSMV) (genus *Hordeivirus*) (Hein *et al.* 2005). PVX carrying an antisense fragment of the gene for PDS showed a characteristic white-leaf phenotype in infected diploid and tetraploid *Solanum* species and *ca.* 80% reduction in host PDS mRNA levels (Faivre-Rampant *et al.* 2004). Similar silencing of the PDS gene was achieved in *Nicotiana* species with a satellite DNA associated with *Tobacco curly shoot virus* (family *Geminiviridae*) (Qian *et al.* 2006), in the legume *Pisum sativum* (pea) with *Pea early browning virus* (genus *Tobravirus*) (Constantin *et al.* 2004), and in monocotyledonous hosts with BMV (Ding *et al.* 2006).

VIGS can also be used to target a transgene. For example, Gammelgård *et al.* (2007) observed a loss of GFP expression in upper leaves of *gfp* transgenic *N. benthamiana* soon after the plants were infected with a *gfp*-carrying PVA. Similarly, *Poplar mosaic virus* (genus *Carlavirus*) carrying the GFP-encoding sequence silenced the GFP transgene in *N. benthamiana* (Naylor *et al.* 2005). The authors expressed their intention to use the vector virus for VIGS in future studies in poplar (genus *Populus*).

In the following sections the main emphasis is given to the aspects of development and use of vector-viruses as heterologous protein overexpression tools, which has also been the main focus of this thesis.

Plant virus-based expression of heterologous proteins for industrial uses

Plants as heterologous protein production platforms

The obvious advantages of plants as 'bioreactors' for heterologous protein production, as compared to animal/insect cell-cultures and bacterial liquid cultures, are the scalability of the system together with the absence of animal pathogens. Furthermore, the materials, facilities and maintenance needed for

the plants are relatively cheap. These and other aspects of heterologous protein production in plants have been discussed by Twyman *et al.* (2003) and Ma *et al.* (2003). Transient expression of proteins from vector-virus and *Agrobacterium* infiltrated to leaves (Gleba *et al*. 2007, Fisher *et al.* 1999), protein expression in transgenic microalgae (León-Bañares *et al.* 2004, Franklin & Mayfield 2004) and moss (Decker & Reski 2004), and secretion of the proteins from plant roots to liquid culture medium (Borisjuk *et al.* 1999) have been reported. In addition, approaches for higher-level protein production (Streatfield 2006), differences in posttranslational protein modification between plants and animals (Gomord & Faye 2004), and safety issues concerning the genetically modified organisms (Mascia & Flavell 2004) have been reviewed.

The differences in protein *N*-glycosylation patterns between animals and plants is seen as one of the main problems in production of mammalian glycoproteins, especially immunoglobulins, in plants. The plant-specific N-glycan residues xylose and α-1,3-fucose (reviewed in Lerouge *et al.* 1998) bound to human proteins expressed in plants might be immunogenic when injected to humans. On the other hand, the animal-specific protein side-chains such as sialic acid and ß-1,4-galactose are missing from proteins expressed in plants. Attemps at 'humanizing' *N*-glycosylation of proteins expressed in plants have been undertaken. Bakker *et al.* (2006) showed that human immunoglobulins produced in tobacco plants transgenic for human ß-1,4-galactosyltransferase had galactose residues and low levels of xylose and fucose residues associated with them, whereas immunoglobulins produced in wt tobacco plants had no galactose but contained high levels of xylose and fucose. Cox *et al.* (2006) produced human antibodies with low amounts of incorporated xylose and fucose in duckweed (*Lemna minor*) by using RNA silencing to lower expression of the plant enzymes responsible for incorporating xylose and fucose to proteins.

Comparison of plant virus-based expression vectors and transgenic plants in heterologous protein production

Agrobacterium tumefaciens is used to deliver heterologous coding sequences inserted in the bacterial T-DNA (an expression cassette) to the plant cell nucleus, where the T-DNA is incorporated into the genome (Chahal & Gosal 2002). Plant cells can be transformed to express heterologous coding sequences by two *A. tumefaciens*-mediated approaches. Regeneration of plants from transformed cells results in transgenic plants that have the foreign sequence incorporated to all cells (Chahal & Gosal 2002). The foreign protein encoded by the transgene will be produced in all cells, depending on the promoter used. Obtaining transgenic plants takes a longer time than cloning the heterologous coding sequence in a vector-virus and using the vector virus as a vehicle for heterologous protein expression. However, *Agrobacterium* cells containing the expression cassette can also be delivered into mature leaves of non-transgenic plants (agroinfiltration), leading to expression of the heterologous protein only in the cells of the targeted leaf area (Kapila *et al.* 1997). While most of the T-DNA molecules are not incorporated into the genome, they are transcriptionally active for a few days (reviewed by Fisher *et al.* 1999). The difference of this method as compared to vector-viruses is that the expression cassettes do not replicate and spread cell-to-cell or systemically.

Typical yields of heterologous proteins in transgenic plants are *ca.* 0.1-1% of total soluble proteins (Ma *et al.* 2003, Abranches *et al.* 2005). Occasionally higher levels are reached. The phytase of *Aspergillus niger* expressed in the transgenic legume *Medicago truncatula* amounted to *ca.* 6.5% of total soluble proteins (Abranches *et al.* 2005). The amounts achieved with most vector-viruses are within the range of 1-10% of the total soluble proteins, and GFP levels as high as 50% of the total soluble proteins of GFP have been reported when expressed from a TMV-based vector (Gleba *et al.* 2007). On the other hand, in plants that have the transgene incorporated into the chloroplastic (or mitochondrial) genome, the yields of heterologous proteins

can reach 46% of the total soluble proteins (De Cosa *et al.* 2001) (reviewed by Daniell *et al.* 2002). Protein expression in chloroplasts is similar to that in bacteria, which means that most of the eukaryotic-type post-translational modifications of proteins will not take place (Twyman *et al.* 2003). The advantages of the approaches of using vector-viruses and transgenic plants can be combined in plants transformed with a viral replicon (infectious clone of the virus), discussed below.

DNA viruses as overexpression vectors

The first plant virus converted into an overexpression vector for production of heterologous proteins was CaMV (Fig. 1). Brisson *et al.* (1984) replaced 461 nt of the viral ORF II encoding an aphid transmission factor (479 nt) of CaMV with a bacterial dihydrofolate reductase encoding sequence (234 nt), conferring resistance to methotrexate in *E. coli*. The vector-virus spread systemically in turnip leaves (*Brassica rapa* cv. 'Just Right') and expression of target protein was observed. With the same vector, De Zoeten *et al.* (1989) obtained accumulation of active human interferon αD in infected turnips. In this case, the heterologous sequence was slightly longer (100 nt) than the deleted part of ORF II. Expression of heterologous sequences longer than 561 nt from this cloning site in CaMV was not successful (Fütterer *et al.* 1990).

Fig. 1. Genome organization of *Cauliflower mosaic virus* (family *Caulimoviridae*), consisting of a circular double-stranded DNA of 8024 nt shown here in a linear manner. The thick black horizontal line represents the DNA and the gray boxes represent the open reading frames / proteins encoded by them. The hatched box on the left is a promoter for the polycistronic 35S RNA containing open reading frames VII → V. The hatched box on the right is a promoter for the 19S RNA encoding a translational activator (TAV) needed for translation of the other genes. Heterologous sequences can be placed between the open reading frames I and III by replacing the insect transmission factor (ITF) encoding sequence with the heterologous coding sequence. MP, movement protein; CP, coat protein; RT, reverse transcriptase/RNaseH.

Virus species from another DNA virus family, the *Geminiviridae*, have also been used as vectors to produce heterologous proteins. For example, Hayes *et al.* (1988) showed expression of the bacterial neomycin phosphotransferase from a *Tomato golden mosaic virus* (Fig. 2) (genus *Begomovirus*) vector in systemically infected leaves. A *Wheat dwarf virus* (genus *Mastrevirus*) replicon expressing transposons was constructed by Laufs *et al.* (1990). Recently, an interesting geminivirus-based vector from a *Tomato yellow leaf curl virus* (genus *Begomovirus*) has been made (Peretz *et al.* 2007). The replicative dsDNA form of the virus genome, which moves cell-to-cell and systemically in host plants, is produced by the host without the help of any viral encoded polypeptides. Thus, the gene essential for viral ssDNA replication (rep, Fig. 2) could be interrupted with an insert of at least 5 kb, which is considerably longer than the viral monopartite genome (2781 nt).

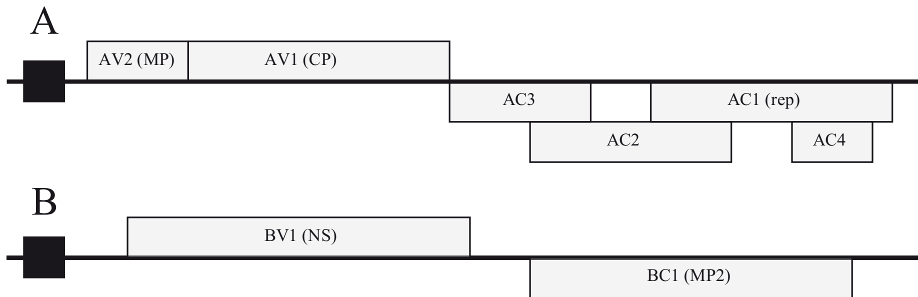


Fig. 2. Genome organization of *Tomato golden mosaic virus*, a bipartite begomovirus (family *Geminiviridae*), that consists of circular single-stranded DNAs of 2588 (A) and 2508 (B) nt shown here in linear manner. Monopartite begomoviruses have only DNA A. The thick horizontal line represents the DNA and the gray boxes represent the open reading frames / proteins encoded by them. In DNA A, two genes are encoded by the virion-sense strand (AV) and four genes on the complementary-sense strand (AC). The AC genes are all involved in replication. In DNA B, a single gene is encoded from each strand. The black box represents an almost identical region in the two DNA molecules from where bi-directional transcription begins. Heterologous sequences have been used to replace the CP gene, or inserted into the rep gene. MP, movement protein; CP, coat protein; rep, replication initiation protein; NS, nuclear shuttle protein.

Injection of this plasmid to plants caused systemic spread of the replicon. Removing a part of the coat protein gene attenuated disease symptoms considerably. The target protein yields were *ca.* 6% of the total soluble leaf proteins, which is more than that achieved with most vector-viruses. The vector-virus was able to replicate and spread in all the plant species tested including monocots and dicots, and woody plants, but it was mostly limited to the phloem cells in many hosts.

RNA viruses as overexpression vectors

Most of the described vectors for expression of heterologous proteins in plants are based on RNA viruses. Viruses with isometric and rod/filamentous-shaped virions have been used. Cloning of an RNA virus as an infectious cDNA copy was necessary for this invention and was first published by Ahlquist *et al.* (1984) for BMV (Fig. 3). In the first vector-viruses a part or all of the CP gene of the virus was replaced by a region encoding a foreign protein, which was the case in BMV (Fig. 3) (French *et al.* 1986), TMV (Fig. 4) (Takamatsu *et al.* 1987), BSMV (Joshi *et al.* 1990) and PVX (Fig. 5) (Chapman *et al.* 1992). Since the CP is the most abundant protein produced by these viruses, it was reasoned that similar high amounts of heterologous proteins should be produced from the CP replacement vector-viruses. However, the problem was that the vector-viruses were unable to enter the phloem for a fast systemic spread. To obtain high yields of heterologous protein, each leaf would have to be inoculated separately. Subsequently, the heterologous sequence was placed between the MP and CP genes of TMV under a duplicated subgenomic promoter for CP, which allowed phloem-assisted movement (Dawson *et al.* 1989). However, these vector viruses soon lost their insert when inoculated into plants, due to recombination between the duplicated homologous CP promoters. To avoid this problem, the subgenomic CP promoter of a related virus (*Odontoglossum ringspot virus*) was succesfully used to drive the expression of the foreign protein encoding sequence in the vector-TMV (Donson *et al.* 1991).

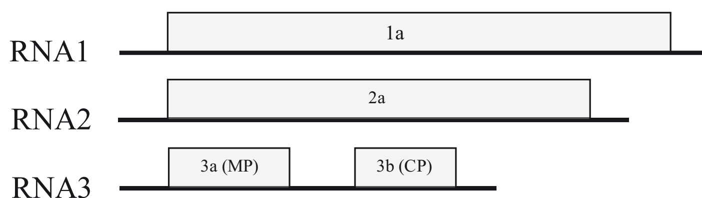


Fig. 3. Genome organization of *Brome mosaic virus* (family *Bromoviridae*), that has (+)-sense single-stranded RNAs of 3234 (RNA1), 2865 (RNA2) and 2117 (RNA3) nt. The thick horizontal lines represents viral RNAs and the gray boxes represent the open reading frames / proteins encoded by them. Genome organization of *Cucumber mosaic virus* (family *Bromoviridae*, genus *Cucumovirus*) is essentially similar, with one additional gene on RNA2 (2b, not shown) encoding an RNA silencing suppressor. Heterologous sequences have been placed either after the CP gene on RNA3, or used to replace the 2b gene on RNA2 of *Cucumber mosaic virus*. Open reading frames 1a and 2a encode polypeptides forming the replicase complex. MP, movement protein; CP, coat protein.

Fig. 4. Genome organization of *Tobacco mosaic virus* (genus *Tobamovirus*), that has a (+)-sense single-stranded RNA of 6395 nt. The thick horizontal line represents the RNA and the gray boxes represent the open reading frames / proteins encoded by them. The 126K protein contains the methyl transferase and helicase motifs of a replicase. The 183K protein (the complete replicase) is produced occasionally when the stop codon of the 126K gene (dashed line) is ignored. The movement protein, 30K (MP), and the coat protein (CP) are produced from subgenomic RNAs. Heterologous sequences have been placed under a duplicated CP promoter either between the MP and CP genes, or in place of the CP gene.

Fig. 5. Genome organization of *Potato virus X* (family *Flexiviridae*), that has a (+)-sense single-stranded RNA of 7568 nt. The thick horizontal line represents the RNA and the gray boxes represent the open reading frames / proteins encoded by them. Three separate subgenomic RNAs are made, from which the TGB1 and TGB3, TGB2, and CP are produced, respectively. Heterologous sequences have been placed under a duplicated CP promoter either between the TGB3 and CP genes or in place of the three TGB and the CP genes. Rep, replicase; TGB, triple-gene-block protein; CP, coat protein.

Using a similar engineering strategy, PVX was converted to a systemically moving overexpression vector, in which the heterologous protein encoding sequence was placed under a duplicated subgenomic promoter of CP that was located between the triple gene block 3 and CP genes (Fig. 5) (Chapman *et al.* 1992). Expression of GUS was detected in systemically infected leaves at 13 days post-inoculation (dpi). However, complete and partial deletions of the sequence encoding GUS were observed in a northern blot analysis, attributable to recombination between the homologous sequences of the duplicated subgenomic promoters (81 nt). Additional examples of PVX as a vector virus are shown in Table 1. In another potexvirus-based expression construct (*Zygocactus virus X*), the subgenomic promoter of CP directing the transcription of the heterologous coding sequence and most of the CP gene was replaced with corresponding sequence from a related virus (*Schlumbergera virus X*) (Koenig *et al.* 2006). Deletions were nevertheless observed in the heterologous *Beet necrotic yellow virus* (genus *Benyvirus*) CP gene expressed from the vector in infected plants. TMV-based vectors expressing either GFP or human growth hormone (hGH) were stable for a period of three years in *N. benthamiana* roots that were maintained in a liquid culture (subcultured every 6 weeks) (Skarjinskaia *et al.* 2008). The same GFP-expressing vector-virus generated deletions in the *gfp* sequence already four weeks post-inoculation when the vector virus multiplied in the aerial parts of tobacco (Rabindran & Dawson 2001).

Whole-plant agroinfiltration, by dipping the aerial parts of a plant in *Agrobacterium*-containing liquid and applying a weak vacuum, can be used to instantaneously spread a vector-virus to all parts of a host plant. This gives a simultaneous start for the vector-virus replication and heterologous protein production in all leaves of the plants (termed 'magnifection') (Marilloinnet *et al.* 2005). This method allows the replacement of the MP and/or CP genes of vector viruses with heterologous sequences, at least in some virus species. TMV lacking the CP gene has been used to produce large amounts of foreign

proteins (Marilloinnet *et al.* 2005, Gils *et al.* 2005, Dorokhov *et al.* 2007) (Table 1). TMV MP amounts in infected plants were increased 10-fold when the CP gene was deleted (Lehto *et al.* 1990), whereas increases of 8-fold and 4-fold in the MP amounts were observed in plants infected with TMV-constructs where 470 and 207 nt of the CP gene were deleted, respectively (Culver *et al.* 1993). The authors suggested that the closer a gene was to the 3′ terminus of TMV, the higher its expression levels. Another interesting example of a replacement virus vector is a PVX-construct where all viral genes except the replicase were removed and replaced with heterologous sequence (Komarova *et al.* 2006). The expression levels of GFP from this vector were *ca.* 2.5-fold higher than with PVX where the GFP encoding sequence was placed between the triple gene block and the CP genes (Fig. 5). Removal of the CP prevents systemic movement of TMV and PVX (Takamatsu *et al.* 1987, Chapman *et al.* 1992, Marilloinnet *et al.* 2005, Komarova *et al.* 2006) and thus spread of the genetically modified virus, which may be positive from the biosafety point of view. Expression of heterologous proteins from the vector-viruses can be further enhanced by co-inoculation with constructs to express heterologous RNA-silencing suppressor proteins, which increases the quantity of intact mRNAs (Lindbo 2007). Expression of GFP was enhanced 100-fold from a vector-TMV when *Tomato bushy stunt virus* silencing suppressor p19 was co-expressed in the plants.

TMV-based expression vectors have been commonly used, partly because of the high yields of target proteins obtained (Table 1). Over 50 proteins of different origins with potential pharmaceutical use have been expressed succesfully using the magnifection system alone (Klimyuk *et al.* 2005, Gleba *et al.* 2007).

In recent years the infectious cDNA clones of RNA viruses in other families/genera have been modified to express heterologous polypeptides, of which CPMV (Fig. 6) and CMV (Fig. 3) appear particularly successful (Table 1). The vector-CMV may have plenty of useful applications, given that it has over 1200 host plants in more than 100 families (Douine et al. 1979,

Edwardson & Christie 1991, Palukaitis & García-Arenal 2003). In addition, TRV, which is widely used for VIGS studies, can be used to express heterologous proteins (Fig. 7) (Table 1).

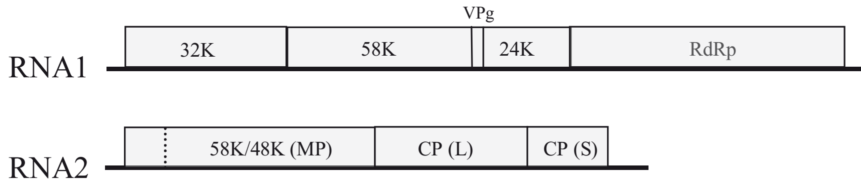


Fig. 6. Genome organization of *Cowpea mosaic virus* (family *Comoviridae*), that has (+)-sense single-stranded RNAs of 5889 (RNA1) and 3481 (RNA2) nt. The thick horizontal lines represent the RNAs and the gray boxes represent the genes / proteins encoded by them. The protein encoding regions from each RNA molecule are first translated as a major polyprotein that is subsequently processed into mature proteins, a strategy similar to that of members of the family *Potyviridae*. Heterologous sequences have been placed on RNA2 either between the MP and CP (L) genes, or in place of the MP, CP (L) and CP (S) genes. RNA1: 32K, 32 kDa cysteine proteinase; 58K, 58 kDa protein of unknown function; VPg, viral genome-linked protein; 24K, 24 kDa main viral proteinase; RdRp, RNA dependent RNA polymerase. RNA2: 58K/48K(MP), two proteins with overlapping cistrons – 58 kDa protein of unknown function and 48 kDa movement protein; CP (L), large coat protein subunit; CP (S), small coat protein subunit.

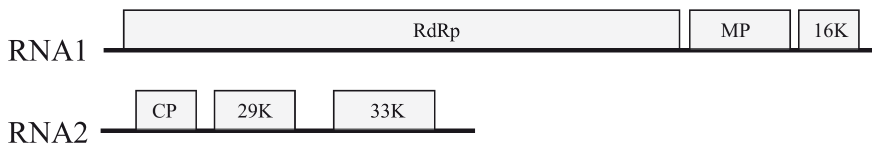


Fig. 7. Genomic organization of *Tobacco rattle virus* (genus *Tobravirus*), that has (+)-sense single-stranded RNAs of 6791 (RNA1) and 3855 (RNA2) nt. The thick horizontal lines represents the RNAs and the gray boxes represent the open reading frames / proteins encoded by them. The movement protein (MP), the 16K protein (an RNA silencing suppressor) and the coat protein (CP) are produced from different subgenomic RNAs. The 29K and 33K gene products from RNA2 are associated with vector transmissibility, and missing from some isolates. They can be replaced with heterologous sequences. RdRp, RNA dependent RNA polymerase.

Table 1. Vector-viruses based on virus species from genera other than *Potyvirus*.

Virus genus & species[1]	Cloning site, or the viral gene(s) replaced	Heterologous coding sequence(s) used	Target protein expression, or vector accumulation levels	Notes	Reference
Carlavirus					
PopMV	triple gene block and CP replacement	*gfp*	-	Intended to be used as a VIGS vector.	Naylor *et al.* 2005
Comovirus					
BPMV	RNA-2 (movement protein (MP) / large coat protein (L CP)	*Aequorea victoria gfp*, *Discosoma* DsRed, *Streptomyces* phosphinothricin acetyltransferase (*bar*), various RNA-silencing suppressors	GFP: ~1% of soluble plant proteins	Genome organization essentially the same as with CPMV (Fig. 6).	Zhang & Ghabrial 2005
CPMV	RNA-2 (MP/ L CP)	*gfp*		Genome organization in Fig. 6	Verver *et al.* 1998
CPMV	RNA-2 (MP/ L CP)	*gfp*, *Cowpea aphid-borne mosaic virus* HC-Pro	GFP: 1-2% of soluble plant proteins	GFP::CP fusion: GFP-tagged virions produced.	Gopinath *et al.* 2000
CPMV	RNA-2 (MP, L&S CP replacement)	*gfp*	At least 0.6% of soluble plant proteins in RNA-2-GFP transgenic plants.	CP-deficient RNA-2. Transgenic and transient expression.	Cañizares *et al.* 2006
CPMV	RNA-2 (MP, L&S CP replacement)	antibody heavy chain, antibody light chain	mature antibody: up to 74 μg/g leaf (FW)	Simultaneous production from separate RNA-2 vectors in inoculated leaves.	Sainsbury *et al.* 2007
Cucumovirus					
CMV	RNA-3 (after CP)	human α_1-antitrypsin	1.7 ± 0.5% of soluble proteins (70% in active form)	Under an inducible promoter.	Sudarshana *et al.* 2006
CMV	RNA-2 (2b replacement)	human acidic fibroblast growth factor (aFGF)	5-8% of soluble proteins in *N. benthamiana*, ~2.5% in soybean and ~1.5% in *Arabidopsis* tissues.	Shows deletion of the silencing suppressor 2b is possble. No severe disease symptoms inflicted as with wt CMV.	Matsuo *et al.* 2007
Hordeivirus					
BSMV	γRNA (after γb)	*gfp*	-	Used mainly as a VIGS vector.	Holzberg *et al.* 2002
Luteovirus					
PLRV	ORF5 3'-fusion	*gfp*	-	Aphid transmissible. *Gfp* sequence highly unstable.	Nurkiyanova *et al.* 2000
Potexvirus					
PVX	TGB3/CP	*UidA*	-	Duplicated CP promoter. Genome organization in Fig 5.	Chapman *et al.* 1992
PVX	TGB3/CP	TEV P1:HC-Pro duplex fused to *gfp*	-	First time 3 heterologous proteins expressed from a single vector-virus locus.	Anandalakshmi *et al.* 1998
PVX	TGB3/CP	rotavirus VP6	50 μg / g leaf (FW)	-	O'Brien *et al.* 2000
PVX	TGB3/CP	*Wasabia japonica* defensin	0.4 μg / g leaf (FW) (purified)	-	Saitoh *et al.* 2001
PVX	TGB3/CP	human *proinsulin*, murine *interleukin-10*, HIV-1 *nef*, petunia *expansin-1*, human *gad65*	-	Insert length 261 – 1758 nt, all intact in original plants 12 dpi, but not in passage 1 plants 12 dpi.	Avesani *et al.* 2006
PVX	triple gene block + CP replacement	*gfp*	~2.5 fold more GPF produced than with PVX where *gfp* is situated between the 8K and CP.	Only the replicase left of viral proteins. Needs agrobacterium delivery like all replacement vectors. PVX strain UK3.	Komarova *et al.* 2006

PVX	triple gene block + CP replacement	*UidA*	1.5-2 fold more GUS produced from PVXdt_Tula_GUS than from PVXdt_UK3_GUS. A similar difference was observed in CP levels in plants infected with corresponding wt strains.	Shows that by using a more virulent strain for making a vector-virus, higher heterologous protein yields can be attained.	Ravin *et al.* 2008
ZVX	TGB3/CP	CP of *Beet necrotic yellow vein virus* and *Soil-borne cereal mosaic virus*	-	CP and its promoter from a related *Schlumbergera virus X* to avoid homologous recombination.	Koenig *et al.* 2006
Sequivirus					
ALSV	RNA-2 (42KP/Vp25)	*gfp*, *Apple chlorotic leaf spot virus* ORFs	GFP: ~5 μg / g leaf (FW)	-	Li *et al.* 2004
Tobamovirus					
TMV	CP replacement	bacterial chloramphenicol acetyltransferase	~10 U / g leaf (FW)	-	Takamatsu *et al.* 1987
TMV	MP/CP	dihydrofolate reductase, neomycin phosphotransferase		1st systemically spreading plant RNA virus-based expression vector. CP and its promoter from a related *Odontoglossum ringspot virus*.	Donson *et al.* 1991
TMV	MP/CP	*Trichosanthes kirilowii* α-trichosantin	2% of soluble proteins	-	Kumagai *et al.* 1993
TMV	MP/CP	human papillomavirus capsid protein L1	20-37 ng / g leaf (FW)	-	Varsani *et al.* 2006
TMV	CP replacement	tuberculosis antigens Ag58B & ESAT6	Ag58B: 800 μg / g leaf (FW) ESAT6: 2 μg / g leaf (FW)	Co-expressing P19 RNA-silencing suppressor protein.	Dorokhov *et al.* 2007
TMV	CP replacement	*gfp*	4000 μg / g leaf (FW) in *N. benthamiana*. 2500 μg / g leaf (FW) in *N. tabacum*.	40% and 25% of soluble leaf proteins, respectively.	Marilloinnet *et al.* 2005
TMV	CP replacement	human growth hormone	1000 μg / g leaf (FW)	10% of soluble proteins.	Gils *et al.* 2005
TMV	MP/CP	*gfp*	600-1200 μg / g leaf (FW)	Co-expressing P19 RNA-silencing suppressor protein.	Lindbo 2007
TMV	MP/CP	*gfp*, human growth hormone (hGH)	GFP: 50-120 μg / g leaf (FW) hGH: 3-6 μg / g leaf (FW)	Expression in *N. benthamiana* liquid root cultures. Vectors stable for 3 years.	Skarjinskaia *et al.* 2008
TMV + PVX	CP replacement (in both vectors)	IgG antibody heavy chain & light chain	500 μg / g leaf (FW)	Co-infection with non-related vectors to achieve wide-spread double-infection of individual cells.	Giritch *et al.* 2006
Tobravirus					
TRV, PEBV, PepRSV	RNA-2 (2b & 2c replacement)	*gfp*, *Galanthus nivalis* lectin	TRV produced lectin: ~10 μg/g of roots (FW)	Vector-TRV is the most used vector virus for VIGS.	MacFarlane & Popovich 2000
Tritimovirus					
WSMV	NIb/CP	neomycin phosphotransferase (NPT), *UidA*	NPT: 300 μg/g soluble leaf protein	*UidA* unstable.	Choi *et al.* 2000

[1] ALSV, *Apple latent spherical virus*; BPMV, *Bean pod mottle virus*; BSMV, *Barley stripe mosaic virus*; CMV, *Cucumber mosaic virus*; CPMV, *Cowpea mosaic virus*; PEBV, *Pea early-browning virus*; PepRSV, *Pepper ringspot virus*; PLRV, *Potato leafroll virus*; PopMV, *Poplar mosaic virus*; PVX, *Potato virus X*; TMV, *Tobacco mosaic virus*; TRV, *Tobacco rattle virus*; WSMV, *Wheat streak mosaic virus*; ZVX, *Zygocactus virus X*.

RNA viruses of the genus Potyvirus *as overexpression vectors*

Potyviruses (family *Potyviridae*) have a single-stranded (+)-sense RNA genome of *ca.* 9500-10000 nucleotides encapsidated by *ca.* 2000 copies of a single species of CP into a filamentous virus particle. Approximately one-third of known plant viruses belong to this genus (Fauquet *et al.* 2005). The biological and molecular properties of potyviruses have been reviewed by Shukla *et al.* (1994) and Rajamäki *et al.* (2004).

The infection cycle of a potyvirus begins with entry of virions into plant cells, usually mediated by an aphid that probes leaf epidermal cells for finding a good feeding position. Aphids carry potyviral virions in a non-persistent manner in the tip of the stylet (Wang *et al.* 1996). Following entry to the cell, virions disassemble, a process that has been little studied in potyviruses. Co-translational disassembly, shown in TMV particles (Wu *et al.* 1990), is probably true also for potyviruses. The RNA genome associates with ribosomes and is translated into a large polyprotein. It is processed yielding up to ten mature proteins by three viral proteinases (reviewed by Riechmann *et al.* 1992). The first (P1) and the second (HC-Pro) protein cleave the polyprotein *in cis* at their respective C-termini. The other seven protein junctions are cleaved by the C-terminal domain of NIa-Pro *in cis* or *in trans* (Fig. 8). The cleavage site (marked with /) of P1 has been determined in TVMV (F_{274}/S_{275}) (Mavankal & Rhoads 1991), TEV (Y_{304}/S_{305}) (Verchot *et al.* 1992) and PVY (P_{284}/S_{285}) (Yang *et al.* 1998). Following alignment of the corresponding amino acid sequence area of 35 other potyvirus species, the consensus amino acid sequence surrounding the cleavage site of P1 was observed to be P or Y/S (Adams *et al.* 2005). The cleavage site of HC-Pro has been determined in TEV (G_{763}/G_{764}) (Carrington *et al.* 1989). The consensus amino acid sequence surrounding the cleavage site (YXVG/G, X marking a non-conserved residue) was observed to be extremely conserved in 38 potyvirus species analysed (Adams *et al.* 2005). NIa-Pro-mediated protein cleavage and its recognition sites in a substrate have been extensively studied (reviewed by Adams *et al.* 2005). The crystal structure of NIa-Pro of TEV

bound to a substrate showed that the amino acids of the substrate at postitions P6, P4, P3, P2, P1 and P1′ (using the nomenclature of Schechter & Berger 1967) made contact with the active site of the enzyme (Phan *et al.* 2002). This result agreed with a consensus amino acid sequence obtained from an alignment of amino acid sequences surrounding the 343 NIa-Pro cleavage sites in sequenced species of *Potyviridae* (Adams *et al.* 2005). Amino acids at positions P6, P4, P2, P1, P1′, and additionally P2′ and P3′ have 50% consensus, whereas 80% consensus is observed at position P4 (V or I), P2 (H, L or F), P1 (G) and P1′ (A, S or G).

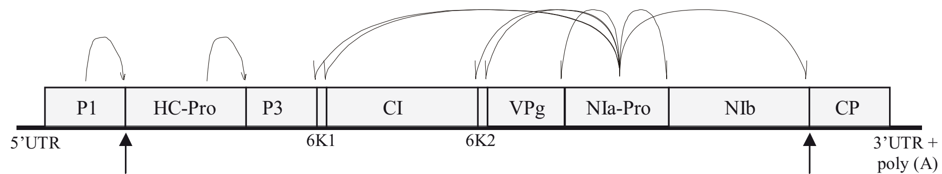


Fig. 8. Genomic organization of potyviruses that have a (+)-sense single-stranded RNA genome. The thick horizontal line represents the RNA and the gray boxes represent the polyprotein-encoding region and the mature proteins. The RNA encodes a large polyprotein which is subsequently processed into individual mature proteins by the P1, HC-Pro and NIa-Pro proteins. The arrows above the drawing point at the sites where the proteinases cleave the polyprotein. The arrows below the drawing point at the locations where heterologous sequences have successfully been placed in various potyvirus species (Table 3). 5′UTR, 5′-untranslated sequence; P1, protein 1 (proteinase); HC-Pro, helper component proteinase; P3, protein 3; 6K1, 6 kDa protein 1; CI, cylindrical inclusion protein; 6K2, 6 kDa protein 2; VPg, viral genome-linked protein; NIa-Pro, proteinase; NIb, polymerase; CP, coat protein; 3′UTR, 3′ untranslated region; poly(A), polyadenine tail.

Viral proteins and possibly some host proteins form a replication complex synthesizing the negative-strand complementary RNA that is used as a template for new genomic RNA copies (Wang & Maule 1995). Cell-to-cell movement through plasmodesmata and loading into sieve elements are assisted by the multifunctional viral proteins HC-Pro, CI, 6K2, VPg and CP (Table 2) (Dolja *et al.* 1994, Rojas *et al.* 1997, Nicholas *et al.* 1997, Rodríguez-Cerezo *et al.* 1997, Roberts *et al.* 1998, Carrington *et al.* 1998, Spetz & Valkonen 2004). Long-distance movement in the sieve elements (SE) follows the route of

photoassimilates between the maturing and growing tissues (the source-sink relationship) (Van Bel 2003). In the sink tissues, the virus is unloaded from SE into phloem cells (Ding *et al.* 1998, Rajamäki & Valkonen 2003), followed by invasion of mesophyll and epidermal cells by replication and cell-to-cell movement.

Table 2. Known functions of potyviral proteins and non-coding sequences. In addition, all of them have more or less an effect on viral replication and disease symptoms.

5'UTR	5'-untranslated region, enhancement of translation (Carrington & Freed 1990)
P1	Proteinase (Verchot *et al.* 1991) and an enhancer of RNA silencing suppression activity of HC-Pro (Pruss *et al.* 1997)
HC-Pro	'Helper component' in aphid transmission (Govier & Kassanis 1974), proteinase (Carrington *et al.* 1989), suppressor of RNA silencing (Anandalakshmi *et al.* 1998, Kasshau & Carrington 1998, Brigneti *et al.* 1998), cell-to-cell movement (Rojas *et al.* 1997), vascular movement (Cronin *et al.* 1995)
P3	Protein 3, associated with CI structures in cytosol (Rodríguez-Cerezo *et al.* 1993), associated with NIa in the nucleus and nucleoli (Langenberg & Zhang 1997)
6K1	6 kDa protein 1, symptom development (Riechmann *et al.* 1995)
CI	Cylindrical inclusion protein, RNA helicase (Laín *et al.* 1990), cell-to-cell movement (Carrington *et al.* 1998)
6K2	6 kDa protein 2, vascular movement (Spetz & Valkonen 2004), an integral membrane protein proposed to anchor the viral replication complex to ER membranes (Schaad *et al.* 1997a)
VPg	Viral genome-linked protein (Murphy *et al.* 1991), cell-to-cell movement (Nicolas *et al.* 1997), vascular movement (Schaad *et al.* 1997b)
NIa-Pro	Nuclear inclusion protein a, proteinase (Carrington & Dougherty 1987)
NIb	Nuclear inclusion protein b, RNA-dependent RNA polymerase (Hong & Hunt 1996)
CP	Coat protein, encapsidation of viral RNA (McDonald & Bankcroft 1977), aphid transmission (Atreya *et al.* 1990), cell-to-cell movement (Dolja *et al.* 1994), vascular movement (Dolja *et al.* 1995)
3'UTR	3'-untranslated region, symptom induction (Rodríguez-Cerezo *et al.* 1991a)
poly(A)	Polyadenine tail of 15-500 residues (Hari 1981, Láin *et al.* 1988)

The first vector-potyvirus was made of TEV (Fig. 8) that carried the GUS encoding sequence between the P1 and HC-Pro encoding sequences (Dolja *et al.* 1992). Since then at least eight other potyvirus species have been used to produce heterologous proteins (Table 3). Vector-virus design in potyviruses differs from most other plant viruses as there are no subgenomic promoters from which the heterologous sequence could be expressed. Instead, the target protein is translated as part of the polyprotein and cleaved from it using the original and the engineered novel recognition sites for the viral proteinases. Introduction of a novel pentapeptide cleavage site for NIa-Pro was first done in a TEV-GUS vector, where it was placed between the GUS and the viral HC-Pro encoding sequences (Carrington *et al.* 1993). The GUS and HC-Pro were separated from each other by the *in trans* action of NIa-Pro in infected protoplasts. One or more virus-derived amino acids remain at the termini of the expressed heterologous proteins after separation from the polyprotein.

Table 3. Vector-viruses based on virus species from genus *Potyvirus* (family *Potyviridae*).

Cloning site[1]	Potyvirus[2]	Heterologous coding sequence(s) used[3]	Target protein expression / vector accumulation levels	Notes	Reference
P1/HC-Pro	TEV	*UidA*	-	GUS:HC-Pro fusion proteins.	Dolja *et al.* 1992
	TEV	*Beet yellows virus* ORFs	From 1 to 70% of parent virus in plants, but from 50 to 100% in protoplasts, depending on the ORF.	The same vector as above.	Dolja *et al.* 1997
	TEV	*Streptomyces* phosphinothricin acetyltransfease *bar*, *Streptomyces* cytocrome $P450_{SU1}$, *UidA*	-	-	Whitham *et al.* 1999
	PPV	*UidA*	Wt virus CP levels, although it took twice the number of days.	-	Guo *et al.* 1998
	LMV	*gfp*, *E. coli UidA*	Wt virus CP levels and ~25% of wt levels with GFP and GUS expressing vectors, respectively.	Fusions to HC-Pro. Both inserts stayed intact longer than in most studies.	German-Retana *et al.* 2000
	LMV	*gfp*, *UidA*	Wt virus CP levels when GFP separated or fused to HC-Pro with one virus strain. With another, wt levels only when insert separated.	-	German-Retana *et al.* 2003
	PSbMV	*UidA*	-	GUS:HC-Pro fusion.	Johansen *et al.* 2001
	ZYMV	CMV coat protein, WMV coat protein	INF: ~2 µg/g leaf (FW)	Both inserts unstable.	Arazi *et al.* 2001
	ZYMV	*gfp*, dust mite allergen Der p 5	GFP: 3.7 µg/g leaf (FW) (purified), Der p 5: 1.5 µg/g leaf (FW) (purified)	-	Hsu *et al.* 2004
	ZYMV	tospoviral nucleocapsid proteins	12 – 25 µg/g leaf (FW) (purified)	-	Chen *et al.* 2005
NIb/CP	ClYVV	*Aequorea victoria gfp* and soybean glutamine syntethase (both in the same vector)	GFP: 20-50 µg/g leaf (FW), which is close to wild-type (wt) virus CP amounts	Simultaneous heterologous protein production.	Masuta *et al.* 2000
	PPV	*UidA*	-	-	Varrelmann & Maiss 2000
	PPV	*gfp*, *Rabbit hemorrhagic disease virus* VP60 protein	GFP: 250 µg/g leaf (FW), PPV CP: 390 µg/g leaf (FW)	-	Fernández-Fernández (2001)
	ZYMV	CMV coat protein, *gfp*, *UidA*, human interferon α2 (INF)	INF: ~2 µg/g leaf (FW)	*UidA* unstable. Vector-infected plants lack the severe symptoms seen in wt infected plants.	Arazi *et al.* 2001
	ZYMV	*Streptomyces* phosphinothricin acetyltransferase (*bar*)	Close to wt virus CP levels.	The same vector as above.	Shiboleth *et al.* 2001
	ZYMV	antiviral and antitumour proteins MAP30 and GAP31	-	The same vector as above.	Arazi *et al.* 2002
	ZYMV	*Trichoderma-induced MAPK* of cucumber, in sense and antisense orientation	-	The same vector as above.	Shoresh *et al.* 2006
	PVA	*gfp*	-	-	Ivanov *et al.* 2003
	TVMV	*gfp*, *Discosoma* DsRed	-	-	Dietrich & Maiss 2003
P1/HC-Pro + NIb/CP	TuMV	P1/HCpro: *gfp*, *UidA* NIb/CP: *gfp*, *UidA*	-	Simultaneous production. *UidA* unstable.	Beauchemin *et al.* 2005

[1]P1, protein 1 (proteinase); HC-Pro, helper component proteinase; NIb, replicase; CP, coat protein.

[2]ClYVV, *Clover yellow vein virus*; LMV, *Lettuce mosaic virus*; PSbMV, *Pea seed-borne mosaic virus*; PPV, *Plum pox virus*; PVA, *Potato virus A*; TEV, *Tobacco etch virus*; TuMV, *Turnip mosaic virus*; TVMV, *Tobacco vein mottling virus*; ZYMV, *Zucchini yellow mosaic virus*.

Epitope/peptide presentation vectors

It is not always necessary to produce a full-size functional protein. For example, short peptides can be incorporated into the viral CP at a location displayed on the surface of a mature virion. In most applications, the short peptide has been an antigen of an animal pathogen, and (partially) purified virions have been used as candidate vaccines (reviewed by Grill *et al.* 2005, Cañizares *et al.* 2005). The maximum length of the peptide that could be expressed was thought to be less than 25 amino acids in tobamoviruses until a peptide of 133 amino acids was succesfully displayed on the surface of a virion (Werner *et al.* 2006). This was made possible by using a flexible 15 amino acid linker peptide consisting mostly of glycine residues.

The first succesful epitope presentation system with plant viruses was achieved with CPMV (Usha *et al.* 1993). A *Foot-and-mouth disease virus* (FMDV) VP1 epitope displayed on the surface of mature virions reacted with FMDV-specific antiserum. Hence, it was possible to display heterologous antigenic peptides on the surface of plant virus virions, which opened the possibility to use epitope presentation vectors and their virions as vaccines. Many candidate vaccines produced in this manner induce specific antibody accumulation in animals (reviewed by Cañizares *et al.* 2005 and Grill *et al.* 2005). In some cases, protection against a lethal dose of a pathogen has been reported. For example, swine were immunized with *Bamboo mosaic virus* (genus *Potexvirus*) particles displaying FMDV epitopes, and they survived a subsequent exposure to a lethal dose of infectious FMDV (Yang *et al.* 2007). The vaccine candidates from the vectors releasing the antigen to the plant cell cytosol also have been shown to elicit immunoresponse in animals (Pérez Filgueira *et al.* 2003, Wagner *et al.* 2004) and humans (Reddy *et al.* 2002).

Vector viruses as stable transgenes

The infectious vector-virus DNA clone can be integrated into the host plant genome and, hence, virus inoculation of each new generation of plants avoided. This strategy for heterologous protein production is handy once the

replication, stability and expression of a desired insert-vector construct have been properly tested and found satisfactory. The procedure allows continuous large-scale heterologous protein production over a longer time. The first examples were in BMV (Mori *et al.* 1993) and PVX (Angell & Baulcombe 1997). However, the first vector-viruses tested as stably expressed transgenes mostly failed to give constant yields in regenerated plants and in subsequent plant generations. The reason for these difficulties was proposed to be transgene-induced transcriptional gene silencing associated with the transgene promoter and/or coding region methylation (Wassenegger & Pélissier 1998, Matzke *et al.* 2001). The second generation amplicons were therefore placed under inducible promoters. With a BMV-based replicon 30-230 fold higher transgene mRNA amounts were detected in *N. benthamiana* plants after glucocorticoid induction as compared to similar transgenic amplicons placed under the constitutively expressed 35S promoter (Mori *et al.* 2001). In transgenic potato plants transformed with an ethanol-inducible *Bean yellow dwarf virus* (family *Geminiviridae*) based vector, an 80-fold increase in mRNA levels and a 10-fold increase in translation products were observed (Zhang & Mason 2005). A variant of this approach was presented by Hull *et al.* (2005). The transgene inducer (yeast GAL4 DNA binding domain :: *Herpex simplex virus* VP16 transcription factor fusion protein) was expressed from a TMV-based vector in infected *UidA*-transgene harboring plants. The GAL4 DNA binding domain recognized and bound to the GAL4 specific sequence on the transgene promoter area, and the VP16 domain activated *UidA* transcription.

Potato virus A

PVA was first reported in *Solanum tuberosum* (family *Solanaceae*) in Ireland (Murphy & McKay 1932). It is found all over the potato-growing areas of Europe and North America (Hooker 1981). The known hosts are limited to the plant family *Solanaceae* (Bartels 1971). PVA is transmitted in a non-persistent manner by at least seven aphid species (including *Aphis frangulae*, *A. nasturtii*, *Myrzus persicae*), and can also be transmitted by mechanical inoculation, but

not in seed (Bartels 1971). Five strains of PVA (Ali, B11, Her, U, and TamMV) have been fully sequenced (Puurand *et al.* 1994, Kekarainen *et al.* 1999). The genome length is 9565 nt in isolates Ali, B11, and U, and 9567 nt in isolate Her, and 9672 nt in isolate TamMV, excluding the poly(A) tail. Isolate B11 originates from potato but has been propagated in various *Nicotiana* species since the early 1980s (Rajamäki *et al.* 1998). It can no longer systemically spread in most potato cultivars and is not transmitted by aphids (Andrejeva *et al.* 1999, Kekarainen *et al*. 1999). Strain B11 originates in Hungary, and strain U in North America (Valkonen *et al.* 1995). The most closely related potyvirus appears to be *Tobacco vein mottling virus* (TVMV) (Kekarainen *et al.* 1999, Nishiguchi *et al.* 2006). The particles of PVA are flexuous filaments of *ca.* 730 nm in length and *ca.* 11 nm in width (Brandes & Paul 1957).

In the only previous report of PVA as a vector-virus, nucleotide sequence encoding GFP was cloned in between the NIb and CP encoding sequences of PVA (Ivanov *et al.* 2003). The GFP-expressing PVA with amino acid substitutions at putative phosphorylation sites within the CP was observed under UV-light to be restricted to single cells in *N. benthamiana*, whereas the PVA-GFP without the substitutions was able to move cell-to-cell and systemically.

AIMS OF THE STUDY

The overall aim was to develop an infectious PVA clone into a versatile heterologous protein expression vector that could be used for research purposes and other applications in plants.

The specific aims were:

1) To test whether a full-size heterologous protein encoding sequence can be inserted into a novel putative cloning site inside the P1 encoding sequence and viral functions of replication and systemic movement retained.

2) To investigate whether certain human proteins can be expressed in active form in plants from the PVA vector.

3) To combine several cloning sites in a single vector-PVA, and to test simultaneous expression of up to three heterologous proteins.

MATERIALS & METHODS

The PVA-based expression vectors

All the constructed vector-viruses were based on the infectious clone of PVA strain B11 (Puurand *et al.* 1996). The clone was originally under the bacteriophage T7 RNA polymerase promoter, which was later changed to CaMV 35S DNA polymerase promoter, to allow inoculation of the infectious cDNA clone of the virus by using particle bombardment (unpublished). Detailed description of the construction of the vector-viruses (Table 4) (Fig. 9) can be found from the publications referred to in table 4. A brief overview is given below.

The cloning site within the P1 encoding region

Kekarainen *et al.* (2002) made a *Mu*-transposon insertion (15 nt) library from the PVA B11 clone. The transposon inserted a 15-nt sequence to the cDNA of the virus randomly and the clones in the resultant mutant library contained only a single insertion. The inserted heterologous sequence did not change any of the viral amino acids, nor did it change the reading-frame. One mutant of this library contained an insertion in the P1 encoding region (genomic postion 235) (Fig. 9 A) and was able to spread systemically in tobacco plants (Kekarainen *et al.* 2002). The insertions in the mutant library contained a unique recognition sequence for the *Not*I endonuclease. Hence, the insertion in P1 could be used as a cloning site (CS1). In this study, the *Not*I site at CS1 was used to insert a GFP encoding sequence (714 nt) into the mutant. In one subclone, the sequence encoding a novel heptapeptide cleavage site for the NIa-Pro proteinase was added to the 3'end of the GFP encoding sequence. Following proteolytic cleavage, the first 25 amino acids of P1 remained attached to the N-terminus of the expressed GFP. In another construct, the sequence encoding the NIa-Pro site was added at both sides of the *gfp* for separation of the GFP from the viral protein (Fig. 9B). In this construct a total of 768 non-viral nt were incorporated into the P1 encoding sequence.

Whenever GFP expression from P1 (CS1) is discussed in further parts of this thesis, it refers to this subclone, unless otherwise indicated. The detailed amino acid composition flanking the GFP within the P1 is presented in Fig. 9B. In addition, a subclone with a partial *gfp* gene (the first 123 nt of the coding sequence) was made without adding the sequences encoding the NIa-Pro mediated cleavage sites.

The cloning site between the P1 and HC-Pro encoding regions

The P1/HC-Pro cloning site (CS2) was located at the genomic position 1062 between the third and fourth codons of HC-Pro (Fig. 9). The sequence consisting of the first three codons of HC-Pro was duplicated at the 3'end of CS2 to allow full-length HC-Pro production in infected plants (Fig. 9B). The coding sequence of sea anemone *Renilla reniformis* luciferase gene (*Rluc*) (933 nt) was inserted into this site. To enable separation of the heterologous protein from the HC-Pro, a 21-nt sequence encoding the heptapeptide NIa-Pro cleavage site was incorporated at the 3'end of the luciferase sequence (Fig. 9 B). Consequently, luciferase was expressed either as a fusion to HC-Pro or as a free protein.

The cloning site between NIb and CP encoding regions

The third cloning site (CS3) between the replicase (NIb) and CP (Fig. 9) was initally used for testing the expression of human proteins with a vector-PVA. Proteins expressed from this vector contained 32 and 27 additional amino acids at their N- and C-termini, respectively (Fig. 9 B), most of which were derived from a duplicated viral sequence (132 nt, genomic sequence 8478-8346). The coding sequences for soluble catechol-*O*-methyltransferase (S-COMT) (663 nt) or sorcin (597 nt), encoding a Ca^{2+}-binding protein, were inserted into this cloning site using the previously engineered *Bfr*I and *Mlu*I endonuclease sites in CS3 (Ivanov *et al.* 2003). Also the coding sequence of *E. coli UidA* (1809 nt) encoding β-glucuronidase (GUS) was cloned into this site.

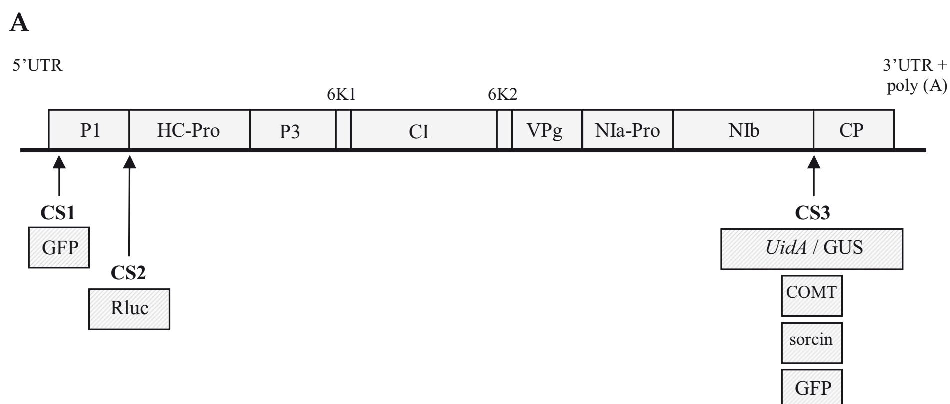
Subsequently, two insertion sites for heterologous sequences were combined in the vector-virus clones (CS1 and CS2, CS1 and CS3, CS2 and CS3) (Table 4) (III). Finally, a clone including all three sites was made (PVA-3i) (III). The *gfp*, *Rluc*, and *UidA* coding sequences were inserted in CS1, CS2, and CS3 of these four vector-viruses, respectively (Fig. 9A).

Table 4. Heterologous protein expressing vector-virus plasmid constructs created.

Plasmid name	Plasmid name used in the reference	Heterologous coding sequence/protein expressed and its source	Reference
PVA-CS1(*gfp*)	M14-pGFPp, pG00	green fluorescent protein, *Aequorea victoria*	I, III
PVA-CS2(*Rluc*)	pRluc, p0L0	luciferase, *Renilla reniformis*	III
PVA-CS3(*gfp*)	vPVA-GFP(59)	see above	II
PVA-CS3(*S-COMT*)	PVA-^{his}S-COMT(59)	soluble catechol-*O*-methyltransferase, *Homo sapiens*	II
PVA-CS3(*UidA*)	p00G	β-glucuronidase, *Escherischia coli*	III
PVA-CS3(*sorcin*)	PVA-^{his}sor(59)	sorcin, a Ca^{2+}-binding protein, *Homo sapiens*	II
PVA-CS3(*sorcin*-NIa_v2.0)	PVA-^{his}sor(9)	see above	II
PVA-CS3(*gfp*-NIa_v3.0)	-	see above	unpubl.
PVA-CS1(*gfp*)+CS2(*Rluc*)	pGL0	see above	III
PVA-CS1(*gfp*)+CS3(*UidA*)	pG0U	see above	III
PVA-CS2(*Rluc*)+CS3(*UidA*)	p0LU	see above	III
PVA-CS1(*gfp*)+CS2(*Rluc*)+CS3(*UidA*) or PVA-3i for short	pGLU	see above	III

Optimization of the NIb/CP cloning site

To minimize the number of additional amino acids remaining in the heterologous proteins when they are expressed from CS3, two modified versions of it were made. In CS3(NIa_v2.0) two unique hexanucleotide endonuclease recognition sites (*Bfr*I and *Mlu*I) were added at the genomic position 8535 that is between the first and the second codon of the CP gene (Fig. 9B) (II). Following the endonuclease sites, a 15-nt sequence was added, encoding a pentapeptide (VYFQ/A) that allows NIa-Pro mediated cleavage (Fig. 9B). In CS3(NIa_v3.0), three unique hexanucleotide endonuclease recognition sites (*Sac*II, *Xma*I and *Avr*II) were added at the genomic position 8514 that is between the 508th and 509th codon of the NIb gene (Fig. 9B) (unpublished). A 21-nt sequence, encoding a heptapeptide (DMVYFQ/A) that allows NIa-Pro mediated cleavage, was included between the *Sac*II and *Xma*I recognition site sequences (Fig. 9B).



B

Cloning site	Virus clone	Amino acid sequence
CS1(P1)	wt PVA	APVAAI
	PVA-CS1(*insert*)	APV**AAAADMVYFQ/A -insert- DMVYFQ/AAAAV**AAI
CS2(P1/HCpro)	wt PVA	HHY/STGDVF
	PVA-CS2(*insert*)	HHY/STG**LK -insert- DMVYFQ/ATR***STG*DVF
	PVA-CS2(*insert*:HC-Pro)	HHY/STG**LK -insert- TR***STG*DVF
CS3(NIb/CP)	wt PVA	EEDDMVYFQ/AETL
	PVA-CS3(*insert*)	EEDDMVYFQ/A***ETLDASEALAQKSEGRKKEGESNSSTAGLLKSG* - insert- PRTRRNHKCGH*SIQFDEQMDEED*DMVYFQ/A**ETL
	PVA-CS3(NIa_v2.0)	EEDDMVYFQ/A**LK -insert- TR*VYFQ/A***ETL
	PVA-CS3(NIa_v3.0)	EED**PR*DMVYFQ/A*PG -insert- PR**DMVYFQ/AETL

Fig. 9. Schematic presentation of *Potato virus A* (PVA) based vector-viruses and the structure of the cloning sites. **A)** The horizontal black line represents the RNA genome, and the gray boxes represent the protein-encoding regions and the corresponding mature proteins. Arrows point at the cloning sites. Hatched gray boxes below them represent the expressed heterologous sequences / proteins. Transcription of the cDNA clone was driven by the *Cauliflower mosaic virus* 35S promoter. GFP, *Aequorea victoria* green fluorescent protein; Rluc, *Renilla reniformis* luciferase; GUS, *E. coli* β-glucuronidase; COMT, human (soluble) catechol-*O*-methyl transferase. **B)** Amino acid sequences flanking the heterologous proteins at the three insertion sites. The added amino acids are in bold, the duplicated viral amino acids are in italics, the heptapeptide NIa-Pro recognition sites are boxed, and the NIa-Pro cleavage site is marked with a slash. Wt, wild-type.

Methods for virus inoculation and detection, and for analysis of expressed heterologous proteins

The experimental methods applied in the study are listed in Table 5. Detailed descriptions of the methods can be found from the publications referred to in the table.

Table 5. Various methods used during the study.

Method	Reference
Agrobacterium-assisted protein expression cassette delivery into plants	I
Affinity purification of proteins	II
Double antibody sandwich – enzyme linked immunosorbent assay (DAS-ELISA)	I, II, III
Electroporation of tobacco protoplasts	I
Enzyme activity assay – β-glucuronidase	I, III
Enzyme activity assay – luciferase	III
Enzyme activity assay – S-COMT *	II
Fluorometric GFP quantitation	Remans *et al.* 1999
Immunocapture – reverse transcription – polymerase chain reaction (IC-RT-PCR)	I, II, III
Microscopy – stereomicroscope	I, II, III
Microscopy – immunosorbent electron microscopy (ISEM) *	III
Nucleic acid spot hybridization (NASH)	I
Photography of plants under UV-light	I, II
Plant growing – conditions and fertilization	I, II, III
Protein blotting & immunodetection	I, II, III
Protoplast isolation (tobacco)	I
Quantitation of proteins – PAGE and SYPRO Ruby staining	II
Real-time – polymerase chain reaction	I
RNA blotting and RNA/DNA probe based detection – mRNA	I
RNA blotting and RNA/DNA probe based detection – siRNA	I
Standard cloning and related RNA/DNA manipulation	I, II, III
Statistical data analysis	I, III
Virus inoculation into plants – plasmid-coated microprojectile bombardment	I, II, III
Virus inoculation into plants – plant sap	I, II, III

*not conducted by the author

RESULTS & DISCUSSION

Infectivity of the PVA-based vectors in *N. benthamiana*

All PVA-based expression vectors with an insert at CS1, CS2 or CS3, with inserts in two of them, or with inserts in all three cloning sites (Fig. 9) were able to spread systemically in *N. benthamiana* plants. Practically all the plants inoculated with the different vector clones with one or two inserts by particle bombardment were systemically infected.

All plants were infected when PVA-CS3(*gfp*) was inoculated to *N. benthamiana* leaves using particle bombardment, although usually only two or three GFP expressing infection foci per shot were observed on the inoculated leaves under a UV-microscope at 4 dpi. Previously, particle bombardment with wt PVA on leaves of potato hybrid 'A6' resulted in 10-20 necrotic lesions due to hypersensitive response when the optimized conditions were used (Kekarainen & Valkonen 2000). The number of initial infection sites observed in this study was lower, which could be explained by the use of a different host plant, higher sensitivity of the response in 'A6', or a possibly decreased infection capacity of the PVA-CS3(*gfp*) as compared to the wt virus.

A vector virus with three inserts, *gfp*, *Rluc*, and *UidA* coding sequences (PVA-3i) inserted at CS1-CS3, respectively, spread systemically in only 19% of *N. benthamiana* plants inoculated by particle bombardment. However, mechanical inoculation with PVA-3i resulted in systemic infection in all 15 *N. benthamiana* plants. The inoculum used was leaf sap from a systemically infected *N. benthamiana* plant. The reverse-transcription-PCR test indicated that all 3 inserts were intact in the PVA-3i in the inoculum and in the systemically infected plants (III).

CS1 was a novel cloning site not previously tested in potyviruses. A previous attempt to express GUS as an N-terminal fusion to P1 in TEV was unsuccesful (Dolja *et al.* 1992). Applicability of CS2 was known to be variable depending on the potyvirus species. In ZYMV (Arazi *et al.* 2001), inserts that were found to be labile at the P1/HC-Pro site were stable at the NIb/CP site.

With *Turnip mosaic virus* (TuMV) vector constructs, the outcome was insert-specific. *UidA* was more stable at the NIb/CP site, whereas *gfp* was tolerated similarly at both cloning sites (Beauchemin *et al.* 2005). However, in another study, *gfp* was better suited for the NIb/CP site in all six host species tested, while a sequence encoding a dust mite allergen worked similarly well at both cloning sites (Chen *et al.* 2007). CS3 has been succesfully applied to express heterologous proteins in many potyviral expression vectors (Table 3).

Disease symptoms and accumulation of the vector-viruses in *N. benthamiana*

The wt PVA isolate B11 caused deformation and chlorosis of the systemically infected leaves and slight stunting of growth in *N. benthamiana* plants. Later on, dark-green islands developed producing mosaic-like pattern in the leaves (Fig 2B in I). A positive correlation between the severity of symptoms and high virus titers was observed in the plants infected with wt PVA and PVA-based vectors (Table 6). The likely cause of this is the diversion of host metabolism to the production of the viral nucleic acids and proteins, and the other effects of the viral proteins. The potyviral HC-Pro is capable of suppressing the host RNA silencing mechanism (Llave *et al.* 2000, Mallory *et al.* 2002) by sequestering the double-stranded small interfering RNA molecules (siRNAs) generated from the silencing-inducing dsRNA, for example, the viral RNA. HC-Pro prevents incorporation of the siRNAs into RISC, which prevents amplification of RNA silencing (Lakatos *et al.* 2006). In addition, HC-Pro can interfere with the host micro-RNA (miRNA) species that have a role in the post-transcriptional regulation of host gene expression (Kasschau *et al.* 2003, Chapman *et al.* 2004). Many miRNAs control transcription factors involved in developmental processes (Voinnet 2005). In the infection front where *Pea seed-borne mosaic virus* (family *Potyviridae*) replication is highly active in cotyledons of pea, the mRNAs for nine starch biosynthesis enzyme genes are down-regulated and the accumulation of corresponding enzymes tested within the infection front is suppressed, when

compared to healthy cotyledons (Wang & Maule 1995). Behind the infection front, gene expression recovers and higher enzyme amounts than in healthy cotyledons are detected. The levels of mRNAs for heat shock protein 90, polyubiquitin and glutathione reductase 2 are transiently upregulated at the infection front, while expression of actin, tubulin, and pea heat shock transcription factor genes show no change (Aranda *et al.* 1996, Escaler *et al.* 2000).

Attenuation of symptoms can be useful in research applications in which visual observations are needed. It can also be beneficial in target protein production as growth retardation and necrosis of cells can be avoided. A natural mutant causing only mild symptoms was isolated from a ZYMV-infected cucurbit plant that did not display the usual severe symptoms and was subsequently used as the parent for a vector-virus (Arazi *et al.* 2001). A single amino acid substitution in the middle region of HC-Pro was found to cause attenuation of symptoms. A heterologous sequence in the viral genome can also often cause attenuation of symtoms, for example in a vector-CMV expressing human acidic fibroblast growth factor in *N. benthamiana* and soybean (*Glycine max*) (Matsuo *et al.* 2007).

Influence of the inserts within the P1-encoding region in single-insert vectors (I, III)

The full-size coding sequence of *gfp* was tolerated at CS1 located in the P1 encoding region of PVA (I). Symptoms and accumulation levels of PVA with insertions of 15 nt (the insert from the transposon, Kekarainen *et al.* 2002) or 168 nt (fragment of the *gfp*) in CS1 were similar to those observed with the wt PVA in systemically infected leaves of *N. benthamiana* plants (I). The symptoms included chlorosis, dark-green islands and severe deformation of the leaves (Fig. 2A in I). The vector carrying a substantially larger insert (*gfp*, 714 nt) caused no leaf deformation, less chlorosis and a new symptom of green vein banding in the leaves. When the plants were grown under stronger light (250 µE $m^{-2}s^{-1}$ instead of 100 µE $m^{-2}s^{-1}$), chlorosis was almost absent and no green vein banding developed. The effects of light intensity on vector-

virus accumulation and alteration of symptoms was not further studied. In four experiments, the amounts of CP detected in the leaves systemically infected with PVA-CS1(*gfp*) were 39-55% of the wt PVA levels (Table 6). The difference in accumulation of these viruses was similar also in inoculated leaves of *N. benthamiana* and in transfected protoplasts of *N. tabacum* (Table 2 in I). The results with protoplasts indicated a somewhat impaired replication of PVA-CS1(*gfp*) as compared to wt PVA. To study the rate of systemic movement of PVA-CS1(*gfp*), inoculated leaves of *N. benthamiana* plants were excised at various periods of time post-inoculation (I). From 11 to 22% of the plants became systemically infected when a PVA-CS1(*gfp*) inoculated leaf was removed compared to 67-78% of the plants when a wt PVA inoculated leaf was removed 48 hours post-inoculation (three experiments). The rate of systemic spread of PVA-CS1 was similar with the 15-nt transposon insertion and with the 168-nt *gfp* fragment as with wt PVA.

Influence of inserts at the P1/HC-Pro site in single-insert vectors (III)

Previous attempts to engineer a cloning site between the first and second codon of HC-Pro of PVA were not successful (Andres Merits, personal communication). This was unexpected because a cloning site exactly between the P1 and HC-Pro in PPV (Guo *et al.* 1998) and a site between the first and second codons of HC-Pro in LMV (German-Retana *et al.* 2000) had been exploited successfully. Also, no other N-proximal amino acid of HC-Pro other than the serine directly after the cleavage site was observed to be conserved in 38 potyvirus species analysed, and, hence, thought to be required for the P1-mediated proteolysis to occur in potyviruses (Adams *et al.* 2005). However, a cloning site (CS2) between the third and fourth amino acid of HC-Pro was applicable in LMV (German-Retana *et al.* 2003), and this site was found to work also for PVA (H. Vihinen & K. Mäkinen, personal communication).

The sequence encoding *R. reniformis* luciferase was inserted at CS2 (III). The effect of adding of a sequence (21 nt) encoding a novel NIa-Pro proteinase cleavage recognition site at the 3’end of luciferase in CS2 was tested. Two

vector-virus subclones were made. From one of them, luciferase would be produced as a fusion with HC-Pro. From the other one, luciferase would be produced as a free protein following separation from the viral polyprotein at the engineered proteolytic site for NIa-Pro. Both of these viral constructs spread systemically in plants. The amounts of CP detected in the systemically infected leaves of four plants were 25% and 54%, respectively, of the insertless PVA. Both vector versions induced identical mild chlorosis symptoms in the systemically infected leaves but no severe leaf malformation, in contrast to the wt PVA. The vector-virus with the engineered NIa-Pro site, PVA-CS2(*Rluc*), was used in further studies. Its accumulation in systemically infected leaves was 40–75% of that of the insertless virus PVA in different experiments (III) (Table 6).

Influence of inserts at the NIb/CP site in single-insert vectors (II, III)

Four heterologous coding sequences (*gfp*, *UidA*, *S-COMT* and *sorcin*) were inserted at CS3, resulting in expression vectors PVA-CS3(*gfp*), PVA-CS3(*UidA*), PVA-CS3(*S-COMT*) and PVA-CS3(*sorsin*). The *gfp*, *S-COMT* and *sorcin* sequences are about the same size (597-714 nt), while *UidA* (1809 nt) is almost three times larger. All four vector-viruses caused similar chlorosis and malformation symptoms in the systemically infected leaves of *N. benthamiana* plants. Leaf malformation was not as severe as with wt PVA.

When potato poly(A)-binding protein (PABP) was expressed from CS3, systemically infected leaves in *N. benthamiana* displayed a striking vein chlorosis symptom with little or no leaf deformation that was not observed with wt PVA or any other vector-PVA (data not shown). It is tempting to speculate that the symptom was a result of RNA silencing activated against the *N. benthamiana* PABP sequence by the potato PABP coding sequence in the vector-virus. No other heterologous sequence used in this study was homologous to genes in *N. benthamiana.* The partial sequences of *N. benthamiana* PABP gene available show several identical areas of 21-nt or longer as compared with the potato PABP gene sequence.

Accumulation of the vector-viruses was similar irrespective of the heterologous sequences inserted into CS3. The CP titers reached almost those of the wt virus CP levels, even with the long *UidA* insert (Table 6). The titers of the viruses with an insert at CS3 were consistently higher than those with PVA-CS1(*gfp*) and somewhat higher than those acquired with PVA-CS2(*Rluc*) (Table 6). However, more precise comparison of the vectors with CS2 to those with CS1 or CS3 in terms of virus accumulation would require, for example, the *gfp* to be cloned into CS2 and expressed as a free protein as already is the case in CS1 and CS3.

Influence of multiple inserts (III)

Double-insert vectors

Constructs with the three possible double-insert combinations were made (Table 4). The same heterologous sequences in the same CSs as in PVA-3i were used. Symptoms caused by the double-insert vector-viruses were similar or somewhat milder than those caused by the single-insert vectors with the same inserts. Accumulation of the viral CP with the double-insert vectors was approximately half of the amount obtained with the single-insert vectors (Table 6; Fig. 10, top panel). The vector containing inserts in CS1 and CS2 accumulated to lower titers than the vector carrying inserts in CS2 and CS3 (Table 6, Fig. 10). The vector carrying inserts in CS1 and CS3 accumulated to amounts that were between those of the aforementioned vectors.

The triple-insert vector

The construct PVA-3i had three cloning sites, CS1, CS2 and CS3, combined into a single vector-virus, and *gfp*, *Rluc*, and *UidA* coding sequences inserted into them, respectively. Simultaneous insertion of three heterologous sequences encoding a full-size protein in three genomic locations in a single plant virus vector has not been reported previously. From this vector, the three proteins were produced theoretically in equimolar amounts in the same

cell. PVA-3i caused no disease symtoms in *N. benthamiana* plants, probably due to the lower virus titers as compared to the insertless vector (~15% in two experiments) (Table 6). Systemic spread of PVA-3i and accumulation in the upper leaves was no different from, *e.g.*, PVA-CS1(*gfp*), as indicated by the even distribution of GFP throughout the leaves observed under UV-light.

Expression of three heterologous proteins, GFP and the TEV proteins P1 and HC-Pro, from a single insert has been previously achieved with a PVX vector (Anandalakshmi *et al.* 1998). A polyprotein consisting of the three proteins was produced, and subsequently separated into three proteins by the P1 and HC-Pro mediated cleavage at their respective C-termini. Production of two proteins using a potyviral vector has been reported previously with vectors based on *Clover yellow vein virus* (ClYVV) (Masuta *et al.* 2000) and TuMV (Beauchemin *et al.* 2005). The proteins are expressed either from the same site (P1/HC-Pro) and subsequenly separated by NIa-Pro cleaving at an engineered site between the two heterologous proteins (Masuta *et al.* 2000), or from two separate sites (P1/HC-Pro and NIb/CP) (Beauchemin *et al.* 2005). It is not known whether there are benefits of using separate cloning sites instead of a single site for expression of multiple heterologous proteins from potyviruses. A long heterologous sequence pushes the flanking viral sequences apart from each other, which might cause alterations in folding of RNA or the polyprotein and subsequently cause problems in replication and polyprotein processing. The successful use of a single site to express two or several foreign proteins may be dependent on which cloning site is used.

Previously, two heterologous proteins have been expressed within the same cell using co-infection with a TMV- and PVX-based vector-virus (Giritch *et al.* 2006). This approach might suffer from two problems. Firstly, co-infection probably does not occur in all cells. Secondly, if the vector-viruses are related, an unknown mechanism inhibits a wide-spread double-infection of the same cells. This phenomenon is observed with at least TMV, CPMV and potyviruses (Dietrich & Maiss 2003, Giritch *et al.* 2006, Sainsbury *et al.* 2006). Dietrich & Maiss (2003) studied three potyviruses (PPV, TVMV, and ClYVV)

all expressing different reporter proteins, and co-inoculated plants with them in different combinations. Doubly virus-infected cells were rarely observed with potyviruses, whereas they were common when the potyviruses were co-infecting the plants with PVX-GFP. Co-inoculation of leaves with two TMV constructs, one expressing GFP and the other one DsRed, and preparation of protoplasts from infected leaf tissue resulted in only 5% of protoplasts that were expressing both reporter proteins, i.e. were co-infected, as observed under a UV-microscope (Giritch *et al.* 2006). When these reporter molecules were expressed from a TMV and a PVX construct, 85-95% of the protoplasts are doubly infected. When CPMV RNA1 was co-inoculated with two separate RNA2 constructs each expressing a different reporter protein, co-localization was detected in the inoculated leaves but not observed in the systemically infected leaves (Sainsbury *et al.* 2006). It was concluded that in the systemically infected leaves, the vector-viruses segregated, as evident from the patchy distribution of the two reporter proteins. Monoclonal antibodies assembled in plant cells when the heavy and light chain polypeptides were expressed from separate TMV vector constructs (Verch *et al.* 1998). However, the same group used transgenic tobacco plants in their subsequent studies instead of using the vector-virus approach for producing the antibodies (Ko *et al.* 2005). The reason may have been the low yield of mature antibodies obtained with the vector-virus approach, attributable to rare co-infections of cells. However, Alamillo *et al.* (2006) produced the heavy and light chain polypeptides from separate PVX vector constructs, fed the neutralizing IgA-containing plant material to piglets, and showed decreased transmissible gastroenteritis virus titers in the guts of the piglets. While there was no mention of IgA amounts produced in these plant tissues, the assembly of IgA apparently was successful.

Recently, three novel cloning sites for heterologous sequences were found at the junctions of HC-Pro/P3, 6K1/CI and NIa-Pro/NIb (Fig. 11) in a TuMV clone (Chen *et al.* 2007). Using vector-TuMV with these three novel sites, and also the P1/HC-Pro and NIb/CP sites, they expressed either GFP or

a dust mite allergen. However, in their study only a single heterologous protein was expressed at a time. From P1/HC-Pro and NIb/CP sites both the heterologous proteins were produced in high quantities (*ca.* 1-2% of total soluble leaf proteins) in six different host plants (Chen *et al.* 2007). From the three novel sites GFP was expressed in similar or somewhat lower amounts than from the two aforementioned sites. The dust mite allergen could not be produced from the 6K1/CI site (Chen *et al.* 2007). From the NIa-Pro/NIb site it was produced only in 2 of the 6 host plants. In addition, the amounts of the mite allergen when expressed from the HC-Pro/P3 and NIa-Pro/NIb sites was much less than the expression from the P1/HC-Pro and NIb/CP sites in most hosts (Chen *et al.* 2007). If the three new cloning sites for heterologous sequences in TuMV function also in PVA, they could be used to find out the most suitable combination of cloning sites for simultaneous production of two or more heterologous proteins from PVA. Yet another recently found cloning site in PVA might be the VPg/NIa-Pro junction. A 45-nt sequence encoding a histidine-hemagglutinin double affinity tag was cloned to the 3'end of VPg to be expressed in fusion to the VPg (Hafrén & Mäkinen 2008). The clone was infectious in *N. benthamiana* and *N. tabacum*. However, addition of a sequence encoding GFP between the VPg and the double tag did not produce an infectious PVA clone (Hafrén & Mäkinen 2008). It was not tested if a heterologous protein, e.g. the GFP, could be expressed from the VPg/NIa-Pro site when it is proteolytically separated from VPg.

Table 6. Disease symptoms, virus accumulation, and insert stability of *Potato virus A* based heterologous protein expression vectors in *Nicotiana benthamiana* at 14 days post-inoculation (I, II, III).

Number of inserts	The cloning site(s) used and the heterologous sequence(s) therein[1]	Disease symtoms[2]	Amounts of viral coat protein as compared to insertless PVA (%)[3]	Insert stability (days)[4]
A single insert	CS1(*gfp*)	+	39 – 55	~14
	CS2(*Rluc*)	+	40 – 75	> 14
	CS3(*gfp*)	++	86 – 114	~30
	CS3(*S-COMT*)	++	86 – 91	> 14
	CS3(*UidA*)	++	77 – 89	< 14
	CS3(*sorcin*)	++	61 - 112	~30
	CS3(*sorcin*-NIa_v2.0)	+	4 – 17	~40
	CS3(*gfp*-NIa_v3.0)	++	85	> 17
Two inserts	CS1(*gfp*) + CS2(*Rluc*)	+	9 – 31	~14; >14
	CS1(*gfp*) + CS3(*UidA*)	+	24 – 41	~14; <14
	CS2(*Rluc*) + CS3(*UidA*)	+	34 – 46	>14; <14
Three inserts	CS1(*gfp*) + CS2(*Rluc*) + CS3(*UidA*)	-	12 – 15	~14; >14; <14
No inserts	Insertless vector-virus	+++	100 (1.4–3.3% or 19-36 ng/µg of soluble leaf proteins)	-

[1]CS1, within P1 encoding region; CS2, between the third and fourth codon of HC-Pro encoding region; CS3, between the first and second codon of CP encoding region; *gfp*, *Aequorea victoria* green fluorescent protein; *Rluc*, *Renilla reniformis* luciferase; *S-COMT*, human soluble catechol-*O*-methyltransferase; *UidA*, *Escherichia coli* β-glucuronidase; *sorcin*, human calcium ion binding protein.

[2] -, no symptoms; +, leaf chlorosis; ++, leaf chlorosis and leaf malformation; +++, more severe cases of leaf chlorosis and leaf malformation.

[3]Amounts of coat protein (µg/mg of leaf, fresh weigth) were estimated by DAS-ELISA including known amounts of purified PVA virions for comparison, and the values were compared to amounts of insertless PVA within every experiment.

[4]The number of days after inoculation after which at least in two plants a deletion event was detected with a reverse-transcription-PCR method.

Optimizations of the NIb/CP cloning site (II, unpublished)

The vector-viruses used in studies II and III produced heterologous proteins with 32 and 27 additional amino acids at their N- and C-termini, respectively (Fig. 9B). Therefore, the CS3 was further engineered for future uses by minimizing the number of additional amino acid residues to remain at the termini of the heterologous proteins expressed from it.

In the first optimized vector-virus, PVA-CS3(NIa_v2.0), a nucleotide sequence encoding a pentapeptide motif (VYFQ/A) was added to allow NIa-Pro mediated cleavage between the heterologous polypeptide and viral CP junction (Fig. 9B) (II). The amino acid sequence is identical to the NIa-Pro cleavage site located between NIb and CP of PVA (Merits *et al.* 2002). The positions P4 and P1 in the pentapeptide are highly conserved among the seven NIa-Pro sites within the PVA polyprotein (Merits *et al.* 2002), and also in 34 of 37 other potyviruses (Adams *et al.* 2005). Any heterologous protein expressed from PVA-CS3(NIa_v2.0) would have three and six additional amino acids at its N- and C-terminus, respectively (Fig. 9B). PVA-CS3(NIa_v2.0) expressing sorcin spread systemically in *N. benthamiana* plants (II). Its titers were *ca.* 10% of those of the wt PVA in systemically infected leaves (Table 6). The titers were also considerably lower than those of the unoptimized PVA-CS3(*sorcin*) vector in the infected plants (Table 6). Therefore, this approach was considered to compromise the target protein yields too heavily, although it substantially lowered the number of extra amino acids that would be incorporated in a heterologous protein.

In the next construct, PVA-CS3(NIa_v3.0), a heptapeptide NIa-Pro cleavage site consisting of amino acids at positions P6-P1′ was engineered to have the heterologous polypeptide separated from the viral polypeptide (unpublished) (Fig. 9B). A cleavage site of this length has been succesfully used in other vector potyviruses (Arazi *et al.* 2002, Choi *et al.* 2000, Dietrich & Maiss 2003). Titers of PVA-CS3(*gfp*-NIa_v3.0) (A_{405} 1.38 ± 0.14) were close to those of wt PVA (A_{405} 1.49 ± 0.05) in similarly diluted samples from systemically infected leaves of *N. benthamiana* at 17 dpi (Table 6). The vector

virus behaved similarly to PVA-CS3(*gfp*) in infected *N. benthamiana* plants in terms of systemic spread, symptoms and GFP expression. When expressed from this optimized vector, the target protein would have three and eight additional amino acids at the N-terminus the C-terminus, respectively. These numbers could be reduced to one and six if the restriction endonuclease recognition sites used for cloning in the heterologous sequence were situated differently. Instead of the current construct in which the restriction sites flank the insert (EED**PR***DMVYFQ/A***PG**-insert-**PR**DMVYFQ/AETL) (restriction sites in bold) these sites could flank the NIa-Pro recognition sites (EED**PR***DMVYFQ/A*-insert-DMVYFQ/A**PG**ETL) (NIa-Pro sites boxed). The NIa-Pro sites would then have to be included in the primers used for cloning of the heterologous sequences.

Similar NIa-Pro recognition sequences with seven amino acids were also used at the other two cloning sites (CS1 & CS2) to separate the heterologous proteins from the viral polyprotein. Various silent point mutations were introduced to the sequences for the added NIa-Pro cleavage site to reduce the chance of homologous recombination during virus replication.

Variability in virus titers

The titers of vector-virus constructs and wt PVA in inoculated *N. benthamiana* plants were primarily measured 14 dpi in all experiments. However, it was suspected that 14 days was not always long enough for systemic spreading with all vector-viruses. Therefore, PVA-CS3(*gfp*-NIa_v3.0) and wt PVA were tested by DAS-ELISA (for CP), western blot analysis (for CP and GFP) and a fluorometric quantitation assay (for GFP) in *N. benthamiana* at 14, 17 and 20 dpi (unpublished). All assays indicated that the peak of the CP and GFP accumulation in PVA-CS3(*gfp*-NIa_v3.0) infected plants was at 17 dpi, whereas the wt PVA reached its maximal CP amounts by 14 dpi. The CP amounts of PVA-CS3(*gfp*-NIa_v3.0) were 62% of those of the wt virus at 14 dpi, while at 17 and 20 dpi the titers of the two viruses were similar. This experiment was not repeated, but it showed that 14 dpi is not always an

optimal sampling timepoint for PVA-based vector-viruses in *N. benthamiana*. Varying environmental conditions and the age and physical condition of the plants at the time of inoculation possibly influences the rate of systemic spreading.

Testing vector-viruses in *N. tabacum* cv Samsun nn (I, III)

Mutant viruses with inserts larger than 168 nt in CS1 exhibited lower virulence than the wt PVA in tobacco plants (Table 1 in I). Only half of the tobacco plants inoculated were infected with vector-viruses carrying a full-size *gfp* in CS1, and their systemic spread was low (I).

However, when tobacco plants were inoculated with *N. benthamiana* leaf sap from leaves infected with PVA-CS1(*gfp*), PVA-CS2(*Rluc*), PVA-CS3(*UidA*) or the double-insert vector PVA-CS1(*gfp*)-CS3(*UidA*), all plants became systemically infected (III). However, none of the eight plants inoculated with the PVA-3i-containing leaf sap became systemically infected. No symptoms were seen in any tobacco plants infected with any of the aforementioned constructs. The GFP-expressing vector-viruses caused typical fluorescent spots in the non-inoculated leaves (similar to those in Fig. 2 D in I) at 9 dpi. During the following days the diameter of the spots increased but the intensity of green fluorescence faded, so that at 17 dpi the fluorescence was barely detectable. In leaves infected with PVA-CS1(*gfp*)-CS3(*UidA*), the number of fluorescent spots was noticeably less than in leaves infected with PVA-CS1(*gfp*). GFP expression was detected in the PVA-CS1(*gfp*)-CS3(*UidA*) inoculated tobacco plants only in the two first systemically infected leaves at 9 dpi, which indicated partially compromised systemic spreading of the construct in *N. tabacum*. DAS-ELISA revealed high amounts of CP in wt PVA infected tobacco plants at 24 dpi, but only trace amounts in the plants infected with the single-insert vectors and no virus in the plants inoculated with the double-insert vector. Tobacco plants, at least cv Samsun nn, seemed to be more resistant to PVA-based vector-viruses than *N. benthamiana* plants.

The cDNA clones of vector-viruses can be modified to increase infectivity. For example, 1000-fold more tobacco protoplasts were infected with a cDNA of a TMV-based vector after 16 plant introns were added and silent mutations to thymine-rich regions were made to increase the GC content (Marilloinnet *et al.* 2005). Similar modifications to the clones of PVA-3i and the other vector-viruses might increase their virulence in *N. tabacum*.

Testing of the vector-viruses in *S. tuberosum* (unpublished)

PVA strain B11, which was the basis for the expression vectors in I, II and III, is not able to spread systemically from the inoculated leaves in potato cultivars tested so far (Valkonen *et al.* 1995). The infectious clone of strain U, however, spreads systemically in two diploid potato lines (v2-134, v2-51) as does a chimeric subclone (pBUIII) where the first two-thirds of PVA are from strain B11 and the rest (from genomic position 6371 onwards) from strain U (Paalme *et al.* 2004). pBUIII accumulates to similar amounts as PVA strain B11 in inoculated potato leaves (Paalme *et al.* 2004). In this study, higher titers were observed with pBUIII (A_{405} 2.32 ± 0.21) than with B11 (A_{405} 1.82 ± 0.11) in similarly diluted samples of systemically infected leaves (DAS-ELISA). New vector-viruses based on pBUIII were made that had either the CS1 or CS3 and contained the *gfp* coding sequence. However, these GFP-expressing vector-viruses could not spread in potato cv. Pito or in the diploid potato line v2-134 (DAS-ELISA). They were also not detected in the roots. The GFP-expressing vectors infected the inoculated leaves but could not exit them. Only the insertless clones (with 15 nt and 29 nt of non-viral sequence in CS1 and CS3, respectively) moved systemically in both hosts according to DAS-ELISA and IC-RT-PCR tests.

Stability of the chimeric viruses during infection (I, II, III)

For most applications of vector viruses, it is important to express full-size proteins. Furthermore, for some applications, such as expression of proteins intended for medical use, the uniform full-size product is essential.

All the vector virus - insert combinations made in this study were tested for stability during the infection of *N. benthamiana* plants. In general, the observed stability was dependent on both insert and cloning site. The growing subpopulations of deletion mutants within the virus populations in the infected plants could be observed from an increased disease severity. The plants that eventually contained only deletion mutants had disease symptoms as severe as those of plants infected by wt PVA, whereas mild, if any, symptoms were observed in plants infected with vector-viruses with intact inserts.

The *gfpuv* (Crameri *et al.*1996) inside the CS1 began to disintegrate after two weeks of infection (Fig. 3 in I, III), while the coding sequence of the same gene (although a different variant) was stable for four weeks within CS3 (II) (Table 6). A putative recombination hot-spot was identified within the *gfpuv* sequence, and was proposed to cause the lability (I). The *mgfp5* (Haseloff *et al.* 1997) variant of *gfp* used in CS3 in this study was more stable and differed from *gfpuv* at the proposed recombination hot-spot by only one nucleotide. In some cases also, the viral sequence adjacent to the insert was lost in the deletion process (I), which has been shown to occur also with other vector-potyviruses (Dolja *et al.* 1993, German-Retana 2000).

Inserts in CS2 were not extensively tested for stability, but the sequence encoding luciferase was intact in all tested plants at 14 dpi in three experiments (III).

Multiple heterologous inserts were tested in CS3, of which only *UidA* was clearly more prone to deletions than the others (Table 6) (II, III). Lability of *UidA* in vector-potyviruses has been observed before (Arazi *et al.* 2001, Beauchemin *et al.* 2005), although this sequence can be retained intact over 120 days when serial passaging is done every 4-6 days (Dolja *et al.* 1993). In eight PVA-CS3(*sorcin*-NIa_v2.0) infected *N. benthamiana* plants, the *sorcin* sequence was still intact at 40 dpi (Table 6) (II), and at 60 dpi in seven of the eight. In comparison, PVA-CS3(*sorcin*) was stable for about one month (Table 6) (II). The difference may be due to the shorter duplicated viral sequences flanking

the insert in PVA-CS3(*sorcin*-NIa_v2.0) and hence a lower rate of homologous recombination.

Stability of the heterologous inserts in the multi-insert vectors was analysed in all infected *N. benthamiana* plants. The results indicated that the stability of the inserts was similar to that in the single-insert constructs (Table 6) (III).

Stability of inserts in three single-insert vectors was tested in *N. tabacum*. Inserts in PVA-CS1(*gfp*) and PVA-CS2(*Rluc*) were intact in systemically infected leaves at 24 dpi, whereas *UidA* was completely lost from PVA-CS3(*UidA*) at this time (III).

Recombination, both homologous and non-homologous, is one of the driving forces in RNA virus evolution (Simon & Bujarski 1994) and has been shown to occur in potyviruses (Cervera *et al.* 1993, Ohshima *et al.* 2007). The deletions detected in the inserts and PVA in this study were assumed to occur via homologous and non-homologous recombination. It is generally found that vector-viruses sooner or later lose the inserted heterologous sequences, although there always exists the theoretical chance that the added sequence improves the survival of the virus and would thus be retained. Another major class of mutations occuring in RNA viruses, point mutations, are produced mainly by the viral RNA-dependent RNA polymerases (RdRp) that lack proofreading capacity. The error rate of RdRps during replication is estimated to be 10^{-3}–10^{-4} per nucleotide (Hull 2002). The mutation rate in heterologous dihydrofolate reductase and neomycinphosphotransferase II encoding sequences in a vector-TMV was estimated to be $\leq 10^{-4}$ per nucleotide per passage through *N. benthamiana* (Kearney *et al.* 1993). Point mutations within heterologous sequences in vector viruses are proposed to cause no disadvantage for the virus (Kearney *et al.* 1993). Hence, all point mutations will be retained, which will eventually lead to heterologous sequence that does not encode a functional heterologous protein. The point mutations generated in the vector viruses during infection in plants are generally overlooked. One reason for this is probably that the deletions that occur in the

heterologous sequences in vector viruses cause production of intact heterologous protein to end anyway usually within a few weeks after inoculation (see most of the examples in tables 1 and 2, and the results of this study). However, in a TMV-based vector expressing either GFP or hGF in *N. benthamiana* roots kept in a liquid culture (subcultured every six weeks) no deletions in the heterologous sequences were observed during a three year period (Skarjinskaia *et al.* 2008). GFP and hGF proteins of expected sizes were seen on a western blot analysis, but neither their functionality nor the nucleotide sequence of the inserts in the vector viruses was shown on the roots after the three year period.

Serial passaging by inoculation to new hosts at fixed time intervals and a prolonged propagation time within the same plant have been used for testing insert stability in vector-viruses. The inserts in potyviruses seem to be stable for a much longer total time (months instead of weeks) when serial passaging is used (Dolja *et al.* 1992 & 1993, German-Retana *et al.* 2000 & 2003), especially then the intervals are kept short to minimize systemic spreading, as compared to stability during prolonged propagation in the same host. The cause is not clear, but perhaps in the serial passaging most of the deletion mutants are lost by chance since they are not allowed to have sufficient time to replicate extensively and consequently outcompete the parent vector virus. In this study, only the prolonged propagation strategy was applied since it was considered to be more informative about the stability of the constructs.

Heterologous protein expression and accumulation in *N. benthamiana* plants (I, II, III)

The amounts of the heterologous proteins positively correlated with the amounts of the viral CP detected in systemically infected leaves of *N. benthamiana* plants, as expected (Fig. 10).

Expression of the jellyfish GFP (I, II, III)

GFP was produced in a functional form in the leaves of the two *Nicotiana* species as observed under UV-light. The observed intensity of the emitted green fluorescence positively correlated with the measured vector-virus titer in each individual plant. GFP amounts were not quantified. According to Leffel *et al.* (1997), an amount of 0.1% of soluble leaf proteins is the unambiguous detection level of GFP variant mGFP4 by a naked eye under a hand-held UV-light device in transgenic tobacco plants. If true also in this study, this means that even in the case of the PVA-3i-infected *N. benthamiana* plants, which had the lowest titers of a GFP-expressing vector-PVA (Table 6) (Fig. 10), the amount of 0.1% of soluble leaf proteins was reached. Systemically infected leaves in those plants were observed to be pale green under a hand-held UV-light. In both PVA-CS1(*gfp*) and PVA-CS3(*gfp*) infected inoculated leaves of potato, GFP was detected only with a fluorescence microscope, which indicated that GFP expression levels and the vector-virus titers in potato were lower than in the *Nicotiana* species.

Two versions of PVA-CS1(*gfp*) were made (M14-pGFPp and M14-GFPp) (I). The difference between them was a sequence encoding a heptapeptide NIa-Pro recognition site (indicated by a lower-case letter 'p') included in the 5′-part of the *gfp* in M14-pGFPp. In systemically infected leaves of *N. benthamiana*, distinct differences were observed both in the intensity of green fluorescence and the amounts of GFP detected in a western blot analysis (Fig. 2 F in I), both of which were significantly higher with M14-pGFPp than with M14-GFPp. Furthermore, in three experiments, M14-pGFPp exited the inoculated leaves on average a day earlier than M14-GFPp (I). One explanation for these results could be that the first 25 amino acids of P1 fused to the N-terminus of GFP in the M14-GFPp interfered with proper folding of the GFP polypeptide. Another explanation could be accelerated turnover of P1 in infected cells (Hinrichs-Berger *et al.* 2003, Rodrígues-Cerezo & Shaw 1991b).

Expression of the seapansy luciferase (III)

The amounts of activities of luciferase expressed from CS2 correlated with titers of PVA CP (Fig. 10). Western blot analysis revealed that a minor portion of luciferase was fused to the HC-Pro protein in the samples (middle panel in Fig. 3 in III), which indicated slow processing of the novel NIa-Pro cleavage site engineered between the luciferase and HC-Pro. This was unexpected, since similar novel NIa-Pro sites were processed quickly when they were used with inserts at CS1 and CS3 to separate expressed heterologous proteins from the polyprotein (I, II, III). Alterations in the structure of the viral polyprotein caused by the luciferase polypeptide might have occurred and partially interfered with accessibility of the engineered novel cleavage site to NIa-Pro. When GUS or GFP was produced from an identical location from LMV-based vector with a similar engineered heptapeptide NIa-Pro site, the heterologous proteins were efficiently separated from HC-Pro (German-Retana *et al.* 2003).

Yields of human soluble S-COMT and bacterial GUS (II, III)

For quantification of S-COMT, the total proteins of infected leaves were isolated and the polyhistidine-tagged S-COMT (partially) purified (II). The partially purified S-COMT fraction was analysed together with known amounts of bovine serum albumin on a polyacrylamide gel (Fig. 4 in II) using a quantitative staining of the proteins. Results indicated that S-COMT accumulated to up to 0.8% of total leaf proteins in *N. benthamiana* (II).

The GUS activity was measured and compared to the activity of known amounts of recombinant GUS in extracts from leaves of healthy plants (Fig. 10) (III). The yield of GUS was up to 0.7% of soluble leaf proteins in *N. benthamiana* (III). Considering the lability of the *UidA* nucleotide sequence within the vector-viruses, the slightly lower yield of GUS as compared to the yield of S-COMT was not unexpected.

The amounts of CP (1.5-3.5% of total soluble proteins) were similar with the S-COMT- and GUS-expressing constructs and the wt PVA (III). The amount of CP can be considered to equal to the maximum attainable amount

of heterologous protein, since the viral proteins and heterologous proteins are translated and produced in equimolar amounts from a potyvirus. Assuming that PVA CP was stable in *N. benthamiana* cells, the obtained yields of GUS and S-COMT corresponded to ca. 25-50% and 50% of the expected maximum, respectively.

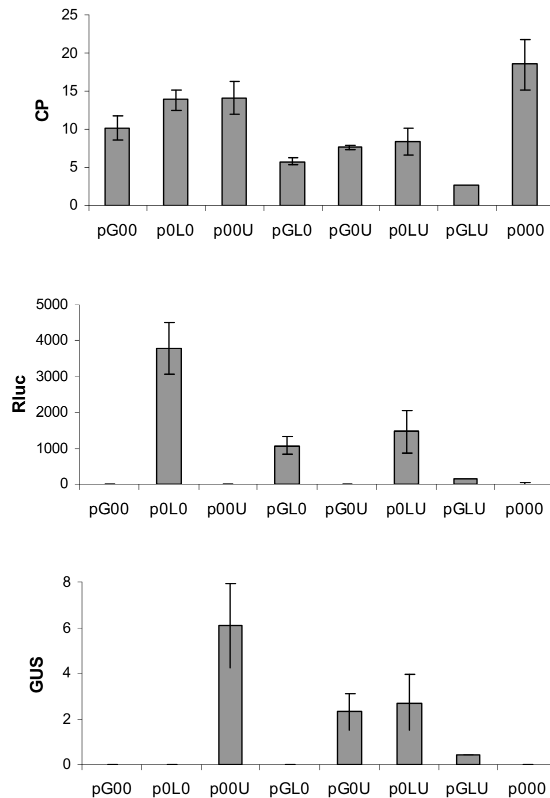


Fig. 10. Amounts of viral coat protein (CP) (ng/µg of soluble leaf proteins), relative activities of *Renilla reniformis* luciferase (Rluc) (relative light units/µg of soluble leaf proteins), and amounts of β-glucuronidase (GUS) (ng/µg of soluble leaf proteins) expressed from the single, double, and triple insert vector-PVA in systemically infected leaves of *N. benthamiana* plants. The experiment was repeated with similar results. pG00, PVA-CS1(*gfp*) expressing GFP; p0L0, PVA-CS2(*Rluc*) expressing luciferase; p00U, PVA-CS3(*UidA*) expressing GUS; pGL0, PVA-CS1(*gfp*)+CS2(*Rluc*) expressing GFP and luciferase; pG0U, PVA-CS1(*gfp*)+CS3(*UidA*) expressing GFP and GUS; p0LU, PVA-CS2(*Rluc*)+CS3(*UidA*) expressing luciferase and GUS; pGLU, PVA-CS1(*gfp*)+CS2(*Rluc*)+CS3(*UidA*) expressing GFP, luciferase and GUS; RLU, relative light units.

The role of P1 in HC-Pro mediated suppression of RNA silencing (I)

Potyviral HC-Pro protects non-host RNAs from degradation via its ability to suppress the host RNA silencing system (Kasshau & Carrington 1998, Johansen & Carrington 2002, Lakatos *et al.* 2006), as a P1/HC-Pro duplex (Kasshau & Carrington 1998) and as HC-Pro (Brigneti *et al.* 1998).

Three different PVA-CS1(*gfp*) clones, M14-pGFPp, M14-GFPp, and M14-delGFP, were used to further investigate the role of P1 in HC-Pro mediated RNA silencing suppression (I). From these vectors the sequences encoding P1, the heterologous polypeptide or HC-Pro were cloned into a binary vector for 35S promoter driven expression from *A. tumefaciens*. Similarly, from wt PVA the sequences encoding both P1 and HC-Pro, or P1 or HC-Pro alone, were cloned into a binary vector. In all constructs, the viral 5'UTR was included, which is a translation enhancer (Carrington & Freed 1990). These constructs were used to express the aforementioned proteins in tobacco leaves. Following expression from these constructs, P1 was expected to separate itself from HC-Pro, but the GFP or the fragment of GFP would be retained in the P1.

Expression of mRNA from the wt P1/HC-Pro and the P1(delGFP)/HC-Pro constructs was 10-30 fold higher than from the HC-Pro expressing construct in the agroinfiltration experiments (Fig. 4B in I). These data suggested that the P1/HC-Pro polyprotein enhanced the accumulation of the corresponding mRNA more than was observed with HC-Pro. P1 produced alone had no effect on the mRNA levels, as compared to HC-Pro, as had been shown previously (Brigneti *et al.* 1998). The highest GUS activities were observed when *UidA* was co-expressed with P1/HC-Pro in four experiments. These data suggested that P1/HC-Pro protected mRNAs against silencing to a higher extent than was observed with HC-Pro alone. It is possible or even likely that not all cells were doubly infected when leaves were co-infiltrated with the two constructs. This could explain the observed weaker protection effect of the *UidA* mRNA than the mRNA expressing the P1/HC-Pro polyprotein. Pruss *et al.* (1997) have shown that expression of TEV P1/HC-Pro

duplex from PVX in tobacco protoplasts led to a substantial increase in levels of PVX (-)strand RNA, but not the (+)strand RNA, as compared to corresponding RNA levels when TEV HC-Pro was produced alone. Their study provided the first indication that P1/HC-Pro is an efficient RNA silencing suppressor, and more efficient than HC-Pro alone. Hence, P1 seems to enhance the silencing suppression activity of HC-Pro.

Sizes of virions of the PVA-based vectors (III)

PVA-3i has a genome that is *ca.* 38% longer than the wt PVA. Immunosorbent electron microscopy (ISEM) was used to study whether PVA-3i and other vector-viruses had retained viral RNA encapsidation and particle formation. Previously, PPV carrying the *UidA* coding sequence had been found to form virions in infected *N. benthamiana* leaves (Varrelmann & Maiss 2000). Also, there is indirect evidence that TEV-*UidA* forms virus particles, as partially purified virions from vector-virus and wt TEV infected leaves contained comparable amounts of viral RNA as detected by northern blot analysis (Dolja *et al.* 1997).

Virion formation of potyviruses has been proposed to start at or near the 5′ end of the genome, and to proceed rapidly towards the 3′ end once initiated (Wu & Shaw 1998). It is still somewhat unclear whether potyviruses move cell-to-cell and systemically as virions or as other ribonucleoprotein complexes (Dolja *et al.* 1994, 1995; Rodrígues-Cerezo *et al.* 1997). TEV CP mutants, in which the highly conserved amino acids in the core region are substituted with alanine, are defective in movement and encapsidation (Dolja *et al.* 1994, 1995). However, these mutants could also fail in various protein-protein or protein-RNA interactions that might be needed for formation of the putative ribonucleoprotein transport complex other than virions. Hence, it is also not clear whether the vector-potyviruses need to form virions in order to cause a systemic infection. Immunogold labeling and *in situ* hybridization experiments have revealed TVMV CP and RNA inside cones that are formed by the viral cylindrical inclusion (CI) protein and span the cell walls and

membranes of adjacent cells (probably through plasmodesmata) (Rodrígues-Cerezo *et al.* 1997). However, whether the viral CP and RNA were assembled to virions remained unclear.

PVA virions were captured from systemically infected leaves of *N. benthamiana* plants 10 dpi with PVA-3i, PVA-CS1(*gfp*)-CS3(*UidA*), PVA-CS1(*gfp*), PVA-CS3(*UidA*), or wt PVA. Grids coated with a monoclonal antibody recognizing an epitope near the N-terminus of PVA CP (Rajamäki *et al.* 1998) were floated upside down on drops of sap from infected leaves to trap virus particles (III). ISEM revealed virions with all four vector-viruses tested (Fig. 4 A in III). The genome lengths positively correlated with the observed virion lengths and regression analysis indicated that 82% of the variation in the observed virion lengths could be explained by genome lengths ($p<0.001$) (Fig. 4B in III). In contrast, in a similar regression analysis with a data-set consisting of the virion and genome lengths of 38 wt potyvirus species from other studies, only a weak correlation was found and only 21% of the virion length variation was explained by genome length variation (Fig. 4B in III). Measurement of the sizes of potyvirus virions is difficult and also affected by many possible sources of error. The viral poly(A) tail encapsidated in virions can vary substantially in length and is *ca.* 15-500 nt in TEV and PPV (Hari 1981; Laín *et al.* 1988). Virions can break during the capturing and staining procedures and form end-to-end fusions with themselves or with full-size virions. In addition, other technical aspects influence the results, for example divalent cations that may introduce up to 20% variability in virion lengths during virion capture and staining procedures (Govier & Woods 1971).

CONCLUSIONS

Heterologous sequences were inserted into three locations in an infectious clone of *Potato virus A* (PVA) without greatly compromising vital viral functions such as replication and systemic movement. One of the cloning sites was novel for potyviruses. In addition to vector-PVA constructs containing single inserts, vector-viruses were made with two or three heterologous sequences placed at different positions at the viral genome. They simultaneously expressed several heterologous reporter proteins in systemically infected leaves of *N. benthamiana* plants. These vector-viruses formed virions, the lengths of which correlated with genome lengths. Hence, direct evidence could be obtained on the dependence of virion length on genome size in potyviruses.

The future application of all these vector-viruses could be in protein overexpression, and as research tools. The flexibility of potyviruses for both of these purposes is underlined by a recent study that reported three novel cloning sites for heterologous sequences in a TuMV clone (Chen *et al.* 2007) (Fig. 11).

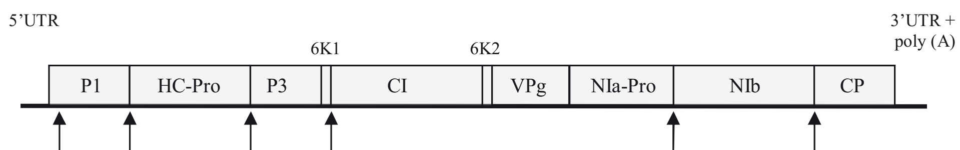


Fig. 11. A schematic map of a potyvirus genome/polyprotein that illustrates its versatility as a heterologous protein expression vector. The arrows point at the locations within the genome identified so far as tolerating foreign sequences encoding full-size proteins. The site within the P1 encoding region was characterized in this study. The P1/HC-Pro and NIb/CP sites have been used in many studies (Table 3). The HC-Pro/P3, 6K1/CI, NIa-Pro/NIb sites (Chen *et al.* 2007) have been found recently.

The results of Chen *et al.* (2007), the ones listed in Table 3, and those of this study all indicate that foreign proteins can be expressed from a vector-potyvirus in plants in high amounts. Nevertheless, there are differences in how different foreign sequences are tolerated by potyviruses and/or potyvirus – host species combinations, and by the different cloning sites in potyviral genomes.

The aim to have a vector-virus capable of infecting potato plants was not completely successful. All the constructs were able to infect the inoculated leaves but none of them could spread systemically. Further development of the vector-PVA for use in potato could include testing other cloning sites than the P1 or NIb/CP sites used in this study. Also, other potato cultivars or wild relatives of potato that are susceptible to PVA could be tested. However, it is also possible that co-evolution of PVA and its host, potato, has led to a situation where addition of a fairly large heterologous sequence to most, if not all, places within the virus genome is not possible.

ACKNOWLEDGEMENTS

I thank A. Merits (University of Tartu), H. Vihinen (Institute of Biotechnology, Helsinki), R. Gabrinaite-Verkhovskya, A. Hafrén and K. Mäkinen (University of Helsinki) for sharing unpublished data, and Viikki Graduate School in Molecular Biosciences for travel and other grants and the privileged opportunity to participate in high quality courses organized by them and other graduate schools in Helsinki.

I especially thank my supervisors who kept pushing me to jump over the hedge where it is the highest, and all the personnel and fellow pHd students of the Department of Applied Biology for the help and advise they have given me during these years.

REFERENCES

Adams MJ, Antoniw JF & Beaudoin F (2005). Overview and analysis of the polyprotein cleavage sites in the family *Potyviridae*. *Mol Plant Pathology* **6**, 471-487.

Ahlquist P, French R, Janda M & Loesch-Fries LS (1984). Multicomponent RNA plant virus infection derived from cloned viral cDNA. *Proc Natl Acad Sci USA* **81**, 7066-7070.

Alamillo JM, Monger W, Sola I, García B, Perrin Y, Bestagno M, Burrone OR, Sabella P, Plana-Durán J, Enjuanes L, Lomonossoff GP & García JA (2006). Use of virus vectors for the expression in plants of active full-length and single chain anti-coronavirus antibodies. *Biotechol J* **1**, 1103-1111.

Anand A, Vaghchhipawala Z, Ryu CM, Kang L, Wang K, del-Pozo O, Martin GB & Mysore KS (2007). Identification and characterization of plant genes involved in *Agrobacterium*-mediated plant transformation by virus-induced gene silencing. *Mol Plant-Microbe Interact* **20**, 41-52.

Anandalakshmi R, Pruss GJ, Ge X, Marathe R, Mallory AC, Smith TH & Vance VB (1998). A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci USA* **95**, 13079-13084.

Andrejeva J, Puurand Ü, Merits A, Rabenstein F, Järvekülg L & Valkonen JPT (1999). Potyvirus helper component-proteinase and coat protein (CP) have coordinated functions in virus--host interactions and the same CP motif affects virus transmission and accumulation. *J Gen Virol*, **80**, 1133 - 1139.

Angell SM & Baulcombe DC (1997). Consistent gene silencing in transgenic plants expressing a replicating potato virus X RNA. *Plant J* **20**, 357-362.

Aranda MA, Escaler M, Wang D & Maule AJ (1996). Induction of HSP70 and polyubiquitin expression with plant virus replication. *Proc Natl Acad Sci USA* **93**, 15289-15293.

Arazi T, Slutsky SG, Shiboleth YM, Wang Y, Rubinstein M, Barak S, Yand J & Gal-On A (2001). Engineering zucchini yellow mosaic potyvirus as a non-pathogenic vector for expression of heterologous proteins in cucurbits. *J Biotech* **87**, 67-82.

Arazi T, Huang PL, Huang PL, Zhang L, Shiboleth YM, Gal-On A & Lee-Huang S (2002). Production of antiviral and antitumor proteins MAP30 and GAP31 in cucurbits using the plant virus vector ZYMV-AGII. *Biochem Biophys Res Commun* **292**, 441-448.

Atreya CD, Raccah B & Pirone TP (1990). A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus. *Virology* **178**, 161-165.

Avesani L, Marconi G, Morandini F, Albertini E, Bruschetta M, Bortesi L, Pezzotti M & Porceddu A (2007). Stability of *Potato virus X* expression vectors is related to insert size: implications for replication models and risk assessment. *Transgenic Res* **16**, 587-597.

Bakker H, Rouwendal GJ, Karnoup AS, Florack DE, Stoopen GM, Helsper JP, van Ree R, van Die I & Bosch D (2006). An antibody produced in tobacco expressing a hybrid ß-1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc Natl Acad Sci USA* **103**, 7577-7582.

Bartels R (1971). *Potato virus A*. Entry 54 in DPVweb. Association of Applied Biologists, Adams MJ & Antoniw JF (2006). DPVweb: a comprehensive database of plant and fungal virus genes and genomes**.** *Nucleic Acids Research* 34, Database issue, D382-D385. http://www.dpvweb.net/index.php

Baulcombe DC, Chapman SN & Santa Cruz S (1995). Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J* **7**, 1045-1053.

Baulcombe D (2005). RNA silencing. *Trends Biochem Sci* **30**, 290-293.

Beauchemin C, Bougie V & Laliberté JF (2005). Simultaneous production of two foreign proteins from a potyvirus-based vector. *Virus Res* **112**, 1-8.

Borisjuk NV, Borisjuk LG, Logendra S, Petersen F, Gleba Y & Raskin I (1999). Production of recombinant proteins in plant root exudates. *Nature Biotechnol* **17**, 466-469.

Brandes J & Paul HL (1957). Das elektronenmikroskop als hilfmittel bei der diagnose pflanzlicher virosen – betrachtungen zur vermessung faden und stabchenformiger virusteilchen. *Archv für Mikrobiol* **26**, 358-368.

Brigneti G, Voinnet O, Li WX, Ji LH, Ding SW & Baulcombe DC (1998). Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana benthamiana*. *EMBO J* **17**, 6739-6746.

Brisson N, Paszkowski J, Penswick JR, Groenenborn B, Potrykus I & Hohn T (1984). Expression of a bacterial gene in plants by using a viral vector. *Nature* **310**, 511-514.

Cañizares MC, Lomonossoff GP & Nicholson L (2005). Development of cowpea mosaic virus-based vectors for the production of vaccines in plants. *Expert Rev Vaccines* **4**, 687-697.

Cañizares MC, Liu L, Perrin Y, Tsakiris E & Lomonossoff GP (2006). A bipartite system for the constitutive and inducible expression of high levels of foreign proteins in plants. *Plant Biotech J* **4**, 182-193.

Carrington JC, Cary SM, Parks TD & Dougherty WG (1989). A second proteinase encoded by a plant potyvirus genome. *EMBO J* **8**, 365-370.

Carrington JC & Dougherty WG (1987). Small nuclear inclusion protein encoded by plant potyvirus genome is a proteiase. *J Virol* **61**, 2540-2548.

Carrington JC & Freed DD (1990). Cap-independent enhancement of translation by a plant potyvirus 5'nontranslated region. *J Virol* **64**, 1590-1597.

Carrington JC, Haldeman R, Dolja VV & Restrepo-Hartwig MA (1993). Internal cleaveage and tras-proteolytic activities of the VPg-proteinase (NIa) of tobacco etch potyvirus in vivo. *J Virol* **67**, 6995-7000.

Carrington JC, Jensen PE & Schaad MC (1998). Genetic evidence for an essential role for potyvirus CI protein in cell-to-cell movement. *Plant J* **14**, 393-400.

Cervera MT, Riechmann JL, Martín MT & Garvía JA (1993). 3'-terminal sequence of the plum pox virus PS and o6 isolates: evidence for RNA recombination within the potyvirus group. *J Gen Virol* **74**, 329-334.

Chahal GS & Gosal SS (2002). *Genetic Transformation and Production of Transgenic Plants*. In *Priciples and Procedures of Plant Breeding – Biotechnical and Conventional Approaches*. Alpha Science International Ltd., Pangbourne, UK. pp 486-508.

Chapman S, Kavanagh T & Baulcombe DC (1992). Potato virus X as a gene vector for gene expression in plants. *Plant J* **2**, 549-557.

Chapman EJ, Prokhnevsky AI, Gopinath K, Dolja VV & Carrington JC (2004). Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. *Genes & Development* **18**, 1510.

Chen CC, Chen TC, Raja JAJ, Chang CA, Chen LW, Lin SS & Yeh SD (2007). Effectiveness and stability of heterologous proteins expressed in plants by *Turnip mosaic virus* vector at five different insertion sites. *Virus Res* **130**, 210-227.

Chen TC, Hsu HT, Jain RK, Huang CW, Lin CH, Liu FL & Yeh SD (2005). Purification and serological analyses of tospoviral nucleocapsid proteins expressed by *Zucchini yellow mosaic virus* vector in squash. *J Virol Methods* **129**, 113-124.

Choi IR, Stenger DC, Morris J & French R (2000). A plant virus vector for systemic expression of foreign genes in cereals. *Plant J* **23**, 547-555.

Constantin GD, Krath BN, MacFarlane SA, Nicolaisen M, Johansen IE & Lun OS (2004). Virus-induced gene silencing as a tool for functional genomics in a legume species. *Plant J* **40**, 622-631.

Cox KM, Sterling JD, Regan JT, Gasdaska JR, Frantz KK, Peele CG, Black A, Passmore D, Moldovan-Loomis C, Srinivasan M, Cuison S, Cardarelli PM & Dickey LF (2006). Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat biotechnol* **24**, 1591-1597.

Crameri A, Whitehirn EA, Tate E & Stemmer WPC (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol* **14**, 315-319.

Cronin S, Verchot J, Haldeman-Cahill R, Schaad MC & Carrington JC (1995). Long-distance movement factor: a transport function of the potyvirus helper component proteinase. *Plant Cell* **7**, 549-559.

Culver JN, Lehto K, Close SM, Hilf ME & Dawson WO (1993). Genomic position affects the expression of tobacco mosaic virus movement and coat protein genes. *Proc Natl Acad Sci USA* **90**, 2055-2059.

Daniel H, Khan MS & Allison L (2002). Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. *Trends Plant Sci* **7**, 84-91.

Dawson WO, Lewandowski DJ, Hilf ME, Bubrick P, Raffo AJ, Shaw JJ, Grantham GL & Desjardins PR (1989). A tobacco mosaic virus-hybrid expresses and loses an added gene. *Virology* **172**, 285-292.

De Cosa B, Moar W, Lee SB, Miller M & Daniell H (2001). Overexpression of the Bt cry2Aa2 operon in chloroplasts leads to formation of insecticidal crystals. *Nat Biotechnol* **19**, 71-74.

De Zoeten GA, Penswick JR, Horisberger MA, Ahl P, Schultze M & Hohn T (1989). The expression, localization and effect of a human interferon in planst. *Virology* **172**, 213-222.

Decker EL & Reski R (2004). The moss bioreactor. *Curr Opin Plant Biol* **7**, 166-170.

Dietrich C & Maiss E (2003). Fluorescent labelling reveals spatial separation of potyvirus populations in mixed infected *Nicotiana benthamiana* plants. *J Gen Virol* **84**, 2871-2876.

Ding XS, Carter SA, Deom CM & Nelson RS (1998). Tobamovirus and potyvirus accumulation in minor veins of inoculated leaves from representatives of the Solanaceae and Fabaceae. *Plant Physiol* **116**, 125-136.

Ding XS, Schneider WL, Chaluvadi SR, Mian MAR & Nelson RS (2006). Characterization of a *Brome mosaic virus* strain and its use as a vector for gene silencing in monocotyledonous hosts. *Mol Plant Microbe Interact* **19**, 1229-1239.

Dolja VV, Haldeman R, Robertson NL, Dougherty WG & Carrington JC (1994). Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. *EMBO J* **13**,1482-1491.

Dolja VV, Haldeman-Cahill R, Montgomery AE, Vandenbosch KA & Carrington JC (1995). Capsid protein determinants involved in cell-to-cell and long-distance movement of tobacco etch potyvirus. *Virology* **206**, 1007-1016.

Dolja VV, Hong J, Keller KE, Martin RR & Peremyslov VV (1997). Suppression of potyvirus infection by coexpressed closterovirus protein. *Virology* **234**, 243-252.

Dolja VV, McBride HJ & Carrington J (1992). Tagging of plant potyvirus replication and movement by insertion of β-glucuronidase into the viral polyprotein. *Proc Natl Acad Sci USA* **89**, 10208-10212.

Donson J, Kearney CM, Hilf ME & Dawson WO (1991). Systemic expression of a bacterial gene by a tobacco mosaic virus-based vector. *Proc Natl Acad Sci USA* **88**, 7204-7208.

Dorokhov YL, Sheveleva AA, Frolova OY, Komarova TV, Zverevea AS, Ivanov PA & Atabekov JG (2007). Superexpression of tuberculosis antigens in plant leaves. *Tuberculosis* **87**, 218-224.

Douine L, Quiot JB, Marchoux G & Archange P (1979). Index of plants susceptible to cucumber mosaic virus (CMV) – Bibloigraphical study. *Annales de Phytopathologie* **11,** 439.

Edwardson JR & Christie RG (1991). *CRC Handbook of viruses infecting legumes***.** Boca Raton, CRC Press, 293.

Escaler E, Aranda MA, Roberts IM, Thomas CL & Maule AJ (2000). A comparison between virus replication and abiotic stress (heat) as modifiers of host gene expression in pea. *Mol Plant Pathol* **1**, 159-168.

Faivre-Rampant O, Gilroy EM, Hrubikova K, Hein I, Millam S, Loake GJ, Birch P, Taylor M & Lacomme C (2004). Potato virus X-induced gene silencing in leaves and tubers of potato. *Plant Physiol* **134**, 1308-1316.

Fauquet CM, Mayo MA, Maniloff J, Desselberger U & Ball LA (2005). *Virus taxonomy; Eight Report of the International Committee on Taxonomy of Viruses.* San Diego, CA: Elsevier Academic Press.

Fernández-Fernández ME, Mouriño M, Rivera J, Rodríguez F, Plana-Durán J & García JA (2001). Protection of rabbits against rabbit hemorrhagic disease virus by immunization with the VP60 protein expressed in plants with a potyvirus-based vector. *Virology* **280**, 283-291.

Franklin SE & Mayfield SP (2004). Prospects for molecular farming in the green alga *Chlamydomonas reinhardtii*. *Curr Opin Plant Biol* **7**, 159-165.

French R, Janda M & Ahlquist P (1986). Bacterial gene inserted in an engineered RNA virus: Efficien expression in monocolyledonous plant cells. *Science* **231**, 1294-1296.

Fütterer J, Bonneville JM & Hohn T (1990). Cauliflower mosaic virus as a gene expression vector for plants. *Physiologia Plantarum* **79**, 154-157.

Gammelgård E, Mohan M & Valkonen JPT (2007). Potyvirus-indiced gene silencing: the dynamic process of systemic silencing and silencing suppression. *J Gen Virol* **88**, 2337-2346.

German-Retana S, Candresse T, Alias E, Delbos RP & Le Gall O (2000). Effects of green fluorescent protein of β-glucuronidase tagging on the accumulation and pathogenicity of a resistance-breaking *Lettuce mosaic virus* isolate in susceptible and resistant lettuce cultivars. *Mol Plant Microbe Interact* **13**, 316-324.

German-Retana S, Redondo E, Tavert-Roudet G, Le Gall O & Candresse T (2003). Introduction of a NIa proteinase cleavage site between the reporter gene and HC-Pro only partially restores the biological properties of GUS- or GFP-tagged LMV. *Virus Res* **98**, 151-162.

Gils M, Kandzia R, Marilloinnet S, Klimuyk V & Gleba Y (2005). High-yield production of authentic human growth hormone using a plant virus-based expression system. *Plant Biotech J* **3**, 613-620.

Giritch A, Marilloinnet S, Engler C, van Eldik G, Botterman J, Klimyuk V & Gleba Y (2006). Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. *Proc Natl Acad Sci USA* **103**, 14701-14706.

Gleba Y, Klimyuk V & Marilloinnet S (2007). Viral vectors for the expression of proteins in plants. *Curr Opin Plant Biol* **18**, 134-141.

Gopinath K, Welling J, Porta C, Taylor KM, Lomonossoff GP & van Kammen A (2000). Engineering cowpea mosaic virus RNA-2 into a vector to express heterologous proteins in plants. *Virology* **267**, 159-173.

Gomord V & Faye L (2004). Posttranslational modification of therapeutic proteins in plants. *Curr Opin Plant Biol* **7**, 171-181.

Govier DA & Kassanis B (1974). Evidence that a component other than the virus particle is needed for aphid transmission of potato virus Y. *Virology* **57**, 285-286.

Govier DA & Woods RD (1971). Changes induced by magnesium ions in morphology in some plant viruses with filamentous particles. *J Gen Virol* **13**, 127-137.

Grill L, Palmer KE & Pogue GP (2005). Use of plant viruses for production of plant-derived vaccines. *Crit Rew Plant Sci* **23**, 309-323.

Guo HS, López-Moya JJ & García JA (1998). Susceptibility to recombination rearrangements of a chimeric plum pox potyvirus genome after insertion of a foreign gene. *Virus Res* **57**, 183-195.

Hafrén A & Mäkinen K (2008). Purification of viral genome-linked protein VPg from potato virus A infected plants reveals several post-translationally modified forms of the protein. *J Gen Virol*, accepted for publication.

Hagiwara Y, Peremyslov VY & Dolja VV (1999). Regulation of closterovirus gene expression examined by insertion of a self-processing reporter and by northern hybridization. *J Virol* **73**, 7988-7993.

Hari V (1981). The RNA of tobacco etch virus: Further characterization and detection of protein linked to RNA. *Virology* **112**, 391-399.

Haseloff J, Siemering KR, Prasher DC & Hodge S (1997). Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proc Natl Acad Sci USA* **94**, 2122-2127.

Hayes RJ, Petty ITD, Coutts RHA & Buck KW (1988). Gene amplification and expression in plants by a replicating genimivirus vector. *Nature* **334**, 179-182.

Hein I, Barciszewska-Pacak M, Hrubikova K, Williamson S, inesen M, Soenderby IE, Sundar S, Jarmolowski A, Shirasu K & Lacomme C (2005). Virus-induced gene silencing-based functional characterization of genes associated with powdery mildew resistance in barley. *Plant Physiol* **138**, 2155-2164.

Hinrichs-Berger J, Berger S & Buchenauer (2003). The P1 protein of *Potato virus Y* is transiently accumulated in systemically infected leaves of tobacco plants. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz* **110**, 568-571.

Hooker WJ (1981). *Compendium of potato diseases*. The American Phytopathological Society, St Paul, MN.

Holzberg S, Brosio P, Gross S & Pogue GP (2002). *Barley stripe mosaic virus*-induced gene silencing in a monocot plant. *Plant J* **30**, 315-327.

Hong Y & Hunt AG (1996). RNA polymerase activity catalyzed by a potyvirus-encoded RNA-dependent RNA polymerase. *Virology* **226**, 146-151.

Hsu CH, Lin SS, Liu FL, Su WC & Yeh SD (2004). Oral administration of a mite allergen expressed by zucchini yellow mosaic virus in cucurbit species downregulates allergen-induced airway inflammation and IgE synthesis. *J Allergy Clin Immunol* **113**, 1079-1085.

Hull AK, Yusibov V & Mett V (2005). Inducible expression in plants by virus-mediated transgene activation. *Transgenic Res* **14**, 407-416.

Hull R (2002). Virus replication. In *Matthew's Plant Virology* (4th edition), pp 352-353. Academic Press, Bath, Great Britain.

Ivanov K, Puustinen P, Rönnstrand L, Valmu L, Vihinen H, Gabrinaite R, Kalkkinen N & Mäkinen K (2003). Phosphorylation of the potyvirus capsid protein by plant protein kinase CK2 play an essential role in virus infection. *Plant Cell* **15**, 2124-2139.

Johansen IE, Lund OS, Hjulsager CK & Laursen J (2001). Recessive resistance in *Pisum sativum* and potyvirus pathotype resolved in a gene-for-cistron correspondence between host and virus. *J Virol* **75**, 6609-6614.

Johansen LK & Carrington JC (2001). Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. *Plant Physiology* **126**, 930-938.

Joshi RL, Joshi V & Ow DW (1990). BSMV genome mediated expression of a foreign gene in dicot and monocot plant-cells. *EMBO J* **9**, 2663-2669.

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

Kasschau KD & Carrington (1998). A counterdefensive strategy of plant viruses: Suppression of posttranscriptional gene silencing. *Cell* **95**, 461-470.

Kasschau KD, Cronin S & Carrington (1997). Genome amplification and long-distance movement fuctions associated with the central domain of tobacco etch potyvirus helper component-proteinase. *Virology* **228**, 251-262.

Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA & Carrington JC (2003). P1/HC-Pro, a viral suppressor of RNA silencing, interferes with *Arabidopsis* development and miRNA function. *Developmental Cell* **4**, 205-217.

Kearney CM, Donson J, Jones GE & Dawson WO (1993). Low level of genetic drift in foreign sequences replicating in an RNA virus in plants. *Virology* **192**, 11-17.

Kekarainen T, Merits A, Oruetxebarria I, Rajamäki ML & Valkonen JPT (1999). Comparison of the complete sequences of five different isolates of *Potato virus A* (PVA), genus *Potyvirus*. *Arch Virol* **144**, 2355-2366.

Kekarainen T, Savilahti H & Valkonen JPT (2002). Functional genomics on *Potato virus A*: Virus genome-wide map of sites essential for virus propagation. *Genome Res* **12**, 584-594.

Kekarainen T & Valkonen JPT (2000). Inoculation of viral RNA and cDNA to potato and tobacco plants using the Helios™ Gene Gun. *Technical Note 2351*, BIO-RAD Laboratories, USA. 4pp.

Klimyuk V, Marilloinnet S, Knaeblein J, McCaman M & Gleba Y (2005). Production of recombinant proteins in plants. In *Modern Biopharmaceuticals*. Edited by Knaeblein J. WILEY-VCH Verlag GmbH & Co. KGaA; 893-917.

Ko K, Steplewski Z, Glogowska M & Koprowski H (2005). Inhibition of tumor growth by plant-derived mAb. *Proc Natl Acad Sci USA* **102**, 7026-7030.

Koenig R, Lesemann DE, Loss S, Engelmann J, Commandeur U, Deml G, Schiemann J, Aust H & Burgermeister W (2006). Zygocactus virus X-based expression vectors and formation of rod-shaped virus-like particles in plants by the expressed coat proteins of *Beet necrotic yellow vein virus* and *Soil-borne cereal mosaic virus*. *J Gen Virol* **87**, 439-443.

Komarova TV, Skulachev MV, Zvera AS, Schwartz AM, Dorokhov YL & Atabekov JG (2006). New viral vector for efficient production of target proteins in plants. *Biochemistry (Moscow)* **71**, 846-850.

Kumagai MH, Donson J, Della-Cioppa G, Harvey D, Hanley K & Grill LK (1995). Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. *Proc Natl Acad Sci USA* **92**, 1679-1683.

Kumagai MH, Turpen TH, Weinzettl N, Della-Cioppa G, Turpen AM, Donson J, Hilf ME, Grantham GL, Dawson WO, Chow TP, Piatak M & Grill KL (1993). Rapid, high-level expression of biologically active α-trichosanthin in transfected plants by an RNA viral vector. *Proc Natl Acad Sci USA* **90**, 427-430.

Laín S, Riechmann JL & García JA (1990). RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Res* **12**, 584-594.

Laín S, Riechmann JL, Méndez E & García JA (1988). Nucleotide sequence of the 3′ terminal region of plum pox potyvirus RNA. *Virus Res* **10**, 325-342.

Lakatos L, Csorba T, Pantaleo V, Chapman EJ, Carrington JC, Loi YP, Dolja VV, Fernández-Calvino L, López-Moya J & Burgyán J (2006). Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J* **25**, 2768-2780.

Langenberg WG & Zhang L (1997). Immunocytology shows the presence of tobacco etch virus P3 protein in nuclear inclusions. *J Struct Biol* **118**, 243-247.

Laufs J, Wirzt U, Kammann M, Matzeit V, Schaeffer S, Schell J, Czernilofsku AP, Baker B & Gronenborn (1990). Wheat dwarf virus Ac/Ds vectors: Expression and excision of transposable elements introduced into various cereals by a viral replicon. *Proc Natl Acad Sci USA* **87**, 7752-7756.

Leffel SM, Mabon SA & Stewart CN (1997). Applications of green fluorescent protein in plants. *BioTechniques* **23**, 921-918.

Lehto K, Grantham GL & Dawson WO (1990). Insertion of sequences containing the coat protein subgenomic RNA promoter and leader in front of the tobacco mosaic virus 30K ORF delays its expression and causes defective cell-to-cell movement. *Virology* **174**, 145-157.

León-Bañares R, González-Ballester D, Galván A & Fernández E (2004). Transgenic microalgae as green cell-factories. *Trends Biotechnol* **22**, 45-52.

Lerouge P, Cabanes-Macheteau M, Rayon C, Fischette-Lainé AC, Gomord V & Faye L (1998). N-glycoprotein biosynthesis in plants: recent developments and future trends.

Li C, Sasaki N, Isogai M & Yoshikawa N (2004). Stable expression of foreign proteins in herbaceous and apple plants using Apple latent spherical virus RNA2 vectors. *Arch Virol* **149**, 1541-1558.

Lindbo JA (2007). High-efficiency protein expression in plants from agroinfection-compatible *Tobacco mosaic virus* expression vectors. *BMC Biotech* **7**, 52

Liu YL, Schiff M & Dinesh-Kumar SP (2004). Involvement of MEK1 MAPKK, NTF6 MAPK, WRKY/MYB transcription factors, COI1 and CTR1 in N-mediated resistance to tobacco mosaic virus. *Plant J* **38**, 800-809.

Llave C, Kasschau KD & Carrington JC (2000). Virus-encoded suppressor of posttranscriptional gene silencing targets a maintenance step in the silencing pathway. *Proc Natl Acad Sci USA* **97**, 13401-13406.

Ma JKC, Drake PMW & Christou P (2003). The production of recombinant pharmaceutical protein in plants. *Nat Rev Genet* **4**, 794-805.

MacFarlane SA & Popovich AH (2000). Efficient expression of foreign proteins in roots from tobravirus vectors. *Virology* **267**, 29-35.

Mallory AC, Reinhart BJ, Bartel D, Vance VB & Bowman LH (2002). A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro-RNAs in tobacco. *Proc Natl Acad Sci USA* **99**, 15228-15233.

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

Mascia PN & Flavell RB (2004). Safe and acceptable strategies for producing foreing molecules in plants. *Curr Opin Plant Biol* **7**, 189-195.

Masuta C, Yamana T, Tacahashi Y, Uyeda I, Sato M, Ueda S & Matsumura T (2000). Development of clover yellow vein virus as an efficient, stable gene-expression system for legume species. *Plant J* **23**, 539-546.

Matsuo K, Hong JS, Tabayashi AI, Masuta C & Matsumura T (2007). Development of *Cucumber mosaic virus* as a vector modifiable for different host species to produce therapeutic proteins. *Planta* **225**, 277-286.

Matzke MA, Matzke AJM, Pruss GJ & Vance VB (2001). RNA-based silencing strategies in plants. *Curr Opin Genet & Develop* **11**, 221-227.

Mavankal G & Rhoads RE (1991). *In vitro* cleavage at or near the N-terminus of the helper component protein in tobacco vein mottling virus polyprotein. *Virology* **185**, 721-731.

McDonald JG & Bancroft JB (1977). Assembly studies of potato virus Y and its coat protein. *J Gen Virol* **35**, 251-263.

Merits A, Rajamäki ML, Lindholm P, Runeberg-Roos P, Kekarainen T, Puustinen P, Mäkeläinen K, Valkonen JPT & Saarma M (2002). Proteolytic processing of potyviral proteins and polyprotein processing intermediates in insect and plant cells. *J Gen Virol* **83**, 1211-1221.

Mori M, Kaido M, Okuno T & Furusawa I (1993). mRNA amplification system by viral replicase in transgenic plants. *FEBS Lett* **336**, 171-174.

Mori M, Fujuhara N, Mise K & Furusawa I (2001). Inducible high-level mRNA amplification system by viral replicase in transgenic plants. *Plant J* **27**, 79-86.

Murphy JF, Rychlik W, Rhoads RA, Hunt AG & Shaw JG (1991). A tyrosine residue in the small nuclear inclusion protein of tobacco vein mottling virus links the VPg to the viral RNA. *J Virol* **65**, 511-513.

Murphy PA & McKay R (1932). *Scient. Proc. R. Dubl. Soc.* **20,** 347.

Nicolas O, Dunnington SW, Gotow LF, Pirone TP & Hellmann GM (1997). Variations in the VPg protein allow a potyvirus to overcome *va* gene resistance in tobacco. *Virology* **237**, 452-459.

Nishiguchi M, Yamasaki S, Lu XZ, Shomoyama A, Hanada K, Sonoda S, Shimono M, Sakai J, Mikoshiba Y & Fujisawa I. (2006). Konjak mosaic virus: the complete nucleotide sequence of the genomic RNA and its comparison with other potyviruses. *Arch Virol* **151**, 1643-1650.

Nurkiyanova KM, Ryabov EV, Commandeur U, Duncan GH, Canto T, Gray SM, Mayo MA & Taliansky ME (2000). Tagging *Potato leafroll virus* with the jellyfish green fluorescent protein gene. *J Gen Virol* **81**, 617-626.

O'Brien GJ, Bryant CJ, Voogd C, Greenberg HB, Gardner RC & Bellamy AR (2000). Rotavirus VP6 expressed by PVX vectors in *Nicotiana benthamiana* coats PVX rods and also assembles into viruslike particles. *Virology* **270**, 444-453.

Ohshima K, Tomitaka Y, Wood JT, Minematsu Y, Kajiyama H, Tomimura K & Gibbs AJ (2007). Patterns of recombination in turnip mosaic virus genomic sequences indicate hotspots of recombination. *J Gen Virol* **88**, 298-315.

Paalme V, Gammelgård E, Järvekülg L & Valkonen JPT (2004). *In vitro* recombinants of two nearly identical potyviral isolates express novel virulence and symptom phenotypes in plants. *J Gen Virol* **85**, 739-747.

Palukaitis P & García-Arenal F (2003). *Cucumber mosaic virus*. Entry 400 in DPVweb. Association of Applied Biologists, Adams MJ & Antoniw JF (2006). DPVweb: a comprehensive database of plant and fungal virus genes and genomes**.** *Nucleic Acids Research* 34, Database issue, D382-D385. http://www.dpvweb.net/index.php

Pantaleo V, Szittya G & Burgyán J (2007). Molecular bases of viral RNA targeting by viral small interfering RNA-programmed RISC. *J Virol* **81**, 3797-3806.

Perez Y, Mozes-Koch R, Akad F, Tanne E, Czosnek H & Sela I (2007). A universal expression/silencing vector in plants. *Plant Physiol* **145**, 1251-1263.

Pérez Filqueira DM, Zamorano PI, Domínguez MG, Taboga O, Del Médico Zajac MP, Puntel M, Romera SA, Morris TJ, Borca MV & Sadir AM (2003). Bovine herpes virus gD protein produced in plants using a recombinant tobacco mosaic virus (TMV) vector possesses authentic antigenicity. *Vaccine* **21**, 4201-4209.

Phan J, Zdanov E, Evdokimov AG, Tropea JE, Peters HK, Kapust RB, Li M, Wlodawer A & Waugh DS (2002). Structural basis for the substrate specifity of tobacco etch virus protease. *J Biol Chem* **277**, 50564-50572.

Pruss G, Ge X, Shi XM, Carrington JC & Bowman Vance V (1997). Plant viral synergism: The potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *Plant Cell* **9**, 859-868.

Puurand U, Mäkinen K, Paulin L & Saarma M (1994). The nucleotide sequence of potato virus A genomic RNA and its sequence similarities with other potyviruses. *J Gen Virol* **75,** 457-461.

Qian Y, Mugiira RB & Zhou X (2006). A modified viral satellite DNA-based gene silencing vector is effective in association with heterologous begomoviruses. *Virus Res* **118**, 136-142.

Rabindran S & Dawson WO (2001). Assessment of recombinants that arise from the use of a TMV-based transient expression vector. *Virology* **284**, 182-189.

Rajamäki ML, Merits A, Rabenstein F, Andrejeva J, Paulin L, Kekarainen T, Kreuze JF, Forster RLS & Valkonen JPT (1998). Biological, serological and molecular differences among isolates of potato A potyvirus. *Phytopathology* **88**, 311-321.

Rajamäki ML, Mäki-Valkama T, Mäkinen K & Valkonen JPT (2004). Infection with potyviruses. In: N. Talbot (Ed.), *Plant-Pathogen Interactions*, Blackwell Publishing, Oxford, pp. 68.91.

Rajamäki ML & Valkonen JPT (2003). Localization of a potyvirus and the viral genome-linked protein in upper non-inoculated leaves at an early stage of systemic infection. *Mol Plant Microbe Interact* **16**, 25-34.

Ravin NV, Mardanova ES, Kotlyarov RY, Novikov VK, Atabekov JG & Skryabin KG (2008). Complete sequencing of Potato virus X new strain genome and construction of viral vector for production of target proteins in plants. *Biochemistry (Moscow)* **73**, 44-49.

Reddy SA, Czerwinski D, Rajapaksa R, Reinl S, Garner SJ, Cameron TI, McCormick AA, Barret J, Novak J, Tusé D, Holtz RB & Levy R (2002). Plant derived single-chain Fv idiotype vaccines are safe and immunogenic in patients with follicular lymphoma: results of a phase I study. *Blood* **100**: 163a, In: Proceedings of the 44th Annual Conference of the American Society of Hematology.

Remans T, Schenk PM, Manners JM, Grof CPL & Elliot AR (1999). A protocol for the fluorometric quantification of mGFP5-ER and sGFP(S65T) in transgenic plants. *Plant Mol Biol Reporter* **17**, 385-395.

Riechman JL, Laín S & García JA (1992). Highlights and prospects of potyvirus molecular biology. *J Gen Virol* **73**, 1-16.

Riechman JL, Cervera MT & García JA (1995). Processing of the plum pox virus polyprotein at the P3-6K1 junction is not required for virus viability. *J Gen Virol* **76**, 951-965.

Roberts DJ & Harrison BD (1989). *Tobacco rattle virus*. Entry 346 in DPVweb. Association of Applied Biologists, Adams MJ & Antoniw JF (2006). DPVweb: a comprehensive database of plant and fungal virus genes and genomes**.** *Nucleic Acids Research* 34, Database issue, D382-D385. http://www.dpvweb.net/index.php

Roberts IM, Wang D, Findlay K & Maule AJ (1998). Ultrastructural and temporal observations of the potyvirus cylindrical inclusions (CIs) show that the CI protein acts transiently in aiding virus movement. *Virology* **245**, 173-181.

Rodrígues-Cerezo E, Klein PG & Shaw JG (1991a). A determinant of disease symptoms severity is located in the 3'-terminal noncoding region of the RNA of a plant virus. *Proc Natl Acad Sci USA* **88**, 9863-9867.

Rodrígues-Cerezo E & Shaw JG (1991b). Two newly detected nonstructural viral proteins in potyvirus-infected cells. *Virology* **185**, 572-579.

Rodrígues-Cerezo E, Ammar ED, Pirone TP & Shaw JG (1993). Association of the non-structural P3 viral protein with cylindrical inclusions in potyvirus-infected cells. *J Gen Virol* **74**, 1945-1949.

Rodrígues-Cerezo E, Findlay K, Shaw JG, Lomonossoff GP, Qui SG, Linstead P, Shanks M & Risco C (1997). The coat and cylindrical inclusion proteins of a potyvirus are associated with connections between plant cells. *Virology* **236**, 296-306.

Rojas MR, Zerbini FM, Allison RF, Gilbertson RL & Lucas WJ (1997). Capsid protein and helper component-proteinase function as potyvirus cell-to-cell movement proteins. *Virology* **237**, 283-295.

Sainsbury F, Lavoie PO, D'Aoust MA, Vézina LP & Lomonossoff GP (2008). Expression of multiple proteins using full-length and deleted versions of cowpea mosaic virus RNA-2. *Plant Biotech J* **6**, 82-92.

Saitoh, H, Kiba A, Nishihara M, Yamamura S, Suzuki K & Terauchi R (2001). Production of antimicrobial defensin in *Nicotiana benthamiana* with a potato virus X vector. *Mol Plant Microbe Interact* **14**, 111-115.

Schaad MC, Jensen PE & Carrington JC (1997a). Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein. *EMBO J* **16**, 4049-4059.

Schaad MC, Lellis AD & Carrington JC (1997b). VPg of tobacco etch potyvirus is a host genotype-specific determinant for long-distance movement. *J Virol* **71**, 8624-8631.

Schechter I & Berger A (1967). On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **25**, 157-162.

Schmitz I & Rao ALN (1996). Molecular studies on bromovirus capsid protein: I. Characterization of cell-to-cell movement-defective RNA3 variants of brome mosaic virus. *Virology* **226**, 281-293.

Scholthof BH, Scholthof KBG & Jackson AO (1996). Plant virus gene vectors for transient expression of foreign proteins in plants. *Annu Rev Phytopathol* **34**, 299-323.

Shiboleth YM, Arazi T, Wand Y & Gal-On A (2001). A new approach for weed control in a cucurbit field employing an attenuated potyvirus-vector for herbicide resistance. *J Biotech* **92**, 37-46.

Shingu Y, Yokomizo S, Kimura M, Ono Y, Yamaguchi I & Hamamoto H (2006). Conferring cadmium resistance to mature tobacco plants through metal-adsorbing particles of tomato mosaic virus vector. *Plant Biotech J* **4**, 281-288.

Shoresh M, Gal-On A, Leibman D & Chet I (2006). Characterization of a mitogen-activated protein kinase gene from cucumber required for Trichodema-conferred plant resistance. *Plant Physiol* **142**, 1169-1179.

Shukla DD, Ward CW & Brunt AA (1994). *The Potyviridae*. C.A.B International, Wallingford, UK. 516p.

Simon AE & Bujarski JJ (1994). RNA-RNA recombination and evolution in virus-infected plants. *Annu Rev Phytopathol* **32**, 337-362.

Skarjinskaia M, Karl J, Aurajo A, Ruby K, Rabindran S, Streatfield SJ & Yusibov V (2008). Production of recombinant proteins in clonal root cultures using episomal expression vectors. *Biotechnol Bioeng*, DOI: 10.1002/bit.21802.

Spetz C, Valkonen JPT (2004). Potyviral 6K2 protein long-distance movement and symptom-induction functions are independent and host-specific. *Mol Plant Microbe Interact* **17**, 502-510.

Streatfield SJ (2006). Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnol J* **5**, 2-15.

Sudarshana MR, Plesha MA, Uratsu SL, Falk BW, Dandekar AM, Huang TK & McDonald KA (2006). A chemically inducible cucumber mosaic virus amplicon system for expression of heterologous proteins in plant tissues. *Plant Biotech J* **4**, 551-559.

Takamatsu N, Ishikawa M, Meshi T & Okada Y (1987). Expression of bacterial chloroamphenicol acetyltransferase gene in tobacco plants mediated by TMV-RNA. *EMBO J* **6**, 307-311.

Twyman RM, Stoger E, Schhillberg S, Christou P & Fisher R (2003). Molecular farming in plants: host systems and expression technology. *Trends Biotechnol* **21**, 570-578.

Usha R, Rohill JB, Spall VE, Shanks M, Maule AJ, Johnson JE & Lomonossoff GP (1993). Expression of an animal virus antigenic site on the surface of a plant virus particle. *Virology* **197**, 366-374.

Valkonen JPT, Puurand Ü, Slack SA, Mäkinen K & Saarma M (1995). Three strain groups of Potato A potyvirus based on hypersensitive responses in potato, serological properties, and coat protein sequences. *Plant Dis* **79**, 748-753.

Van Bel AJE (2003). The phloem, a miracle of ingenuity. *Plant, Cell and Environment* **26**, 125-149.

Varrelmann M & Maiss E (2000). Mutations in the coat protein gene of *Plum pox virus* suppress particle assembly, heterologous encapsidation and complementation in transgenic plants of *Nicotiana benthamiana*. *J Gen Virol* **81**, 567-576.

Varsani A, Williamson AL, Stewart D & Rybicki EP (2006). Transient expression of Human papillomavirus type 16 L1 protein in *Nicotiana benthamiana* using an infectious tobamovirus vector. *Virus Res* **120**, 91-96.

Verch T, Yusibov V & Koprowski H (1998). Expression and assembly of a full-length monoclonal antibody in plants using a plant virus vector. *J Immunol Methods* **220**, 69-75.

Verchot J & Carrington JC (1995). Evidence that the potyvirus P1 proteinase fuctions in trans as accessory factor for genome amplification. *J Virol* **69**, 3668-3674.

Verchot J, Herndon KL & Carrington JC (1992). Mutational analysis of the tobacco etch potyviral 35-kDa proteinase: identification of essential residues and requirements for autoproteolysis. *Virology* **190**, 298-306.

Verchot J, Koonin EV & Carrington JC (1991). The 35-kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* **185**, 527-535.

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

Voinnet O (2005). Induction and suppression of RNA silencing: Insights from viral infections. *Nat Rev Genet* **6**, 206-221.

Wagner B, Hufnagl K, Radauer C, Wagner S, Baier K, Scheiner O, Wiedermann U & Breiteneder H (2004). Expression of the B subunit of the heat-labile enterotoxin of *Eschericia coli* in tobacco mosaic virus-infected *Nicotiana benthamiana* plants and its characterization as mucosal immunogen and adjuvant. *J Immunol Methods* **287**, 203-215.

Wang RY, Ammar ED, Thornbury DW, Lopez-Moya JJ & Pirone TP (1996). Loss of potyvirus transmissibility and helper component activity correlate with non-retention of virions in aphid stylets. *J Gen Virol* **77**, 861-867.

Wang D & Maule AJ (1995). Inhibition of host gene expression associated with plant virus replication. *Science* **267**, 229-231.

Wassenegger M & Pélissier T (1998). A model for RNA-mediated gene silencing in higher plants. *Plant Mol Biol* **37**, 349-362.

Waterhouse PM, Graham MW & Wang MB (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci USA* **95**, 13959-13964.

Werner S, Marilloinnet S, Hause G, Klimyuk V & Gleba Y (2006). Immunoabsorbent nanoparticles based on a tobamovirus displaying protein A. *Proc Natl Acad Sci USA* **103**, 17678-17683.

Whitham SA, Yamamoto ML & Carrington JC (1999). Selectable viruses and altered susceptibility mutants in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **96**, 772-777.

Wu X, Beachy RN, Wilson MA & Shaw JG (1990). Inhibition of uncoating of tobacco mosaic virus particles in protoplasts from transqenic tobacco plants that express the viral coat protein gene. *Virology* **179**, 893-895.

Wu X & Shaw JG (1998). Evidence that assembly of a potyvirus begins near the 5′ terminus of the viral RNA. *J Gen Virol* **79**, 1525-1529.

Yang CD, Liao JT, Lai CY, Jong MH, Liang CM, Lin YL, Lin NS, Hsu YH & Liang SM (2007). Induction of protective immunity in swine by recombinant bamboo mosaic virus expressing foot-and-mouth disease virus epitopes. *BMC Biotech* **7**, 62.

Yang LJ, Hidaka M, Masaki H & Uozumi T (1998). Detection of potato virus Y P1 protein in infected cells and analysis of its cleavage site. *Biosci Biotechnol Biochem* **62**, 380-382.

Zhang C & Ghabrial SA (2006). Development of *Bean pod mottle virus*-based vectors for stable protein expression and sequene-specific virus-induced gene silencing in soybean. *Virology* **344**, 401-411.

Zhang X & Mason H (2005). Bean yellow dward virus replicons for high-level transgene expression in transgenic plants and cell cultures. *Biotechnol Bioeng* **93**, 271-279.