MOLECULAR CHARACTERIZATION OF A NOVEL TOBAMOVIRUS INFECTING HIBISCUS

KANDHADAYAR GOPALAN SRINIVASAN

M.Sc (Trichy), M.Sc (Notts.)

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2003

ACKNOWLEDGEMENTS

I am immensely indebted to my supervisor, Associate Professor Sek Man Wong, for his professional guidance, understanding and for urging me on by way of his untiring support. My special thanks go to Prof. Peter Palukaitis, Head of Virology Division, Scottish Crop Research Institute, U.K. He has been a very special mentor and I thank him for sharing with me his vast knowledge in the field of virology.

I would like to acknowledge and extend my appreciation to my collaborator and good friend, Dr. M.V.Skulachev, Senior Researcher, Dept. of Virology, Faculty of Biology, Moscow State University, Russia for contributing towards technical content of this work. I would also like to thank Prof. K.H. Ryu, Seoul women's University, South Korea for spending time money & personnel and Miss. B.E. Min for her continued efforts towards furthering this research project.

I would like to express my gratitude to Mr. P.L. Chong, Mdm. C. F. Liew, Mdm. G.L. Loy, people who stand out most notably in my mind for their excellent technical assistance.

Special thanks to all members of Plant Virology Lab, NUS, particularly, K.C. Lee and A.A. Prabha for being with me during the most difficult times. I wish to acknowledge my friends Naren, Vinod, Ram, Baskar, Rao, Bhinu, Dhinoth and NUS cricketing fraternity for their role in my life in Singapore.

I would like to thank my parents for shaping up my life and for showing unlimited patience and belief in me. I wish to dedicate this work to the memory of my beloved mother.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

Srinivasan KG, Narendrakumar R and Wong SM (2002) Hibiscus Virus-S is a new subgroup II tobamovirus: Evidence from its unique coat protein and movement protein sequences. Archives of Virology, Aug;147(8):1585-98.

Srinivasan KG, Min BE, Ryu KH and Wong SM Determination of complete nucleotide sequence of HVS, a novel tobamovirus infecting hibiscus: Evidence for the presence of an internal poly-A tract within 3' untranslated region. **(Manuscript submitted to Archives of Virology)**.

Srinivasan KG, Skulachev MV and Wong SM Characterization of internal ribosome entry sites in Hibiscus Virus-S. **(Manuscript in preparation)**.

Srinivasan KG and Wong SM (2003). Cloning and Characterization of a new tobamovirus infecting Hibiscus rosa-sinensis) L. Annual Meeting and Symposium of Korean Society for Plant Pathology.

Srinivasan KG, Narendrakumar R and Wong SM (2001) Cloning and characterisation of a novel tobamovirus infecting Hibiscus rosa-sinensis L.
Sixth International Symposium on Positive Strand RNA Viruses, Paris, France.

LIST OF ABBREVIATIONS

Viruses

AMV	Alfalfa mosaic virus
BMV	Brome mosaic virus
BSMV	Barley stripe mosaic virus
CGMMV-SH	Cucumber green mottle mosaic virus
CMV	Cucumber mosaic virus
CPMV	Cowpea mosaic virus
CymMV	Cymbidium mosaic virus
EMCV	Encephalomyocarditis virus
FHV	Florida hibiscus virus
FMDV	Foot and mouth disease virus
GRV	Groundnut rosette virus
HAV	Hepatitis A virus
HCRSV	Hibiscus chlorotic ringspot virus
HCV	Hepatitis C virus
HVS	Hibiscus virus, Singapore isolate
IBV	Infectious bronchitis virus
KGMMV-Y	Kyuri green mottle mosaic virus, strain Yodo
M-MuLV	Moloney murine leukemia virus
ORSV-S1	Odontoglossum ringspot virus, Singapore isolate
PMMV-S	Pepper mild mottle virus, strain S
PVX	Potato virus –X
SHMV	Sunn-hemp mosaic virus

TEV	Tobacco etch virus
TMGMV	Tobacco mild green mosaic virus
TMV- Cr	Tobacco mosaic virus, strain crucifer
TMV-Ob	Tobacco mosaic virus, strain Obuda
ToMV	Tomato mosaic virus
TRV	Tobacco rattle virus
TVCV	Turnip vein clearing virus
ZGMMV	Zucchini green mottle mosaic virus

COMMON TERMS

aa(s)	Amino acid(s)
Avr	Avirulence gene
CP	Coat protein
dRNA	Defective RNA
eIF(s)	Eukaryotic translation initiation factors
GFP	Green fluorescent protein
GUS	β-Glucuronidase
HR	Hypersensitive response
IRES(s)	Internal ribosome entry site(s)
MP	Movement protein
MT-Hel	Methyltransferase helicase
nt(s)	Nucleotide(s)
OAS	Origin of virion assembly
ORF(s)	Open reading frame(s)
PABP	Poly-A binding protein
PARS	Poly- A rich sequences
PME	Pectin methylesterase
PTB	Pyrimidine tract binding protein
RdRp	RNA dependent RNA polymerase
Rubisco	Ribulose bisphosphate carboxylase / oxygenase
sgRNA	Subgenomic RNA
upORF(s)	Upstream ORF(s)
UTR (s)	Untranslated region(s)

LIST OF FIGURES

LIST OF TABLES

SUMMARY

Full-length sequence of a new tobamovirus infecting *Hibiscus rosa-sinensis L* was determined. The genome consists of a 5'UTR, a RNA dependent RNA polymerase (RdRP) encoding for methyltransferase-helicase, a movement protein (MP), a coat protein (CP) and a unique 3' UTR. Phylogenetic trees generated based on the nucleotide and amino acid sequences of its ORFs indicate that this new virus belongs to subgroup II of tobamoviruses. The unique feature in this new tobamovirus is the presence of a 3'UTR that is very different from other tobamoviruses. Usually the 3' UTR region consists of a pseudo-knotted domain followed by a tRNA-like structure at the 3' terminus. In HVS, the pseudo-knotted domain is replaced by a poly-A tract of 77-96 nt in length. This poly-A tract may perform the function of pseudoknot which acts as a translational enhancer.

Several attempts were made to construct biologically active cDNA clones of HVS. Initially due to the lack of information on 3' UTR apart from the poly-A tract, full-length clones of HVS without 3' UTR and with poly $A_{(40)}$ region immediately downstream of CP gene were constructed, tested and were found to be non-infectious. HVS clones with 3' UTRs of other tobamoviruses and with the authentic HVS 3' UTR were constructed and tested for infectivity in kenaf protoplasts, *C. quinoa* and *N. benthamiana*. *In vitro* transcripts of these clones were unable to replicate in any of these systems.

For tobamoviruses, ORFs that are located at the 3' end of the genomes are expressed after synthesis of sgRNAs. Studies directed towards understanding

the translational strategies involved in expression of 3’ proximal genes in HVS led to the identification of two putative IRES regions. These were located immediately upstream of MP and CP ORFs. Using deletion constructs, we have identified two polypurine tracts responsible for the activity of CP IRES. Region driving the expression of MP by cap-independent mechanism was shown to possess stem loop structures that are responsible for IRES-like activity. Preliminary investigations show that both MP and CP IRESs are active *in vivo* in agro-infiltrated tissues of *N. benthamiana*.

CHAPTER I

LITERATURE REVIEW

1.1 Introduction

Tobamoviruses are one of the very well characterized group of viruses and they have been a very useful tool for understanding the very basic processes of virus infection, multiplication and their survival. The advent of modern molecular biological techniques have helped us to gain insights into the genome organization, expression strategies of different viruses in general and tobamoviruses in particular, which in turn have lead to the discovery of methods to overcome the crop losses resulting from virus epidemics as well their exploitation as vectors for expressing therapeutic proteins (Hamamoto et al., 1993; Wu et al., 2003).

The genus tobamoviruses consists of 19 definitive species and 4 tentative species (Brunt et al., 1996). They can be classified into 3 sub-groups based on their host range and genome organization. Complete sequences of several tobamoviruses have been obtained (Goelet et al., 1982; Ohno et al., 1984; Solis and Garcia-Arenal., 1990; Ikeda et al., 1993; Alonso et al., 1991;Chng et al., 1996; Lartey et al., 1995; Meshi et al., 1981; Silver et al., 1996;Tan et al., 2000; Ugaki et al., 1991) and partial sequences of tobamoviruses are also being reported (Srinivasan et al., 2002, Adkins et al., 2003). Crop losses due to various viral epidemics have been documented in spite of the fact that viral infections may result in direct and indirect damages to crop plants. There are various compilations of crop loss data which may help in predicting global estimates of crop losses (Waterworth and Hadidi

1998). Tobamovirus -related reduction in crop yields have been estimated to be 30 -35% in tobacco and 15 -30% in cucurbits (Sutic et al., 1999).

1.2 Tobamovirus taxonomy

Viruses are classified based on various characteristics such as morphology, physicochemical properties, genome organization, serological relatedness, mode of transmission, host range and symptomology. For classification in a higher taxa more stable properties such as particle morphology, type of nucleic acid and genome organization have been used. To distinguish between viruses falling within the same group, characteristics such as serological relatedness, mode of transmission, host range, symptomology may serve the purpose (Fauquet 1999).

Tobamoviruses are a group of plant viruses that have rod shaped virions, encapsidating a single-stranded positive sense RNA genome. Their genome is capped at the 5' terminus and has a tRNA like structure at the 3' end. *Tobacco mosaic virus* (TMV) is the type member of tobamoviruses and its genome encodes for at least 5 different polypeptides that are expressed at different stages of the infection cycle. Based on sequence motifs in RNA-dependent RNA polymerase (RdRp) positive-strand RNA viruses have been classified into three 'Supergroups' of which tobamoviruses belongs to Supergroup III (Koonin and Dolja., 1993).

Initially tobamoviruses were divided into two subgroups, based on locations of their virion origin of assembly sequence (OAS) (Fukuda et al 1981). For subgroup I tobamoviruses the OAS is located within the movement protein (MP) open reading frame (ORF); for subgroup II tobamoviruses it is located within the coat

protein (CP) ORF (Fukuda 1981). It was argued that this subdivision is also supported by organization of MP and CP ORFs (Lartey et al 1996). In subgroup I viruses, for example in TMV, these ORFs are separated by two nucleotides whereas, in the genomes of subgroup II viruses such as *Cucumber green mottle mosaic virus*, strain SH (CGMMV-SH) and *Sunn-hemp mosaic virus* (SHMV) the MP and CP ORFs overlap by 25 and 28 nucleotides, respectively. There are exceptions. In other subgroup II viruses such as *Kyuri green mottle mosaic virus* (KGMMV) and *Zucchini green mottle mosaic virus* (ZGMMV) the MP and CP ORFs do not overlap even though the OAS is located within the CP ORF, which is a subgroup II characteristic. Crucifer-infecting tobamoviruses are unique in the sense that their OAS is located within the MP ORF which is more of a subgroup I characteristic but the MP and CP ORFs overlap by 77 nucleotides which is more like subgroup II viruses. They were classified as subgroup III viruses by Lartey and his co-workers (1996). Moreover, the genome organization of crucifer-infecting tobamoviruses suggests that they probably have evolved from the other two subgroups rather than being 'ancestral' (Gibbs, 1999).

Results of phylogenetic clustering based on CP and MP amino acids clearly separates subgroup I (most of which infect solanaceous plants), subgroup II (non solanaceous infecting viruses) and subgroup III (crucifer-infecting viruses) viruses. *Odontoglossum ringspot virus* (ORSV) has its OAS located within the MP ORF and the genome organization is similar to subgroup I viruses. Phylogenetic clustering based on MP and CP amino acids place ORSV among crucifer-infecting subgroup III, whereas they seem to be clustering with subgroup I viruses, when RdRp and MT-Hel domains are compared (Lartey et al.,1996). Hence

ORSV is thought to be a recombinant virus (Gibbs, 1999). In our study we found that Hibiscus Virus Singapore Isolate (HVS) clusters with subgroup II viruses in general and Adkins et al (2003) have found that HVS and *Florida hibiscus virus* (FHV) form a distinct cluster within this subgroup. With the cloning and sequencing of several other possible tobamoviruses it will be interesting to locate the root of tobamovirus lineage. Currently, it is thought to be in between SHMV and CGMMV lineages (Gibbs 1999). This conclusion is supported by the results of phylogenetic analysis of rubisco complexes of their respective host plants since, tobamoviruses are thought to have co-evolved with their hosts (Gibbs, 1986). In conclusion, division of tobamoviruses into different subgroups based on phylogenetic analysis of their CPs seems to be an unambiguous method of classifying them, not to mention that analysis of other ORFs and genome organization will certainly help in establishing the tobamoviral inter-relationships.

1.3 Genome organization

Tobamoviruses share common features such as the 5' cap , 3' UTR, readthrough expression strategy, and expression of 3' proximal ORFs by sub-genomic RNA (sgRNA) with Supergroup III viruses (Strauss et al., 1996).

To date sequences of several tobamoviruses have been determined. The genome length of tobamoviruses vary between 6311 (Lartey et al., 1995) and 6611 (Ikegami et al.,1995) nucleotides. The genome encodes for at least 5 different polypeptides.

Starting from the 5' end, typically a tobamovirus genome consists of a 5' untranslated region (UTR), RdRP gene, MP gene, CP gene and 3' UTR (Fig. 1.1).

TOBAMOVIRUS : GENOME ORGANISATION

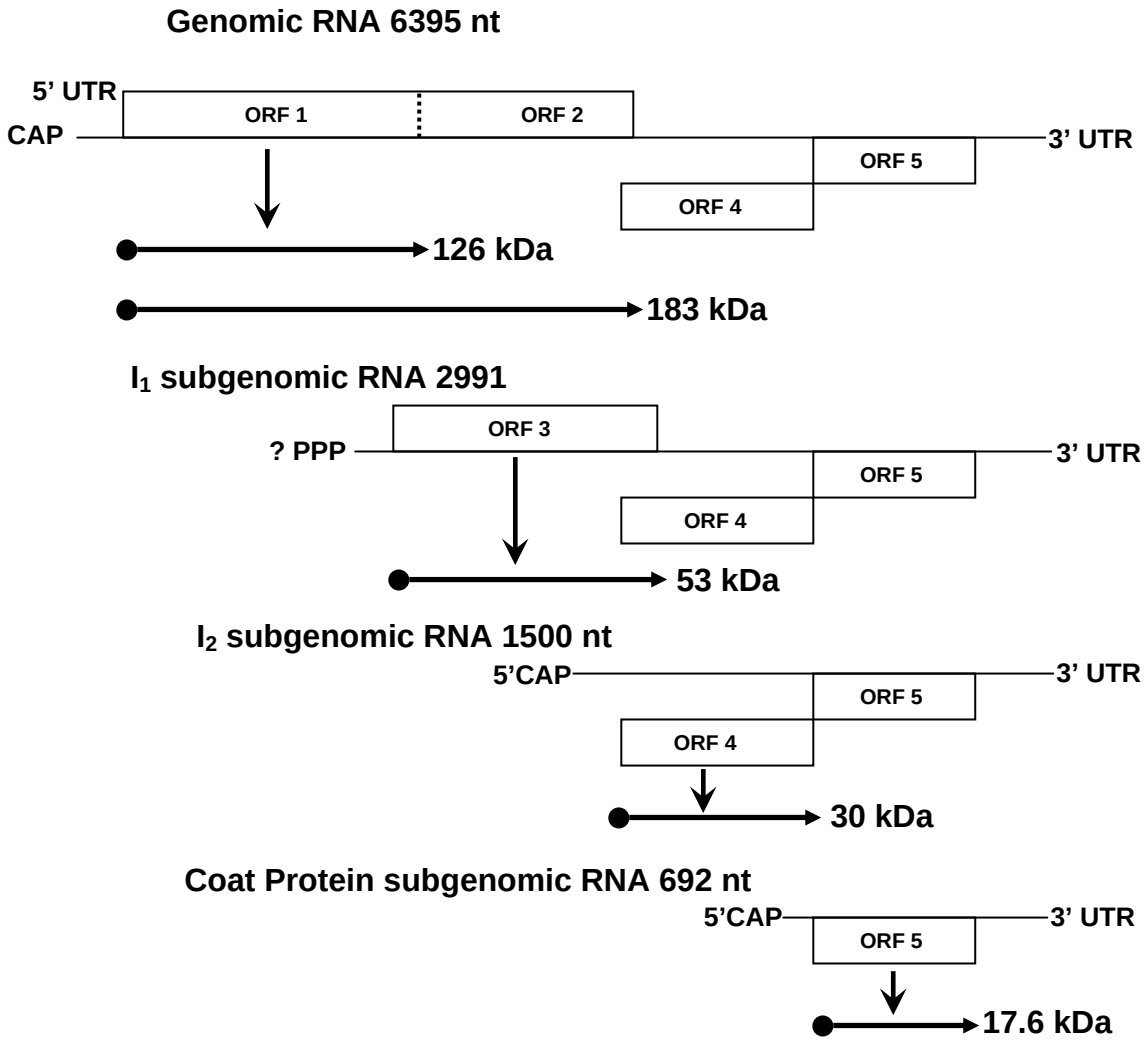


Fig 1.1 Genome organization and translation strategy of TMV-vulgare.
Cistrons are represented by open rectangles.
Protein functions:126kDa methyl transferase, helicase and readthrough product 183kDa , polymerase. 54kDa, not known. 30kDa , cell to cell movement. 17.6 kDa, Coat protein.

(Adapted from Bustamante and Hull, 1998)

The first ORF begins around 60-70 nucleotides downstream of 5' terminus and terminates in an amber stop codon, UAG. This region encodes for a Methyltransferase helicase (MT-Hel) or 126 kDa protein. Readthrough of this amber stop codon results in a 183 kDa protein. This is followed by the MP and CP genes. Apart from the ORFs mentioned above, it has been shown that there is a sgRNA for a 54 kDa protein (Sulzinski et al., 1985) the function of which is yet to be understood. The 54 kDa ORF is located at the 3' end of the 126 kDa and it runs into 3' end of 183 kDa ORF. The sgRNA encoding the 54 kDa protein has been found to be associated with polyribosomes, suggesting that it is being translated in infected tissues (Sulzinski et al.,1985).

Yet another protein (4 kDa) is encoded by ORF X which begins at the 3' end of MP and extends into the 5' region of CP (Morozov et al.,1993). It was found that this 4 kDa protein was associated with 50 kDa protein in vitro translation systems. This protein is shown to bind with elongation factor EF-1-α in *in vitro* translation systems (Fedorkin et al., 1995). However it is not proven as to whether this 4 kDa protein is associated with translation factors or whether it has a role in translation of viral genes *in planta*. In sub group II viruses such as cucumber green mottle mosaic virus (CGMMV) and sunn-hemp mosaic virus (SHMV) ORF-X lacks the initiator codon (Morozov et al.,1993).

1.4 Expression strategies

The 126 kDa (MT-Hel) and 183 kDa proteins (RdRp) located at the 5' end of the genome are expressed directly from the genomic RNA (Siegel et al.,1978) by scanning of ribosomes from 5' end by cap dependent translation (Kozak,1991).

The expression of 183 kDa protein is by readthrough strategy whereby the UAG amber stop codon is bypassed (Pelham, 1978). It was proposed that anticodon-codon interactions involving tyrosine tRNA could form the basis for readthrough mechanism (Beier et al.,1984). Later experiments showed that the readthrough mechanism was influenced by the sequences flanking the UAG stop codon (Skuzeski et al.,1991). Deletion of the amber stop codon in *Tomato mosaic virus* (ToMV) led to poor replication rates in protoplasts and the virus reverted back to wild type in plants (Ishikawa et al., 1986, 1991). Both 126 kDa and 183kDa proteins are required for efficient replication rates and it was proposed that they function as heterodimers (Watanabe et al., 1999).

The internally located MP and CP genes are expressed by a similar mechanism from sgRNAs that are 3' co terminated (Beachy and Zaitlin,1977; Jackson,1971; Hunter et al.,1976; Siegel et al., 1976). Several mechanisms by which the sgRNAs are generated in positive strand RNA viruses have been proposed. In internal initiation mechanism, (+) strand sgRNA is synthesized by viral replicase through binding internally on the (-) strand copy of genomic RNA (Miller et al., 1985). Another mechanism is through premature termination of (-) strand synthesis from genome length (+) strand which subsequently serves as template for (+) strand sgRNA synthesis (Sit et al.,1998). Discontinuous transcription may also result in the production of sgRNAs. Variations of discontinuous transcription has been reported in *Mouse hepatitis virus* (Zhang and Lai, 1994; Sawicki and Sawicki,1990). In *Potato virus X* (PVX), it has been shown that mutation of sequences at the 5' terminal that base pairs with an octanucleotide segment upstream of transcription start sites of sgRNAs abolishes production of sgRNA

(Kim and Hemenway,1996,1999). Various mechanisms involved in the synthesis of sgRNA in positive-strand RNA viruses have been reviewed in detail (Miller and Koev, 2000).

In TMV, MP and CP gene products are synthesized from sgRNAs but they are expressed at different times of the viral infection cycle. MP is expressed early in infection (Watanabe et al., 1984b) in low amounts (Ooshika et al.,1984). CP is expressed late and at extremely high levels (Siegel et al.,1978). Sequences upstream of CP and sequences within MP were shown to direct sgRNA synthesis (Lehto et al.,1990a; Deom et al.,1994; Giesman-Cookmeyer et al., 1995). Fine mapping has led to clear identification of minimal promoter elements required for sgRNA synthesis in TMV. Using deletion analysis, the borders for fully active promoter in TMV-U1 was localized between -95 to +40 for MP and -157 to + 54 for CP relative to the sgRNA transcription start site (Grdzelishvili et al.,2000). Initially, levels of sgRNA transcription were thought to be affected by their position relative to the 3'end of the genome (Culver et al.,1993). Later it was found that such a control on sgRNA production was exerted by the pseudoknot regions present in the 3'UTR (Shivaprasad et al.,1999).

CP sgRNA has a short leader and it has a 5' 7 methyl guanosine cap. MP sgRNA was reported to be uncapped (Hunter et al.,1983), but later work supports the theory that MP sgRNA is capped (Watanabe et al.,1984a, Grdzelishvili et al.,2000).

Other than the 5' scanning mechanism, initiation of translation in various RNA viruses takes place through binding of the 40S subunit of ribosomes to sequences

called internal ribosome entry sites (IRES) (Pelletier and Sonenberg,1988; Jang et al.,1988 ; Poole et al.,1995).

Among tobamoviruses, TMV-Cr has been shown to employ a cap independent translation mechanism for expressing both MP and CP genes from sgRNA (Skulachev et al.,1999; Ivanov et al.,1997,). The sequences upstream of both MP and CP genes were identified and shown to drive the expression of a 3' proximal GUS gene from bicistronic constructs *in vitro* and *in vivo* (Skulachev et al., 1999). Their experiments also identified a 228bp segment upstream of the TMV-U1 MP gene which promoted internal initiation *in vitro*. Although the TMV-U1 MP IRES was shown to be functional in bicistronic constructs, it was not functional in the context of its respective full length genomic RNA. It was proposed that extensive secondary structures in full length genomic RNA could interfere with the binding of ribosomes and as a result the expression of MP gene product via internal initiation would not be possible. Conformation of the TMV-Cr MP IRES segment was shown to be the reason for its IRES-like activity (Skulachev et al., 1999). Within the TMV-Cr CP IRES, polypurine (A) rich sequences (PARS) were identified to be responsible for recruiting the translation machinery and promoting internal initiation. Further to that, it was proposed that the PARS when found in other mRNAs such as *Nicotiana tabacum* heat shock factor mRNA, will show IRES-like activity (Dorokhov et al., 2002).

1.5 Functions of encoded proteins

1.5.1 RNA dependent RNA polymerase

The RdRp gene is a readthrough product of the first ORF and it is involved in viral replication. The suppression of the UAG codon happens approximately 10 % of

the time resulting in production of RdRp protein (Lewandowski and Dawson, 2000). The 183 kDa product has methyltransferase (MT), helicase (Hel) and polymerase domains. The 126 kDa product consists of MT and Hel like domains and has guanylyltransferase and MT activities *in vitro* (Dunigan and Zaitlin,1990; Merits et al., 1999). Template dependent RdRp extracts isolated from ToMV-infected plants have been found to contain both 126 and 183 kDa proteins as heterodimers (Osman and Buck, 1996; Watanabe et al.,1999). Reasons for the absence of defective RNA (dRNA) in TMV infection was investigated by Knapp et al 2001. That work led to an interesting discovery that 126 kDa protein could be involved in movement. This hypothesis is supported with the finding that TMV-U1 defective in helicase domain was deficient in cell to cell movement (Hirashima and Watanabe,2001). TMV-Ob strain which has a single nucleotide mutation in the 126 kDa protein is able to evade induction of a hypersensitive response (HR) in tobacco carrying the N gene (Padgett and Beachy 1993). Thus it is hypothesized that the 126 kDa region could induce an HR. A protein of 50 kDa within the helicase domain of the 126 kDa protein is sufficient to induce an HR even if the ATPase like activity of this p50 protein is abolished (Erickson et al., 1999).

1.5.2 Movement protein

In tobamoviruses, MP is typically expressed from a sgRNA. MP is expressed transiently in the early part of the infection cycle. The protein is expressed between 12-24 hrs post–inoculation and it undergoes little turnover (Lehto et al 1990b). MP can bind to single-stranded RNA and DNA in a non specific manner (Citovsky et al., 1990, 1992). Expression of TMV-MP in transgenic plants supports the cell to cell movement of movement defective *Cucumber mosaic virus* (CMV)

(Cooper et al 1996). *Barley stripe mosaic virus* (BSMV) with its MP replaced by TMV MP could infect host plants that are common to both parental viruses (Soloyev et al., 1996).

There are several stages in which the MP works to translocate the viral RNA from cell-to-cell. Firstly, it binds to RNA and forms a ribonucleoprotein complex that can enter the plasmodematal gate (Citovsky et al.,1990, 1992), interacts with the cytoskeleton to aid the transport of P30-RNA complexes to plasmodesmata (Heinlein et al.,1995), and increases the plasmodesmatal size exclusion limit (Wolf et al.,1989). Recent work also has shown that MP interacts with a protein kinase which in turn phosphorylates the MP, downregulating the ability to expand the plasmodesmata in tobacco plants (Waigmann et al.,2000). A host encoded enzyme, pectin methylesterases (PME) was implicated in the cell-to-cell movement of TMV (Dorokhov et al 1999; Chen et al 2000). Using a reverse genetics approach, it was demonstrated that PME interacts with MP and such an interaction is required for the systemic movement of TMV. In particular this interaction was essential for exit of TMV from the vasculature into adjoining non-vascular tissues (Chen and Citovsky, 2003). Within the MP there are conserved domains I and II (Saito et al., 1988) which play a role in stabilizing MP and its interaction with RNA (Citovsky et al., 1990) and host factors. The interaction between MP with microtubules and actin was shown by tagging MP with GFP and it was co-localized along the periphery of protoplasts forming a filamentous network (Mclean et al.,1995).

Transgenic tobacco plants expressing MP showed an increase in plasmodesmatal size exclusion limit allowing dextrans up to 10 kDa to pass through, compared to

wild type tobacco plants, which can traffic 0.75 to 1.0 kDa dextrans (Wolf et al., 1989). Such an increase in size exclusion limit could allow passage of MP-TMV RNA complexes (Citovsky et al., 1992).

Other than its interaction with the cytoskeleton MP was shown to be phosphorylated by a 38 kDa protein kinase at the carboxyl terminal serine-258, threonine-261 and serine-265 residues. When these amino acids were replaced with non-phosphorylatable alanines the MP was not phosphorylated but the ability to bind to p38 was retained (Citovsky et al.,1993). This shows that the receptor and binding activities are located in different domains.

In ToMV, phosphorylation of serine-37 or its presence is essential to preserve the stability of MP (kawakami et al., 1999). When (threonine) Thr–104 in the MP was replaced with neutral alanine, the ability of the virus to move cell-to-cell was not affected. When replaced with aspartic acid, the virus lost its ability to induce local lesions in Xanthi nc. Assuming that aspartic acid would mimic the effect of threonine phosphorylation, it was proposed that phosphorylation of Thr 104 could be a plant defense mechanism to prevent the spread of the virus (Karger et al.,2003). It is yet to be proven whether Thr-104 is phosphorylated *in vivo*.

The MP gene expression in transgenic *Nicotiana tabacum* Xanthi nc induces resistance by interfering with the wild type MP from increasing plasmodesmatal size exclusion limit (Deom et al.,1991). Two amino acid substitutions at the N terminal of ToMV MP allowed the virus to overcome Tm-2 resistance in tomato (Meshi et al.,1989). Added to that, two amino acid substitutions at the C-terminus of ToMV MP allowed the virus to overcome Tm-2^2 resistance in tomato (Weber et

al., 1993). Furthermore, MP could also be involved in host specificity (Fenczik et al.,1995).

1.5.3 Coat protein

The CP gene product is synthesized from a sgRNA (Guilley et al 1979). As discussed earlier, the CP sgRNA is synthesized from CP sg promoter (Lehto et al., 1990a; Grdzelishvili et al., 2000). The CP sgRNA has a 9 nucleotide leader sequence and a 7-methylguanosine (m^7G) cap (Guilley et al.,1979). CP is expressed late and at reaches extremely high levels 70 hr post infection (Siegel et al.,1978).

The structure of the TMV CP and its protein–protein and protein-RNA interactions have been determined in detail (Wittman-Liebold and Wittmann,1967; Champness et al., 1976; Bloomer et al., 1978; Goelet et al., 1982; Namba and Stubbs, 1986; Namba et al., 1989). Multiple alignment of CP amino acids of tobamoviruses show three putative RNA binding domains (Altschuh et al.,1987).

The CP of plant viruses perform multiple functions such as encapsidation and protection of viral nucleic acids, aiding long distance movement of viruses (Dawson et al.,1988), vector transmission (Gal-On A et al.,1992), symptom modulation and induction of host response (Saito et al.,1989). They have also been attractive targets for engineering virus resistance (Abel et al., 1986) and expression of proteins that are of therapeutic value (Hamamoto et al.,1993; Wu et al.,2003).

Tobamoviruses require CP for long distance movement but its role in cell-to-cell movement still remains to be answered. CGMMV is translocated through the phloem as viral particles, which suggests that CP is involved in long distance

movement (Simon-Buela and Garcia-Arenal, 1999). Using a TMV chimera expressing ORF3 of *Groundnut rosette virus* (GRV) and other umbraviruses in place of the CP gene, it was shown that the chimeric TMV could move systemically (Ryabov et al., 1999; Ryabov et al., 2001). Further investigation of this phenomenon showed that the ORF3 encoded protein formed stable ribonucleoprotein complexes which may help in long distance movement of chimeric TMV (Taliansky et al., 2003). This data suggests that formation of virions may not be a prerequisite for systemic movement in this case.

The evidence showing that formation of virions may not be the only requirement for establishing a systemic infection comes from studies involving TMV chimera expressing ORSV CP. This chimeric TMV replicated, formed virions and moved cell to cell but was incapable of moving out of the inoculated leaves of *Nicotiana tabacum* (Hilf and Dawson,1993). It is unclear whether it is ORSV CP that is failing to evade the *N. tabacum* defense response or the ORSV CP is unable to interact with host factors that are necessary for systemic movement.

Interaction of CP with MP was thought to be essential for cell to cell movement and systemic spread of viruses, but there was no direct evidence to demonstrate whether such an interaction exists. In earlier work done using transgenic tobacco plants expressing a mutant TMV CP, it was shown that these plants were able to confer stronger resistance to TMV infection as compared to transgenic tobacco plants expressing wild type CP (Bendahmane et al., 1997). The mutant CP had tryptophan in position 42 replaced by threonine. To understand the underlying mechanism, transgenic tobacco protoplasts expressing mutant CP were inoculated with wild type TMV RNA. After inoculation, the wild type TMV

accumulated low levels of MP. Inoculation of BY2 protoplasts expressing wild type CP with TMV showed elevated levels of MP as compared to non transgenic protoplasts (Bendahmane et al.,2002). These results showed that wild type CP regulated production and accumulation of MP. The mutant CP accumulated very low levels of MP and as a consequence, it interfered with cell-to-cell movement and spread of the virus.

CP has been implicated in various other roles such as host range determination, as was shown with a TMV chimera expressing the *Alfalfa mosaic virus* (AMV) CP (Spitsin et al.,1999), and symptom modulation (Banerjee et al.,1995). AMV CP or its sub-genomic transcript is required for AMV genomic RNAs to be infectious (Bol et al.,1971) and it was implicated in 'genome activation'. Later addition of CP was thought to protect the genomic RNA from ribonucleases rather than acting as a trigger in activating infection cycle. Such a possibility was ruled out by Houwing and Jaspars (2000), and it was later shown that the earlier messenger release model for genome activation could be the mechanism (Jaspars and Houwing, 2002).

Viral CP acts as an elicitor in the induction of a HR plant defense mechanism. This process has been understood at the molecular level with the discovery of plant resistance genes (*R*) and their interaction with avirulence gene product (*Avr*) elicitors, in this case, the viral CP. *Pepper mild mottle virus* (PMMoV) induces HR in *Capsicum frutescens* and other *Capsicum* species containing the *R* genes. PMMoV CP was implicated as the factor inducing HR in *Capsicum frutescens* (Berzal Herranz et al., 1995; de la Cruz et al.,1997).

1.6 The origin of assembly

Assembly of TMV particles *in vitro* starts with recognition of internal sequence of RNA by a 20S CP disk aggregate (Butler and Klug,1971; Okada and Ohno,1972). This internal recognition site in the viral RNA is called the OAS which is located at the 3' half of the genome. In TMV, the OAS is located in the MP gene, whereas in CGMMV-SH and SHMV, it is located in the CP gene. Based on the location of OAS, tobamoviruses have been sub-divided into subgroup I and subgroup II viruses, respectively (Fukuda et al.,1981). From the OAS the virion assembly proceeds towards the 5' end and then to the 3' end bidirectionally (Lebeurier et al.,1977; Bloomer and Butler, 1986).

The OAS of TMV contains a characteristic AAGAAGUCG (GXXGXXG) sequence exposed as a target for CP disk aggregate binding. The OAS contains three stable hairpin loops among which loop I has been shown to be crucial, by mutational analysis (Turner and Butler,1986; Turner et al.,1988). The OAS of TMV-vulgare RNA could encapsidate chimeric single-stranded RNA *in vitro* and *in vivo* (Sleat et al., 1986, 1988; Hwang et al., 1994). Folding and presentation of RNA plays an important role in facilitating selective protein-RNA recognition. Altering the RNA folding close to the apex of the loop 1 stem or shortening the stem reduces the rate of disk binding. Altering the RNA sequence making up the hairpin does not seem to affect the protein-RNA interaction as long as it can present the apex signal sequence in a special conformation (Turner et al., 1988).

1.7 Untranslated regions (UTR)

Tobamoviruses have a 5' and a 3' UTR, which play an important role in replication and translational enhancement. The 5' UTR is 60-80 nucleotides in length and it consists of a leader sequence (Ω) which enhances the expression of chimeric mRNAs (Gallie et al.,1987). The 5' end of the leader contains a $m^7G^5ppp^5G$ cap which stabilizes the mRNA and enhances the binding of 49S ribosomal subunits to the 5'end (Shatkin, 1976; Kozak, 1983). Removal of the 5' cap affects ribosomal binding and translation. The cap binding protein complex is involved in the melting of mRNA secondary structures, thereby facilitating ribosome binding and migration. The cap is required for infectivity when RNA is used for plant inoculation (Dawson et al., 1986). During translation of replicase, the coat protein is stripped off the virion which is also called 'co-translational disassembly'. At the 5'end the coat protein is loosely bound to the mRNA (Mundry et al., 1991) which facilitates this process. The Ω fragment translates free RNAs and neither the viral sequence nor the viral protein is required for its activity (Gallie et al.,1987). The 5' leader sequence contains AUU binding sites for a second ribosome upstream of the 126kDa start codon (Gallie et al.,1987), providing a putative second in-frame translation initiation site (Tyc et al.,1984; Schmitz et al.,1996). Mutation of AUU to CUU had no effect on the enhancement effect given by the full length leader sequence (Gallie et al., 1988).Three copies of an eight base direct repeat and a (CAA_n) region represents the two motifs found in leaders of several TMV strains. Two copies of CAA_n have been shown to be sufficient to enhance expression. Along with the direct repeats then provide a moderate enhancement effect (Gallie and Walbot, 1992), which shows that these two motifs are functionally redundant.

The 5’ UTR can act alone as a translational enhancer and it does not require the 3’UTR for performing this function (Gallie et al., 1987).

The role of 5’UTR in replication has been shown by Takamatsu et al., (1991). In ToMV it has been shown that deletion of nucleotides 2-8 in the leader sequence abolished any detectable viral replication. Interestingly, this mutant RNA is able to drive the expression of 130 kDa protein in the rabbit reticulocyte system.

Instead of a poly(A) tail at the 3’ end, tobamoviruses have a 3’ UTR-like structure which has several roles to play in viral infection cycle. TMV 3’ UTR performs functions similar to the poly(A) tail and can substitute for poly(A) tails present in plant and animal mRNAs (Gallie and Walbot,1990). Typically a TMV UTR is 204 base pairs in length with a tRNA-like structure at the 3’ extremity and a stretch of 3 consecutive pseudoknots upstream of it (van Belkum et al.,1985). All tobamoviral 3’ UTRs were found to contain pseudoknots and tRNA-like structures (van Belkum et al., 1985; Pleij et al.,1987; Garcia-Arenal, 1988; Isomura et al.,1991). In detail, the 3'-UTR can be folded into a tRNA-like structure consisting of a 3' pseudoknotted domain (D1), which acts as a tRNA acceptor branch ending in an unpaired CCA sequence and a domain (D2), which looks similar to a tRNA anticodon branch (Felden et al., 1994; van Belkum et al., 1985). Upstream of the tRNA-like structure is domain D3, containing three pseudoknots, each of which contains two double-helical segments. Domains D1, D2, and D3 are connected by a central pseudoknotted structure, C.

The tRNA-like structures are excellent substrates for aminoacylation (Mans et al., 1991) which can be catalyzed by specific aminoacyl-tRNA synthetases. The tRNA

-like structures in tobamoviruses can accept histidine in general, except for SHMV which is valinated (Oberg and Philipson, 1972).

Several functions have been proposed for plant viral 3'UTRs in general and tRNA -like structures in particular (Haenni et al., 1982; Florentz et al., 1984). To study the functions of the 3' UTR, several ToMV mutants with deletions in the 3' UTR were tested in tobacco plant and protoplast systems. Deletion of double-helical segments II to V in the central pseudoknot region D3 resulted in a reduction in viral replication, associated with loss of symptom development. Double-helical segment I, upstream of the tRNA-like structure, is indispensable for viral replication and double-helical segment VI is not essential for viral replication (Takamatsu at al., 1990).

A further detailed analysis of the ToMV 3' UTR regions using template dependent RdRP extracts has revealed that several double-helical regions, which form the pseudoknot and stem-loop structures in domains D1, D2, and D3 and the central core, C, are necessary for high template efficiency. Domain D2 and central core C can bind to RNA polymerase with high affinity, whereas domains D1 and D3 showed comparatively lesser affinity towards binding RNA polymerase. Mutation of the 3' terminal CCCA identified the 3'-terminal CA as crucial for minus-strand synthesis. Maximum transcriptional efficiencies were achieved when the 3' terminal sequence ended with CCCA or GGCA (Osman et al., 2000).

Pseudoknots are involved in translational regulation (Tang and Draper,1989) and also have been implicated in ribosomal frameshifting in coronaviruses (Brierley et al.,1989). TMV has 3 pseudoknots immediately downstream of the CP gene. The

TMV 3’ UTR when fused with a chimeric mRNA enhanced the expression levels to several fold and increased the stability of chimeric mRNA (Gallie et al., 1991). Enhanced expression of 5'-capped RNAs has been attributed to improved translational efficiency, which is due the synergistic interaction between 3'-UTR and 5'-cap and to a smaller extent due to increased mRNA stability (Leathers et al., 1993; Gallie and Kobayashi,1994).Translational enhancement by the TMV 3' UTR is primarily due to the pseudoknot structure that is upstream of the tRNA -like structure. The tRNA-like structure has been shown to enhance mRNA stability (Gallie and Walbot, 1990).

1.8 Mechanisms of translation initiation

Translation of proteins from most eukaryotic mRNAs involves the binding of translation initiation factors to the 5’ cap structure, followed by scanning of mRNA by 40S ribosomal subunit (pre-initiation complex) until it reaches an AUG codon. Subsequently, the larger 60S ribosomal subunit associates itself with the preinitiation complex to form the 80S ribosome, which deciphers the genetic code and translates it into a protein until it reaches the termination codon (Kozak, 1986; Kozak,1989).

A favourable sequence context for the ribosomes to initiate protein synthesis requires the presence of a purine (R) at -3 position and a G residue at position +4 (GCCRCCaugG) (Kozak, 2002). If the first AUG codon resides in a weaker context lacking R in -3 position and G in +4 position, the 40S subunits may continue to scan further downstream and initiate when it encounters an AUG

codon, which is also known as the leaky scanning mechanism. Such a mechanism has been shown to operate in *Rice tungro virus* (Futterer et al., 1997). Termination-reinitiation is another mechanism by which ribosomes gain access to downstream AUGs after initiating in upstream ORFs (upORFs). The size of upORFs seems to be a limiting step for reinitiation to be operational. The size of upORFs could be from 10-12 codons (Kozak, 2001) and on some occasions re-initiation occurs following a 24 codon upORF (Luukkonen et al.,1995). Other variations in translation initiation include the utilization of a non-AUG codon (Sasaki and Nakashima, 2000) and a 'shunting' mechanism (Futterer et al., 1993; Yueh and Schneider, 2000).

1.8.1 Internal ribosome entry sites

Translation initiation of eukaryotic mRNAs predominantly occurs via a 5' cap dependent mechanism. Various cellular and viral RNAs utilize a non-orthodox mechanism for expression of internal genes, namely the IRES dependent translation. Cellular mRNAs express a wide range of proteins such as homeotic genes (Vagner et al., 1995a), oncogenes (Johannes et al., 1999), transcription factors (Chappell et al., 2000) and translation initiation factors (Gan and Rhoads, 1996) by IRES driven translation.

An IRES element is found in the picornaviral 5' UTR (Jang et al.,1988; Pelletier and Sonenberg,1988), which confers internal initiation. IRES elements are also reported in *Hepatitis C virus* (HCV) (Tsukiyama-Kohara et al.,1992), *Moloney murine leukemia virus* (M-MuLV) (Vagner et al.,1995b) and *Infectious bronchitis virus* (IBV) (Liu and Inglis, 1992).

Among plant viruses, *Tobacco etch virus* (TEV) (Levis and Astier-Manifacier,1993), TMV-Cr (Ivanov et al., 1997; Skulachev et al.,1999), *Cowpea mosaic virus* (CPMV) (Thomas et al.,1991) and *Hibiscus chlorotic ringspot virus* (HCRSV) (Koh et al., 2003) have been reported to possess IRES elements.

IRES elements have the ability to promote translation initiation independent of initiation factors that are required for cap dependent translation (Pestova et al., 1998). During cellular stresses and viral infections, the initiation factor eIF4F is modified which results in the inhibition of protein synthesis. Under such conditions, cap-dependent initiation is severely compromised (Thompson and Sarnow, 2000). IRES-dependent translation will still be functional and confers an advantage to the invading virus over the host, which relies heavily on cap-dependent translation. Apart from that, internal initiation helps in translation of mRNAs that are constrained by numerous upstream AUGs or RNA secondary structures in their 5' leader sequence (Chappell et al., 2001; Le Quesne et al., 2001).

Viral IRESs present in the 5' UTR could be folded into complex secondary structures and contain multiple conserved AUGs. Secondary structural elements within the IRESs are essential for activity and they form part of double-stranded regions (Jang and Wimmer,1990; Hoffman and Palmenberg,1996) or sequences located in apical or internal loops (Lopez de Quinto and Martinez-Salas, 2000). The *Foot and mouth disease virus* (FMDV) IRES contains a conserved GNRA tetraloop which is involved in long distance RNA interactions (Lopez de Quinto and Martinez-Salas, 1997).

In a plant virus, the TMV-Cr CP IRES contains PARS elements that are responsible for recruiting the translation machinery and promoting internal initiation (Dorokhov et al., 2002). The MP IRES segment of TMV-Cr could promote internal initiation in bicistronic constructs by virtue of its secondary structure (Skulachev et al., 1999).

In contrast to viral IRESs, short non contiguous sequences within cellular IRES elements (Stoneley et al.,1998; Chappell et al., 2000) and short sequences that are complementary to the 3' end of 18S rRNA have been found to be involved in IRES activity (Hu et al.,1999; Tranque et al.,1998). Rbm3 mRNA has a 720 nucleotide (nt) IRES segment which could be dissected into small regions, namely 22 nt IRES region, a 10 nt enhancer and two inhibitory sequences showing the modular nature of a cellular IRES (Chapppell and Mauro, 2003).

Using the yeast system, it was demonstrated that synthetic short oligonucleotides could enhance the expression of *HIS3* in bicistronic constructs (Zhou et al., 2003). Another finding in this study was that these short oligonucleotides could base pair with relatively unstructured regions of the 18S rRNA thereby binding to the 40S ribosomal subunits and initiate translation.

The extent of diversity in IRES sequences suggests that internal initiation of translation can function by different mechanisms. The requirement of initiation factors and other trans-acting factors by different IRESs varies to a great extent. The *Cricket paralysis virus* IRES, requires only 40S ribosomal subunits for translation initiation (Pestova and Hellen, 2003) whereas *Hepatitis A virus* (HAV) IRES requires all of the initiation factors that are involved in promoting initiation of translation of capped mRNAs by the scanning ribosome mechanism (Ali et al.,

2001). IRES elements can bind to initiation factors and various IRES trans-acting factors (ITAFs), which facilitates the recruiting of ribosomes for translation (Kolupaeva et al.,1996; Pilipenko et al., 2000). Interaction of translation initiation factor eIF4G with FMDV and *Encephalomyocarditis virus* (EMCV) IRESs has been shown to be an essential step in recruitment of translational machinery *in vivo* (Lopez de Quinto and Martinez-Salas, 2000; Kolupaeva et al.,1998).

Among ITAFs the pyrimidine tract binding protein (PTB) has been found to interact with IRES elements of all picornaviruses (Kolupaeva et al.,1996; Pilipenko etal.,2000). PTB and ITAF (45) promote binding of eIF4G/4A to FMDV IRES and are thought to act as RNA chaperones that control the functional state of IRES (Pilipenko et al., 2000). Poly(rC) binding protein promotes the stability of poliovirus mRNA by binding to 5' cloverleaf RNA. Addition of ribohomopoly (C) RNA competitor rendered wild-type poliovirus mRNA unstable in HeLa S10 translation-replication reactions (Murray et al., 2001). Owing to the fact that the 5' cloverleaf RNA sequences are conserved among various picornaviruses, it is suggested that interactions between Poly (rC) binding protein and 5' cloverleaf RNA might promote the stability of polioviruses, coxsackieviruses, echoviruses, and rhinoviruses.

Using IRESs it has become possible to select transfected cell lines expressing genes of interest (Gaines et al.,1999) and they also are being used for synthesizing several proteins of interest from a multicistronic mRNA which will find application in gene therapy (Wen et al., 2001).

Hibiscus Virus-S (HVS)

1.9 General characteristics

The genus tobamovirus consists of 13 definitive species and four tentative species (Brunt et al., 1996). The natural host range includes solanaceous crops, orchids, brassicas and cucurbits. Previously, the occurrence of a rod-shaped virus in *Hibiscus rosasinensis* L. has been reported. Based on the symptoms and virion morphology, the rod-shaped virus was named *Hibiscus yellow mosaic tobamovirus* (Kashiwazaki et al., 1982). It has been placed as a tentative species of the genus tobamovirus (Brunt et al., 1996). While isolating HCRSV from hibiscus plants from a garden in the National University of Singapore, HVS co-purified along with HCRSV. Hibiscus leaves were collected from these plants and HVS was purified by 5 successive single lesion transfers in *Chenopodium quinoa* Willd. The virus was subsequently maintained in kenaf (*Hibiscus cannabinus* L.) plants.

Transmission electron microscopic observation of unfractionated purified HVS showed rigid rod shaped particles of varying sizes. Single-stranded RNA profile revealed the presence of a 6.5 kb genomic RNA band, together with a 1.5 kb and a 0.7 kb sub-genomic RNAs of MP and CP, respectively.

1.10 Objectives of this research project

Owing to the characteristics such as broad host range, compact genome, rapid replication rates, tobamoviruses are one of the most studied plant virus groups. With the availability of biologically active cDNA clones of TMV, it had become

possible to express proteins of therapeutic value. TMV based vectors are also being exploited for study of plant gene functions by virus induced gene silencing (VIGS). At least one tobamovirus has been shown to express internally located genes by cap independent translation.

Here we describe a novel tobamovirus infecting hibiscus and an attempt was made to synthesize biologically active cDNA clones of this virus. Implications of other interesting features of this virus, such as cap-independent translation of 3' proximal ORFs via and a unique 3' UTR, are discussed.

The aims of this project was

1. To determine the complete nucleotide sequence of HVS
2. Analyze the sequence and assign the virus to a particular taxonomic group
3. To synthesize biologically active cDNA clones of HVS
4. To study translation mechanisms involved in the expression of 3' proximal ORFs
5. To isolate and characterize regions involved in internal initiation
6. To establish an efficient *in vivo* translation system for testing HVS IRESs

CHAPTER 2

MATERIALS AND METHODS

2.1 Cloning and sequencing of HVS

2.1.1 Bacterial strains and plasmids

Escherichia coli cell strains used for propagating plasmid clones were XL1-Blue XL10-Gold^{TM,} JM 109 and DH 5α. *Agrobacterium tumifaciens*, strain EHA 105 was used for transient GUS assays. Bacterial glycerol stocks were prepared by adding 150µl of 100% glycerol to 850 µl of liquid culture, frozen in liquid nitrogen and stored at -80^{o}C. Fresh working stock plates were made from frozen glycerol stocks when required and stored at 4^{o}C.

2.1.2 Cloning vectors

pBluescript II SK(+) (Stratagene,La Jolla, CA), pCR-Script^{TM} SK (+), pGEM $^{®}$-TEasy (Promega Corp., Madison, WI), pSPORT I (Gibco BRL, USA), pCAMBIA 1300 and 1301 were used as cloning vectors.

2.1.3 Bicistronic constructs

hGFP-I-GUS, provided by ICON Genetics, GmbH (06120 Halle, Germany)

pFF19G (kindly provided by Dr.M.V.Skulachev, Moscow State University, Russia).

2.1.4 Media

LB medium: 1% Bacto$^{®}$ - tryptone, 0.5% Bacto$^{®}$ - yeast extract, 0.5% NaCl, pH 7.5

LB agar: LB medium with 1.5% Bacto$^{®}$ - agar, pH 7.5

SOC medium: 2% Bacto$^{®}$ - tryptone, 0.5% Bacto$^{®}$ - yeast extract, 10mM NaCl, 2.5mM KCl. The medium was autoclaved at 121^{o}C for 20 min and cooled at room

temperature. Filter sterilized $MgCl_2$ and $MgSO_4$ and glucose were added to final concentrations of 10 mM, 15 mM and 20 mM, respectively.

2.1.5 HVS purification

Virus was isolated from hibiscus leaves collected from a garden in the National University of Singapore, Republic of Singapore, by five successive single lesion transfers in *Chenopodium quinoa* L. The virus was subsequently maintained in kenaf (*Hibiscus cannabinus* L.) plants. Mechanical inoculation was carried out by grinding the leaves in 0.2 M borate buffer (pH 8), with a mortar and pestle. The extract was inoculated onto leaves dusted with Carborandum.

Virus was purified from fresh kenaf leaves by homogenizing the tissues in 3 volumes of 0.2 M borate buffer (pH 8), containing 0.2 M diethyl dithiocarbamic acid (DIECA). The homogenate was centrifuged at 12,000 x g for 15 min. The supernatant was clarified with an equal volume of butanol/chloroform (1:1), filtered and centrifuged at 35,000 x g for 2.5 hr. The pellet was resuspended overnight in 0.2 M borate buffer (pH 8), and subjected to centrifugation at 9,000 x g for 15 min. The supernatant was dialysed against water for 9 hr and centrifuged in 30% CsCl at 40,000 x g for 16 hr at 20^{o}C. The centrifugation steps and overnight resuspension of the pellet were carried out at 4^{o}C. The virus band was collected and the yield was quantified spectrophotometrically using an extinction coefficient of 3.3.

2.1.6 Antiserum production

Antiserum against the virus was prepared by injecting rabbits intramuscularly with a suspension of 1 mg purified virus emulsified with an equal volume of Freund's Incomplete Adjuvant (Sigma, USA). Subsequently, three boosters of 1 mg of

purified virus mixed with Freund's Incomplete Adjuvant was given each time. Antiserum was purified from blood, stored with 0.5% sodium azide at 4^{o}C and used for western blot and enzyme linked immunosorbent assay (ELISA).

2.1.7 Ultra-thin sectioning and staining

N. benthamiana leaves infected with HVS were cut into 1mm strips and treated with fixative solution (0.2M cacodylate 10 ml, 10% paraformaldehyde 4ml, 25% glutaraldehyde 2ml, distilled water 2ml) for 1hr. The specimen was rinsed four times in 0.1M cacodylate buffer pH 6.8 for 15 min. The tissue was treated with 0.1M cacodylate and post-fixed with an equal volume of 2% osmium tetroxide (0.25g OsO_4 and 12.5 ml de-ionized water) for 2 hr. The tissues were dehydrated in a series of ethanol solutions, increasing in concentration (30%, 50%, 70%, 85%, 95%, 100%) for 15 min. The tissue was left in 100% ethanol for 30 min, twice. The tissue was treated with 1:1 propylene oxide and absolute alcohol for 30 min. Then fresh propylene oxide was added to the samples and was left for 20 min. Then the samples were soaked in 1:1 propylene oxide and resin for 30 min. Subsequently, the samples were treated with 4:1 resin and propylene oxide overnight or 1 hr. The samples were drained and infiltrated with full resin for 1 hr, twice.

Samples were embedded in full resin, dried for 20 hr and transferred to a dessicator. Ultra-thin sections of embedded tissue were prepared with a Reichert-Jung Ultracut E ultramicrotome.

Ultrathin sections were mounted on copper grids (200 mesh) pre-treated with nitric acid. The sections were stained with 0.2% uranyl acetate for 30 min and rinsed three times with distilled water. The sections were then stained with lead citrate for 10 min and rinsed with 0.02 N NaOH followed by three rinses in distilled

water. Sections were examined using a Philips CM -10 transmission electron microscope.

2.1.8 Electron microscopy of HVS particles

Purified HVS virions were deposited onto formvar coated copper grids and stained with 2% uranyl acetate. The specimen was examined using a Philips CM 10 transmission electron microscope. Virion lengths were measured using the step-by-step procedure for distance measuring, according to the Philips CM-10 user's manual. End-to-end aggregates or multimers of 34 nm particles were counted as monomers. Lengths of 500 virions were measured and a virion length histogram was plotted.

2.1.9 Determination of coat protein molecular weight using mass spectrometry

The viral CP was purified using the acetic acid degradation method (Frankel-Conrat, 1957). To one volume of purified virus suspension, two volumes of glacial acetic acid were added and incubated on ice for 1hr. Insoluble RNA was removed by a centrifugation step at 14,000 x g for 30 min at 4^{o}C. The molecular weight of CP was determined using a Perkin-Elmer Sciex API 300 triple quadrupole instrument equipped with ionspray interface. The ionspray voltage was set at 4,600 V and the orifice voltage at 30 V. Nitrogen was used as a curtain gas with a flow rate of 0.6 L/min, while compressed air was used as nebulizer gas. The samples were infused at a flow rate of 5 μl/min.

2.1.10 Isolation of viral RNA

To the purified viral suspension, equal volume of water saturated phenol pH 4.0, was added and vortexed vigorously for 2 min. The sample was centrifuged at

14000 rpm for 3 min. The upper aqueous layer was collected and re-extracted twice with an equal volume of water saturated phenol and chloroform. To the upper aqueous layer, 1/10 of the volume of 3M Sodium acetate pH 5.2 and 2.5 volumes of 95% ethanol or add 2 volumes of 100% ethanol were added. RNA was precipitated by incubating at -20^{o}C overnight. Alternatively, the reaction can be incubated at -20^{o}C for 1 hour or at -80^{o}C for 30 min. The sample was centrifuged at 14,000 rpm for 20 min to precipitate the RNA. The resultant RNA pellet was washed with 200 µl of 70% ethanol and vacuum dried. The pellet was re-suspended in nuclease free water and the concentration was measured at A 260/280 nm.

2.1.11 Polyadenylation of RNA

Purified HVS RNA (5 µg) was polyadenylated in 50 mM Tris-HCl, pH7.9, 10 mM $MgCl_2$, 2.5 mM $MnCl_2$, 0.25M NaCl , 0.25 mM ATP, 0.5 mg/ml BSA and 4 units of poly(A) polymerase (Gibco- BRL) at 37^{o}C for 1.5 hr. The polyadenylated RNA was purified by phenol-chloroform extraction and ethanol precipitation.

2.1.12 cDNA synthesis

Purified HVS RNA was used as template to synthesize cDNA using degenerate primers spanning methyltransferase-helicase (SYVHCC) and coat protein region (FQTQQA) (Table 2.1). The resultant product was cloned into pGEM®-T Easy vector and sequenced using vector specific oligonucleotide primers.

Subsequent cDNA synthesis and cloning was done by Gubler and Hoffman (1983) based Gibco-BRL cDNA synthesis kit using oligo d(T) or gene specific primers as per manufacturer's instructions. The resultant double stranded cDNA was cloned into pBluescript ™ II SK (+) (Stratagene, La Jolla, CA).

2.1.12.1 Purification of PCR fragments

PCR fragments were purified directly if a single specific product was obtained using the Qiaquick gel extraction kit. In the event of several non-specific products being present in the PCR reaction, the fragments were electrophoresed on 1% agarose gels. The desired fragment was excised and purified using QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions.

2.1.12.2 End-filling of DNA fragments

Following restriction digestion, the DNA was purified and re-suspended in TE buffer pH 8.0. Typically a 50 µl reaction contained 2 µg of restricted plasmid vector, 5 µl of 10 X Klenow buffer (Promega), 40 µm of dNTPs and 1µl (5U/µl) of Klenow fragment. The reaction was incubated at 25^{o}C for 30 min. Subsequently, the end filled DNA was recovered by phenol/chloroform extraction and ethanol precipitation.

2.1.12.3 Dephosphorylation of vector

Digested DNA with compatible ends were dephosphorylated using 10U of calf intestinal alkaline phosphatase (New England Biolabs) in 10 X CIAP buffer. The reaction was incubated at 37^{o}C for 1 hr. DNA was purified by phenol/chloroform extraction and ethanol precipitation.

2.1.12.4 Site-directed mutagenesis

PCR based site-directed mutagenesis was carried out to create single nucleotide mutations. This method involved the usage of two pairs of primers. The pair of primers that flank the mutation sites were referred to as primers A (forward primer)

and B (reverse primer). The second pair of primers that carry the mutation were referred to a primers C (forward primer) and D (reverse primer). First round of PCR was performed using primer A/primer D and primer B/primer C, respectively. The resultant PCR products were purified, added together and a fragment that is a fusion of these two PCR products was obtained using primer A/primer B.

2.1.12.5 Bacterial competent cell preparation and transformation

Escherichia coli competent cells were prepared according to Sambrook et al., (1989). A single bacterial colony was transferred to 2 ml LB medium and grown overnight in a shaker-incubator at 37^{o}C, 250 rpm. Fresh LB medium (100 ml) was inoculated with 2ml overnight culture and grown at 37^{o}C, 250 rpm, till the cell concentration reaches 0.3-0.5 $A_{600nm.}$ All subsequent manipulations were carried out at 4^{o}C. Cells were centrifuged at 2300 rpm for 5 min at 4^{o}C. The cell pellet was washed with 10 mm $CaCl_2$ twice. The cell pellet was re-suspended in 2 ml of 50 mm $CaCl_2$ and 15% glycerol. Cells were aliquoted and frozen in liquid nitrogen and stored at -80^{o}C.

The frozen aliquots were thawed on ice and 2µl of ligation mixture was added. The cells were subjected to heat shock (42^{o}C 1 min) and quick chilled on ice for 2 min. The cells were grown in 450 µl of LB medium for 1 hr, then plated onto LB agar containing appropriate selection antibiotic.

2.1.12.6 Determination of the 5' end sequence of viral RNA

First strand cDNA was synthesized from viral RNA using gene specific primer (GSP1), HVS Hel 1 primer (20 picomoles) (Table 2.1) and Superscript^{TM} II, RNase H $^{-}$ reverse transcriptase at 42^{0}C provided with the 5'RACE kit (Gibco–BRL). The mRNA template was removed by digestion with RNase mix at 37^{o}C.

Unincorporated dNTPs and primer were removed by passing the reaction mix through GlassMAX™ spin column. A homopolymeric(C) tail was added to the single-stranded cDNA using terminal transferase. The first round of PCR was done using GSP1 and the abridged anchor primer (AAP). PCR product from this first round of amplification was used as a template for subsequent round of nested PCR using HVS 5' RACE 2 (Table 2.1) and the abridged Universal Anchor Primer (AUAP). The PCR product was gel purified using the Qiaquick gel extraction kit and cloned into pGEM®-T Easy. The 5' extremity of the viral RNA was determined by sequencing the ligated junctions using vector specific primers. At least 10 independent clones were sequenced.

2.1.12.7 Determination of the 3' end sequence of viral RNA

The presence of an internal poly (A) tract in the viral 3' UTR, hampered the determination of 3' end nucleotides using 3' RACE kits. Synthesis of first strand cDNA using oligo d(T) resulted in products that terminated with an A-tail after the CP gene.

Degenerate primer Tobmv-3' end (Table 2.1) was designed to obtain the 3' end sequence (Adkins et al., 2003) [5'- TGG GCC SCW ACC SGS GGT TAG- 3'] and used with HVS CP 1 primer [nt 5800-5822] to carry out RT-PCR. Thermal cycling was carried out in a Perkin-Elmer GeneAmp® PCR system 9700 (Perkin-Elmer Foster City, CA) using the following cycling parameters.

Rapid ramping	50°C		30 min	
Denaturing RT- enzyme	94°C		3 min	
Denaturing cDNA	94°C		10 sec	30 cycles
Primer annealing	49°C		1 min	
Extension	68 °C	(increase 5sec/cycle)	1 min	
Extension	68 °C		7 min	

The PCR product was cloned into pGEM®-T Easy and sequenced using vector specific oligonucleotide primers.

2.2 Construction of full length HVS cDNA clones

A near full-length cDNA clone of HVS was obtained by amplifying the HVS viral RNA by using primers complementary to the 5' and the 3' ends by a two step RT-PCR. The 5' end primer Sma I-T7 HVS (+) contained a Sma I restriction site, T7 promoter sequence and a stretch of 19 nt of HVS 5'end sequence. The 3' end primer, 3' Not I HVS (-) has a Not I site and 18 nt region, complementary to the 3' end of CP gene (nt 6291-6274). The first strand cDNA was synthesized with Superscript II reverse transcriptase (Gibco-BRL) according to the manufacturer's instructions. First strand cDNA derived from viral RNA was used as template for PCR. Near-full length product (6291 nt) was obtained by using Expand long template DNA polymerase (Roche Diagnostics GmBH, Germany) in a Perkin-Elmer GeneAmp® PCR system 9700 (Perkin-Elmer Foster City, CA) with the following cycling parameters (Yoon et al., 2002).

Denaturing cDNA	94°C	2min	
Denaturing cDNA	94°C	20 sec	5 cycles
Primer annealing	55°C	30 sec	
Extension	68 °C	7 min	
Denaturing cDNA	94°C	20 sec	10 cycles
Primer annealing	57°C	30 sec	
Extension	68 °C	8 min	

Denaturing cDNA	94°C	20 sec	10 cycles
Primer annealing	59°C	30 sec	
Extension	68 °C	9 min	
Extension	68 °C	7 min	

The PCR product was digested with Sma I and Not I sites and cloned into Sma I/Not I digested pSPORTI. The resultant plasmid pHVS was linearized with Not I and *in vitro* transcribed using Ambion mMessage mMachine® kit.

Table 2.1

Oligonucleotide primers used for determination of HVS nucleotide sequence

Oligonucleotide	SEQUENCE 5'⟶ 3'	Position
Tobmv-domain SYVHCC	TC(I)TA(ct)GT(I)CA(ct)TG(ct)TG(ct)	Degenerate
Tobmv-domain FQTQQA	(I)GC(tc)TG(ct)TG(I)GT(tc)TG(ag)AA	Degenerate
Tobmv-domain DDATVA	(I)GC(I)AC(I)GT(I)GC(ga)TC(ga)TC	Degenerate
HVS 5'RACE 2	CTTGTAATTCTCGTTTCTGGC	466-487
HVS 5'RACE 1	GCGGTTCACTACGCCAGA	511-528
HVS Hel 1	CCTCGGAGAAATGGAGTGC	775-794
HVS Hel 2	AGGGTGGATTCACCATCG	879-897
HVS Hel 3	GCTGAAGAACATTACTGTCC	1491-1511
HVS Hel 4	CGAACCGACAAGACCCCCTA	1559-1579
HVS Hel 5	GAAAGCTCTGGCACCTCCAC	1648-1668
HVS Hel 6	GGACGTCTGGAGCTATCTGG	1728-1748
HVS Hel 7	TGGTGGGAAACCCGAGC	1920-1937
HVS Hel 8	GCGACAGCGGCCTTCATC	1954-1971
HVS Hel 9	CTGGTGTCGACCATGTGC	2166-2184
HVS Hel 10	GCGGCTGCCTGTTTTCC	2686-2703
HVS Hel 11	TTACACCATTACCGGGCAAG	3137-3157
HVS RdRP 1	GTCGATTATCCCAGCCA	3806-3823
HVS RdRp 2	GTCTATACCGAGACGTTCCC	4349-4369
HVS RdRP 3	AACCGATGTACAAGCCAATGC	4606-4626
HVS MP 1	AATTCCTCGGCGTGTTCCC	4779-4797
HVS MP 2	TACCAGAAATTACTATAGCTA	5190-5211
HVS CP 1	ATGCCTTACCTTAATTTGACACC	5800-5822
HVS –3' UTR(+)	GCGTGTGGATGATGCTACCG	6137-6157
Tobmv -3' end	TGG GCC SCW ACC SGS GGT TAG	Degenerate 3' end
SmaI T7-HVS (+)	TCCCCCGGGTAATACGACTCAC TATAGTATGTTTTTAGTTTGAAC	1-19
3'Not I HVS(-)	AAGGAAAAAAGCGGCCGCTTAC GTTGTAGTAGACGT	6274-6291
M13 forward	GTAAAACGACGGCCAGT	Vector
M13 reverse	GGAAACAGCTATGACCATG	Vector

2.3 Construction of HVS clones with hybrid UTR

HVS clones with UTRs of related tobamoviruses TMV-U1, ORSV-S1, SHMV and a clone with a poly(A) tail were constructed and tested for infectivity. A common strategy was utilized for generating all of these hybrid constructs. The first step involved a PCR for fusing the full-length CP region of HVS and the 3' UTR of respective tobamovirus. Once the fusion product was obtained, it was digested with an Spe I site present in the HVS CP sequence and a Not I site engineered in the 3' UTR reverse primer. The PCR product was cloned into Spe I / Not I digested pHVS plasmid.

2.4 Nucleotide sequencing

Nucleotide sequence was determined by automated sequencing using the ABI PRISM^{TM} BigDye^{TM} Terminator Cycle Sequencing kit (Perkin Elmer). Fluorescence-based dideoxy sequencing reactions were performed according to the manufacturer's instructions. One µg of DNA was added to 3.2 pmol of vector specific primer, 8 µl of terminator reaction mix and sterile water to 20 µl. The cycling conditions were 96^{o}C for 10 s, 50 ^{o}C for 5s, 65^{o}C for 4 min for 25 cycles. The extension products were precipitated with 50 ml of 95% ethanol and 2 ml 3M NaOAc. The sample was centrifuged at maximum speed and resulting pellet was washed with 70% ethanol and vacuum dried. The pellet was resuspended and sequenced using an automated fluorescent DNA sequencer (ABI PRISM 377, Perkin Elmer, USA) according to the manufacturer's instructions. DNA sequences were determined on both strands of the cDNA clones. Ambiguities in the

sequences were removed by re-sequencing using sequence specific primers (Table 2.1).

2.4.1 Phylogenetic analysis

The complete nucleotide sequence of HVS was compared with the following nucleotide sequences retrieved from the GenBank: *Tobacco mosaic virus* (TMV, accession no. V01409), *Tomato mosaic virus* (ToMV, accession no. X02144, *Tobacco mosaic virus –strain Obuda* (TMV-Ob, accession no. D13438), *Tobacco mild green mosaic virus* (TMGMV, accession no. M34077), *Pepper mild mottle virus*-strain S (PMMV-S, accession no. M81413), *Turnip vein clearing virus* (TVCV, accession no. U03387), *Odontoglossum ringspot virus* – Singapore Isolate (ORSV-S1, accession no. U34586), *Kyuri green mottle mosaic virus* (KGMMV-Y AB015145), *Cucumber green mottle mosaic virus*- strain SH (CGMMV-SH, accession no. D12505), *Sunn-hemp mosaic virus* (SHMV, accession nos. J02413, U47034), *Florida hibiscus virus* (FHV, accession no. AY250831), *Tobacco rattle virus* (TRV, accession nos. AF034622, AF034621), and *Barley stripe mosaic virus* (BSMV, accession no. NC003469).

Pairwise sequence comparison between different tobamoviruses was carried out using the OLD DISTANCES program of the Genetic Computer Group (GCG) (Devereux et al., 1984). Sequence identities of nt and aa of HVS with eleven other tobamoviruses were compared.

CLUSTAL X (Thompson et al., 1997) was used to perform multiple sequence alignment excluding gaps and to generate phylip format trees. One thousand bootstrapped data sets were generated using NJ bootstrap option. Phylograms

were visualized using the program TREEVIEW (Page, 1996). *Tobacco rattle virus* (TRV) was used as outgroup for rooting the trees.

2.5 RNA gel electrophoresis

Purified viral RNA was electrophoresed on 1.2% agarose formaldehyde denaturing gels and visualized by ethidium bromide staining using a modified protocol (Lehrach et al., 1977). A 1.2% (w/v) gel was prepared by melting 0.36 g of agarose in 20 ml of DEPC-treated water. Upon cooling 6 ml of 5X MOPS buffer (0.1 M (N-morphalino) propanesulfonic acid, 40mM NaOAc, 5mM EDTA pH 7.0) and 5.5 ml of 37% formaldehyde were added. RNA samples (4.5 µl) were denatured in 2 µl 5 x MOPS, 3.5 µl 37% formaldehyde, 10 µl de-ionized formamide at 65^{o}C for 15 minutes. The samples were mixed with 1 µl of loading dye and electrophoresed in 1X MOPS buffer at 50 volts for 2 hr.

2.5.1 Northern blot analysis

After electrophoresis, the gel was rinsed with DEPC-treated water and soaked in 0.05N NaOH for 20 minutes. This step partially hydrolyzes RNA and helps in efficient transfer of bigger RNA fragments. Subsequently, the gel was soaked in 20 X SSC for 45 minutes. Size fractionated RNAs were transferred to positively charged nylon membrane (Roche Diagnostics GmBH, Germany) by capillary action overnight (Sambrook et al., 1989) or by using vacuum transfer (Vacugene XL, Pharmacia LKB) . The RNA was UV cross linked to the membrane using UVC 500, UV Crosslinker (Hoefer,USA) at energy setting of 120 mjoules / cm^2 according to the manufacturer's instructions. After cross-linking, RNA was visualized by staining the membrane with 0.03% methylene blue in 0.03 M NaOAc.

Hybridization of cRNA probe and detection were performed using DIG system (The DIG System User's Guide, Roche Diagnostics GmBH, Germany).

2.5.2 Generation of DIG - labeled cRNA probes

To generate a DIG-labeled probe, the CP fragment (nt 5800-6291) was subcloned into a plasmid vector. To synthesize 'runoff' transcripts the plasmid was linearized with Bam HI. The linearized template was *in vitro* transcribed with T3 polymerase to generate DIG-labelled antisense cRNA probes using the DIG ^{TM} RNA labeling kit (Roche Diagnostics GmBH, Germany). The labeling reaction was carried out as per the manufacturer's instructions. Hybridization was carried out with a DIG-labelled antisense cRNA probe and signals were detected using colour substrate (NBT/BCIP) according to the manufacturer's instructions (Roche Diagnostics GmBH, Germany).

2.6 Polyacrylamide gel electrophoresis and western blots

The CP of purified virus and proteins extracted from *in vivo* IRES assays were separated on 12% SDS-PAGE gels containing 0.4% sodium dodecyl sulfate (SDS-PAGE) according to Laemmli (1970). Protein bands were visualized by staining with Coomassie brilliant blue or transferred onto polyvinylidine difluoride (PVDF) western blotting membranes (Roche Diagnostics GmBH, Germany) using a mini trans-blot cell apparatus (Bio-Rad, USA). Proteins were electrophoretically transferred from SDS-PAGE gels to PVDF membranes in the presence of transfer buffer (10 mM Tris base pH 8.3, 96 mM glycine, 10% methanol). The membrane was incubated overnight at 4^0C in blocking buffer containing 1% bovine serum albumin (BSA) dissolved in TBST (10 mM Tris pH 8.0, 150 mM NaCl, 0.05%

Tween-20). The membrane was probed with protein specific antibody e.g. polyclonal HVS antibody (1µl of crude antiserum dissolved in 10 ml of TBST) for I hr at room temperature with gentle shaking. Unbound antibodies were removed by washing the membrane three times in TBST buffer with each wash lasting up to 10 min. The membrane was probed with goat anti-rabbit IgG conjugated with alkaline phosphatase, as the secondary antibody (1 µl in 10 ml of TBST) for 1 hr. The unbound antibodes were washed with three changes of TBST buffer 10 min each. Bands were visualized using NBT/BCIP colour development substrate (Promega) in alkaline phosphatase buffer (100 mM Tris pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$).

2.7 *In vitro* transcription

Plasmid DNA was linearized with appropriate restriction enzyme and cleaned up using a Qiaquick gel extraction column and 1 µg was used as template for RNA transcription reaction. The Ambion mMessage mMachine® kit was used for generating *in vitro* transcripts.

Components	Amount
2 x NTP/CAP	10 µl
10 x Reaction buffer	2 µl
Linearized DNA template	1 µg
Enzyme mix (T7 polymerase)	2 µl
Nuclease free water	to 20 µl

The reaction was mixed thoroughly and incubated at 37^{o}C for 2 hr. The integrity of *in vitro* transcripts was checked by agarose gel electropforesis.

2.7.1 Coupled *in vitro* transcription and translation

Proteins from linearized plasmid DNA template were expressed using TnT® Coupled Wheat Germ Extract Systems (Promega) according to the manufacturer's instructions. Typically, a 50 µl reaction mix contains 1 µg of linearized plasmid DNA template, TnT wheat germ extract 25 µl, TnT reaction buffer 2 µl, T7 RNA polymerase 1 µl, 50 µCi of [^{35}S] methionine, amino acid mixture minus methionine (1 mM), 2 µl, RNasin® (40U/µl) 1µl, nuclease free water to 50 µl. The reaction was incubated at 30^{o}C for 2 hr. Template concentrations were optimized independently for different IRES constructs. The translation products were resolved by SDS-PAGE and detected by autoradiography.

Translation reaction supplemented with unlabeled amino acids were used for assaying GUS activity by fluorimetric assay.

2.8 Construction of bicistronic vectors for IRES tests

2.8.1 Constructs for *in vitro* assays

T7 driven bicistronic constructs were assembled in construct hGFP-I-GUS, provided by ICON Genetics, GmbH (06120 Halle, Germany). The bicistronic construct contains a hairpin loop at the 5' end immediately upstream of the first ORF, encoding the jelly fish green fluorescent protein (GFP), blocking the expression of GFP. This is followed by the TMV-Cr CP IRES (Crcp) and then the GUS ORF (Fig. 2.1a). HVS IRES regions were cloned into Hind III and Nco I sites between the GFP and GUS ORFs (Fig. 2.1b). Subsequently the HVS CP IRES construct (phGFP-Icp-GFP) and HVS MP IRES construct (phGFP-Imp-GFP) were linearized with Xba I, translated in TnT® coupled wheat germ extract systems (Promega) supplemented with 50 µCi of [^{35}S] methionine and amino acid mixture minus methionine. The translation products were separated by SDS-PAGE and detected by autoradiography. For GUS fluorimetric assays, amino acids without radioactive label were provided in the reaction mix. The proteins were extracted and assayed for GUS activity. Targeted deletions within HVS CP and MP IRESs were generated by PCR. The PCR products were cloned into Hind III and Nco I sites of plasmid hGFP-I-GUS.

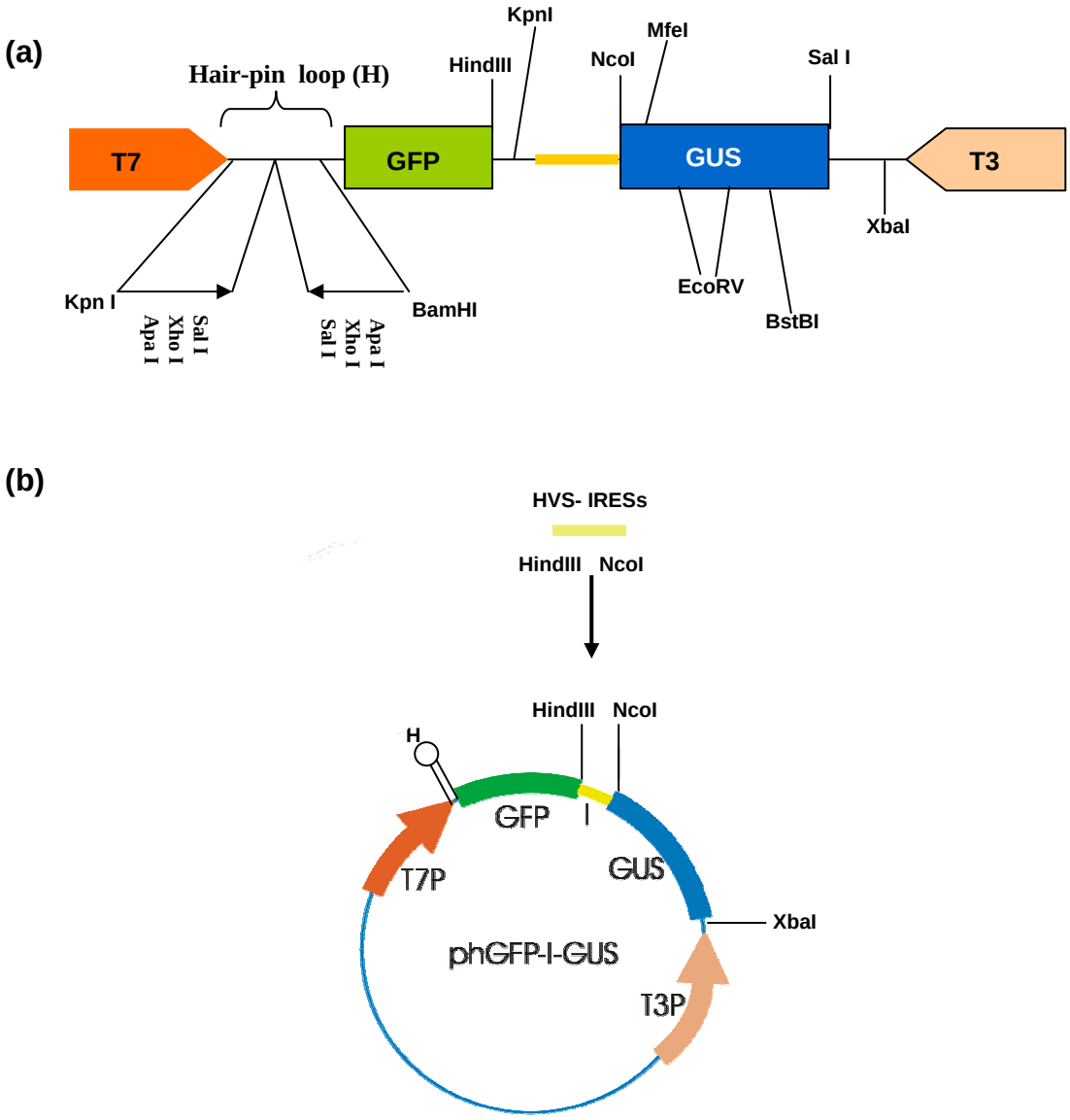


Fig 2.1 **(a)** The hGFP-IRES-GUS cassette, showing details of restriction sites that were utilized in subsequent cloning steps for generating 35S driven constructs for *in vivo* assays.
(b) Strategy used to generate T7 driven, HVS IRES bicistronic constructs (phGFP-I-GUS). T7 P and T3 P denotes T7 and T3 promoters, respectively. Insertion of '**H**' hairpin loop, in front of GFP ORF, blocks 5' ribosome scanning.

2.8.2 Constructs for *in vivo* protoplast assays

Constructs for in vivo assays were built on PFF-19G (Timmermans et al., 1990) backbone (35S-GFP-CrCP-GUS, kindly provided by Dr. M.V. Skulachev, Moscow State University, Russia) for performing functional assays of IRESs in protoplast systems. For testing HVS CP IRES and its deletion clones, plasmid phGFP-Icp-GUS, was digested with Bam H I and Mfe I. The Bam HI site is located at the 3' end of hairpin loop structure and Mfe I is a unique site within the GUS ORF. This fragment was cloned into vector 35S-GFP-I-GUS, containing TMV-Cr CP IRES, digested with Bam H I and Mfe I generating plasmid 35S-GFP- Icp-GUS (Fig. 2.2a).

HVS MP IRES has an internal Mfe I site at nucleotide position 4834. Therefore a different strategy was used for cloning HVS MP IRES. Plasmid phGFP-Imp-GUS was digested with Hind III and BstB I and the resultant 1.4 kb digested product containing HVS MP IRES and a portion of GUS gene was gel purified. Plasmid 35S-GFP-I-GUS was digested with Hind III. The digestion yielded two products, a 1.6 kb fragment containing 35S promoter and full-length GFP ORF and a 5 kb fragment containing CrCP IRES-GUS and the pFF-19 G backbone. Both fragments were gel purified. The larger 5 kb fragment containing CrCP IRES-GUS was digested with Bst BI. This round of digestion yielded a 3.9 kb product consisting of a 3' portion of GUS ORF and the pFF-19G back bone and a 1.1 kb product consisting of the CrCP IRES and the 5' GUS ORF region. The 3.9 kb product (Hind III and Bst BI ends) was gel purified and a ligation was set up with a 1.6 kb fragment containing the 35S promoter and the full length GFP ORF with Hind III compatible ends, plus the 1.4 kb digestion product containing the HVS MP

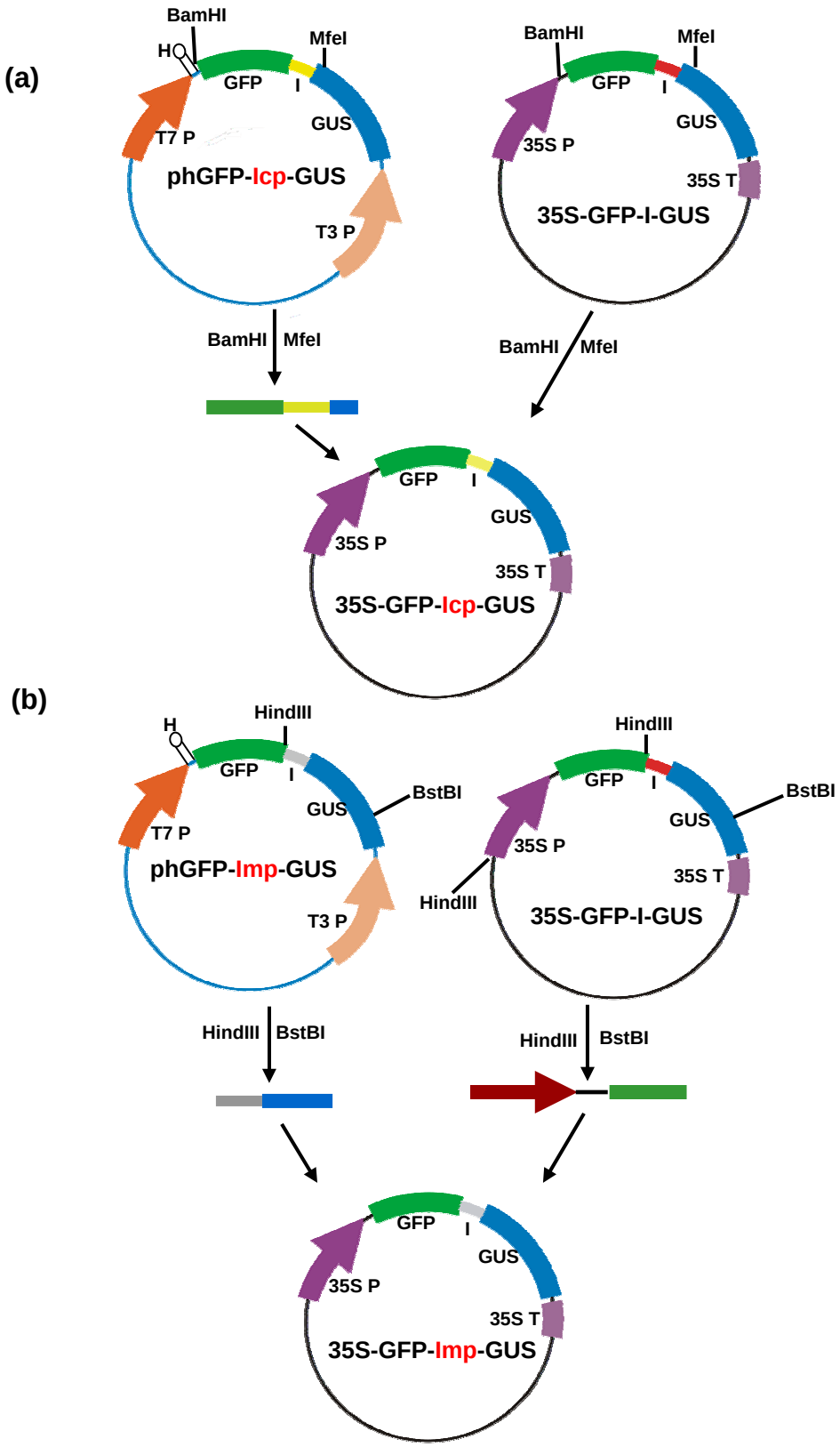


Fig 2.2 Strategy used to generate 35S driven, HVS IRES bicistronic constructs for *in vivo* assays in protoplasts using plasmid pFF19G. 35S P and 35S T denotes 35S promoter and 35S terminator, respectively. Scheme (**a**) was utilized for HVS CP IRES cloning and scheme (**b**) for HVS MP IRES cloning.

IRES and a portion of GUS gene with Hind III and Bst BI compatible ends respectively (Fig. 2.2b). The orientation of the inserts with respect to the 35S promoter in plasmid 35S-GFP-Imp-GUS and their order of arrangement were confirmed with restriction digestion and nucleotide sequencing.

2.8.3 IRES constructs for *Agrobacterium* infiltration assays

Transient GUS assays can be performed by agro-infiltration using constructs harbouring T-DNA borders. For this purpose, the GFP-IRES-GUS cassette has to be cloned into pCAMBIA 1300. In the pFF-19G constructs, the GFP-IRES-GUS cassette is flanked by Xho I sites. pCAMBIA 1300 has Xho I sites flanking a hygromycin (R) gene which is conveniently located between the 35S promoter and the 35S terminator. pCAMBIA 1300 was digested with Xho I, and a 7.8 kb fragment was purified using the Qiaquick gel extraction column. The fragment was de-phosphorylated and used for cloning the HVS IRES cassettes. Similarly, plasmids 35S-GFP-Icp-GUS and 35S-GFP- Imp-GUS were digested with Xho I. The resultant 2.7 kb fragments were ligated with the 7.8 kb fragment from Xho I digested pCAMBIA 1300. The plasmids were named pCAM -GFP- Icp-GUS for CP IRES and pCAM- GFP- Imp-GUS for MP IRES respectively (Fig. 2.3).

Background expression of GUS observed from agrobacterium harbouring the IRES constructs during infiltration assays can interfere with quantifying the expression of GUS from infiltrated plant tissue. Insertion of a plant intron in the GUS gene can alleviate this problem (Ohta et al., 1990). A castor bean catalase intron derived from pCAMBIA 1301 was inserted into the GUS ORF of pCAM -GFP- Icp-GUS and pCAM -GFP- Imp-GUS constructs. pCAMBIA 1301 was

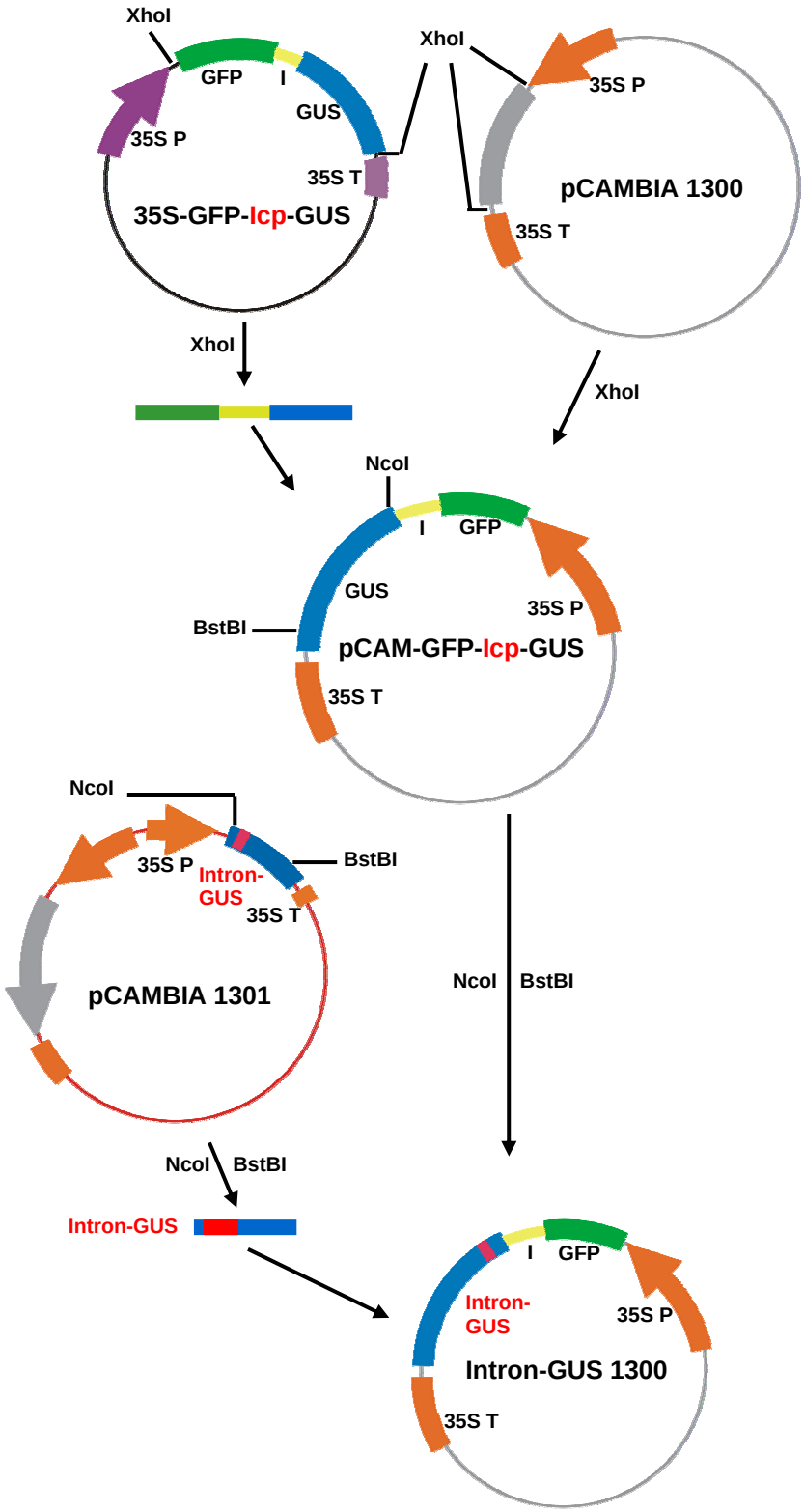


Fig 2.3 Strategy used to generate 35S driven, HVS IRES bicistronic constructs for *In vivo* assays by agro-infiltration. 35S P and 35S T denotes 35S promoter and 35S terminator, respectively. Insertion of castor bean catalase intron into GUS ORF reduces the background GUS expression observed from *Agrobacterium*.

digested with Nco I and Bst BI, which released a 1.3 kb fragment containing the intron and a portion of the GUS gene. This fragment was cloned into Nco I and Bst BI digested pCAM -GFP- Icp-GUS and pCAM -GFP- Imp-GUS constructs (Fig. 2.3).

Promoter-less IRES constructs were also constructed to test the possibility that the IRES itself could act as a promoter in transcribing the 3' proximal GUS gene. Hind III sites were located in the multiple cloning site of pCAMBIA 1300 backbone and immediately downstream of the GFP ORF which was derived from pFF-19G. Promoter-less HVS CP and MP IRES constructs were made by digesting pCAM -GFP- Icp-GUS and pCAM -GFP- Imp-GUS constructs with Hind III and the 8.9 kb fragment containing IRES-GUS and the portion of vector pCAMBIA 1300 without multiple cloning sites, 35S promoter and GFP ORF was re-circularized using T4 DNA ligase (New England Biolabs).

2.8.4 Isolation of protoplasts

Kenaf cultivar, Everglade-41 leaf material was used for isolating protoplasts. Seeds of kenaf cultivars were kindly provided by Dr. B.S.Baldwin, Mississippi State University, USA. Kenaf seedlings were grown at 25^{o}C, 16/8 hours, light/dark cycle. One month old kenaf seedlings at 4 to 6 leaf stage were used for protoplast isolation. The method followed for isolating protoplasts was according to previously published work from this lab (Liang et al., 2002). Leaves were surface sterilized for 10 min with 0.8% Clorox® containing the active ingredient 0.04% sodium hypochlorite. Following that, the leaves were rinsed three times with sterile distilled water, each wash lasting up to 5 min. Leaves were sliced into thin 1mm strips and incubated in filter sterilized enzyme solution. The enzyme mixture

contained 0.2 mM KH_2PO_4, 1 mM KNO_3, 1 mM $MgSO_4$, 1 µM KI, 0.01 µm $CuSO_4$, pH 5.6 (Rottier et al., 1979) and 0.6 M mannitol, 10 mM $CaCl_2$, 0.8% cellulase Onozuka R-10 (Yakult Honsa Co. Ltd), 0.25% macerase R-10 (Yakult Honsa Co. Ltd). Digestions were carried out at 25°C in the dark with shaking at 10 x g/ min (Heidolph Rotamax 120, Germany) for 16 hr. Protoplasts were gently pipetted using a pasteur pipette and released. The protoplast containing solution was passed through a 70µm nylon cell strainer (Becton Dickinson, Franklin Lakes, NJ) to remove the cell debris. The filtrate was later transferred to 50 ml centrifuge tubes and centrifuged at 100 x g for 5 minutes at 4°C. Pellets were three times washed in wash solution containing 0.6M mannitol and 10 mM $CaCl_2$, (pH 5.6),. Protoplast yields were calculated using haemocytometer slide (Marienfield, Germany) and protoplast viability was determined by fluorescein diacetate (FDA) staining (Widholm,1972).

2.8.5 PEG inoculation of protoplasts

Concentrated kenaf protoplasts (4x 10^5 cells) were mixed with 10 µg (30 µl) of *in vitro* transcript and 200µl of 40% PEG 3000 in 3mM $CaCl_2$ for 15 sec. Then protoplast/DNA mixture was diluted with 1.5 ml of wash solution and left on ice for 2 min. The protoplast/DNA mixture was diluted twice with 1.5 ml of wash solution and incubated on ice for another 15 min. The mixture was washed once with 2 ml of wash solution. The protoplast concentration was adjusted to 1x 10^5 cells per ml with MS medium (Murashige and Skoog, 1962) containing 0.6M mannitol and 10 mm $CaCl_2$ and incubated in dark for 24 hr. The protoplasts were harvested and total RNA was isolated for further analysis.

2.8.6 Preparation of electro-competent *Agrobacterium* cells

Agrobacterium EHA 105 (Hood et al., 1993) competent cells were prepared by inoculating a single colony to 5 ml of LB medium supplemented with 50 µg/ml rifampicin and incubated in a shaker incubator at 28^{o}C. The overnight grown bacterial culture was transferred to 100 ml LB medium containing 50µg/ml rifampicin. Cells were harvested when they reached a density of 0.5 at $A_{600\ nm}$. Cells were centrifuged at 10,000 x g for 10 min at 4^{o}C. The pellet was washed with 40 ml of 1mM HEPES buffer pH (7.0). The cells were washed again with 40 ml of 1mM HEPES buffer pH (7.0) containing 10% glycerol. Finally the bacterial pellet was resuspended in 2 ml of 1mM HEPES buffer pH (7.0) containing 10% glycerol. Bacterial cells were transferred as 100 µl aliquots, frozen in liquid nitrogen and stored at -80^{o}C.

2.8.7 Electroporation of *Agrobacterium*

After thawing the *Agrobacterium* competent cells, 0.5 µl of miniprep DNA was added and incubated on ice for 10 minutes. The contents were transferred to a pre-chilled Bio-RAD, Gene Pulser® electroporation cuvette (2 mm gap). The electroporation parameters were set as 25 µF capacitance, 400 Ω resistance and 2.5 KV pulse with an 8-9 sec delay. A BIO-RAD GENE PULSER II was used for electroporation. After pulsing the cells, 1 ml of LB medium was added immediately to the DNA/competent cell mixture and chilled on ice for 2 min. The contents were transferred to a sterile tube and the cells were grown at 28^{o}C for 2 hr to allow recovery and marker expression. The bacterial cells (30 µl) were plated onto LB agar supplemented with appropriate selection antibiotics (Kanamycin 50 µg/ ml, Rifampicin 10 µg/ ml). Cells were grown for 2 days at 28^{o}C.

2.8.8 *In vivo* IRES assays

Agrobacterium harbouring bicistronic constructs were streaked on to LB agar plate containing the selection antibiotics (Kanamycin 50 µg/ ml, Rifampicin 10 µg/ ml) and incubated at 28^{o}C. A single colony was inoculated into 2 ml LB medium containing Kanamycin 50 µg/ ml, Rifampicin 10 µg/ ml and grown at 28^{o}C for 2 days with vigorous shaking. One ml of the culture was transferred to 50 ml of LB medium with antibiotics, 10 mM MES (pH 5.6) and 20 µl of 100 mM acetosyringone. The cultures were incubated at 28^{o}C for 16 hr with vigorous shaking till the cell density reached 1.0 at A_{600nm}. The cells were centrifuged at 2000 x g for 10 min. The pellet was re-suspended in 50 ml 0f 10 mM $MgCl_2$ and 75 µl of 100 mM acetosyringone was added. The cells were kept at room temperature for 3 hr without shaking.

Agrobacterium cells harbouring bicistronic constructs were infiltrated with a 1 ml syringe without a needle. Young expanded leaves of *N.benthamiana* were punched with a pipette tip and a syringe containing the bacterial culture was positioned on the hole. The other side of the leaf was blocked with a finger and the culture was infiltrated into the leaf by gently pushing the syringe barrel. In each plant, 2 to 3 leaves were infiltrated and five replicates were done for each construct. Leaves were collected after 72 hr and assayed for GUS activity.

2.8.9 β-Glucuronidase (GUS) fluorimetric assay

To test the IRES activity, GUS fluorimetric assay (Jefferson, 1987) was used. Fluorimetric assay involves quantifying the rate of β-glucuronidase activity in hydrolyzing 4-methylumbelliferyl β-D-glucuronide, (4-MUG, SIGMA # 5664) substrate, to give the breakdown product, 4-methylumbelliferone (4-MU).

Translation sample (20 µl) was diluted with 80 µl of GUS extraction buffer (50 mM sodium phosphate buffer (pH 7.0) containing 10 mM β- mercaptoethanol,10 mM Na-EDTA, 0.1% (v/v) Triton x-100). From the diluted sample 40 µl was taken to which 70µl of GUS extraction buffer was added. To the diluted sample, 1 µl of 10mM 4-MUG solution was added and incubated at 37^{o}C for 1 hr. To stop the reaction and to enhance fluorescence, 1ml of 0.2 M Na_2CO_3 was added. The rate of accumulation of 4-MU was assayed fluorimetrically using excitation and emission wavelengths of 365 and 455nm, respectively.

For assaying *in vivo* GUS activity from *N.benthamiana* leaves, infiltrated areas were excised 72 hr after inoculation, homogenized in 10 volumes (w/v) GUS-extraction buffer. The sample was centrifuged to pellet the debris and to the supernatant (100 µl), 1 µl of 10mM 4-MUG solution was added. The reaction mix was incubated at 37^{o}C for 1 hr. To stop the reaction and to enhance fluorescence, 1ml of 0.2 M Na_2CO_3 was added. GUS activity was determined using excitation and emission wavelengths of 365 and 455, respectively.

For calculating protein content 5 µl of leaf extract was diluted with 155 µl of GUS buffer. To this 40 µl of Bradford reagent was added. Absorbance at 595 nm was measured and the protein concentration was compared against BSA standard curve (Bradford, 1976). Same samples were applied on SDS-PAGE for GFP western blotting.

CHAPTER 3

CLONING AND CHARACTERISATION OF HVS

3.1 Virus purification and maintenance

HVS was found in hibiscus plants that were co-infected with HCRSV. Purification of the virus from these plants yielded a mixed population of HVS and HCRSV. The mixture of viruses was inoculated onto *C. quinoa*. The mixture of viruses induced chlorotic local lesions. Individual lesions were cut out, ground and transferred to another leaf. An aliquot of the sample was examined under TEM to check for the presence of HVS. This was repeated for five successive times and a pure culture of HVS was obtained. Subsequently, HVS was inoculated onto *H. canabinnus* for obtaining high yields of HVS. For HVS to establish itself systemically in *H. canabinnus* it takes at least four weeks. Typically, 100 gm of infected *H. canabinnus* leaves could yield 25- 30 mg of HVS.

3.2 Particle size distribution

Transmission electron microscopic observation of unfractionated purified HVS showed rigid rod shaped particles of varying sizes (Fig. 3.1a). Virion lengths were measured using the step-by-step procedure for distance measuring, according to the user's manual. End-to-end aggregates or multimers of 34 nm particles were counted as monomers. Lengths of 500 virions were measured and a virion size histogram was plotted. Particles that were 34 nm in length or its multimers were found to be predominant (Fig. 3.1c).The shorter particles may contain CP sgRNA as reported in the case of SHMV (Higgins et al., 1976). Particles of 300 to 330 nm

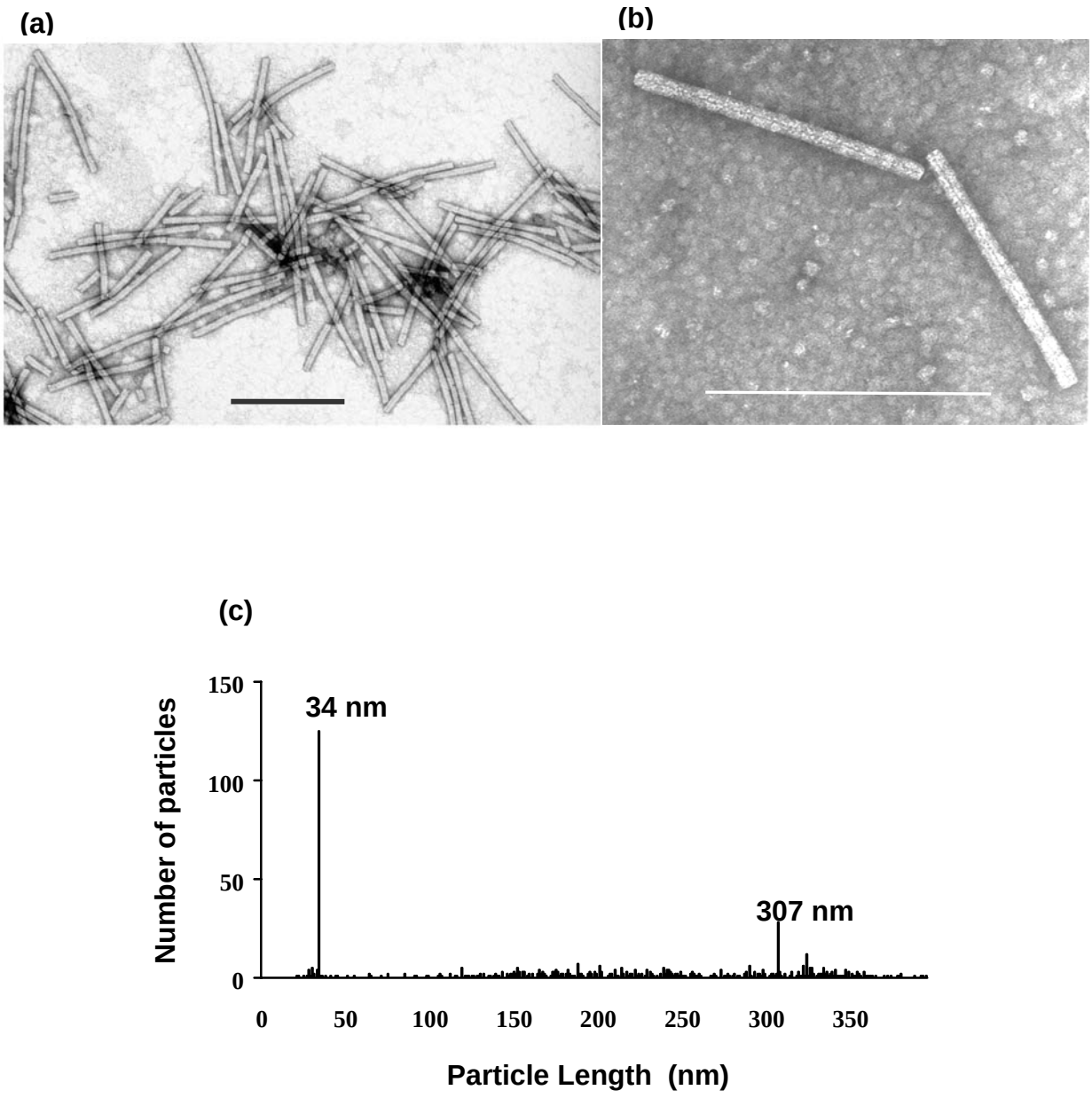


Fig 3.1 Transmission electron micrograph of **(a)** purified HVS particle and **(b)** at high magnification to show the central core. Bars represent 300nm. **(c)** Frequency distribution histogram of unfractionated HVS particles

in length also were found in large numbers. A closer look at the particles at a higher magnification reveals the presence of a central core (Fig. 3.1b).

3.3 Viral RNA purification

RNA was purified using the phenol-chloroform method from HVS particles. Using 1 mg of purified virus, 20 -25 µg of RNA could be obtained. The purified RNA was size fractionated on 1.2% formaldehyde denaturing agarose gels. Single-stranded RNA profile revealed the presence of a ~6.4 kb genomic RNA band, together with a 1.5 kb and a 0.7 kb sgRNAs of MP and CP, respectively. Similarly RNA isolated from purified preparations of SHMV, a sgRNA of 0.7kb could be detected (Fig. 3.2a&b). The presence of CP sub-genomic RNA in purified virions of HVS indicates that CP sg RNA is encapsidated.

3.4 Host range and symptomology

HVS induced local chlorotic lesions in *C. quinoa* (Fig.3.3a) and systemic infection accompanied by wrinkling of leaves in *Nicotiana benthamiana* (Fig.3.3c). HVS could systemically infect several kenaf (*Hibiscus cannabinus* L.) cultivars, but the infected plants remained symptomless throughout and the same has been reported for FHV (Adkins et al., 2003). Although HVS is closely related to subgroup II tobamoviruses, it could not infect primary host plants of SHMV or CGMMV-SH. The inability of HVS to infect legumes, cucurbits and arabidopsis makes its experimental host range looks quite similar to FHV (Adkins et al., 2003). HVS could not infect ladies finger and cotton cultivars that we have tested it against (Table 3.1). FHV causes symptomless systemic

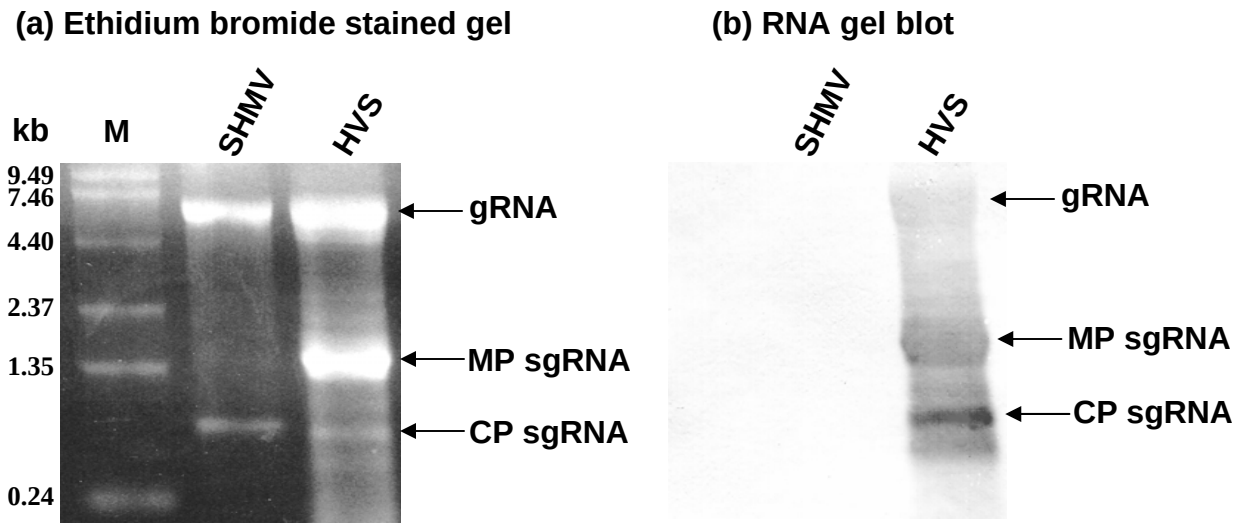


Fig 3.2 (a) Purified viral RNAs of SHMV and HVS size fractionated on a 1.2% formaldehyde denaturing agarose gel. **(b)** Northern blot analysis of an identical gel probed with DIG-labelled HVS CP specific riboprobe

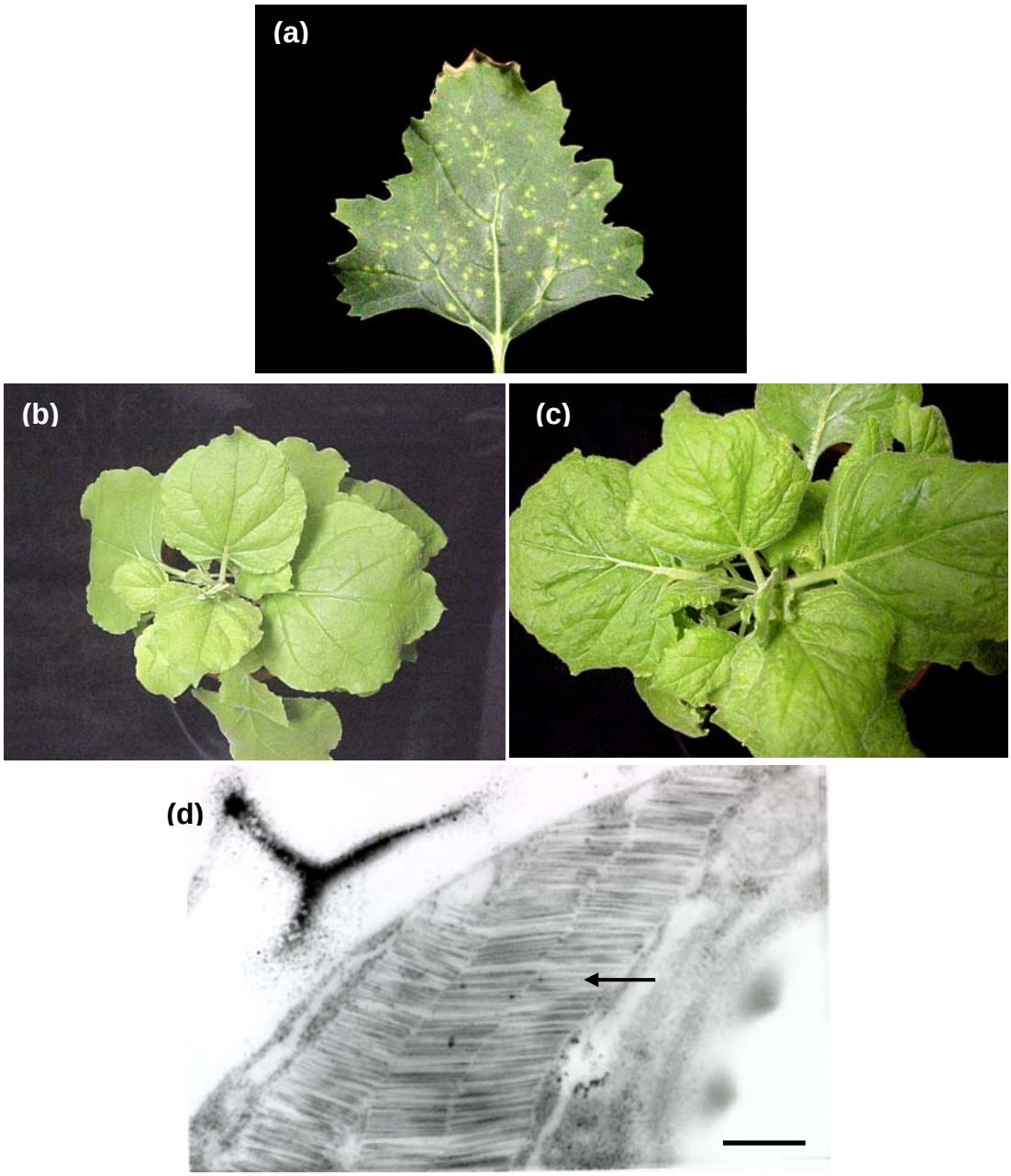


Fig 3.3 Mechanical inoculation of purified HVS particles showing chlorotic local lesions in *Chenopodium quinoa* (**a**) and systemic infection associated with wrinkling of leaves in *N. benthamiana* (**c**) as compared to mock inoculated plant (**b**) Section of *N. benthamiana* leaves infected with HVS showing layered aggregates of HVS particles in cytoplasm (**d**). Bar represents 300nm.

Table 3.1 **Experimental host range and symptoms of Hibiscus Virus-S**

Family	Species	Common name	Symptoms
Brassicaceae	*Arabidopsis thaliana* ecotype Landsberg erecta	Arabidopsis	-/-
Chenopodiaceae	*Chenopodium quinoa* Wild.	Lambsquarters	CLL
Cucurbitaceae	*Cucumis sativus* Bet- α	Cucumber	-/-
Fabaceae	*Phaseolus vulgaris* cv. Early Bush	Bean	-/-
	Phaseolus vulgaris cv. Green Coat	Bean	-/-
Malvaceae	*Abelmoschus esculentus* cv. Pure Luck	Okra	-/-
	Abelmoschus esculentus cv. Lucky Five	Okra	-/-
	Gossipium hirsutum	Cotton	-/-
	Hibiscus cannabinus cv. C-108	Kenaf	NS/NS
	Hibiscus cannabinus cv. COP-2	Kenaf	NS/NS
	Hibiscus cannabinus cv. Everglade-41	Kenaf	NS/NS
	Hibiscus cannabinus cv. MOP-1	Kenaf	NS/NS
	Hibiscus cannabinus cv. SF-459	Kenaf	NS/NS
	Hibiscus cannabinus cv. Tainung-2	Kenaf	-/-
	Hibiscus cannabinus cv. Whitten	Kenaf	-/-
Solanaceae	*Nicotiana benthamiana*		LD/S
	Nicotiana clevelandii		-/-
	Nicotiana tabacum cv. Xanthi nc.	Tobacco	-/-
	Nicotiana tabacum cv. Turkish	Tobacco	-/-

Key to symptoms: - = no infection, NS = infected but no symptoms, CLL = chlorotic local lesions, LD = leaf distortion. S = systemic infection. Systemic infection was confirmed by grinding the upper non - inoculated leaves in 0.2M Borate buffer pH (8.0) and applying the ground sap onto *Chenopodium quinoa* leaves. Infection was confirmed by western blots.

infection in *N. benthamiana* whereas HVS causes systemic infection accompanied by leaf deformation. Ultra-thin sections of HVS infected *N. benthamiana* leaf samples revealed the presence of layered aggregates in the cytoplasm (Fig. 3.3d) that are typical of tobamovirus infections.

3.5 Antiserum production and capsid mass determination

Crude antiserum diluted 10,000 times was able to detect the presence of HVS in infected leaf tissues by ELISA. In 12 % SDS-PAGE, the HVS CP migrated slower than the CP of ORSV and TMV (Fig. 3.4a). Using mass spectrometry, the molecular mass of the HVS CP was determined to be 18,066 Da (Fig. 3.5), as compared to 18,196 Da obtained from theoretical calculation. The difference between the molecular mass determined using mass spectrometry and theoretical molecular mass predicted from deduced aa sequence, could be due to post-translational modifications. Similar discrepancies between observed values from mass spectrometry and calculated molecular mass of purified CP have been reported in the case of TMGMV (Thomas et al., 1998), *Cymbidium mosaic virus* (CymMV) and ORSV (Tan et al., 2000).

3.6 Serological relationships of HVS and other tobamoviruses

Purified HVS CP was analyzed on a 12% SDS-PAGE gel. Coomassie blue stained gels showed two protein bands that migrated very closely (Fig. 3.4a, lane 3). When probed with antiserum against HVS no serological difference was observed between the two proteins (Fig. 3.4d, lane 3). It has been documented previously that purified preparations of SHMV also shows two bands (Whitfield

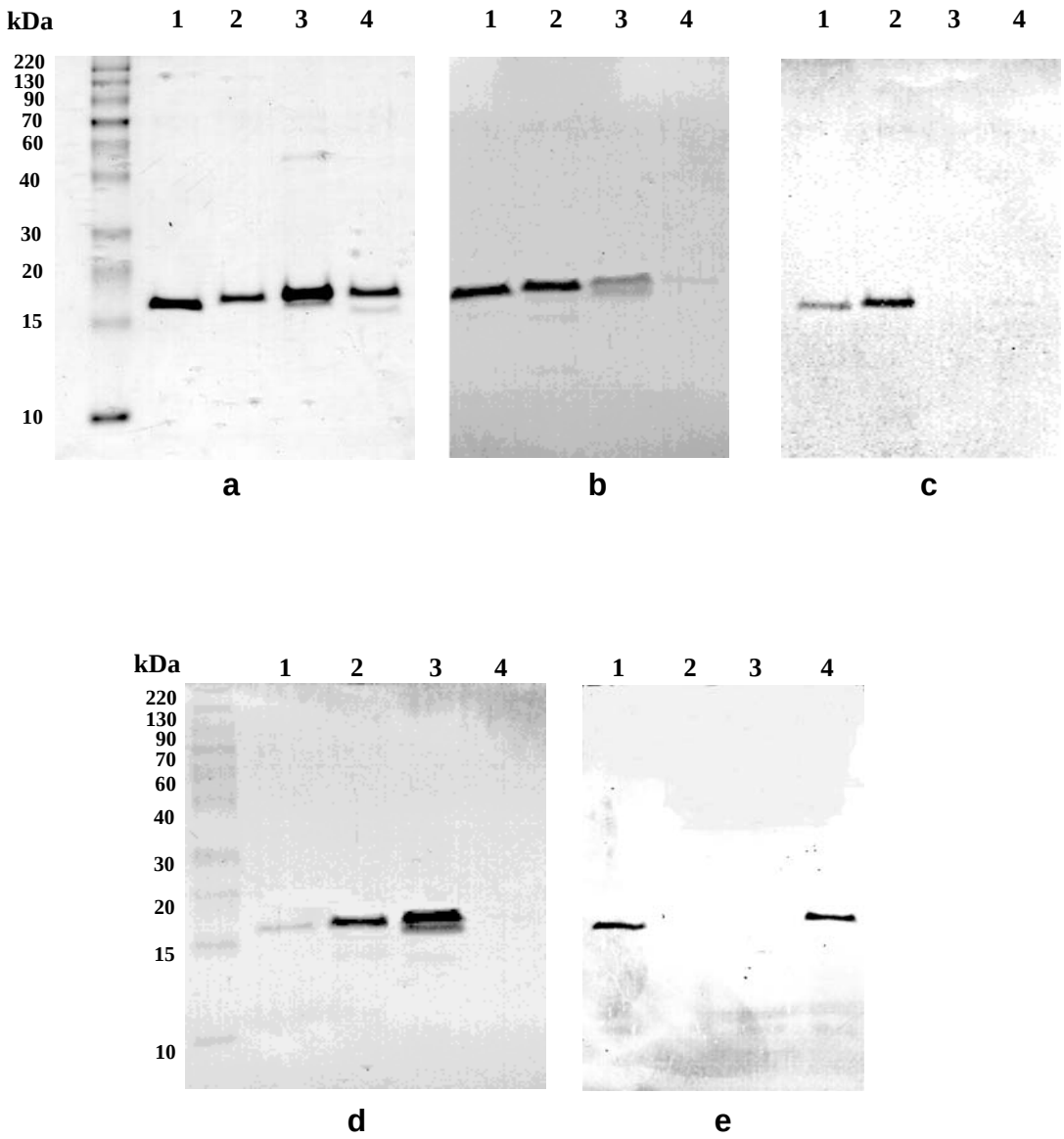


Fig. 3.4 (a) Purified coat proteins of the below given tobamoviruses separated on 12% SDS-PAGE gels and stained with coomassie brilliant blue. Tobacco mosaic virus (TMV, Lane 1), Odontoglossum ringspot virus-Singapore isolate (ORSV-S1, Lane 2), Hibiscus virus S (HVS, Lane 3), Sunn hemp mosaic virus (SHMV, Lane 4). Identical blots were probed with antisera against **b**, TMV; **c**, ORSV-S1; **d**, HVS; **e**, SHMV.

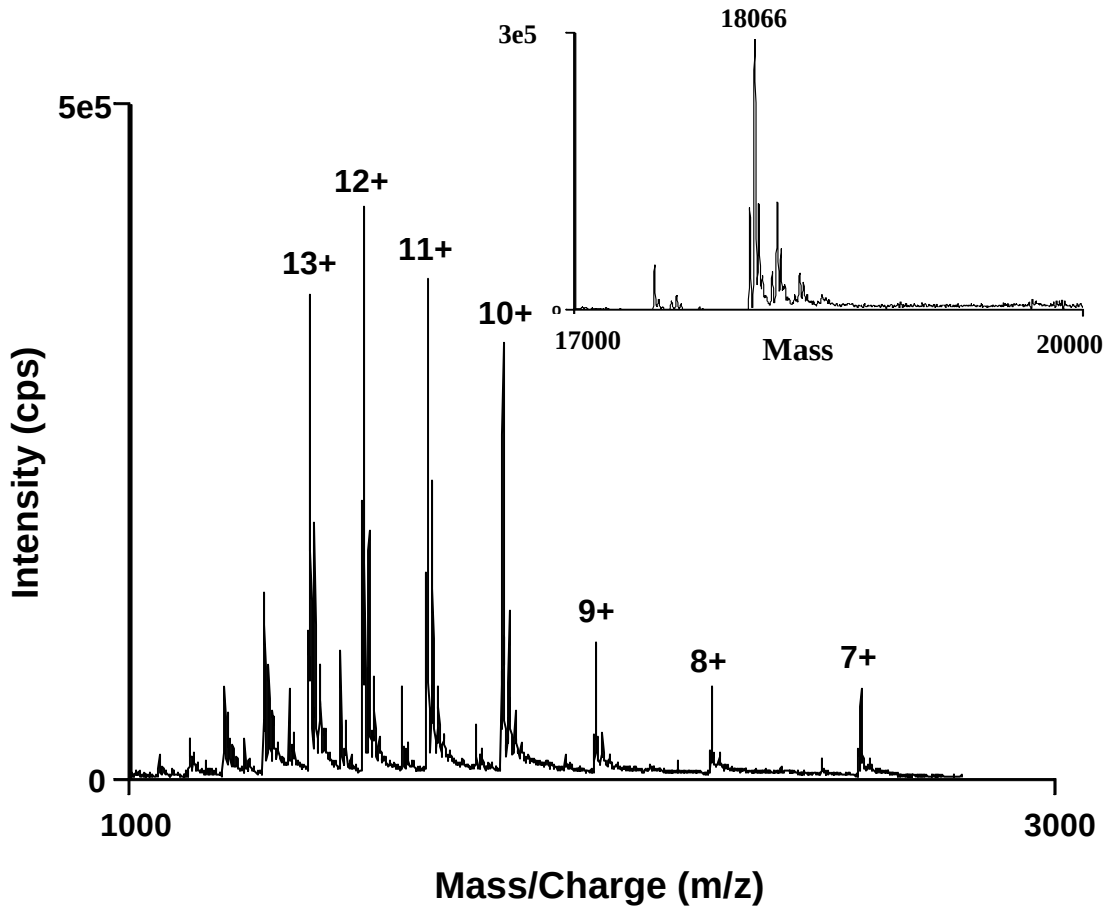


Fig 3.5 Electrospray ionization spectrum of HVS coat protein. The insert shows reconstructed spectrum of HVS coat protein. Molecular mass determined by this method is very close to the theoretical molecular mass calculated from deduced amino acid sequence of coat protein.

and Higgins, 1976). The 18 kDa SHMV CP is believed to be susceptible to plant carboxypeptidases which may account for the presence of the double bands (Varma, 1986).

In western blots, HVS CP cross-reacted with TMV antiserum (Fig. 3.4 b, lane 3), but did not cross-react with ORSV-S1 and SHMV antisera (Fig. 3.4c&e). There were cross reactivities between HVS antiserum with TMV and ORSV-S1 CPs (Fig. 3.4d, lanes 1&2). Although HVS and SHMV shared significantly higher nt and aa sequence identity, HVS antiserum did not react with SHMV CP. This shows that HVS is serologically unrelated to SHMV. These results confirm earlier observations that serological relatedness, sequence similarities and biological properties among tobamoviruses are independent of each other (Van Regenmortel, 1975).

3.7 cDNA synthesis and sequence determination

First strand synthesis using oligo (dT) resulted in clones terminating with the CP coding region. Several clones covering RdRp, MP and CP domains were obtained. The sizes of inserts being cloned ranged from 0.4- 2.2kb, approximately (Fig. 3.6b). Designing of the degenerate primer Tobmv-3' end (Table 2.1) allowed the 3' UTR region to be cloned. The presence of an internal poly-A tract immediately downstream of the CP, is the reason why 3' UTR clones were not obtained using oligo (dT) to prime first strand synthesis. Clones for confirming the 5' end sequence of viral RNA were obtained using the 5' RACE strategy. Several independent clones were sequenced to remove ambiguities at the 5' terminus.

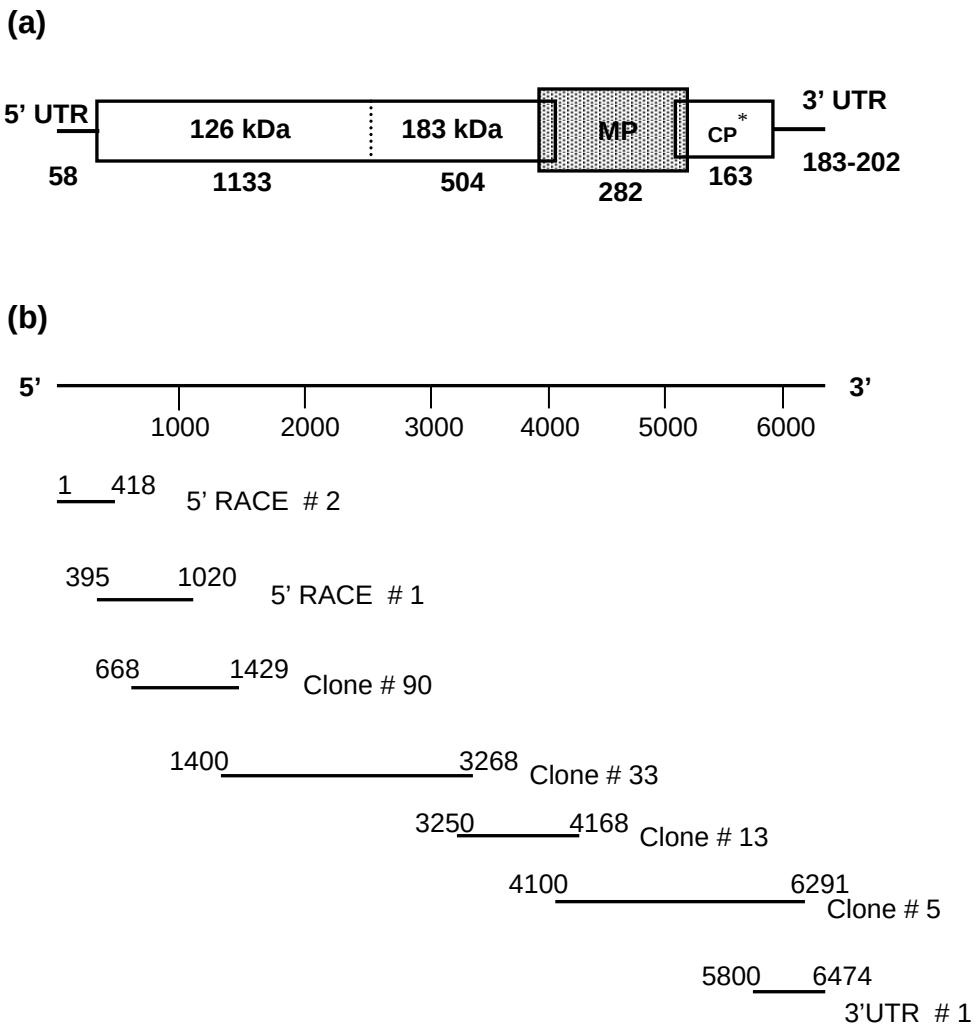


Fig. 3.6 Genome organization of HVS **(a)**. Boxes represent coding regions for methyltransferase-helicase (126 kDa), RdRp (183 kDa), MP (30 kDa) and CP (18 kDa) with their amino acid lengths indicated. Dotted vertical line within the box indicates amber stop codon. Asterik in CP region indicates location of predicted OAS. Untranslated regions at the 5' and 3' ends are represented by horizontal lines with their nucleotide lengths. Length of 3' UTR is thought to be variable. Full length sequence of HVS was determined by sequencing cDNA clones shown **(b)**. Nucleotide positions of the selected clones are specified.

3.7.1 Genome organization of HVS

The genome organization of HVS was compared with 11 other tobamoviruses. The of MP ORF overlapped with RdRp and CP ORFs by 8 and 11 nucleotides (nt) respectively (Fig. 3.6a). This and the presence of the OAS within the CP ORF, is a similarity that HVS shares with subgroup II viruses, such as SHMV and CGMMV-SH. Comparison with FHV was not possible since the complete nucleotide sequence of FHV is not available yet.

3.7.2 RNA dependent RNA polymerase

The RdRp gene product is primarily essential for viral replication apart from which, a non-conserved region within the RdRp has been proposed to act in concert with MP in facilitating movement (Hirashima and Watanabe, 2003). RdRp also acts as a suppressor of post transcriptional gene silencing (PTGS) (Kubota et al., 2003).

The initiation codon for the RdRp gene is located at nt 59 and extends up to nt 3457 encoding a protein which is composed of 1133 amino acids, with a theoretical molecular mass of 128,234 Da. The readthrough product which terminates at nt 4972 results in a protein consisting of 1637 amino acids with a calculated mass of 185,835 Da. In TMV, the readthrough of the amber stop codon is influenced by the consensus sequence UAG-CAA-UUA (Skuzeski et al., 1991) which is found in the HVS sequence (nt 3458-3466) (Fig. 3.7). Multiple sequence alignment of tobamoviral RdRps show several conserved motifs that perform defined roles in replication of viral RNA.

Capping of the 5’ end involves removal of phosphate from 5’ end of RNA followed by linking of guanosine to the 5’ end of RNA through guanylyltransferase activity

```
GTA TGT TTT TAG TTT GAA CAT TTC AAC AAC ATT CAA CTA CGA AAC AGC AAC AAC  54

AAA TATG AC TAC AGA CAT TAA ACA ACA AAT CAC AAA CGT ACG CGA CGA CGC ATG 108
      M   T   T   D   I   K   Q   Q   I   T   N   V   R   D   D   A   C  17

CCA AGG GGC ATC GTC GTT AAT TAA AGG TTT GGC CAC CCG TCG TGT GTA CGA TGA 162
  Q   G   A   S   S   L   I   K   G   L   A   T   R   R   V   Y   D   E  35

GGC GGT ACA AAC CTT AAC AGT GTT GGA CAA AAG GCC GAA ATT CAC CTT TTC TAA 216
  A   V   Q   T   L   T   V   L   D   K   R   P   K   F   T   F   S   N  53

TAT GGT TAC AGC CGA TCA GGC GAA ACT TGT CAC TGA GGC TTA CCC CGA GTT TAA 270
  M   V   T   A   D   Q   A   K   L   V   T   E   A   Y   P   E   F   N  71

TAT AAG CTA TGT GGG CAA TAG GAA CGC CGT CCA TTC ACT GGC CGG AGG ATT GAG 324
  I   S   Y   V   G   N   R   N   A   V   H   S   L   A   G   G   L   R  89

GAA GTT GGA AAT GGA ATA TTT GAT GAC ACT CGT TCC TTA CGG TTC TCC TAC CTA 378
  K   L   E   M   E   Y   L   M   T   L   V   P   Y   G   S   P   T   Y 107

TGA CAT AGG AGG AAA TTA TGC GCA GCA CCT TTT GAA AGG ACG ATC ATA TGT TCA 432
  D   I   G   G   N   Y   A   Q   H   L   L   K   G   R   S   Y   V   H 125

CTG CTG CAA TCC TTG TTT GGA TTT CAG AGA TAT AGC CAG AAA CGA GAA TTA CAA 486
  C   C   N   P   C   L   D   F   R   D   I   A   R   N   E   N   Y   K 143

GGA GAC TTT GGA TTC ATA TCT GAA GCG GTT CAC TAC GCC AGA TTG CGG TTG GTT 540
  E   T   L   D   S   Y   L   K   R   F   T   T   P   D   C   G   W   F 161

CTC CGC GGC GTG CAG AAA GGT TAC TGC AAG ACC TTT GCC TTA TTT TCA ACA GGA 594
  S   A   A   C   R   K   V   T   A   R   P   L   P   Y   F   Q   Q   D 179

TGC ATT CGA ACG GTA TAG ACA AAA TCC GAA AGA CGT CAC CTG CTC AGA ACC CTT 648
  A   F   E   R   Y   R   Q   N   P   K   D   V   T   C   S   E   P   F 197

TCA GGA TTG TTG TAT GCA ACC CCC TAG TAA TCA GGA TTG TTT TGC AAT TTC TGT 702
  Q   D   C   C   M   Q   P   P   S   N   Q   D   C   F   A   I   S   V 215

GCA TTC ACT CTA TGA CAT TCC TGT CGA TGA ATT GGG CGC GGC ACT CTT AAG GAA 756
  H   S   L   Y   D   I   P   V   D   E   L   G   A   A   L   L   R   K 233

GAA CGT AAA AAC TCT TTA CGC ACT CCA TTT CTC CGA GGA GCT TTT AAT GGG CGC 810
  N   V   K   T   L   Y   A   L   H   F   S   E   E   L   L   M   G   A 251
```

```
AAC TGA AGC GCG ACT GGG CAC GAC TGG TGC AAA ATT TTT CCG GGA GGG TGA CGA 864
  T   E   A   R   L   G   T   T   G   A   K   F   F   R   E   G   D   E 269

GGT CAC ATT CGG TTT CGA TGG TGA ATC CAC CCT TCT TTA CAC GCA CTC CTT TAA 918
  V   T   F   G   F   D   G   E   S   T   L   L   Y   T   H   S   F   K 287

AAA TAT TTG TGC TTA TAT TAC TAG GAG TTT TTT TTA TGC AGG AGC TAG ATT TGC 972
  N   I   C   A   Y   I   T   R   S   F   F   Y   A   G   A   R   F   A 305

TTA CAT GAA AGA ATT TTT AGT ACG CCG TGT AGA TAC TGT CTA TTG TAA ATT TAT 1026
  Y   M   K   E   F   L   V   R   R   V   D   T   V   Y   C   K   F   I  323

TAA GAT AGA CAC GCA GTG TTT ATA TAA GTC CGT ATT CCA TAC TGA TAT CGA TAA 1080
  K   I   D   T   Q   C   L   Y   K   S   V   F   H   T   D   I   D   N  341

CGA AAA TAT GAT AAA AGG TAC GGA TGA GGC CTT CAT GTA CAA ACG AGA AGC CGC 1134
  E   N   M   I   K   G   T   D   E   A   F   M   Y   K   R   E   A   A  359

GTT GTT GAA CGC AGC AAG ACC TAT TTT TAA GGA TAA GGC CGC TTT TCG TGT GTG 1188
  L   L   N   A   A   R   P   I   F   K   D   K   A   A   F   R   V   W  377

GTT TCC CAA TTC TGT AGG AAA AGT GAT ATA CCC TTT TTT TAA AGG TAG TTC ATG 1242
  F   P   N   S   V   G   K   V   I   Y   P   F   F   K   G   S   S   C  395

TAC ATC CGC CAA GGT AAA GGT GGA AAA AAG ACT GGT TGA TGA AGA CTT TGT CTT 1296
  T   S   A   K   V   K   V   E   K   R   L   V   D   E   D   F   V   F  413

CAC TGT GCT AAA TCA CAT CCG CAC CTA TTC CGG CAA GCA ACT GTC ATA TGA AAA 1350
  T   V   L   N   H   I   R   T   Y   S   G   K   Q   L   S   Y   E   N  431

TGT TTT ATC GTT TGT CGA GTC CAT CCG ATC GAG GGT GGT AGC TAA CGG TGC TAA 1404
  V   L   S   F   V   E   S   I   R   S   R   V   V   A   N   G   A   N  449

CGT GCG GTC AGA GTG GGA TAT CCC AAA GTC TGA GAT CCA AGA TAT CGC CAT GTC 1458
  V   R   S   E   W   D   I   P   K   S   E   I   Q   D   I   A   M   S  467

ATT GTT TTT GAT CAC ACG GTT AAG AAC TCT GCA GGA CAG TAA TGT TCT TCA GCA 1512
  L   F   L   I   T   R   L   R   T   L   Q   D   S   N   V   L   Q   Q  485

GTT TTC GGT CAA GGG CAA GAG TCT GAC AGA TCT CGT GAA GGA GGC CCT AGG GGG 1566
  F   S   V   K   G   K   S   L   T   D   L   V   K   E   A   L   G   G  503

TCT TGT CGG TTC GAT GTT TGA ACC ACT GAC TCA AGT TTG TAT AGA CCG GGG TTG 1620
  L   V   G   S   M   F   E   P   L   T   Q   V   C   I   D   R   G   W  521

GTT TAA GTT AGC AGA GCA CAC TTT AAA AGT GGA GGT GCC AGA GCT TTC TCA AAC 1674
  F   K   L   A   E   H   T   L   K   V   E   V   P   E   L   S   Q   T  539
```

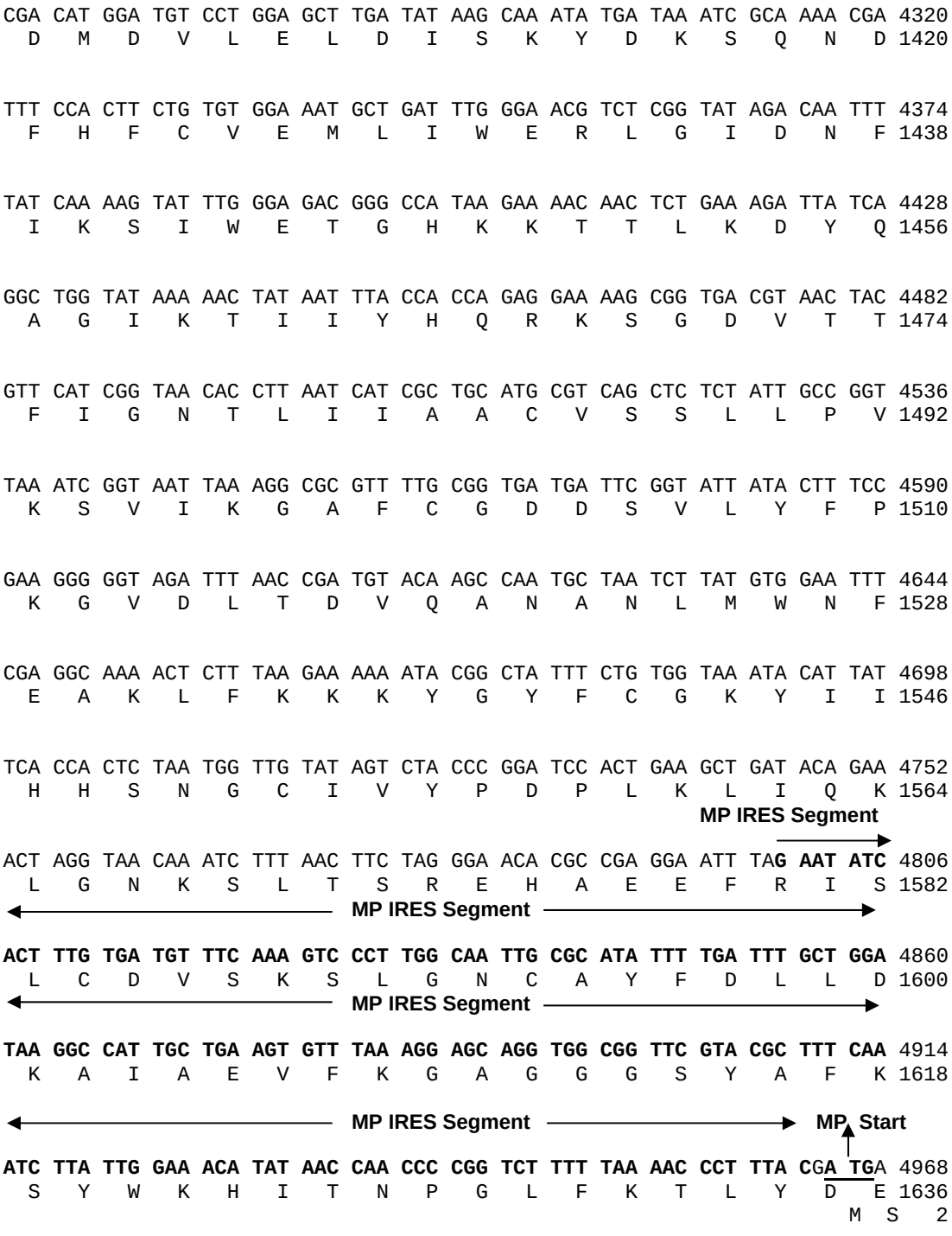
```
CTT CGA GAG TTA TAT GCA AAC GGA GTT TAG TGA GTC GAA AGC AAT CGA AGC ACT 1728
  F   E   S   Y   M   Q   T   E   F   S   E   S   K   A   I   E   A   L  557

GGA CGT CTG GAG CTA TCT GGA TGA GAG CAA TCA GCT TTA CAC CCA TGT CTC AAA 1782
  D   V   W   S   Y   L   D   E   S   N   Q   L   Y   T   H   V   S   K  575

ATT GAT GGA GAA GTA TAC AAT GCC TGA TTT TGA TGT GGA AAA ATT CAG GGA GCT 1836
  L   M   E   K   Y   T   M   P   D   F   D   V   E   K   F   R   E   L  593

CTG CGC GCT ACT AAA GGT GGG ACC GAA CGT TAT ATC AAC GGT CAT TGA AGC AGT 1890
  C   A   L   L   K   V   G   P   N   V   I   S   T   V   I   E   A   V  611

TCT TGA TAA TTC ACT AGG CTT CAC AAT TGT TGG TGG GAA ACC CGA GCA TCA ATA 1944
  L   D   N   S   L   G   F   T   I   V   G   G   K   P   E   H   Q   Y  629

TTC TGA TGA GAT GAA GGC CGC TGT CGC CGT AAG TTC TGA GTA CGG AAA AGG GAA 1998
  S   D   E   M   K   A   A   V   A   V   S   S   E   Y   G   K   G   K  647

AGA ATA TGG CAA CAT CGC CGC TGC GGC TAG CAC CTA CAC ACC TGA AGG CCC GCT 2052
  E   Y   G   N   I   A   A   A   A   S   T   Y   T   P   E   G   P   L  665

TAA AAA GGA GTG GGT TAA AAA AAA CGC TGT GCA AAT GCC TCT ACA GGC TAC AAC 2106
  K   K   E   W   V   K   K   N   A   V   Q   M   P   L   Q   A   T   T  683

GGA GAA GGT TCA ACG CTT TCA TTT GAG CGA GGA GAG CGA GCA TAT TAA CGT GGA 2160
  E   K   V   Q   R   F   H   L   S   E   E   S   E   H   I   N   V   E  701

AAA TCT GCA CAT GGT CGA CAC CAG CGT GTT CAC CAA GGC AAA AGC TTG CTT ACT 2214
  N   L   H   M   V   D   T   S   V   F   T   K   A   K   A   C   L   L  719

CTA CAC AGG CAC GTT AAA AGA GCA GCA GAT GAA GAA TTA CTT AGA TTA TCT CTC 2268
  Y   T   G   T   L   K   E   Q   Q   M   K   N   Y   L   D   Y   L   S  737

AGC ATC TAT CTG CGC CAC TAT AAG CAA CAT CGA AAA AGT CCA ACA TGA TTA CTG 2322
  A   S   I   C   A   T   I   S   N   I   E   K   V   Q   H   D   Y   W  755

GAC AAC TGG TAC TCA ACA GTA TCA GTC CTA CGG TCT GTG GGA TGT GAA GAA AAA 2376
  T   T   G   T   Q   Q   Y   Q   S   Y   G   L   W   D   V   K   K   K  773

GAA ATG GGT GTT GTG TCC ACC CAC GGC AGG TCA TGC ATG GGG GGT CGC GGC TAT 2430
  K   W   V   L   C   P   P   T   A   G   H   A   W   G   V   A   A   I  791

CAA AGG CAA GCA TCA CAT TGT GAT GCT CAA CTT TGA CGG TAA TGA TCA ACC TAT 2484
  K   G   K   H   H   I   V   M   L   N   F   D   G   N   D   Q   P   I  809

CTG TGA TAA GGA GTG GGA TTT AGT TTG TGT CAG CAA TGA TAC AAA ACT ATT TTC 2538
  C   D   K   E   W   D   L   V   C   V   S   N   D   T   K   L   F   S  827
```

```
CAC TAT GAA AAT TTT GGA AAA CCT GCA GGG TTG CAA GAT ACA GGA ACC GAA GGC 2592
  T   M   K   I   L   E   N   L   Q   G   C   K   I   Q   E   P   K   A  845

AAA GAT AAC ATT AGT GGA TGG GGT TCC GGG ATG TGG AAA AAC TGC CGA AAT CTT 2646
  K   I   T   L   V   D   G   V   P   G   C   G   K   T   A   E   I   L  863

GAG TAA GGC AGA TTT CAA GCG TGA TCT GAT TAT CAC GCA AGG AAA ACA GGC AGC 2700
  S   K   A   D   F   K   R   D   L   I   I   T   Q   G   K   Q   A   A  881

CGC TAT GAT AAG AAA AAG GGC AAA CGC CAA TGA TCC TAC GAG ACC AGC AAA CAA 2754
  A   M   I   R   K   R   A   N   A   N   D   P   T   R   P   A   N   N  899

CGA TAA CGT GAG AAC GGT GGA TTC GTT TTT AAT GAA TCC GAA ACC ATT TGA ATA 2808
  D   N   V   R   T   V   D   S   F   L   M   N   P   K   P   F   E   Y  917

CAA AGT CTT GTG GAT CGA TGA AGG GTT AAT GCT GCA CAC TGG TGC AAT CAG TTT 2862
  K   V   L   W   I   D   E   G   L   M   L   H   T   G   A   I   S   F  935

TTG CGT CGC ACT GAG CCA TTG TGA ACA GTG CTA CGT TTT TGG AGA TAC GCA GCA 2916
  C   V   A   L   S   H   C   E   Q   C   Y   V   F   G   D   T   Q   Q  953

AAT TCC CTT TAT TAA CAG AGT CAT GAA TTT TGA TTA TCC GAA GAC CTT GCA AAC 2970
  I   P   F   I   N   R   V   M   N   F   D   Y   P   K   T   L   Q   T  971

GCT GGT CAC GGA TTC TGT AGA GAA ACG GCG GAT TAC ATC GCG GTG TCC GCT AGA 3024
  L   V   T   D   S   V   E   K   R   R   I   T   S   R   C   P   L   D  989

TGT GAC CTG TTA TCT GGC GCA GAG GTA TAA GGG ACC CGT GGT GTC AAC GAA CAA 3078
  V   T   C   Y   L   A   Q   R   Y   K   G   P   V   V   S   T   N   N 1007

TGT TGA AAG GTC CCT TGA CAC AAA GCT CGT GGC TGG TGC CGC TCG CTT TGA ACC 3132
  V   E   R   S   L   D   T   K   L   V   A   G   A   A   R   F   E   P 1025

TCA GCT TAC ACC ATT ACC GGG CAA GGT AAT AAC TTT TAC CCA GGC TGA TAA AGA 3186
  Q   L   T   P   L   P   G   K   V   I   T   F   T   Q   A   D   K   E 1043

GAC ACT TAA AAA ACG CGG ATA TCA AGA TGT GCA TAC TGT GCA CGA GGT GCA AGG 3240
  T   L   K   K   R   G   Y   Q   D   V   H   T   V   H   E   V   Q   G 1061

TGA GAC CTA TAC TGA GGT GTC GTT AGT GCG TCT CAC GCC AAC ACC CCT GAG CAT 3294
  E   T   Y   T   E   V   S   L   V   R   L   T   P   T   P   L   S   I 1079

TAT TTC GCC TGA CTC TCC TCA TGT ATT AGT TAG TTT ATC TAG ACA TAA AAA TAG 3348
  I   S   P   D   S   P   H   V   L   V   S   L   S   R   H   K   N   R 1097

GTT AGT TTA TTA TAC CGT AGT TAA TGA CTG TAT TAC TTA TGC TAT AGG TGA GTT 3402
  L   V   Y   Y   T   V   V   N   D   C   I   T   Y   A   I   G   E   L 1115
```

```
AGC TAA GTT AAA TAA TTT TAT GTT TGA TAT GTA TAC TGT GGA GAA TCA GAA AGC 3456
  A   K   L   N   N   F   M   F   D   M   Y   T   V   E   N   Q   K   A 1133
```

Amber Stop Codon

```
ATA GCA ATT ACA GCA GCA TCG CCT CTA CAT CCA CAA AAA CAT CTT TTG TCC TAC 3510
  *   Q   L   Q   Q   H   R   L   Y   I   H   K   N   I   F   C   P   T 1150

TCC TAA AAC CGG AAC CAT TTC TGA CTT ACA GGA GTA TTA CGA TAG GTG TTT ACC 3564
  P   K   T   G   T   I   S   D   L   Q   E   Y   Y   D   R   C   L   P 1168

TGG GAA TAG CTT CGT CCT TAA TGA TTA CGA TCG AGT GTG TAT GAG ACT CGT TGA 3618
  G   N   S   F   V   L   N   D   Y   D   R   V   C   M   R   L   V   D 1186

CAA CAC GTT CAA CGT TGA ACC ATG CAC ACT GAC TTT GAG TAA AGC TGA CCC TGT 3672
  N   T   F   N   V   E   P   C   T   L   T   L   S   K   A   D   P   V 1204

GGA GGC AGA GAT GCG ATT AAA TAG CAA GAA TTT TTT GAC TCC ATT GTT GCG GAC 3726
  E   A   E   M   R   L   N   S   K   N   F   L   T   P   L   L   R   T 1222

TGC CGT GGA AAA ACC CAG GGT GCC TGG GTT GTT GGA GAA TCT CGT TGC TTT AAT 3780
  A   V   E   K   P   R   V   P   G   L   L   E   N   L   V   A   L   I 1240

TAA GAG GAA TTT TAA TTC CCC TGA TTT GGC TGG GAT AAT CGA CAC CAA GCA GGT 3834
  K   R   N   F   N   S   P   D   L   A   G   I   I   D   T   K   Q   V 1258

TGC AAC TAA TGT AGT CGA TAA ATT CTT TTC TGA TTA CTT TAC CGT AGA AGG TTT 3888
  A   T   N   V   V   D   K   F   F   S   D   Y   F   T   V   E   G   L 1276

AGA TCA TAT AAA TAG ATG TAG AAA TAA TAA GTT TTT GTA TAG TAA AGA GAA CCT 3942
  D   H   I   N   R   C   R   N   N   K   F   L   Y   S   K   E   N   L 1294

TTC TGT ATG GTA CGA TAA ACA GCA GCC GCA GTC CAT TGG GCA GTT GGC TAA TGC 3996
  S   V   W   Y   D   K   Q   Q   P   Q   S   I   G   Q   L   A   N   A 1312

TGA TTT CAT AGA TTT CCC TAG TGT TGA TGA GTA CAA GCA TAT GAT CAA GAC TCA 4050
  D   F   I   D   F   P   S   V   D   E   Y   K   H   M   I   K   T   Q 1330

ACC CAA GGG AAA ACT TGA CTG CAG CAT TCA ATC TGA ATA TCC TGC CCT ACA GAC 4104
  P   K   G   K   L   D   C   S   I   Q   S   E   Y   P   A   L   Q   T 1348

TAT AGT GTA TCA TAG TAA ACA TGT AAA TGC TAT TTT TGG TCC TAT TTT TAG TGA 4158
  I   V   Y   H   S   K   H   V   N   A   I   F   G   P   I   F   S   E 1366

ATT TAC TAG AGT GAT TTT AGA ATC CTT AGA TAC TAA CAA ATT TTT CTT TTA TAC 4212
  F   T   R   V   I   L   E   S   L   D   T   N   K   F   F   F   Y   T 1384

TAG GAA AAC TGT AGA GGA CTT GGA GTT GTT TTT CAG CGA TCT TTC TGA TTC AAG 4266
  R   K   T   V   E   D   L   E   L   F   F   S   D   L   S   D   S   S 1402
```


CGA CAT GGA TGT CCT GGA GCT TGA TAT AAG CAA ATA TGA TAA ATC GCA AAA CGA 4320
D M D V L E L D I S K Y D K S Q N D 1420
TTT CCA CTT CTG TGT GGA AAT GCT GAT TTG GGA ACG TCT CGG TAT AGA CAA TTT 4374
F H F C V E M L I W E R L G I D N F 1438
TAT CAA AAG TAT TTG GGA GAC GGG CCA TAA GAA AAC AAC TCT GAA AGA TTA TCA 4428
I K S I W E T G H K K T T L K D Y Q 1456
GGC TGG TAT AAA AAC TAT AAT TTA CCA CCA GAG GAA AAG CGG TGA CGT AAC TAC 4482
A G I K T I I Y H Q R K S G D V T T 1474
GTT CAT CGG TAA CAC CTT AAT CAT CGC TGC ATG CGT CAG CTC TCT ATT GCC GGT 4536
F I G N T L I I A A C V S S L L P V 1492
TAA ATC GGT AAT TAA AGG CGC GTT TTG CGG TGA TGA TTC GGT ATT ATA CTT TCC 4590
K S V I K G A F C G D D S V L Y F P 1510
GAA GGG GGT AGA TTT AAC CGA TGT ACA AGC CAA TGC TAA TCT TAT GTG GAA TTT 4644
K G V D L T D V Q A N A N L M W N F 1528
CGA GGC AAA ACT CTT TAA GAA AAA ATA CGG CTA TTT CTG TGG TAA ATA CAT TAT 4698
E A K L F K K K Y G Y F C G K Y I I 1546
TCA CCA CTC TAA TGG TTG TAT AGT CTA CCC GGA TCC ACT GAA GCT GAT ACA GAA 4752
H H S N G C I V Y P D P L K L I Q K 1564
MP IRES Segment
ACT AGG TAA CAA ATC TTT AAC TTC TAG GGA ACA CGC CGA GGA ATT TAG AAT ATC 4806
L G N K S L T S R E H A E E F R I S 1582
MP IRES Segment
ACT TTG TGA TGT TTC AAA GTC CCT TGG CAA TTG CGC ATA TTT TGA TTT GCT GGA 4860
L C D V S K S L G N C A Y F D L L D 1600
MP IRES Segment
TAA GGC CAT TGC TGA AGT GTT TAA AGG AGC AGG TGG CGG TTC GTA CGC TTT CAA 4914
K A I A E V F K G A G G G S Y A F K 1618
MP IRES Segment
MP Start
ATC TTA TTG GAA ACA TAT AAC CAA CCC CGG TCT TTT TAA AAC CCT TTA CGA TGA 4968
S Y W K H I T N P G L F K T L Y D E 1636
M S 2
GTC TTA ATC TGG CTG ACC CAA TGA AGT TAT TAG ATG AGT CGA CTG TGT TTA AAA 5022
S * 1637
L N L A D P M K L L D E S T V F K T 20
CTA ACA TTT TAG ATA GGT TTA GGG GTG GAA TGA AAA AGA ACT TGA AGG CCT GCA 5076
N I L D R F R G G M K K N L K A C N 38

```
ACG TCG TCG ATT GCG ACG TGG CTA AGC TTT CAG GTA ATG GTC AGT CAC TAT TTA 5130
   V   V   D   C   D   V   A   K   L   S   G   N   G   Q   S   L   F   S  56

GTG TTA ACC TTA CTT CTA ATT TAA CTT TAG AAC AAA AAT CGT ATA AAT ATG TAC 5184
   V   N   L   T   S   N   L   T   L   E   Q   K   S   Y   K   Y   V   H  74

ATT TGT TAG CTA TAG TAA TTT CTG GTA AAT GGC ATC TGC ATA ACA CGG TTG GTG 5238
   L   L   A   I   V   I   S   G   K   W   H   L   H   N   T   V   G   G  92

GTG GTG CGA CTG TAG CTT TGT TAG ATA AGC GTA TAA CAA ATC CAC AAC AAG CTC 5292
   G   A   T   V   A   L   L   D   K   R   I   T   N   P   Q   Q   A   Q 110

AGT TGG CAA AGA TGG TTG TGG CTG CGA AGG TAG GGG AGT TTC AGG GTA GGC TTA 5346
   L   A   K   M   V   V   A   A   K   V   G   E   F   Q   G   R   L   T 128

CGC CAA ATT ACT CTG TTA CTC TGG AGG ATG CTC TCA AAG ACC CGT GGG AAG TCT 5400
   P   N   Y   S   V   T   L   E   D   A   L   K   D   P   W   E   V   F 146

TCA TAA GCA TTA AGG ATG TGC CGA TAG AAC AAG GTT TTC ATC CTT TGT CAG TTG 5454
   I   S   I   K   D   V   P   I   E   Q   G   F   H   P   L   S   V   E 164

AGA TTG CAG CAT TAC TAA TGT TTT CAG ACA TTT GTA TAC TAA AGT CTC TAC GGG 5508
   I   A   A   L   L   M   F   S   D   I   C   I   L   K   S   L   R   A 182

CAA AGA TGA TAG AAT CAC AAA TCG GTT TTA CTG ATG GAA CCG ACG GCG ACA TAC 5562
   K   M   I   E   S   Q   I   G   F   T   D   G   T   D   G   D   I   P 200

CTT CTG AGA ATA TGG ATG CAT TTG TAC AGA GTG CTT CTA TAA TGC AAG ATT TGA 5616
   S   E   N   M   D   A   F   V   Q   S   A   S   I   M   Q   D   L   R 218
                                                        CP IRES Segment
                                                              ------>
                                                              PPT I
GAA GAC TCA GGA TGG CCA ATA ATT CTC TTG GAA GTG GTG TAA TTG GGG GT[A GAG] 5670
   R   L   R   M   A   N   N   S   L   G   S   G   V   I   G   G   R   D 236

<----------------          CP IRES Segment          ---------------->
            PPT I
[ATA AGA AAG GGA AAA AG]T TTG TAA AGG ATT TTA AGA AGA AAA CTG TGG TTA GTG 5724
   K   K   G   K   K   F   V   K   D   F   K   K   K   T   V   V   S   D 254
<----------------          CP IRES Segment          ---------------->

ATA ATT CCG ATA CTG ATT TTT CAG GAG CGG AGA GTT CGT ATA CTA CAG AAT TCG 5778
   N   S   D   T   D   F   S   G   A   E   S   S   Y   T   T   E   F   D 272

------ CP IRES Segment ------>  CP Start
                PPT II             ^
ACT C[AA GAT CAG AAA CAA GCG] ATG CCT TAC CTT AAT TTG ACA CCT TTG AAT TTA 5832
                                M   P   Y   L   N   L   T   P   L   N   L  11

   S   R   S   E   T   S   D   A   L   P   *                                 282
```

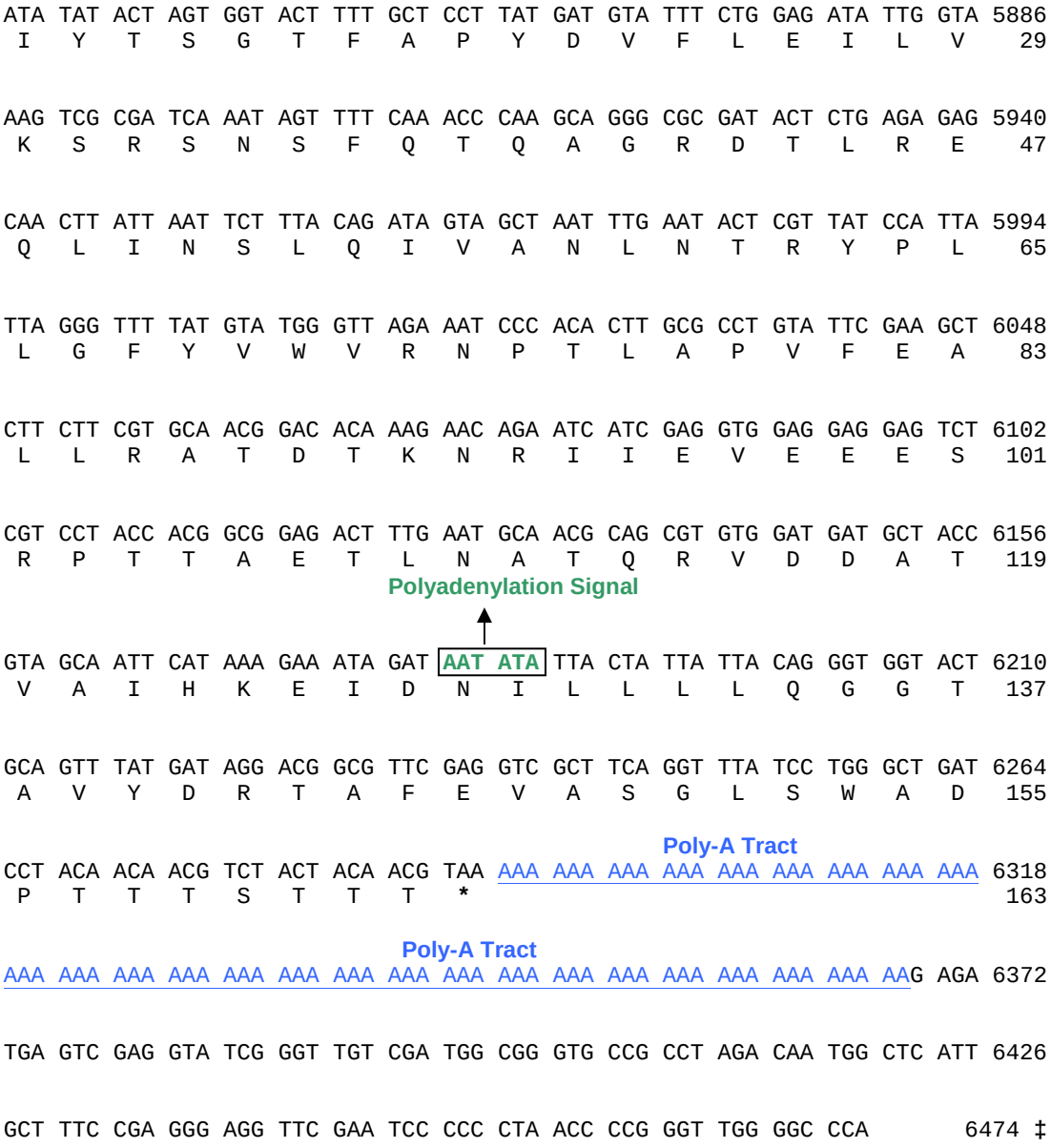


Fig 3.7 Complete nucleotide sequence of HVS genomic RNA and its encoded amino acid sequences for individual ORFs. Nucleotide sequences of 5' UTR, RdRP, MP, CP ORFs and 3' UTR will appear in GenBank database under accession numbers AF 400156, AF 400157, AF395899, AF 395898 and AY 497578 respectively. Underlines and asterisk indicate start and stop codons. PPT I and PPT II are polypurine tracts required for CP IRES activity. Presence of poly- A tract within 3'UTR region, is similar to BSMV genome. ‡ Indicates length varies between 6474 – 6493 due to variation in size of poly-A tract (77 – 96).

of 126 kDa protein. Subsequently, the 5' guanosine is methylated by transfer methyl group from *S*-adenosyl-methionine (Adomet) (Furuichi et al., 1976). The TMV 126 kDa protein was shown to have guanylyltransferase activity (Dunigan and Zaitlin, 1990).

In HVS, the putative MT domain was located at amino acid positions 71- 454. This region has four distinct motifs at amino acid positions 72-94, 131-138, 213-263 and 278-293 (Fig. 3.8a). A putative Adomet binding site was located in motif II and a putative guanylyltransferase activity could be located to motif IV. The presence of an invariant histidine in motif I and Asp-X-X-Arg signature site in motif II at amino acid positions 82 and 135-138 respectively by analogy to nsP1 protein in Sindbis virus and other tobamoviruses may be responsible for MT activity (Alonso et al., 1991). Motifs Ia1 and Ia2 were located at amino acid positions 104-112 and 121-129, respectively.

A putative helicase domain of HVS is located between amino acid positions 849-1095. Helicase domain can be subdivided into six motifs at amino acid positions 849-863, 919-930, 945-956, 982-990,1055- 1063 and 1084-1095, respectively (Fig. 3.8b). Based on the conservation of these six motifs, tobamoviruses have been classified as Sindbis-like viruses (Goldbach and Wellink, 1988). The NTP binding property has been attributed to motifs I and II. In motif I, a GXXXXGKT consensus sequence is located at amino acid positions 852-859 (Fig. 3.8b). Motif II was located at amino acid positions 919-930 with residue D at position 923 a part of the DEAD box (Linder et al., 1989). Recent work done to ascertain the function and requirements of TMV helicase domain have shown that ATPase activity requires Mg^{2+} as a co-factor. Substituting Mg^{2+} with Ca^{2+} or Mn^{2+}

Fig 3.8 (a) METHYLTRANSFERASE MOTIFS

```
                      I                      Ia1              Ia2       II
          72                    94        104     112    121     129 131  138
HVS       ISYVGNRNAVHSLAGGLRKLEMEYLMTLVPYGSPTYDIGGNYAQHLLKGRSYVHCCNPCLDFRDIAR
CGMMV-SH  ..fTATk.s.........L..l..m.MQ......C.......T...F..............Lk.v..
KGMMV-Y   v..T.sALs.........Y..G....MQ......ChY.....s..m...............Lk.v..
ORSV-S1   .tfYNTqL....m.....A..l....MQi.f..I.......fsA..Y...D.....M.N..I..v..
TVCV      ..fTHTqs..........S..l....MQ..f..L.......fsA..F...D.....M.N..V.....
TMV-Type  .tfYNTq...........S..l....MQi....L.......f.S..F...a.....M.N..V...M.
ToMV      .tfYNTq...........S..l....MQi....L.......f.S..F...a.....M.N..V...M.
PMMV-S    .tfYNTq...........S..l....MQi....T.......f.A.mF...D.....M.Nm.L..vM.
TMV-Ob    .tfYNTqL......A...S..l....MQ.....L.......f.A..F...D.....M.N..L...M.
TMGMV     .tfYNTq...........A..l....LQ.............f.A..F...D.....M.N..I...M.
SHMV      .EfT.AS.TS.am.A...G..l...Y......aVs......fPA.mm...........A..A..l..
Consensus itf Ntqnavhslagglrsleleylmmqvpygs tydiggnfaahlfkgr yvhccMpNldiRdi r
```

```
                                   III
          213                                              263
HVS       FAISVHSLYDIPVDELGAALLRKNVKTLYAALHFSEELLMGATEARLGTTGAKF
CGMMV-SH  H.v.l..i....YSSi.P..H....rVC...F....A..l.sPVGN.Ns...q.
KGMMV-Y   y.val..i..v.Ceqi.P......i.V.f..F....d..i.sEFG..PNv..F.
ORSV-S1   y.vGl..i....A..F.......d.HIC...F....N..lET.S.P.DE...T.
TVCV      y.val..i.....e.F.s.........Cf..F..H.Nm.lDCDTVT.DE...T.
TMV-Type  y..al..i....A..F.........H.C...F....N..lEDsYVN.DE.N.C.
ToMV      y..al..i....A..F.........HVC...F....N..lEDsHVN.DE.N.C.
PMMV-S    y.val.......A..F......r..HVC...F....N..lEDsYVS.DD...F.
TMV-Ob    y...l.......A............H.....F..a....lEVsTVE.P...GI.
TMGMV     y.val........H.F....iS..iHVC...SILa.A..lDQ..VT.NE...T.
SHMV      y.T.l.......YQN..P..A..Ri.V.h..F....d..l..s.GL.TQ..GT.
Consensus yavAlhsiydip de gaallrknv vcyaafhfseNllle s   lDEiga f
```

```
                    IV
          278               293
HVS       GESTLLYTHSFKNICAYITRSFFYA
CGMMV-SH  E....H....Le..KLIvM.ty.P.
KGMMV-Y   D....H.....S..RKIv..t..P.
ORSV-S1   N....N.E..y..vIK.vCkt..P.
TVCV      N....N.....S..IK.vCkt..P.
TMV-Type  S....N.C..yS..LK.vCkty.P.
ToMV      S....N.s..yS..LK.vCkty.P.
PMMV-S    A....N....yS.vLK.vCkty.P.
TMV-Ob    N....N.s..yS.lLK.vCkty.P.
TMGMV     D....N.s.Ky...LH.vVk.y.P.
SHMV      D..s.i.....r.vFE.v..t..V.
Consensus NestlNythsy nilkyv ktyfpa
```

```
          409                                                 454
HVS       EDFVFTVLNHIRTYSGKQLSYENVLSFVESIRSRVVANGANVRSEW
CGMMV-SH  A...H.ii...S..DN.A.Vwk..Q.........i...Vs.k...
KGMMV-Y   AT.iH..i...C..DN.A.twR..Q.....FG-LGCm..Vs.....
ORSV-S1   k...y.........QD.A.t.k.............ii..VTA....
TVCV      k...y......k..QA.A.t.A.............ii..VTA....
TMV-Type  k.............QA.A.t.A.............ii..VTA....
ToMV      k...y.........QA.A.t.S.............ii..VTA....
PMMV-S    k...y.........QS.A.t.A.............ii..VTA....
TMV-Ob    k...y.........QA.A.t.N.............ii..VTA....
TMGMV     R...y.........QA.A.t.q.............ii..VTA....
SHMV      k..Yw.A....L..PDGkADfRG.m..l........i..TTTA.q.
Consensus kdfvytvlnhiRtyq kalty nvlsfvesirsrviingvTarsew
```

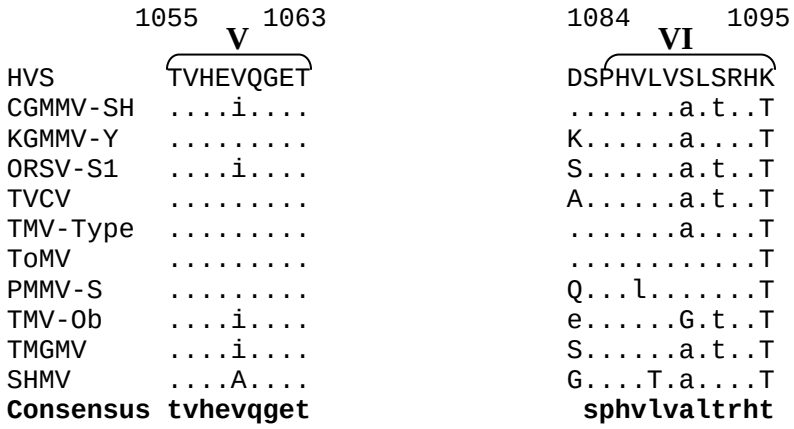
(b) RNA HELICASE MOTIFS

```
                 849            863
                         I
HVS        AKITLVDGVPGCGKTAEIL
CGMMV-SH   VH................i
KGMMV-Y    ..F..i.........Q..i
ORSV-S1    P.v............K...
TVCV       ..vI...........K..i
TMV-Type   ..vV...........K...
ToMV       ..vV...........K...
PMMV-S     ..v............K...
TMV-Ob     ..v............K...
TMGMV      ..mV..........YKGDF
SHMV       .T.............Q...
Consensus  Akvtlvdgvpgcgktkeil
```

```
             919         930
                  II
HVS        KVLWIDEGLMLHTGAISFCVALSHCE
CGMMV-SH   DAvyv.....v...Lln.AlKi.G.k
KGMMV-Y    .T..v.....v...Lln...NiaKVk
ORSV-S1    NK.f..........Cvn.L......R
TVCV       .R.f..........Cvn.LlL..Q.d
TMV-Type   .R.f..........Cvn.L..m.L..
ToMV       .R.f..........Cvn.L.Em.L.d
PMMV-S     .R.fl.......P.Cvn.L.Gm.L.S
TMV-Ob     .R.fv.......P.CvY.L.K..L.N
TMGMV      .R.f..........Cvn.L.Li.G.d
SHMV       .T..v.....v.P.L.n..iNi.CVS
Consensus  krlfideglmlhtgcvnflv ls cE
```

```
             945       956
                 III
HVS        QCYVFGDTQQIPFINRV
CGMMV-SH   kAf....Ak........
KGMMV-Y    eVKi....k........
ORSV-S1    eAM....Ae........
TVCV       VA..y...k.....C..
TMV-Type   IA..y.......y....
ToMV       IA..y.......y....
PMMV-S     eAf.y.......y....
TMV-Ob     eAf.........y....
TMGMV      IA.iy............
SHMV       SV.i...Rk........
Consensus   ayvygdTqqipfinrv
```

```
982      990
    IV
ITSRCPLDV
v.H...R..
TsK...R..
lsL...A..
v.L...A..
T.L...A..
T.L...A..
T.L...A.i
T.L...v..
T.L...G..
T.K......
Ttlrcp dv
```


1055
1063
V
1084
1095
VI
HVS TVHEVQGET DSPHVLVSLSRHK
CGMMV-SH ....i.... ....a.t..T
KGMMV-Y .... K......a....T
ORSV-S1 ....i.... S......a.t..T
TVCV .... A......a.t..T
TMV-Type ....a....T
ToMV ....T
PMMV-S .... Q...l.......T
TMV-Ob ....i.... e......G.t..T
TMGMV ....i.... S......a.t..T
SHMV ....A.... G....T.a....T
Consensus tvhevqget sphvlvaltrht

(c) RNA POLYMERASE MOTIFS

```
                1401      A      1418
HVS          DLS-DSSMDVLELDISKYDKSQ
CGMMV-SH     ...SH.DYei....v.......
KGMMV-Y      T..AKE.Yei....v.......
ORSV-S1      ..DSTVP.e.............
TVCV         ..DSTQa.ei............
TMV-Type     ..DSHVP...............
ToMV         ..DSHVP.......v.......
PMMV-S       ..DSnVP..i............
TMV-Ob       ..DAHQP.e.....v.......
TMGMV        ..DSHVP...............
SHMV         .mV-VGD.ei............
Consensus dlds   mdvleldiskydksq
```

```
              1460        B        1481
HVS          KTIIYHQRKSGDVTTFIGNTLII
CGMMV-SH     ..l..y.............F..
KGMMV-Y      ..v..y.............F..
ORSV-S1      ..Clwy.............v..
TVCV         ..Clwy.............i..
TMV-Type     ..C.wy.............v..
ToMV         ..Clwy.............v..
PMMV-S       ..Clwy.............i..
TMV-Ob       ..Clwy.............v..
TMGMV        ..Clwy.............v..
SHMV         ..v.Ey.............i..
Consensus ktclwyqrksgdvttfigntiii
```

```
          1497   C   1508          1521          D          1546
HVS        KGAFCGDDSVLYFPKGVDLTDVQANANLMWNFEAKLFKKKYGYFCGKYIIHHSNGCIVYPDPLKLIQKLGNKS
CGMMV-SH   .As......li.L...leYP.i..T...v.......r.............a.............S......
KGMMV-Y    .As......li.M.P.leYP.i..T...v.........r........v.................S......
ORSV-S1    .A.......i..I...l..P.i.sG...........yr.r......r.....DR.A...Y......S...C.H
TVCV       .A.......li.I...l..P.i..G...........r........r.v...DR.A...Y......S...C.H
TMV-Type   .........l......CeFP...Hs.............q......r.v...DR.....Y......S...A.H
ToMV       .........l......CeYP.i.QA.............q......r.v...DR.....Y......S...A.H
PMMV-S     .........i......T.FP.i.QG...l........r.r......r.....DR.....Y......S...A.H
TMV-Ob     .A.......lv.L...Ce.Pni.sC.............T......r.v...DR.A...V....i.S...A.H
TMGMV      ...............l.FP.i.sC...........yr.r......r.....DK.A...Y......S...A.H
SHMV       .AG......ii.L.rNlLYP.i.sVs.N...........Lh.....r..lRnGRYLRlL.....i.T...C.a
Consensus k afcgddslly pkgldypdiqs anlmwnfeaklfkkrygyfcgryiihhdRg ivy dplklisklg kh
```

Fig 3.8 Amino acid sequence comparison of HVS RdRp protein with ten other tobamoviruses. CLUSTAL X (Thompson et al., 1997) was used to generate multiple sequence alignments. **(a)** Methyl transferase motifs **(b)** Helicase motifs and **(c)** Polymerase motifs. Dots indicate identical amino acids. Consensus sequence includes highly conserved and partially conserved amino acids

resulted in reduction of ATPase activity. The ssRNA and dsRNA binding activities of this protein was down-regulated by the addition of ATP. The *in vitro* expressed helicase protein was capable of unwinding dsRNA and the helicase molecules interact among themselves to form oligomers, which is a pre-requisite for unwinding of secondary structures thereby facilitating replication (Goregaoker and Culver, 2003).

Membrane-bound RdRp extracts of ToMV infected plants have been shown to replicate template length RNA molecules of ToMV *in vitro* and are template specific (Osman and Buck, 1996). The polymerase activity resides in the carboxy terminal of 183 kDa readthrough product. The polymerase module consists of four domains (A to D) (Poch et al., 1989). These modules are found in all DNA dependent and RNA dependent RNA polymerases and are involved in elongation of nucleotide chains (Quadt and Jaspars 1989). They were located at amino acid positions 1401-1418 (motif A), 1460-1481 (motif B), 1497- 1508 (motif C) and 1521-1546 (motif D) in the HVS readthrough ORF (Fig. 3.8 c). The Gly-Asp-Asp motif (GDD) was located at amino acid position 1502-1504 and is surrounded by hydrophobic residues. The consensus sequence SGXXXTXXXNT is present preceding the GDD box at amino acid positions 1469-1479.

An ORF encoding for a putative 52 kDa protein was located between nucleotides 3605 to 4973. The MP IRES segment (nt 4800- 4965) and a putative MP sg promoter (nt 4816-4950) were located within the 3' end of the RdRP region (Fig. 3.7).

Phylogenetic trees were generated based on nt and aa sequences of 126/183kDa ORFs of HVS and ten other tobamoviruses. HVS always clustered with subgroup

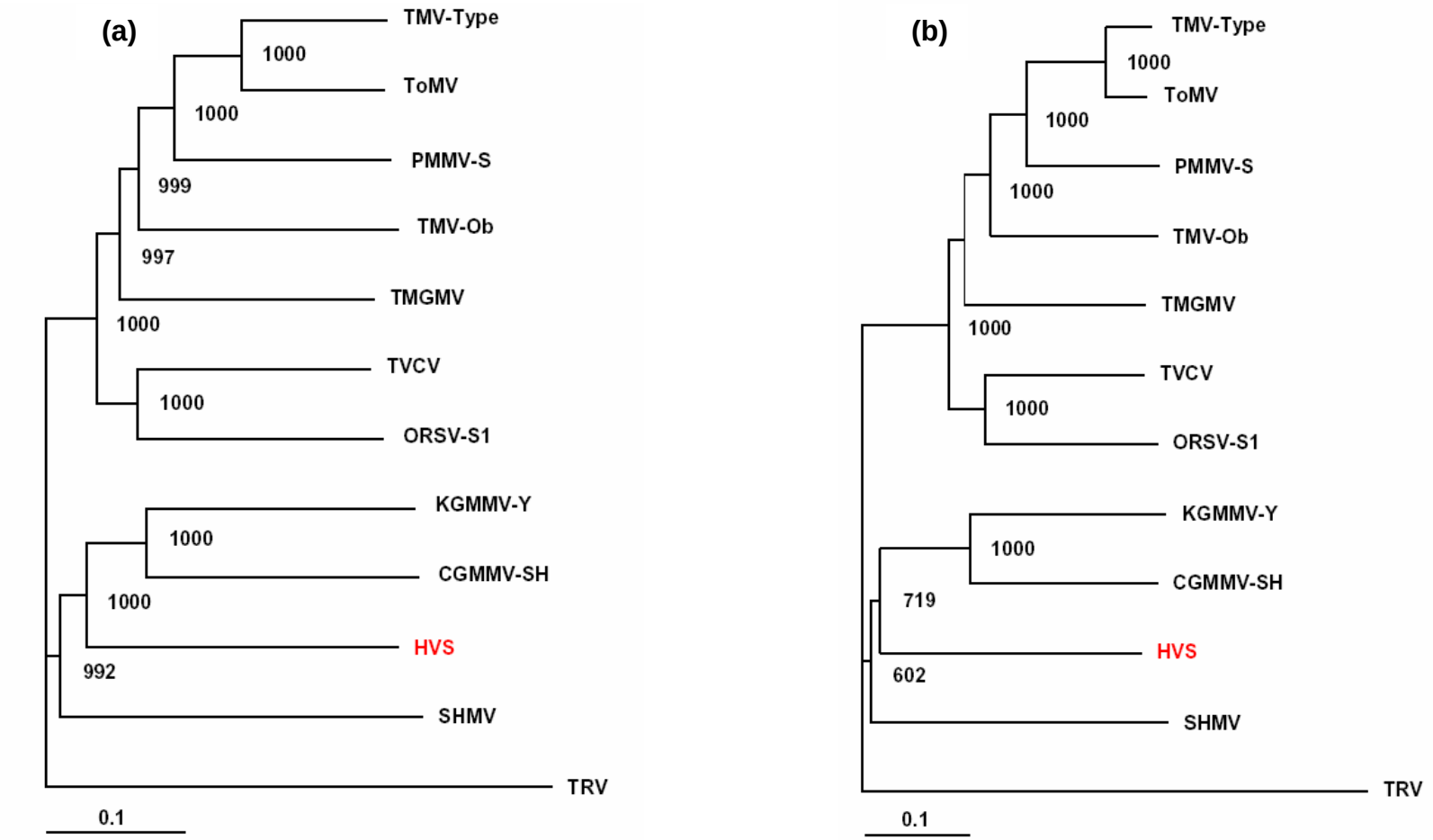


Fig. 3.9 Phylogenetic trees generated from multiple sequence alignment of nucleotide **(a)** and amino acid sequences **(b)** of 126kDa ORF illustrating the position of HVS among the members of tobamoviruses. The percentages of bootstrap replicates are indicated at the nodes. The trees were rooted with TRV as an outgroup.

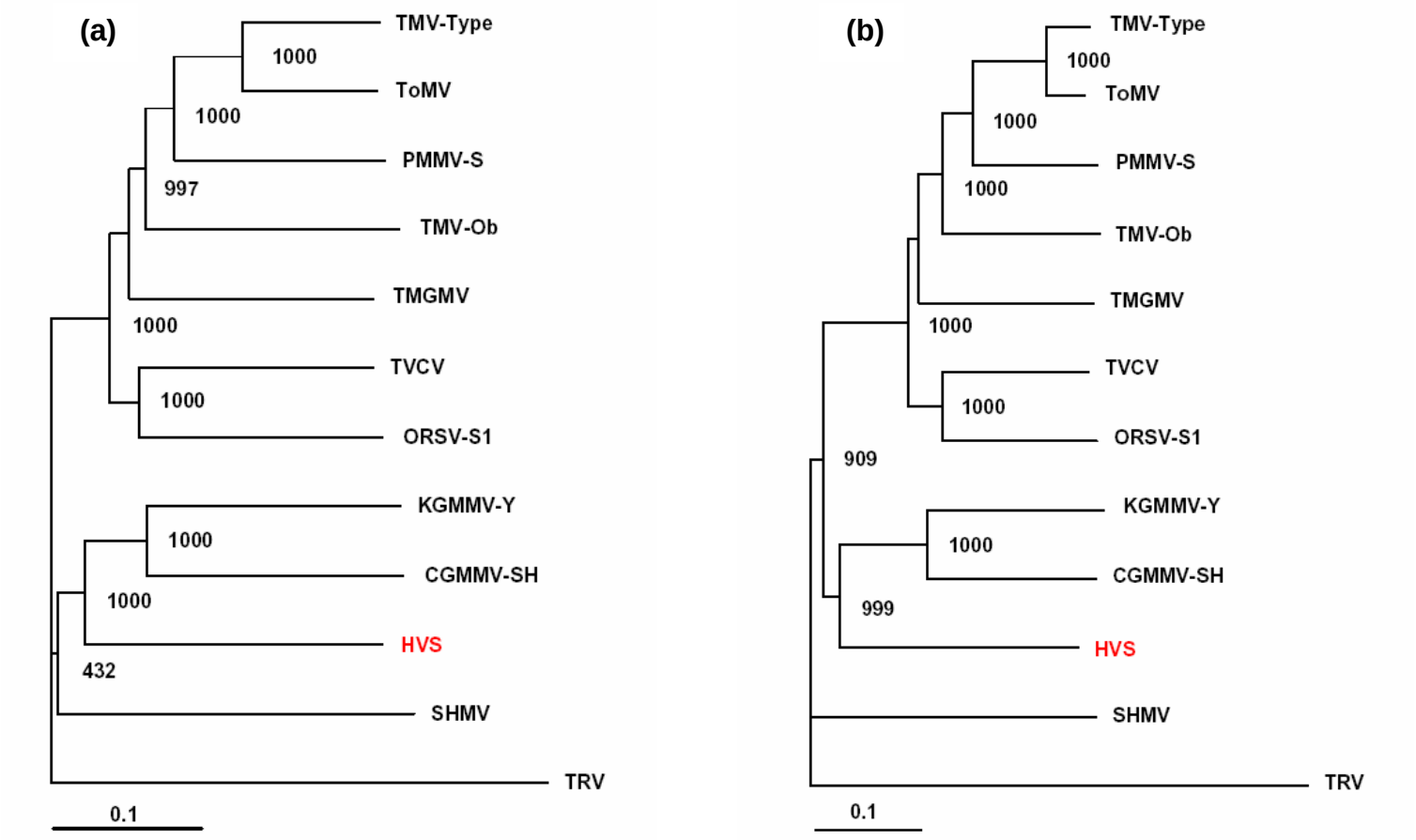


Fig. 3.10 Phylogenetic trees generated from multiple sequence alignment of nucleotide **(a)** and amino acid sequences **(b)** of 183 kDa ORF, illustrating the position of HVS among the members of tobamoviruses. The percentages of bootstrap replicates are indicated at the nodes. The trees were rooted with TRV as an outgroup.

II tobamoviruses (Fig. 3.9a&b, 3.10a&b). Percentage identities of nt and aa sequences of 126 kDa product ranged from 49.14 % (PMMV-S) to 51.08 % (SHMV) and 52.91% (CGMMV-SH & KGMMV-Y) to 55.02% (TMV- type strain (Table 3.2). The percentage identities of nt and aa sequences of 183 kDa product ranged from 51.34 (SHMV) to 53.51% (KGMMV-Y & TMV-Ob) and 56.78% (TMGMV) to 59.62% (KGMMV-Y) (Table 3.2). These results show that HVS is closely related to subgroup II tobamoviruses and its low percentage identities with all other tobamoviruses indicate it is equally distant from solanaceous, legume, cucurbit, crucifer and orchid infecting tobamoviruses.

3.7.3 Movement protein

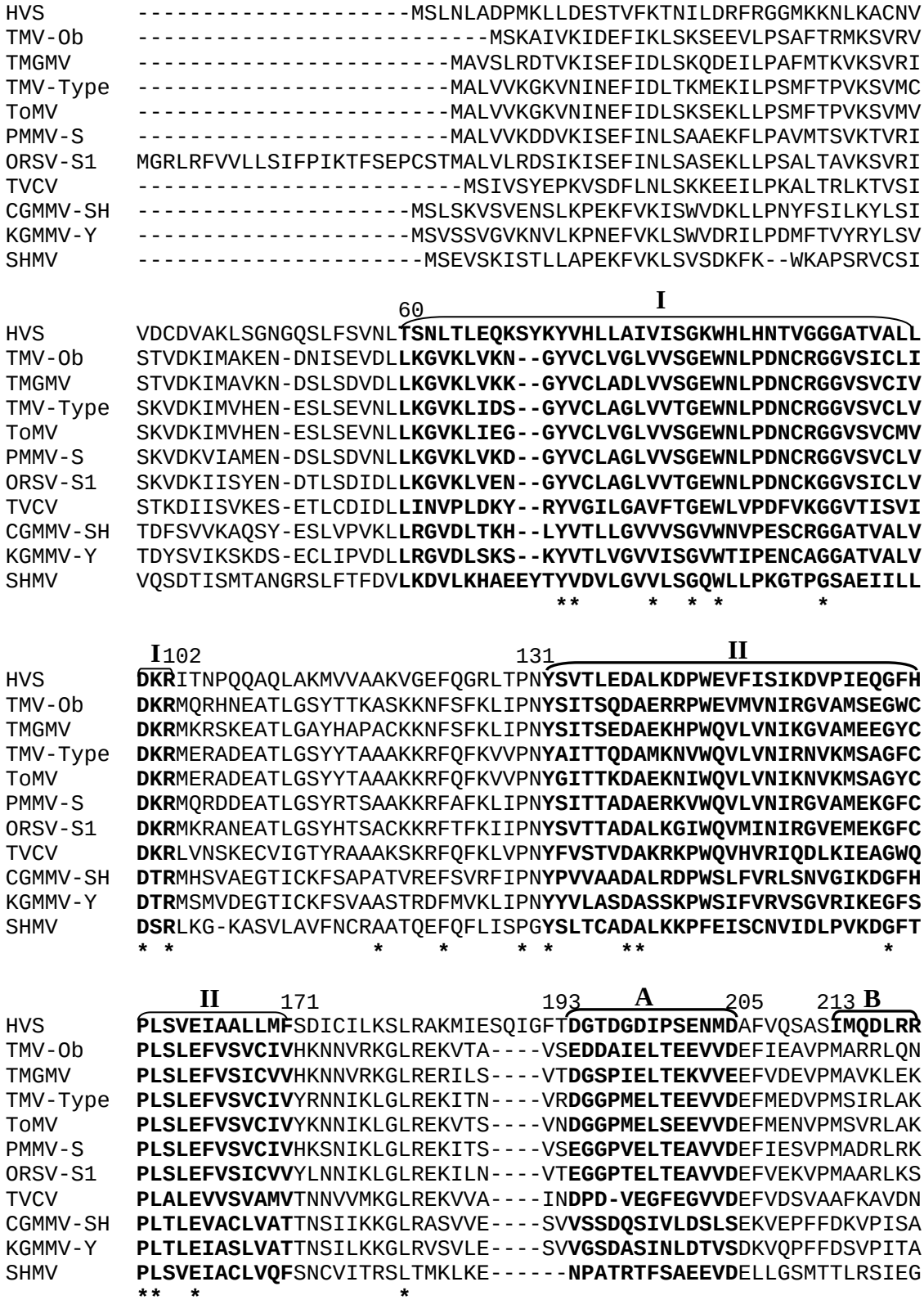
The MP start codon is located at nt 4965 and the termination codon is located at nt 5810. Thus the MP coding region overlaps with the RdRp and CP by 8 and 11 nts, respectively. The MP is composed of 282 amino acids with a calculated molecular mass of 30,891 Da. The 30 kDa protein of tobamoviruses facilitates cell to cell movement by modifying the plasmodesmata size exclusion limit (Wolf et al., 1989). Multiple alignment of tobamoviral 30 kDa proteins revealed 5 conserved regions (Saito et al., 1988). In the HVS MP, motifs I and II are located at amino acid positions 60 -102 and 131-171 respectively (Fig. 3.11). The N-terminal residue G (Glycine) within motif I, was thought to be conserved in the vast majority of tobamoviruses (Koonin et al., 1991). The only exceptions are TMGMV where it is replaced by aspartic acid and in HVS it is replaced by alanine at position 77. Another G residue at position 82 in HVS MP seems to be conserved in the 11 tobamovirus sequences that are aligned. An alanine residue at amino acid position 100 in HVS is conserved among all tobamo-, tobra-, and

Table 3.2 Percentage sequence identities of HVS nucleotide and amino acid sequences with other tobamoviruses

Virus Species	126 kDa nt	126 kDa aa	183 kDa nt	183 kDa aa	MP nt	MP aa	CP nt	CP aa	5' UTR	3' UTR
OAS located in CP										
SHMV	51.08	53.36	51.34	55.71	42.17	34.40	51.63	57.06	31.03	26.74
CGMMV-SH	49.62	52.91	53.22	59.44	43.90	37.88	51.44	52.80	46.55	37.50
KGMMV-Y	49.74	52.91	53.51	59.62	43.09	35.88	48.37	46.63	28.07	32.32
OAS located in MP										
TMV-Type	50.07	55.02	53.23	58.02	39.65	30.97	49.58	47.17	39.66	43.32
PMMV-S	49.14	54.16	53.11	58.19	39.92	32.68	48.73	45.86	44.83	42.25
TMV-Ob	49.82	54.44	53.51	57.74	41.82	31.75	48.15	45.34	39.66	39.04
TMGMV	50.42	55.00	53.48	56.78	43.35	33.20	46.88	49.06	44.83	44.92
ToMV	50.37	54.30	53.45	57.74	42.26	32.20	48.33	46.54	39.66	39.57
ORSV-S1	49.39	53.96	52.82	57.29	39.46	30.85	49.48	46.84	41.38	37.43
TVCV	49.82	54.20	52.12	57.71	38.81	32.21	48.10	48.41	50.00	45.99
*** FHV**	---	---	---	---	---	---	**71.07**	**79.75**	---	---

[Nt] Nucleotide sequence. [aa] Aminoacid sequence.***** Location of OAS not known for FHV. --- indicates data not available. Sequences of viruses were obtained from NCBI GenBank database. SHMV (J02413, U47034), CGMMV-SH (D12505), KGMMV-Y (AB015145), TMV (V01409), PMMV-S (M81413), TMV-Ob, (D13438), TMGMV (M34077), ToMV (X02144), ORSV-S1 (U34586), TVCV (U03387), FHV (AY250831).

Fig 3.11


HVS ----------MSLNLADPMKLLDESTVFKTNILDRFRGGMKKNLKACNV
TMV-Ob ----------MSKAIVKIDEFIKLSKSEEVLPSAFTRMKSVRV
TMGMV ----------MAVSLRDTVKISEFIDLSKQDEILPAFMTKVKSVRI
TMV-Type ----------MALVVKGKVNINEFIDLTKMEKILPSMFTPVKSVMC
ToMV ----------MALVVKGKVNINEFIDLSKSEKLLPSMFTPVKSVMV
PMMV-S ----------MALVVKDDVKISEFINLSAAEKFLPAVMTSVKTVRI
ORSV-S1 MGRLRFVVLLSIFPIKTFSEPCSTMALVLRDSIKISEFINLSASEKLLPSALTAVKSVRI
TVCV ----------MSIVSYEPKVSDFLNLSKKEEILPKALTRLKTVSI
CGMMV-SH ----------MSLSKVSVENSLKPEKFVKISWVDKLLPNYFSILKYLSI
KGMMV-Y ----------MSVSSVGVKNVLKPNEFVKLSWVDRILPDMFTVYRYLSV
SHMV ----------MSEVSKISTLLAPEKFVKLSVSDKFK--WKAPSRVCSI
I
60
HVS VDCDVAKLSGNGQSLFSVNLTSNLTLEQKSYKYVHLLAIVISGKWHLHNTVGGGATVALL
TMV-Ob STVDKIMAKEN-DNISEVDLLKGVKLVKN--GYVCLVGLVVSGEWNLPDNCRGGVSICLI
TMGMV STVDKIMAVKN-DSLSDVDLLKGVKLVKK--GYVCLADLVVSGEWNLPDNCRGGVSVCIV
TMV-Type SKVDKIMVHEN-ESLSEVNLLKGVKLIDS--GYVCLAGLVVTGEWNLPDNCRGGVSVCLV
ToMV SKVDKIMVHEN-ESLSEVNLLKGVKLIEG--GYVCLVGLVVSGEWNLPDNCRGGVSVCMV
PMMV-S SKVDKVIAMEN-DSLSDVNLLKGVKLVKD--GYVCLAGLVVSGEWNLPDNCRGGVSVCLV
ORSV-S1 SKVDKIISYEN-DTLSDIDLLKGVKLVEN--GYVCLAGLVVTGEWNLPDNCKGGVSICLV
TVCV STKDIISVKES-ETLCDIDLLINVPLDKY--RYVGILGAVFTGEWLVPDFVKGGVTISVI
CGMMV-SH TDFSVVKAQSY-ESLVPVKLLRGVDLTKH--LYVTLLGVVVSGVWNVPESCRGGATVALV
KGMMV-Y TDYSVIKSKDS-ECLIPVDLLRGVDLSKS--KYVTLVGVVISGVWTIPENCAGGATVALV
SHMV VQSDTISMTANGRSLFTFDVLKDVLKHAEEYTYVDVLGVVLSGQWLLPKGTPGSAEIILL
I
102
II
131
HVS DKRITNPQQAQLAKMVVAAKVGEFQGRLTPNYSVTLEDALKDPWEVFISIKDVPIEQGFH
TMV-Ob DKRMQRHNEATLGSYTTKASKKNFSFKLIPNYSITSQDAERRPWEVMVNIRGVAMSEGWC
TMGMV DKRMKRSKEATLGAYHAPACKKNFSFKLIPNYSITSEDAEKHPWQVLVNIKGVAMEEGYC
TMV-Type DKRMERADEATLGSYYTAAAKKRFQFKVVPNYAITTQDAMKNVWQVLVNIRNVKMSAGFC
ToMV DKRMERADEATLGSYYTAAAKKRFQFKVVPNYGITTKDAEKNIWQVLVNIKNVKMSAGYC
PMMV-S DKRMQRDDEATLGSYRTSAAKKRFAFKLIPNYSITTADAERKVWQVLVNIRGVAMEKGFC
ORSV-S1 DKRMKRANEATLGSYHTSACKKRFTFKIIPNYSVTTADALKGIWQVMINIRGVEMEKGFC
TVCV DKRLVNSKECVIGTYRAAAKSKRFQFKLVPNYFVSTVDAKRKPWQVHVRIQDLKIEAGWQ
CGMMV-SH DTRMHSVAEGTICKFSAPATVREFSVRFIPNYPVVAADALRDPWSLFVRLSNVGIKDGFH
KGMMV-Y DTRMSMVDEGTICKFSVAASTRDFMVKLIPNYYVLASDASSKPWSIFVRVSGVRIKEGFS
SHMV DSRLKG-KASVLAVFNCRAATQEFQFLISPGYSLTCADALKKPFEISCNVIDLPVKDGFT
II
171
193
A
205
213
B
HVS PLSVEIAALLMFSDICILKSLRAKMIESQIGFTDGTDGDIPSENMDAFVQSASIMQDLRR
TMV-Ob PLSLEFVSVCIVHKNNVRKGLREKVTA----VSEDDAIELTEEVVDEFIEAVPMARRLQN
TMGMV PLSLEFVSICVVHKNNVRKGLRERILS----VTDGSPIELTEKVVEEFVDEVPMAVKLEK
TMV-Type PLSLEFVSVCIVYRNNIKLGLREKITN----VRDGGPMELTEEVVDEFMEDVPMSIRLAK
ToMV PLSLEFVSVCIVYKNNIKLGLREKVTS----VNDGGPMELSEEVVDEFMENVPMSVRLAK
PMMV-S PLSLEFVSVCIVHKSNIKLGLREKITS----VSEGGPVELTEAVVDEFIESVPMADRLRK
ORSV-S1 PLSLEFVSICVVYLNNIKLGLREKILN----VTEGGPTELTEAVVDEFVEKVPMAARLKS
TVCV PLALEVVSVAMVTNNVVMKGLREKVVA----INDPD-VEGFEGVVDEFVDSVAAFKAVDN
CGMMV-SH PLTLEVACLVATTNSIIKKGLRASVVE----SVVSSDQSIVLDSLSEKVEPFFDKVPISA
KGMMV-Y PLTLEIASLVATTNSILKKGLRVSVLE----SVVGSDASINLDTVSDKVQPFFDSVPITA
SHMV PLSVEIACLVQFSNCVITRSLTMKLKE------NPATRTFSAEEVDELLGSMTTLRSIEG

```
                            B                    244        251             C
HVS         LRMANNSLGSGVIGGRDKKGKKFVK-------DFKKKTVVSD------------NSDTDF
TMV-Ob      LRKPKYNKE--KENKNLNNKNSIGVSKPVGLERNKVRSVVRK------------GVRSDS
TMGMV       V--PENKKE--MVGNNVNNK--------------KINNSGKK------------GFKIEE
TMV-Type    FRSRTGKKSDVRKGKNSSNDRSV----------PNKNYRNVK------------DFGGMS
ToMV        FRTKSSK----RGPKNNNNLGKG----------RSGGRPKPK------------SFDEVE
PMMV-S      FRNQSKK----GSNKYVGKR-------------NDNKGLNKE----------------GK
ORSV-S1     FRSVNKKKPS-DSSKFVNGKSRL----------NSRNKLNYENGDSDVGISVVDDIVVGN
TVCV        FRKRKKKVEERDVVSKYKYRPEK---------YAGPDSFNLK-----------EENVLQH
CGMMV-SH    AVMARDPSYRSRSQSVGGRGKRH------------SKPPNRR---------------LDS
KGMMV-Y     AVIARDRSYVSKSGPSRAVRQRP------------QFKGGKK---------------GDA
SHMV        LRKKKEPNDVVQGHLSAEYDVKRSVKRTKSENTPGKRRVNVD------------SVSLGL
```

```
                          C          282
HVS         SGAESSYTTEFDSRSETSDALP----
TMV-Ob      SLGVTDMSQDGSSSEISSDSFI----
TMGMV       ---IEDNVSD-DESIASSSTF-----
TMV-Type    FKKNNLIDDDSEATVAESDSF-----
ToMV        KEFDNLIEDEAETSVADSDSY-----
PMMV-S      LFDKVRIGQNSESSDAESSSF-----
ORSV-S1     GVSDIRIDDDCESFDAQSDPY-----
TVCV        YKPESVPVLRSGVGRAHTNA------
CGMMV-SH    ASEESSSVSFEDGLQSDHT-------
KGMMV-Y     DSAES--MSDGGAYEPPRL-------
SHMV        GKGKSVSAKNEDTESVFDDGILDSDS
```

Fig 3.11 Amino acid sequence comparison of HVS MP with ten other tobamoviruses. CLUSTAL X (Thompson et al., 1997) was used to generate multiple sequence alignments. Asteriks indicate identical amino acid residues. Bold letters indicate conserved regions. Gaps introduced to obtain maximal alignment are shown as dashes.

caulimoviruses (Koonin et al., 1991). Three more conserved regions have been identified in tobamovirus MPs (motifs A, B and C) (Fig. 3.11). Regions A and C are rich in acidic amino acids and region B is is rich in basic amino acids (Saito et al.,1988). Similarly in HVS, motifs A (193- 205) and C (251- 282) were rich in acidic amino acids and motif B (213-244) was rich in basic amino acids. Domain C was not very apparent in HVS. The C- terminal portion of the MP was less conserved. The dissimilar regions in sequences of the MP at the N and C terminals are responsible for interacting with a host protein kinase that is required for phosphorylation of the MP (Citovsky et al., 1993). The MP of tobamoviruses shares a significant similarity with tobraviruses. It is also noteworthy that N-terminal region of the MP of tobamo- and tobraviruses shares a statistically significant similarity with the BL1 protein of geminiviruses, which is involved in movement (Koonin et al., 1991).

Based on nt sequence alignment with the TMV MP sg promoter region (Grdzelishvili et al.,2000) putative MP sgRNA transcription start site of HVS was located at nt 4905. Within the MP ORF an IRES was located (nt 5666-5799) which could drive the expression of CP gene (Fig. 3.7).

The percentage sequence identities of the HVS MP nt and aa sequences with 10 other tobamoviruses were determined. The HVS MP nucleotide sequence was closest to CGMMV-SH (43.90%) and had minimal homology of 38.81% with TVCV. The amino acid sequence identity ranges from 30.85 % (ORSV-S1) to 37.88% (CGMMV-SH) (Table 3.2). Phylogenetic trees generated based on nt and aa sequences of the MP ORF indicated that HVS is positioned within the sub group II viruses (3.12a&b). The MP ORF of HVS seems much different from other

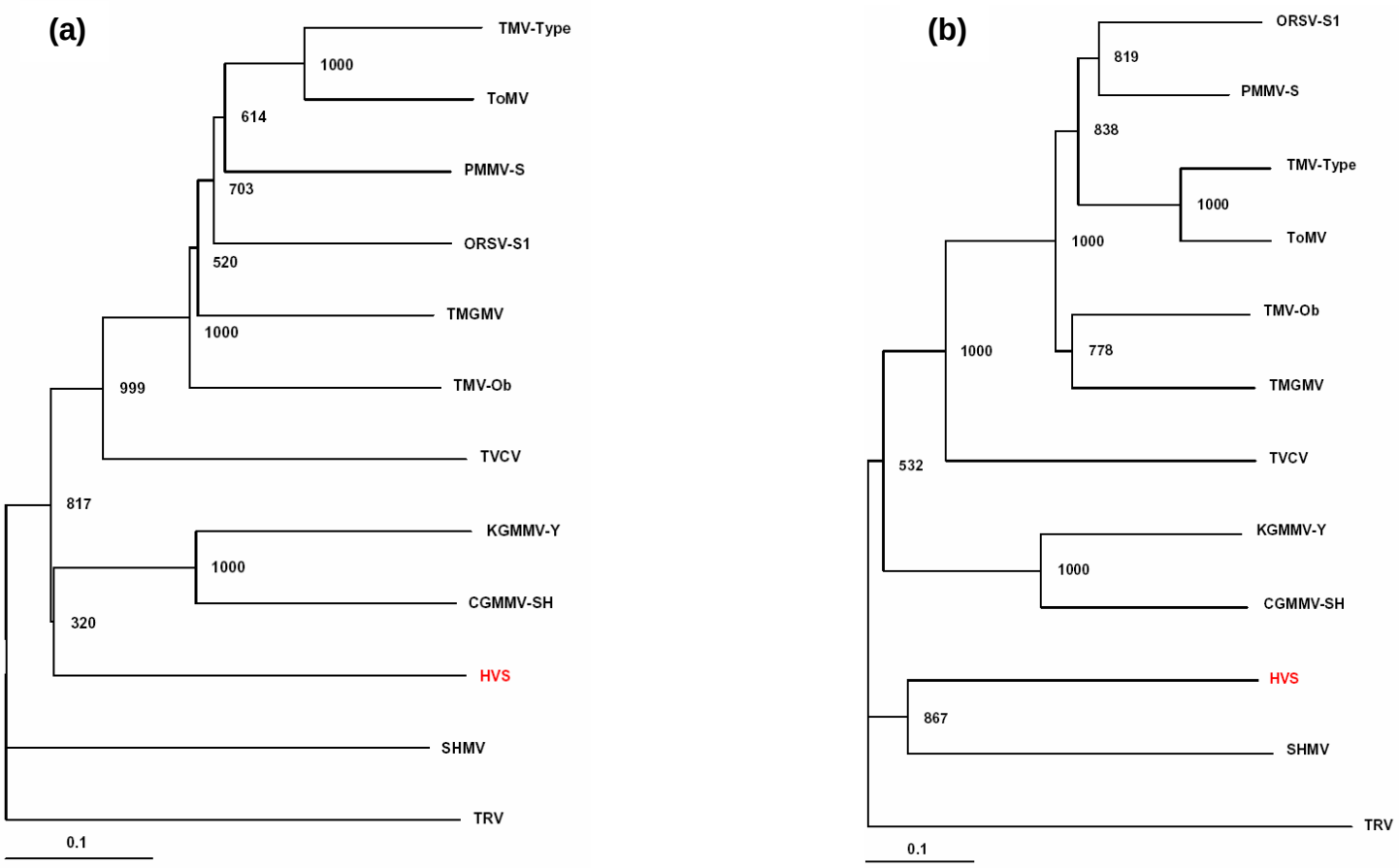


Fig. 3.12 Phylogenetic trees generated from multiple sequence alignment of nucleotide **(a)** and amino acid sequences **(b)** of MP, illustrating the position of HVS among the members of tobamoviruses. The percentages of bootstrap replicates are indicated at the nodes. The trees were rooted with TRV as an outgroup.

tobamoviruses. The MP region shows the lowest percentage identities with other tobamoviruses as compared to percentage identities of other HVS ORFs.

3.7.4 Coat protein

The putative CP sequence in HVS is 492 nt long (nt 5800-6291), encodes for 163 amino acid residues and has a molecular mass of 18,19 kDa. The 5' end of the CP gene overlaps with 3' end of MP gene. The putative CP transcription start site is located at nt 5790. The putative sg promoter of the CP gene extends from nt 5616-5841. In TMV, it has been reported that the 3' end of CP sg promoter runs well into N-terminal portion of CP coding region (Grdzelishvili et al., 2000).

Alignment of the HVS CP gene with 11 other tobamoviruses revealed three highly conserved motifs (Fig. 3.13). These motifs have been reported to be conserved among tobamoviruses (Altschuh et al., 1987). In the HVS CP, these motifs are located at amino acid positions 37-42 (motif I), 89-95 (motif II) and 112- 119 (motif III). These motifs can recognize and bind to viral RNA (Altschuh et al., 1987). There are 25 absolutely conserved amino acid residues within the TMV CP. Ten of the amino acids are neutral and 15 of these residues are charged. In HVS out of this 25 absolutely conserved residues 21 are identical to TMV and 4 amino acids located at positions 26, 63, 91 and 129, have undergone highly conserved changes. One of these mutations was located in conserved motif II (aa 91) whereby arginine is replaced by lysine in HVS and FHV. HVS has a stretch of 6 threonines at the 3' the end of CP with an insertion of serine in the middle. Similarly the 3' end of TMV-Ob has a stretch of 6 threonine residues as their c-terminal amino acids.

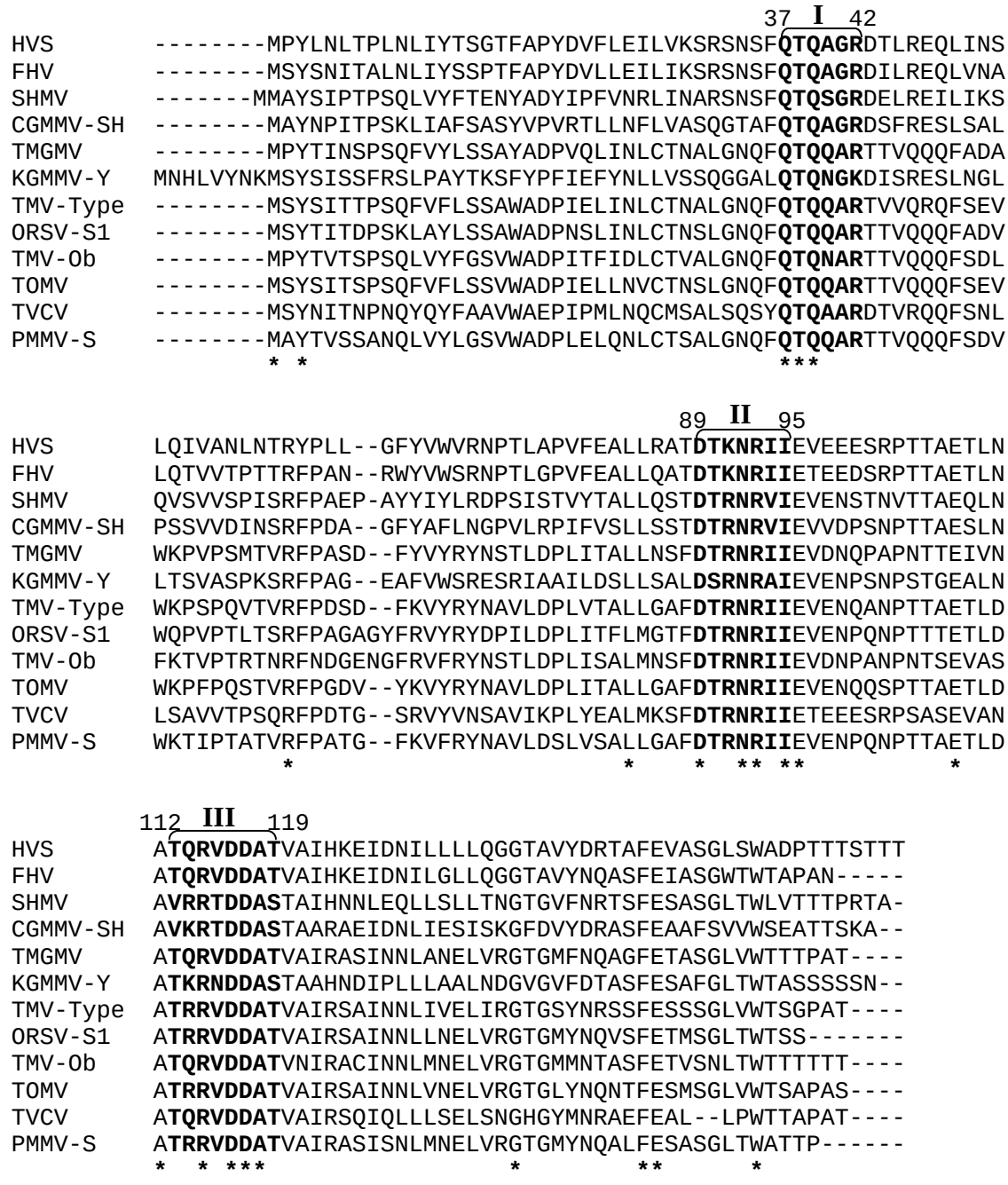


Fig 3.13 Amino acid sequence comparison of HVS MP with eleven other tobamoviruses. CLUSTAL X (Thompson et al., 1997) was used to generate multiple sequence alignments. Asteriks indicate identical amino acid residues. Bold letters indicate conserved regions. Gaps introduced to obtain maximal alignment are shown as dashes.

The HVS CP is longer than the FHV CP by 5 amino acid residues. Generally, the sequences of HVS and FHV are well conserved. There are 18 highly conserved and 8 partially conserved changes in amino acid residues of HVS as compared to FHV. Also there are a few non- conserved changes in HVS CP at positions 4, 16, 44, 55, 65-67, 72, 131, 151 and 155 with respect to the FHV amino acid sequence. None of non-conserved changes are located in RNA binding motifs.

The HVS CP sequence is followed by a poly(A) tract. A putative polyadenylation signal AAUAUA is located at nt 6181- 6186 (Fig. 3.7) near the 3' end of the CP ORF. The percentage identities of CP nt and aa sequences with 11 other tobamoviruses were determined. HVS is closely related to FHV. The nt and aa sequences are 71.07% and 79.75% similar to those of FHV CP (Table 3.2). Phylogenetic trees based on nt and aa sequences show that HVS belongs to subgroup II of tobamoviruses (Fig. 3.14 a&b). HVS and FHV form a distinct cluster within subgroup II viruses.

The OAS was located in the CP cistron. In this region, the target sequence GAUGAUG and the triplet-repeated purine base tract was located. This sequence could be folded into a highly base-paired hairpin loop structure (Fig. 3.15a). The predicted secondary structure resembled that of the OAS of SHMV (3.15b).

3.7.5 Untranslated regions

The 5' UTR of HVS is 58 nt in length and contains five G residues other than the cap structure. Tobamoviral 5' UTRs contain sequences that are responsible for translational enhancement effect. In ORSV-S1, three copies of eight nt direct repeats of ACAATTAC and the $(CAA)_n$ or $(ACA)_n$ region are present (Chng et al., 1996). In TMV one copy of the 8 nt direct repeat and a 25 nt (CAA)n region

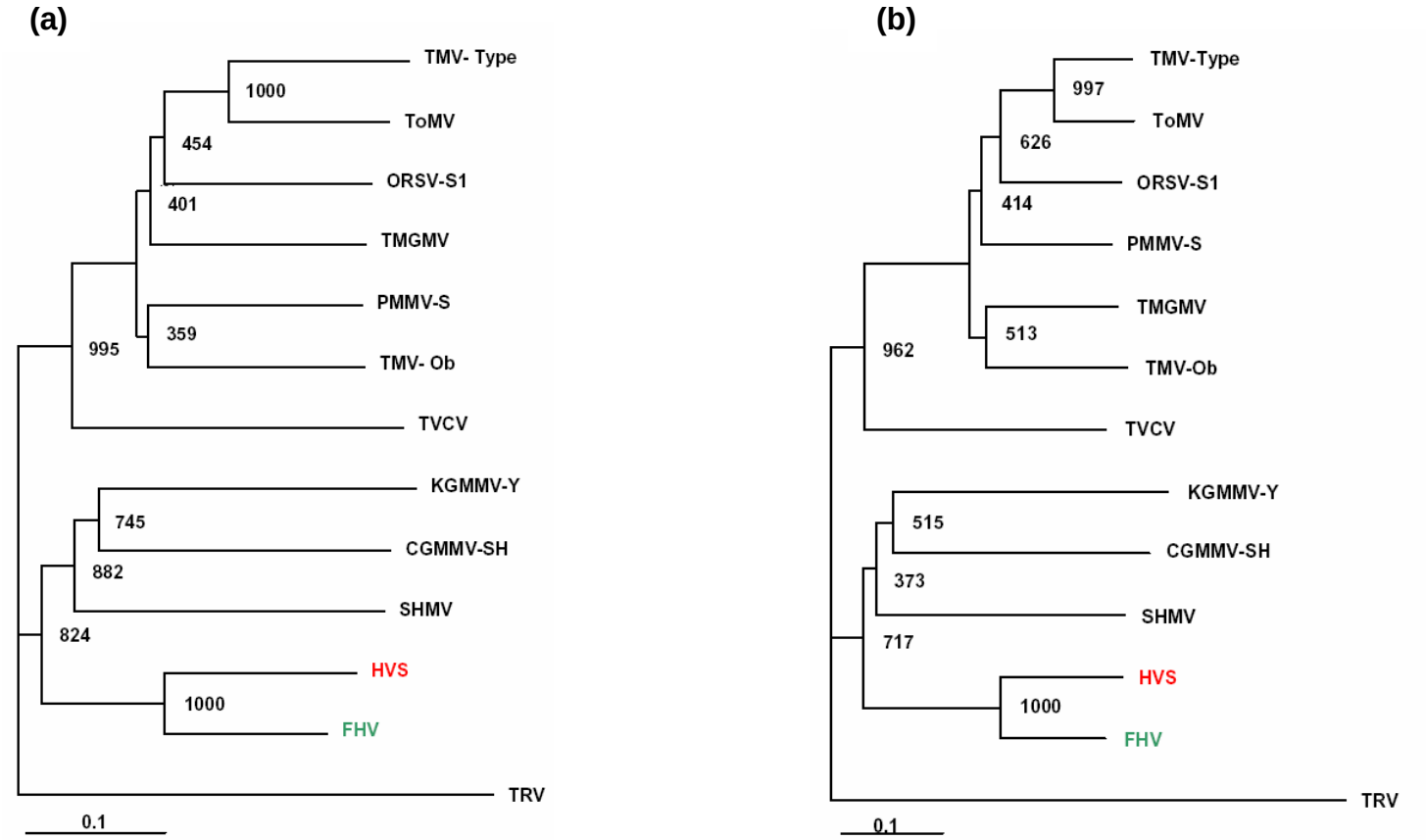


Fig. 3.14 Phylogenetic trees generated from multiple sequence alignment of nucleotide **(a)** and amino acid sequences **(b)** of CP, illustrating the position of HVS among the members of tobamoviruses. The percentages of bootstrap replicates are indicated at the nodes. The trees were rooted with TRV as an outgroup.

Fig. 3.15 **(a)** **(b)**

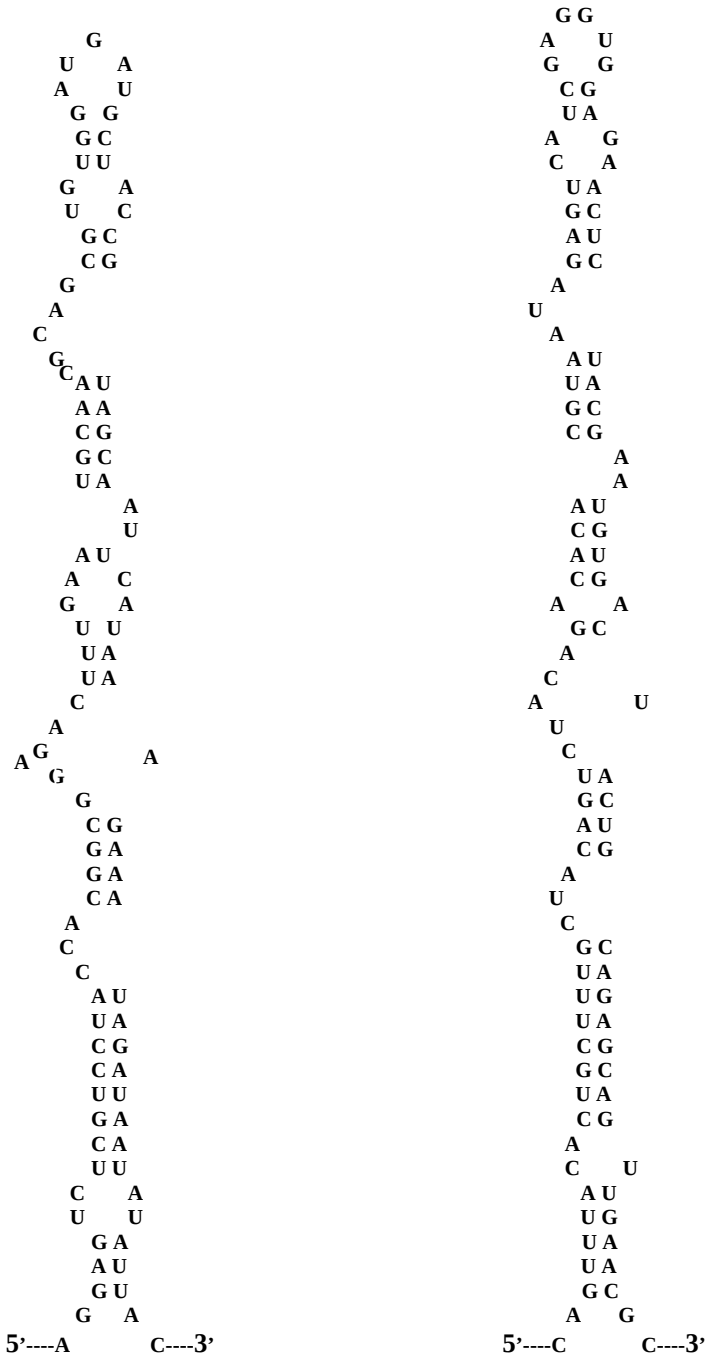
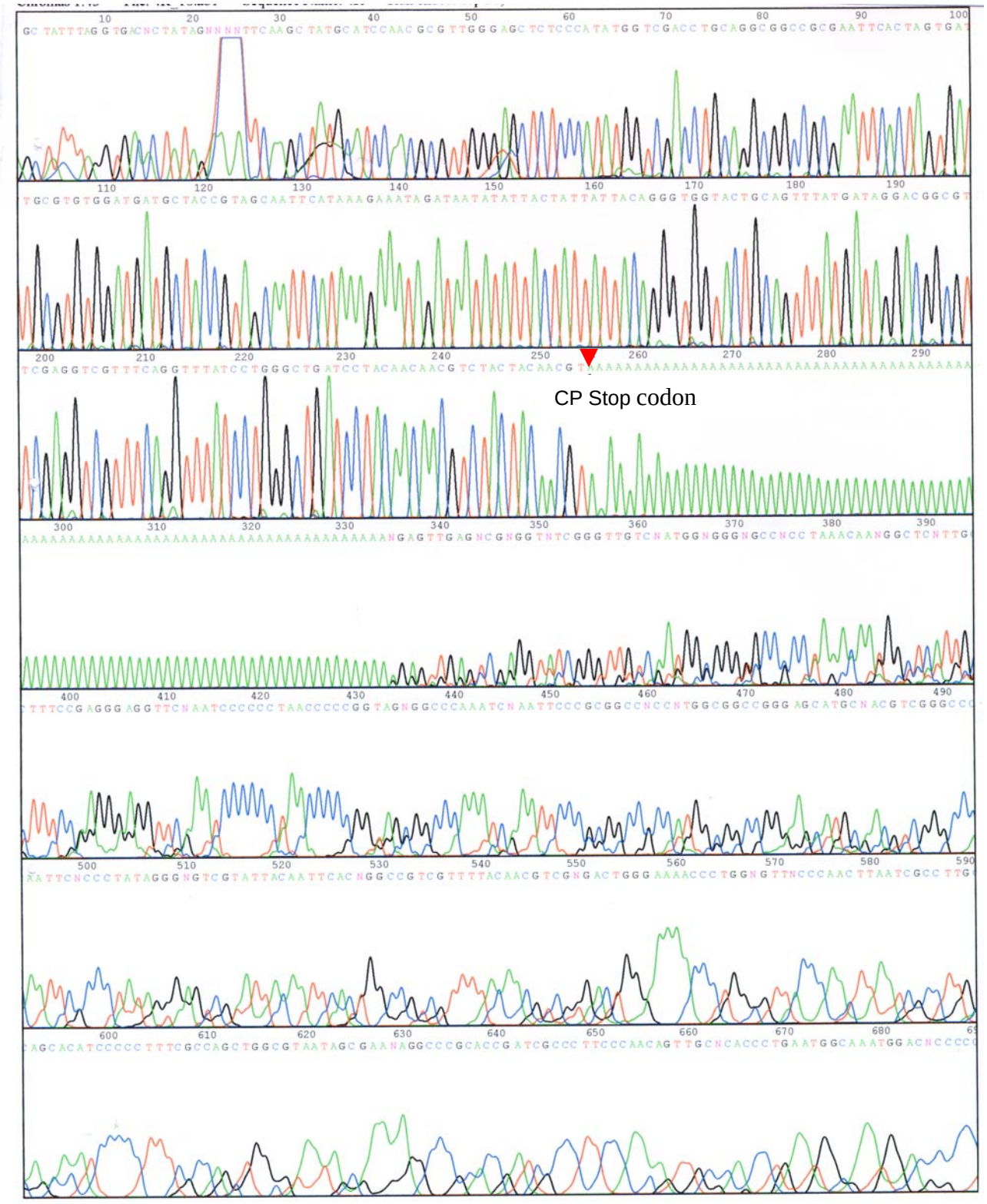


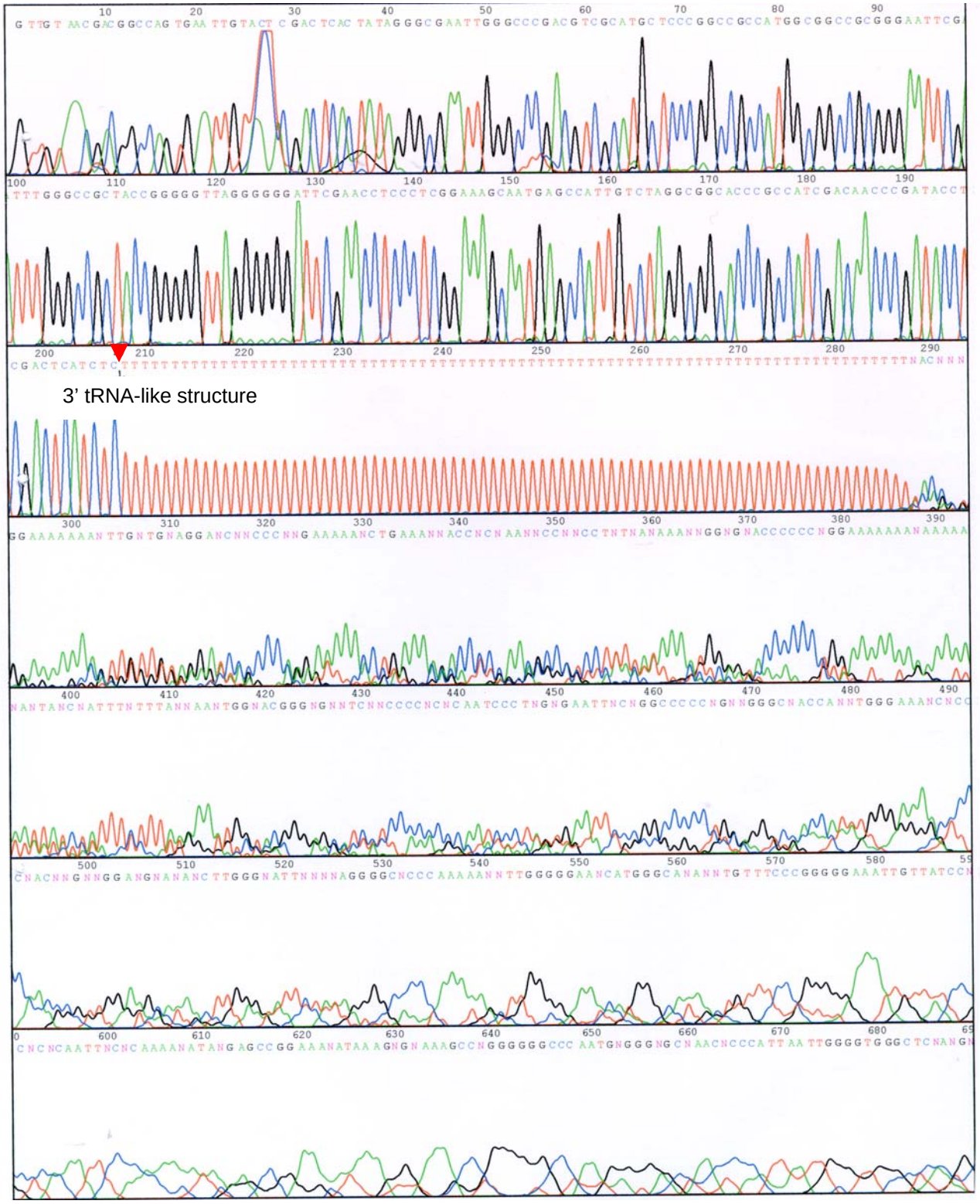
Fig. 3.15 Predicted virion OAS of HVS **(a)** which resembles OAS of SHMV **(b)**.

conferred translational enhancement. The presence of the (CAA)n motif seems to be an absolute requirement for translational enhancement (Gallie and Walbot, 1992). In HVS two direct repeats of AACATT, six copies of CAA and two copies of ACA are present within the 5' UTR region. These sequences form 58% of the 5'UTR region. It was shown that the presence of AUU sequences may play a regulatory role in translational enhancement by the 5' UTR (Gallie et al., 1988). Two AUU ribosome binding sites were found 25 and 36 nt upstream of the AUG in the 5' UTR of HVS. The presence of AUU motifs in such close proximity to AUG, may affect the translational efficiencies by causing steric hindrance to simultaneous binding of two ribosomes (Gallie et al., 1988). In TMGMV, the translational efficiency *in vitro* was not affected by the close proximity of the AUU triplet and AUG (Solis and Garcia-Arenal 1990). Triplets ACA and GCA immediately upstream and downstream of the first AUG codon is conserved in most tobamoviruses. Triplets AAT and ACT were found in HVS at these positions.

The 3' UTR performs multiple functions such as recruitment of host translational machinery thereby effecting translational enhancement (Gallie and Walbot, 1990), binding to RdRp and generation of minus strand RNA (Takamatsu et al., 1990). The tRNA-like structure in the 3' UTR region has been shown to enhance mRNA stability. Substituting of poly(A) tail by a viral 3' UTR enhanced expression of plant and animal mRNAs (Gallie and Walbot, 1990). The tRNA-like structures are excellent substrates for aminoacylation and can be catalyzed by specific aminoacyl-tRNA synthetases (Mans et al., 1991).

Fig 3.16 Electrophoregram obtained from automated sequencing of HVS 3' UTR region. **(a)** CP stop codon followed by poly-A tract **(b)** tRNA-like structure ending in 'GGGT(CCCA)' preceded by T-tract obtained by sequencing with reverse primer


CP Stop codon


3’ tRNA-like structure

The TMV 3' UTR could be folded into a structure comprising a tRNA-like structure which is involved in the synthesis of minus strand RNA and the translation enhancing structures namely, the pseudoknots (Van Belkum et al.,1985). The presence of pseudoknots appears to be a common feature among non-polyadenylated viral RNAs (Dam et al., 1992). HVS 3' UTR is unique in the sense that it has a 106 bp region which could be folded into a t-RNA like structure and a poly(A) tract which is 77-96 bp in length (Fig. 3.17b). The presence of GUG at nt positions 6403-6405 in the anti-codon loop indicates that the 3' UTR could accept histidine. A conserved sequence UGUR is present in histidine accepting tobamoviruses (Garcia-Arenal, 1988). This conserved region, though required for efficient aminoacylation is not found in HVS and ORSV-S1 (Chng et al.1996).

The presence of a poly(A) tract immediately after the CP ORF in HVS is a characteristic observed in BSMV (Agranovsky et al., 1978). The length of the poly(A) tract in BSMV is about 15-43 nt. Recently, a member of alphavirus group, *Chikugunya virus* (CHIK) was reported to have an internal poly-A tract in the 3' UTR region (Khan et al., 2002). The poly(A) tract interacts with poly (A)-binding proteins (PABP) which is involved in mRNA stability and turnover (Bernstein et al., 1989). PABPs require a minimum of 12 adenosine residues for binding to the RNA (Dreher, 1999). In FMDV, it has been shown that the presence of poly(A) tract had a stimulatory effect on IRES dependent translation (Lopez de Quinto et al., 2002). The presence of a long poly(A) tract (Fig. 3.16), and IRES segments suggest a similar role for the HVS poly(A) tract *in vivo*.

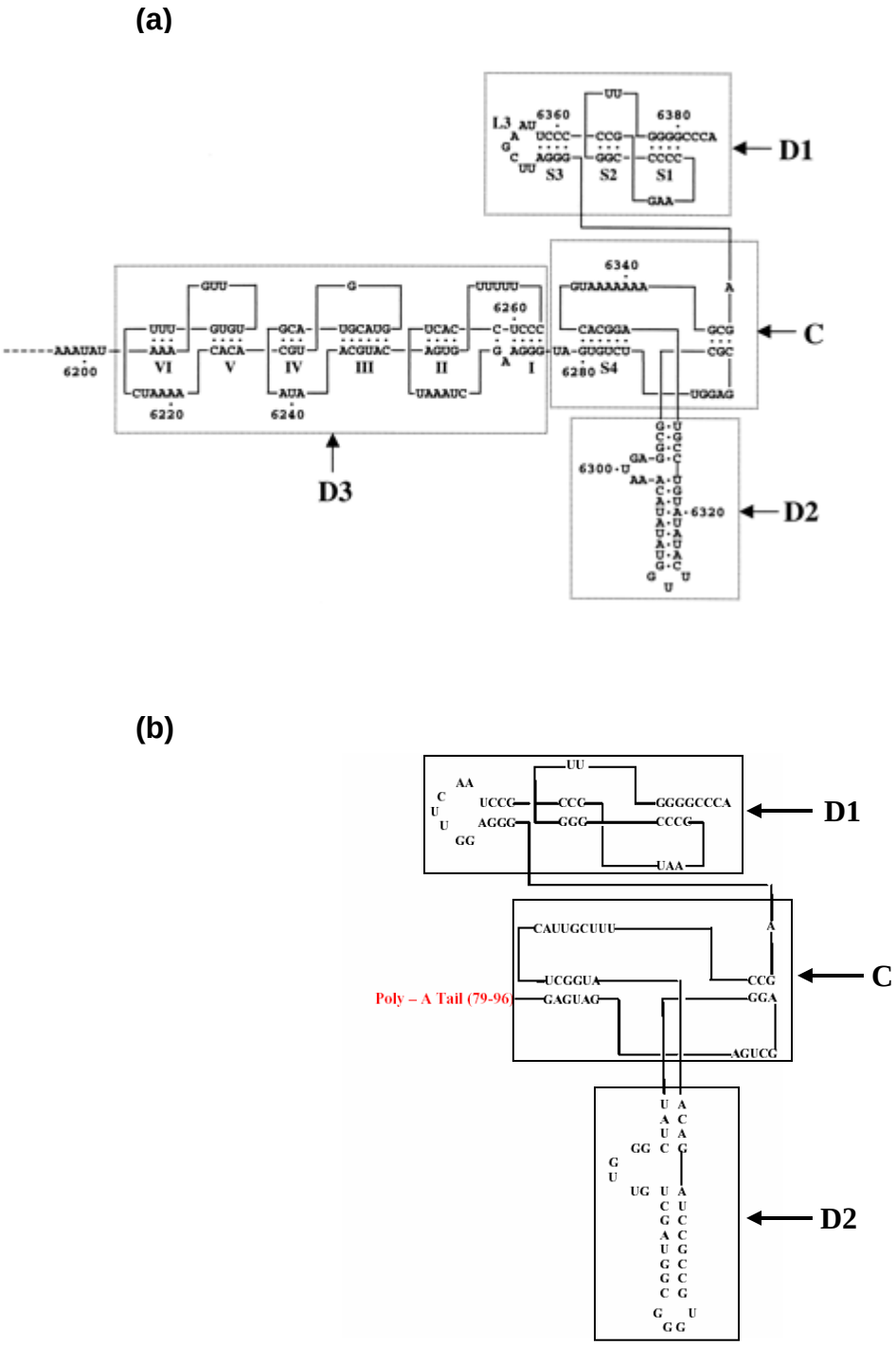


Fig 3.17 Comparison of 3' UTR structures of (**a**) TMV-L (Osman et al., 2000) to that of HVS (**b**). Domain **D3** is replaced by a poly- A tract in HVS.

3.8 Discussion

Hibiscus Virus–Singapore isolate was successfully isolated from hibiscus plants that had a mixed population of HVS and HCRSV. Subsequently high yields of HVS virions could be obtained by propagating the virus in *H. cannabinus* using a protocol which is a modification of the HCRSV isolation method. High yields of HVS could also be obtained from *N. benthamiana* by adopting the ORSV-S1 isolation procedure. The virion morphology is similar to other known tobamoviruses (Fig. 3.1a&b). Particles of different sizes were observed and this is due to the presence of the OAS within the CP gene (Fig. 3.15a). Viral RNA isolated from purified virus when size fractionated on 1.2% formaldehyde denaturing gels revealed the presence of a ~6.4 kb genomic RNA together with a 1.5 kb and a 0.7 kb sgRNAs of MP and CP, respectively. This further shows that the MP and CP sgRNAs are encapsidated which is a subgroup II characteristic (Fig. 3.2a).

We reported HVS causes necrotic local lesions (*nl*) in *C. quinoa* (Srinivasan et al., 2002). FHV causes chlorotic local lesions in *C. quinoa*. It was suggested that *C. quinoa* could serve as a differential host for FHV and HVS (Adkins et al., 2003). Our recent observation is that HVS causes chlorotic local lesions in *C. quinoa*. This could be either due to the mutations accumulated by passaging of virus in different hosts or could be influenced by various environmental factors. The proportion of *nl* variants in TMV populations had been shown to be affected by temperature and the host plant in which the virus is propagated (Aldaoud et al.,

1989). HVS consistently produced systemic infection accompanied by wrinkling of leaves in *N. benthamiana*. FHV infects *N.benthamiana* systemically but the plants remained symptomless (Adkins et al., 2003). *N. benthamiana* could be used as a diagnostic host for distinguishing between HVS and FHV infections in hibiscus plants. HVS was inoculated onto several test plants (Table 3.1) and it was observed that HVS produces symptomless systemic infection in kenaf cultivars. The same has been reported for FHV (Adkins et al., 2003).

Hibiscus yellow mosaic virus, a putative tobamovirus, was reported by a group from Japan and the host plant was co-infected with HCRSV (Kashiwazaki et al., 1982). The yellow mosaic symptom could have been induced by HCRSV rather than the putative tobamovirus. FHV causes symptomless systemic infection in *H.rosasinensis* plants. (Adkins et al., 2003). HVS so far has not been tested against *H.rosasinensis* due to the unavailability of susceptible and virus free cultivars. HVS CP weakly cross-reacted with TMV antiserum in western blots (Fig. 3.4b, lane 3). HVS antiserum cross-reacted with TMV and ORSV CPs (Fig. 3.4d, lanes 1&2). These results suggest that HVS is distantly related to these tobamoviruses. Even though HVS shared a high nt and aa CP sequence identity with SHMV next only to FHV, western blot analysis showed that it is serologically not related to SHMV (Fig. 3.4d&e).

The genome structure of HVS is similar to some members of subgroup II tobamoviruses. The MP gene overlaps with RdRp at the 5' end and CP at 3' end by 8 and 11 nts, respectively. The percentage nt and aa identities of HVS ORFs were compared with ten other tobamoviruses. The HVS nt sequence encoding the 126 kDa protein is similar to that of SHMV (51.08%) but the aa identity of 126 kDa

ORF is more similar to TMV-type strain (55.02%). The MP ORF is similar to subgroup II viruses (Table 3.1). The nt and aa sequences of the CP ORF, as expected were similar to FHV (71.01% & 79.75 %). Even though HVS is very similar to FHV in its CP sequence identities, they are significantly different from one another.

Phylogenetic trees generated based on the nt and aa sequences of the 126 kDa, 183 kDa, MP and CP ORFs indicate that HVS clusters with members of subgroup II of tobamoviruses. Phylogenetic tree based on CP amino acids show that HVS and FHV forms a distinct cluster within subgroup II viruses. With the availability of the FHV full length sequence it will be interesting to observe whether different ORFs of FHV and HVS could form a distinct cluster within subgroup II tobamoviruses.

There are interesting features in HVS which makes it novel among tobamoviruses. Apart from the low percentage of amino acid identities, the 3' UTR of HVS seems to be unique. This region has an internal poly(A) tract which might be a replacement for the pseudoknot domain and a 106 nt region which could be folded into a tRNA-like structure. Earlier, the presence of intercistronic poly(A) tract has been reported in *Brome mosaic virus* (BMV) (Ahlquist et al., 1981). In BMV, a complementary poly-uridylate tract in the minus strand RNA is essential to direct the synthesis of positive strand sgRNA (Adkins et al., 1997). BSMV presents an identical scenario to HVS in that the poly(A) tract is located at the junction between the coding region and the 3' UTR (Agranovsky et al., 1978). The poly(A) tract is required for BSMV to be infectious (Atabekov and Dolja, 1986).

The poly(A) tract determined by direct sequencing of the HVS 3' UTR clones varied in length (77-96 nt). The presence of long stretch of homopolymeric region makes it difficult to determine the exact length of this region by direct sequencing (Fig. 3.16). cDNA clones obtained by the Gubler and Hoffman method (1983) resulted in poly(A) tracts of varying length. Thus the possibility that this homopolymeric region was generated by PCR is ruled out. The mechanism by which poly(A) sequence is added needs to be investigated. The poly(A) tract may bind to PABPs which in turn will interact with translation factors associated with the 5' cap and translate the viral encoded products as a circular complex. Apart from that, the poly(A) tract can also enhance the cap independent translation of 3' proximal ORFs. It is tempting to speculate that the poly(A) tract present in place of the pseudoknot region could compensate for the absence of pseudoknot domains (Fig. 3.17a&b). The structural features of the HVS 3' UTR may need to be probed chemically before functions could be assigned for various sequences within the 3' UTR.

To conclude, based on phylogenetic analysis and percentage sequence identities, HVS is considered a new tobamovirus species. Inclusion of the BSMV CP (hordeivirus) sequence in the phylogenetic analysis clearly indicated that HVS is a tobamovirus and it did not cluster with this hordeivirus. The percentage identities between HVS and FHV are much lower than some tobamoviruses that are considered as distinct tobamovirus species (e.g. TMV and ToMV) (Adkins et al., 2003). Therefore FHV and HVS should be classified as two different tobamoviruses infecting hibiscus.

CHAPTER 4

CONSTRUCTION AND TESTING Of FULL LENGTH AND HYBRID HVS CLONES

4.1 Introduction

The availability of viral infectious cDNA clones has helped in a big way to answer fundamental questions regarding the process of viral multiplication and its interaction with the host. Functions of various proteins encoded by the virus (Ishikawa et al., 1986; Dawson et al., 1988) and the role of structural features within the viral genome (Grdzelishvili et al., 2000; Osman et al ., 2000) have been determined with the help of infectious clones using reverse genetics approaches. Apart from that, it has become possible to express foreign proteins using viral vectors, which confers several advantages over the expression of foreign proteins using transgenic plants (Wu et al., 2003). Recently, RNA mediated defense directed against viral genome by the host, has been exploited for understanding the functions of plant genes (Ruiz et al., 1998). Thus, construction of infectious cDNA clones becomes essential to understand the functional aspects of viral gene products, their exploitation as protein expression vectors and as a tool in functional genomics.

4.2 Construction and testing of near full length HVS cDNA clones

Initial attempts to clone the 3' UTR region resulted in clones ending with a poly(A) tract immediately downstream of the CP coding region. Several attempts were made to clone the 3' UTR using different denaturing agents in the first strand

synthesis reactions. Attempts were also made to clone the 3' UTR by performing a 5' RACE on the minus strand by denaturing dsRNA or the replicating form of virus, and using it as a template. All these attempts resulted in clones with a poly-A tract. To test the hypothesis (i) that HVS either could replicate with a poly-A tract in place of a pseudoknot and tRNA-like structure or (ii) treat this poly-A tract as an artifact resulting from polymerase slippage, near-full length clones with poly-A tract after CP gene and near- full length clones without any sequences after the CP gene were constructed. A near full-length cDNA clone of HVS that ends with the CP gene was obtained by amplifying the HVS viral RNA by using primers complementary to the 5' and 3' ends by a two step RT-PCR as described in the materials and methods. pHVS clones with a poly(A) tail (pHVSA) were constructed by incorporating a stretch of $A_{(40)}$ and a Not I site into the CP - 3' end (-) primer. The PCR products were cloned into pSPORTI under the control of the T7 promoter. Restriction analysis of the ligated products showed the presence of a near full-length genome of HVS (Fig. 4.1a). The resultant plasmids were linearized with Not I and *in vitro* transcribed. To test the infectivity of these constructs, the *in vitro* transcribed RNA was inoculated onto the local lesion host *C. quinoa*, and the systemic host *N. benthamiana* and the transcripts were also PEG-transfected into kenaf protoplasts (Fig. 4.2a&b). Ten days after inoculation on *C. quinoa*, the constructs pHVS and pHVSA, failed to replicate and produce lesions on *C. quinoa*. In *N.benthamiana*. The *in vitro* transcripts of both pHVS and pHVSA did not replicate in the inoculated leaves. The results were confirmed by performing a northern blot by isolating total RNA from the inoculated leaves of *C. quinoa* and *N. benthamiana*, and probing the blot with

Fig 4.1

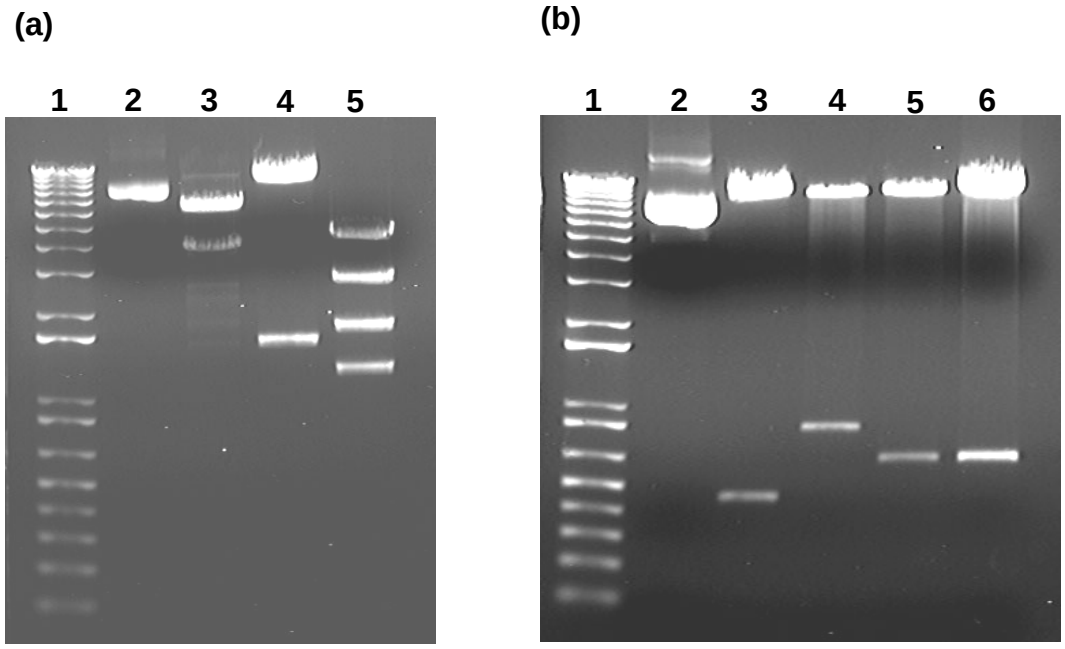
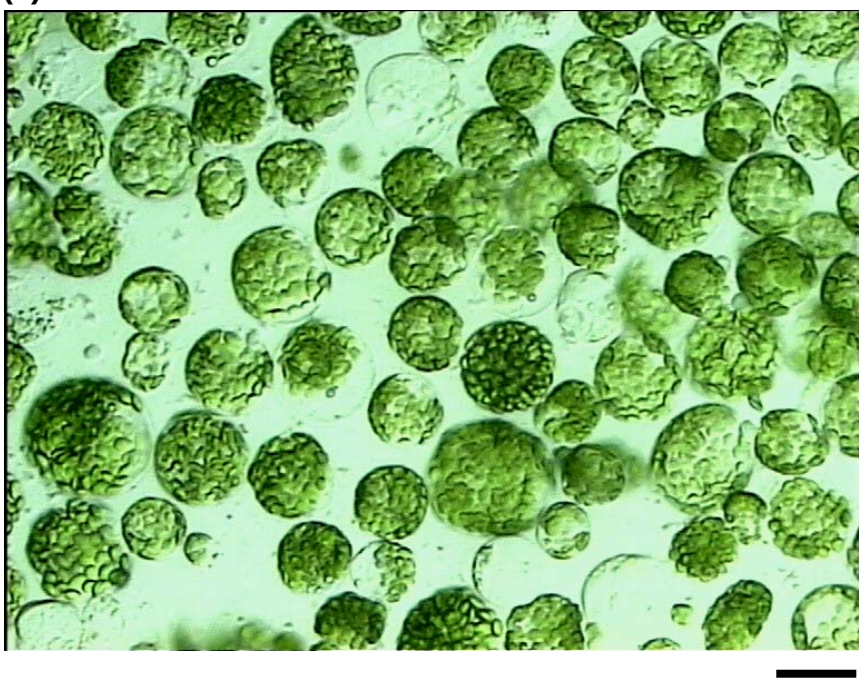


Fig 4.1 (a) Size fractionation of full-length cDNA clone of HVS on 1% TAE agarose gel. Lane 1, DNA molecular weight marker (Gibco-BRL) and lane 2, undigested full length pHVS plasmid clone. pHVS was subjected to restriction digestion using Not I and Sma I (4.0 & 6.3 kb, lane 3), Bam HI (1.6 & 8 kb , lane 4)and Bgl I (1.2, 1.7, 3.0 & 4.4 kb, lane 5) to verify full-length cDNA.

(b) Restriction analysis of pHVS hybrid 3'UTR clones using Spe I and Not I. DNA molecular weight marker (Gibco-BRL) (Lane 1), undigested pHVS (lane 2), pHVS (lane 3), pHVSO (lane 4), pHVSS (Lane 5), pHVST (lane 6).

(a)



(b)

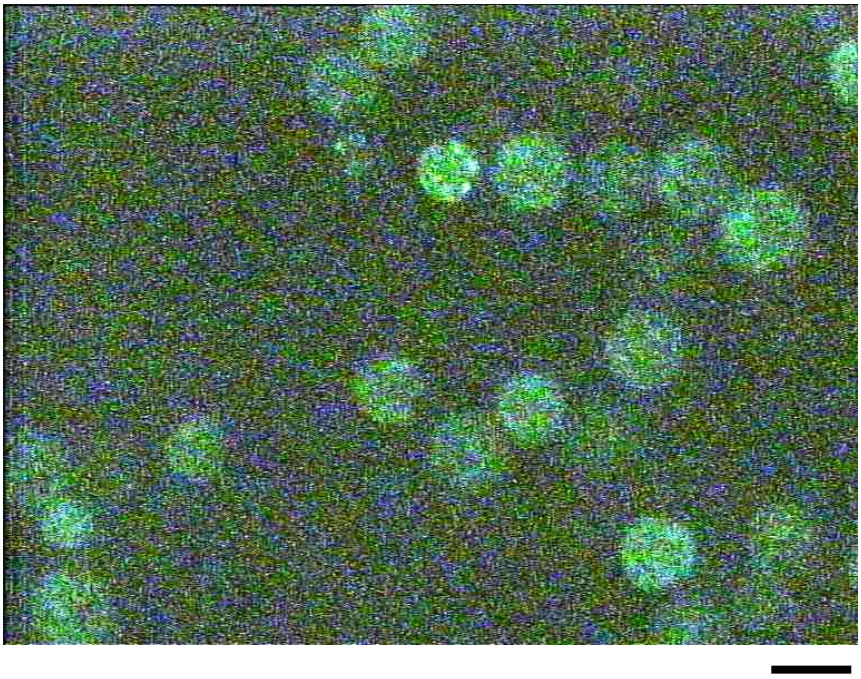


Fig 4.2 (a) Protoplasts isolated from *Hibiscus cannabinus L.* leaves (200X)**. (b)** FDA staining to show the viability of isolated protoplasts (200X). Bar represents 40 µm.

Fig 4.3

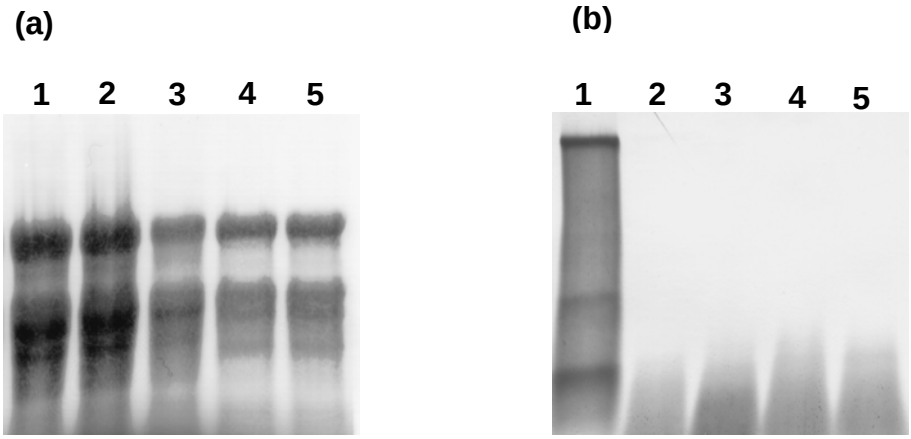


Fig 4.3 Methylene blue stained blot of total RNA isolated from kenaf protoplasts transfected with HVS constructs **(a)**. Protoplasts were transfected with HVS wt- RNA (lane 1), *in vitro* transcribed hybrid 3' UTR constructs HVSO (ORSV UTR, lane 2), HVSS (SHMV UTR, lane 3), HVST (TMV UTR, lane 4), poly-$A_{(40)}$ HVS (lane 5). Identical blot was probed with HVS CP cRNA probe **(b)**.

HVS CP specific DIG-labelled probe. PHVS and pHVSA failed to replicate in kenaf protoplasts as well (Fig. 4.3a&b, lane 5).

4.3 Testing of HVS clones with hybrid 3' UTR and genuine HVS 3' UTR

HVS clones with UTRs of related tobamoviruses ORSV-S1 (pHVSO), SHMV (pHVSS), TMV-U1 (pHVST) were constructed and tested for infectivity. A strategy similar to generating the poly(A) construct was utilized. A PCR for fusing the full-length CP region of HVS and the 3' UTR of the respective tobamovirus was performed. The fusion products were cloned into Spe I / Not I digested pHVS plasmid (Fig. 4.4). The clones were verified by restriction digestion of with Spe I and Not I (Fig. 4.1b, lanes 4,5&6).

Hybrid clones were constructed with a view to identify the reason for the inability of pHVS and pHVSA to replicate in *C. quinoa*, *N. benthamiana* and kenaf. Moreover if these clones could replicate and survive in any of these hosts it would help in de-lineating sg-promoter borders of HVS which will aid in improving the existing viral vector by utilizing HVS MP or CP sg promoters to express proteins of therapeutic value from a ORSV based vector (Lim et al., 2002).

C. quinoa and *N. benthamiana* host plants were inoculated with *in vitro* transcripts of pHVS, pHVSO, pHVSS. Combinations of the aforementioned *in vitro* transcripts with wild type HVS RNA or with wild type RNA from the respective tobamovirus from which the UTR was derived were co-inoculated (Table 4.1). The hybrid constructs could not replicate in *C. quinoa* and *N. benthamiana*. These constructs failed to replicate when co-inoculated wild type HVS RNA or with wild type RNA

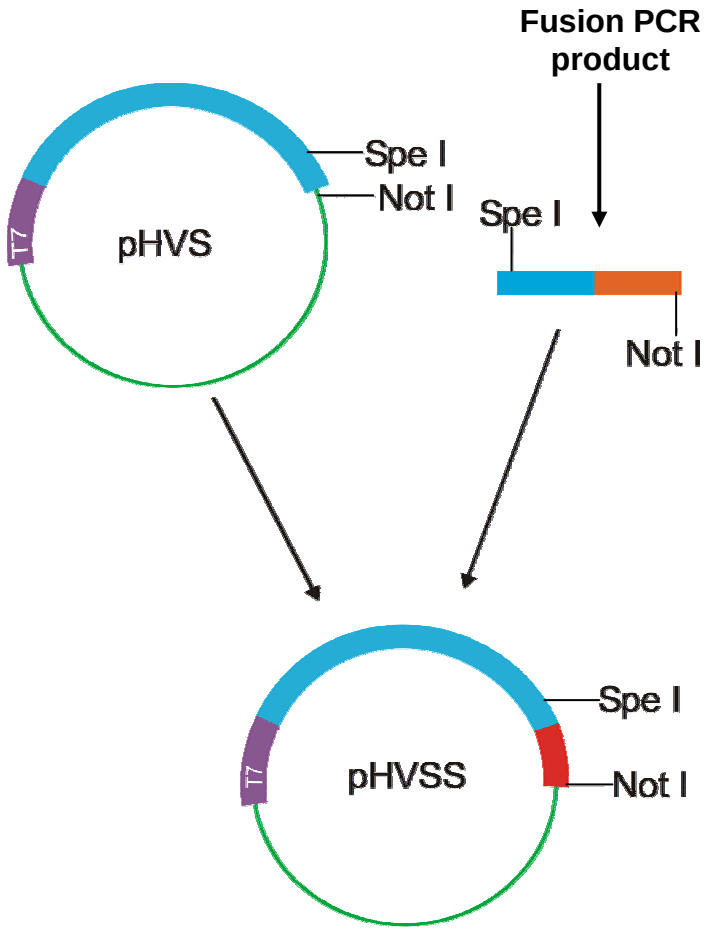


Fig 4.4 General cloning strategy used for generating HVS clones with HVS 3' UTR**,** hybrid 3'UTRs and poly-A tail, driven by T7 promoter. pHVSS denotes pHVS with SHMV 3'UTR.

Table 4.1 Testing of HVS clones without 3' UTR and with hybrid 3'UTRs

Inoculum	Description of RNA Inoculum	Host plants: *C.Quinoa* ,*N.Benthamiana.*
pHVS 1	HVS full length clone without 3'UTR	No replication
pHVS 2	,,	,,
pHVS 3	,,	,,
pHVS 4	,,	,,
pHVSO	pHVS full length clone with ORSV 3'UTR	,,
pHVSS	pHVS full length clone with SHMV 3'UTR	,,
pHVST	pHVS full length clone with TMV 3'UTR	,,
pHVS + ORSV	HVS full length clone without 3'UTR , co-inoculated with ORSV – S1 WT RNA	,,
pHVS + SHMV	HVS full length clone without 3'UTR , co-inoculated with SHMV WT RNA	,,
pHVS + TMV	HVS full length clone without 3'UTR , co-inoculated with TMV WT RNA	,,
pHVSO + ORSV	HVS full length clone with ORSV-S1 3'UTR , co-inoculated with ORSV – S1 WT RNA	,,
pHVSO + HVS	pHVS full length clone with ORSV-S1 3'UTR , co-inoculated with HVS WT RNA	,,
pHVSS + SHMV	pHVS full length clone with SHMV 3'UTR , co-inoculated with SHMV WT RNA	,,
pHVSS + HVS	pHVS full length clone with SHMV 3'UTR , co-inoculated with HVS WT RNA	,,
pHVST + TMV	pHVS full length clone with TMV 3'UTR , co-inoculated with TMV WT RNA	,,
pHVST+ HVS	pHVS full length clone with TMV 3'UTR , co-inoculated with HVS WT RNA	,,

from the respective tobamovirus from which their UTRs were derived. Transcripts of hybrid constructs were PEG transfected into kenaf protoplast. The results of northern blot shows that the hybrid constructs do not replicate in kenaf protoplasts (Fig. 4.3a&b, lanes 2,3&4).

To clone and fuse the HVS 3' UTR with pHVS, degenerate primer Tobmv-3' end (Table 2.1) was designed and the 3' end PCR product was obtained. The PCR product was cloned into Spe I / Not I digested pHVS plasmid to give pHVSH. The clones were transcribed and inoculated on to *C. quinoa, N. benthamiana* and PEG inoculated into kenaf protoplasts. *In vitro* transcripts of pHVSH have not been infectious in any of the above mentioned hosts.

4.4 Discussion

Attempts made to generate infectious transcripts of HVS without 3' UTR and HVS with a poly $A_{(40)}$ were not successful due to the obvious reason that the virus requires a 3' UTR primarily for the replicase to bind and initiate minus strand synthesis (Osman et al., 2000). In positive-strand RNA viruses there have been instances where the 3' UTR is not required for initiation of minus strand whereas, it is essential to promote or regulate minus strand synthesis. This has been shown to be the case in human rhinovirus type 14 (HRV14) and poliovirus type I (PV1). *In vitro* transcripts of HRV14 and PV I lacking the entire 3' UTR was able to replicate in HeLa cell monolayers. It was hypothesized that template selection was determined by the physical proximity of replication initiation complexes and viral RNA, following translation of viral proteins by the host (Todd et al., 1997). Hence an attempt was made to see if HVS could replicate without the 3' UTR. The

non-availability 3’ UTR sequence information other than the poly(A) tract region, at that point of time led us to test the hypothesis that this virus could probably replicate without any 3’ structural elements.

HVS clones with a UTR derived from other related tobamoviruses when inoculated onto *C. quinoa, N. benthamiana* or kenaf protoplasts failed to replicate. The HVS hybrid vectors were not complemented in *C. quinoa* and *N. benthamiana* plants by co-inoculation of hybrid transcripts with wild type HVS RNA. This experiment was set up to provide the replicase protein from wild type HVS RNA. If the replicase component within the hybrid construct is defective, the wild type HVS replicase could perform the function of multiplying the hybrid constructs.

In another set of experiments hybrid transcripts were not complemented with wild type RNA from the respective tobamovirus from which the UTR was derived; e.g., co-inoculation of pHVSS with wild type SHMV RNA. Here the replicase component from SHMV will recognize and bind to the SHMV 3’UTR in the hybrid construct and initiate minus strand synthesis. This set of experiment was set up circumvent the problem whereby HVS replicase component translated from the hybrid transcript, being unable to bind to SHMV 3’ UTR and initiate minus strand synthesis.

The failure of HVS constructs with UTRs derived from related tobamoviruses to replicate could be due to the reason that these hybrid UTRs are not efficient templates for the HVS replicase to bind and promote minus strand synthesis. Insertion of heterologous sequences (SHMV CP and 3’UTR) into viral vector 30B (TTTo/SS-GFP) resulted in very low accumulation of GFP protein and the hybrid

vector RNA accumulated at very low levels in protoplasts and plants. It was proposed that cis-acting sequences of a much closely related tobamovirus may function efficiently with TMV replicase (Shivprasad et al., 1999). Similarly, HVS ORFs are very distantly related to all other tobamoviruses. This could explain the reason for the failure of hybrid constructs to replicate in *C. quinoa, N. benthamiana* and kenaf by themselves or when co-inoculated with wild type HVS RNA. The 3' UTR regions of hybrid clones were sequenced to ensure that there are no mis-incorporations due to PCR. A few HVS clones with ORSV 3' UTR did have mutations in this region. However, many of the hybrid UTR constructs failed to replicate even if they had the correct 3' UTR sequences without any misincorporations.

Inoculation of hybrid constructs with wild type viruses from which the 3' UTR has been cloned, did not result in replication in *C. quinoa* and *N. benthamiana*. The structures within the 5' UTR is required for synthesis of plus strand RNA from minus strand template. Mutations in the 5' UTR region of TMV, was shown to affect the replication of the virus (Takamatsu et al., 1991). Probably the replicase provided by the wild type viruses could initiate the minus strand synthesis, but were unable to synthesis plus strand from minus strand template by binding to the 5' UTR derived from HVS. To know whether the minus strand had indeed been synthesized, the RNA isolated from the inoculated plants should be probed with plus strand probe. This could possibly explain as to which stage of replication the failure occurred.

Full-length infectious clones with HVS 3' UTR were built on the same backbone as the hybrid constructs. The genuine full length HVS constructs did not replicate

in kenaf protoplasts. These results suggest that there could be mutations within the HVS clone which is affecting the replication and survival of HVS. RNA viruses are known to accumulate mutations due to errors during template copying by RdRp (Malpica et al., 2002). This could also be one of the several reasons as to why HVS as well as the hybrid clones failed to replicate.

With the successful construction of HVS infectious cDNA clones it will be possible to investigate various interesting aspects of this novel tobamovirus. For example site specific mutations will help in understanding the functional aspects of poly(A) tract and how it affects translation. The role of poly-A tract in RNA stability, translational enhancement by acting synergistically with IRES regions in the context of full length virus could be investigated.

CHAPTER 5

EVIDENCE FOR THE PRESENCE of IRES-like ELEMENTS WITHIN HVS GENOME THAT ARE FUNCTIONAL *In vitro* and *In vivo*

5.1 Introduction

Translation of host mRNA via cap dependent mechanism can be shut-off by viral proteases upon infection by human picornaviruses. This is achieved by cleavage of the eIF-4G protein, a part of the eIF-4F complex, which brings together the 5' cap and 40S ribosomal subunit (Gradi et al., 1998). Under such conditions, viruses should have an alternate strategy to overcome the barrier and be able to utilize the host cellular machinery for translating its proteins. IRES-dependent translation remains functional in these conditions as it requires the C-terminal portion of eIF-4G which is available after cleavage by viral proteases (Ohlmann et al., 1996). An IRES driven mechanism may also aid in translation of mRNAs that are constrained by numerous upstream AUGs or RNA secondary structures in their 5' leader sequence (Chappell et al., 2001; Le Quesne et al., 2001).

Tobamoviruses express 3' proximal proteins by production of sgRNA (reviewed by Palukaitis and Zaitlin, 1986) and by utilizing an alternative cap independent translational mechanisms (Skulachev et al.,1999; Ivanov et al.,1997,). TMV-Cr possesses IRES elements for driving the expression of MP and CP ORFs, whereas in TMV-U1, a 228bp segment upstream of the MP gene was shown to promote internal initiation *in vitro*. Polypurine (A) rich sequences (PARS) were identified to be responsible for TMV-Cr CP IRES (IRES $_{CP148}^{Cr}$) to be functional. In HCRSV, a carmovirus a small sequence present upstream of its CP ORF, which

is complementary to the 3' portion of 18S rRNA, showed IRES- like activity (Koh et al., 2003). In HVS two IRES-like sequences present upstream of the MP and CP ORFs were identified and characterized and were shown to be active *in vitro* in the WGE systems and *in vivo* in whole plant assays.

5.2 Preliminary investigations for identifying putative IRES elements

Full-length HVS clone pHVS, cloned into pSPORTI, driven by the T7 promoter was linearized with Not I and translated in the WGE systems in the presence of ^{35}S methionine. Analysis of translation products by SDS-PAGE revealed the presence of the bands approximately 30 kDa and 15 kDa, which may correspond to the putative MP and CP gene products (Fig. 5.1a, lane 1). A clone containing MP and CP ORFs along with 365 bp upstream sequence of MP sequence pMPsg was constructed to compare the expression profile in the WGE systems with that of pHVS. The translation products of pMPsg showed bands of comparable sizes to that of pHVS around 30 kDa and 15 kDa in size (Fig. 5.1a, lane 2). This led to the speculation that the 30 kDa and 15 kDa products seen in the translation extracts of pHVS may indeed be putative MP and CP gene products and only an internal initiation event could result in production of these 3' proximal gene products.

The full-length HVS sequence was analyzed with the IRES Predictor program (Designed at the Department of Virology, Moscow State University) to identify putative IRES elements. The program returned with scores of 101.9 and 96 for sequences upstream of CP and MP respectively, indicating that these regions could act as IRES elements. In comparison, scores of 169 and 29 were obtained

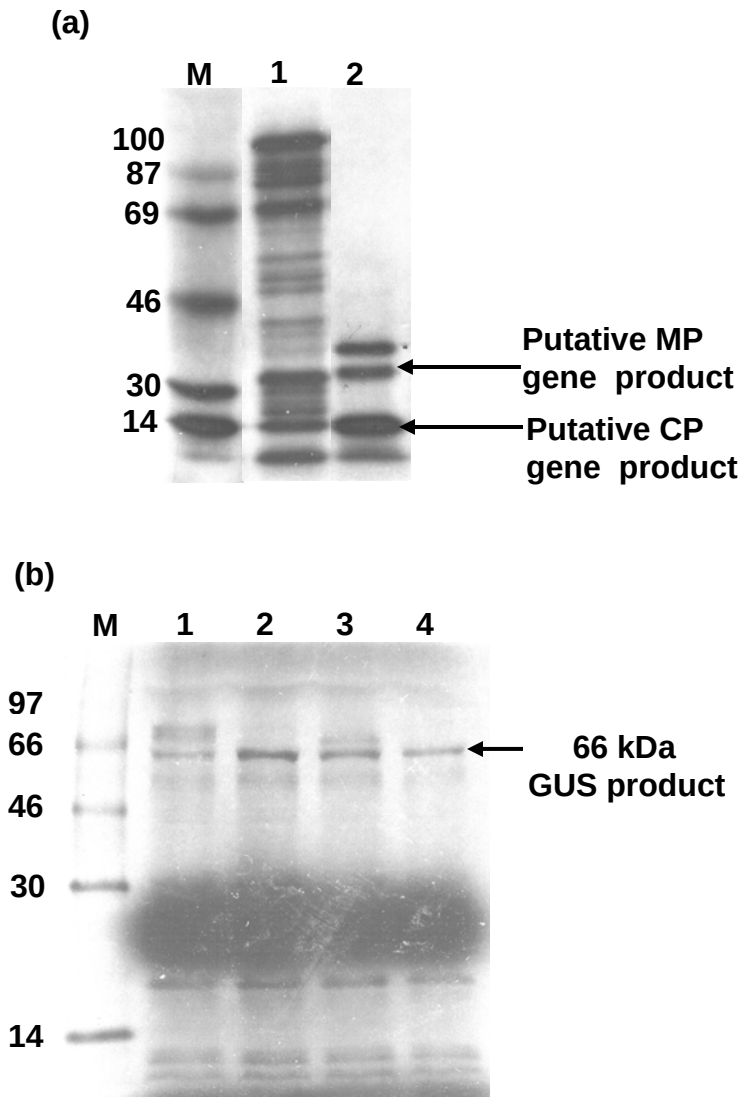


Fig 5.1 Autoradiograph of ^{35}S labeled i*n vitro* translation products derived from linearized plasmids driven by T7 promoter (**a**) full length clone pHVS (**1**), MP subgenomic clone pMPsg (**2**) in WGE system. Presence of 30 kDa translation product in (**1**) and (**2**) indicated the presence of a putative IRES region within the full length pHVS clone.

(b) Autoradiograph of *In vitro* translation products derived from linearized bicistronic IRES constructs driven by T7 promoter using WGE system. Lane (**1**) hGFP- HVS CP IRES 300 - GUS, (**2**) hGFP-CP134 IRES-GUS, (**3**) hGFP- HVS MP IRES 365- GUS, (**4**) hGFP-MP165 IRES-GUS. Shortened CP and MP IRESs of HVS produced a single 66 kDa band (**2**) and (**4**) which correlated with an increase in GUS activity in fluorimetric assays.

for the CP and MP IRESs of TMV-Cr. The scores are primarily based on polypurine(A) rich tracts present within the input sequences. Other elements within the sequence that could confer the ability to attract translation machinery and act as IRES do not seem to affect the scores (Skulachev M.V., pers. comm.). Subsequently, based on the IRES predictor results deletion constructs were created for HVS CP IRES region and tested using bicistronic constructs.

5.3 Testing of putative IRESs using bicistronic constructs *in vitro*

Putative IRES sequences were placed between two ORFs and the expression of the second ORF was assayed (Fig. 5.2a) to ascertain their ability to confer internal initiation. A stem-loop present in front of the first ORF ensures that the second ORF is not expressed by 5' ribosome scanning. Sequences upstream of HVS CP (300bp) and MP (365 bp) were cloned into hGFP-I-GUS bicistronic construct. The translation products of linearized bicistronic constructs harbouring putative CP (hGFP- HVS CPIRES 300-GUS) and MP (hGFP- HVS MPIRES 365-GUS) IRESs revealed the presence of a 66 kDa product which corresponds to the molecular weight of GUS gene. Occurrence of double bands in translation reactions were due to presence of AUGs within the IRES region that were in frame with the GUS ORF (Fig. 5.1b, lanes 1&3). Activity of GUS gene product was examined using fluorimetric assay which confirmed that GUS gene was indeed being expressed. To rule out the possibility that GUS gene could be expressed due to nucleolytic cleavage of the IRESs in the translation reactions, the integrity of the transcripts derived from these bicistronic constructs were assesed after incubating with DIG labeled GUS anti-sense cRNA probes and

compared with that of non- incubated transcripts. The presence of intact transcripts in the translation extracts confirmed the occurrence of internal initiation event resulting in the expression of GUS ORF.

5.4 Deletion analysis of HVS CP IRES i*n vitro*

Having established the presence of IRES element upstream of CP ORF, the next set of experiments were designed to locate the primary sequences or the secondary structures present within the CP IRES that are responsible for recruiting the host translational machinery. The results of the IRES predictor program indicated the presence of two polypurine tracts PPT I and PPT II located between nt 5666-5867 and 5783-5799, respectively. These results indirectly indicated that sequences 134 nt upstream of HVS CP ORF (nt 5666-5799) would be sufficient to drive the expression of GUS ORF in bicistronic constructs.

Preliminary experiments were done with sequences 300nt upstream of CP region. In the following experiments the 300nt IRES region was split up into a 166 bp 5' portion (hGFP-CP Upstream-GUS) and the clone containing the IRES region with PPT I and PPT II (hGFP- CP134-GUS). To dissect the IRES regions further, clones were constructed without the PPT region (hGFP- CP MID-GUS), and without either PPT I (hGFP-PPTI Del-GUS) or PPT II (hGFP-PPT II Del-GUS) (Fig. 5.2b).

GUS fluorimetric assay of translation samples with the HVS CP IRES deletion clones clearly indicated that PPTI and PPT II are essential for IRES-like activity. The clone containing the 300 nt sequence upstream of the CP region (hGFP-HVS CPIRES 300-GUS) contained the PPT regions and hence was able to direct

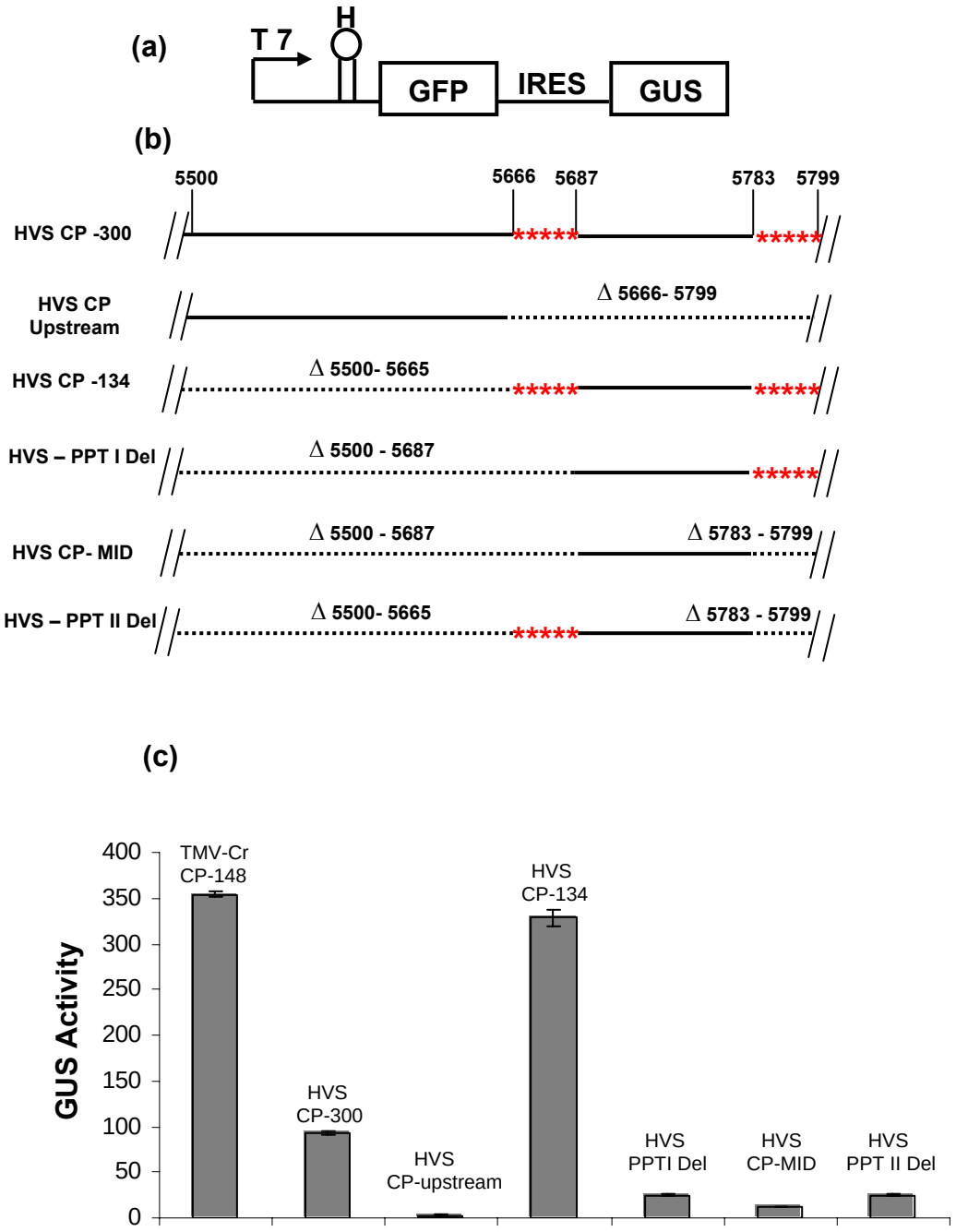


Fig 5.2 (a) Linear schematic of bicistronic construct **(b)** Schematic representation of deletion clones created to identify the region upstream of CP responsible for IRES-like activity **(c)** Deletion analysis of HVS CP- IRES. GUS activity was assayed by fluorimetric method and the values were taken for plotting the histogram. Mean values of five individual WGE translation reactions were taken and plotted with standard error bars. ***** Denotes polypurine tracts (PPT I &II) and **'H'** denotes hairpin loop.

the expression of GUS ORF. The clone (hGFP- CP134-GUS) containing PPT regions I and II gave maximum GUS activity higher than for the clone containing all of the upstream 300nt (hGFP- HVS CP 300-GUS) (Fig. 5.2c, HVS CP 300 vs HVS CP 134). This can be correlated with the disappearance of double bands in SDS-PAGE (Fig. 5.1b, lane 2).

The upstream 166 nt sequence (nt 5500-5665) without the PPT regions I and II showed no GUS activity (Fig. 5.2c, HVS CP Upstream) indicating that 166nt region upstream of CP IRES is not essential for IRES activity. Deletion of either PPT I (hGFP-PPTI Del-GUS) or PPT II (hGFP-PPT II Del-GUS) resulted in significant loss of IRES activity (Fig. 5.2c, HVS PPT I Del & HVS PPT II Del). As expected, the region between the two PPTs also showed a drastically reduced GUS activity (Fig. 5.2c,HVS-CP MID).

Taken together these results indicate unambiguously that PPT I and PPT II confer IRES like activity to HVS CP IRES *in vitro.* This presents an identical scenario to the TMV-Cr CP IRES where the PPT regions were shown to be responsible for the IRES-like activity. The efficiency of HVS CP IRES in driving GUS expression is comparable to that of TMV-Cr CP IRES (Fig. 5.2c, TMV-Cr CP-148 & HVS CP-134). The stability of RNA transcripts in translation extracts showed that the GUS expression was not due to cleavage of HVS IRES sequence by nucleases and that this second ORF was not being translated from degraded HVS CP IRES (Fig. 5.3c, lanes 2&5).

5.5 Deletion analysis of HVS MP IRES *in vitro*

Unlike the CP IRES, the MP IRES of HVS did not have stretches of polypurine tracts that could provide an easy starting point for deletion analysis. The whole 365nt region was taken and several secondary structures were predicted using MFOLD (Zuker,1989). Stable stem loop structures were present in four regions within the 365nt MP IRES region at nucleotide positions 4640-4670, 4760-4790, 4840-4860 and 4935-4955. In TMV-Cr, the MP IRES region was found to be present within the 75nt MP sgRNA leader (Skulachev et al., 1999) however, this IRES has no polypurine tracts. By analogy, sequences from the putative MP sgRNA transcription start site to the MP start codon could function as IRES element in HVS (nt 4905-4964). To avoid any aberrations that might result from designing clones based on such a speculation, sequences overlapping with the sg promoter region to the MP start codon (nt 4800-4964) were cloned into a bicistronic construct (hGFP-MP 165 IRES-GUS). Using multiple sequence alignments the putative MP sg promoter region in HVS was identified between nt 4816 and 4950. This clone results in the deletion of stem loops nt 4640-4670 and nt 4760-4790. Sequences spanning nt 4600-4799 were also tested for IRES activity (HVS MP-Upstream). Deletions clones covering secondary structures 4840-4860 (hGFP-MP 5'Del -GUS) and 4935-4955 (hGFP-MP 3'Del-GUS) were also tested for IRES activity.

GUS fluorimetric assays showed that all the deletion clones were capable of expressing GUS gene product (Fig. 5.3b). Clone hGFP-MP 165 IRES-GUS gave (Fig. 5.3b, HVS MP-165) the highest GUS activity that is comparable to the activity of the TMV-Cr MP IRES (Fig. 5.3b TMV-Cr MP 75). The clone containing

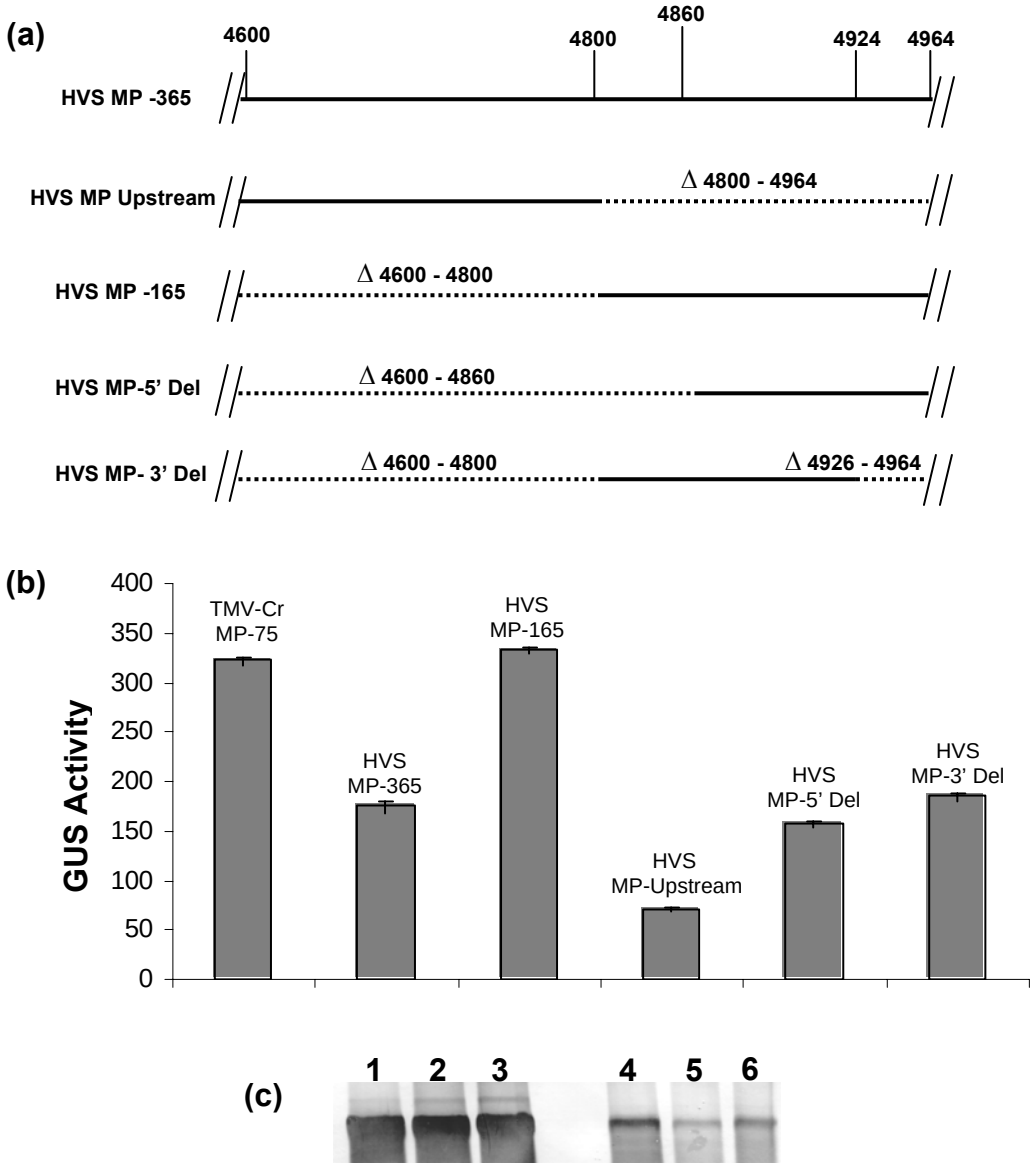


Fig. 5.3 (a) Schematic representation of deletion clones created to identify the region upstream of MP responsible for IRES-like activity. **(b)** Deletion analysis of HVS MP IRES. GUS activity was assayed by fluorimetric method and the values were taken for plotting the histogram. Mean values of five individual WGE translation reactions were taken and plotted with standard error bars.
(c) Northern blot analysis showing stability of bicistronic RNA transcripts in WGE system incubated for 90 min at 30^{o}C. Non incubated RNA controls (lanes 1 to 3) and RNA incubated in WGE (lanes 4 to 6) were hybridized to DIG- labelled GUS antisense probe. Lanes (1&4) hGFP- TMV-Cr CP IRES-GUS, (2&5) hGFP-CP134 IRES-GUS and (3&6) hGFP-MP 165 IRES-GUS.

sequences upstream of the MP sg promoter region (HVS MP-Upstream) was also capable of internal initiation (Fig. 5.3b, HVS-MP Upstream). The stability of RNA transcripts in the translation extracts showed that the HVS MP IRES 165 was not cleaved by nucleases and that the observed GUS expression was due to IRES actvitiy of the HVS MP IRES 165 (Fig. 5.3c, lanes 3&6). The reduced GUS activity in the hGFP-MP 365 IRES-GUS clone might be due to the presence of an ATG codon at nt 4634 (Fig. 5.3b, HVS MP 365). The deletions clones hGFP-MP 5'Del -GUS and hGFP-MP 3'Del-GUS also showed GUS activity but at reduced levels (Fig. 5.3b, HVS MP-5'Del & HVS MP-3'Del). The cumulative activity of these regions would account for GUS activity observed for hGFP-MP 165 IRES-GUS (Fig. 5.3b).

5.6 Functionality of HVS CP and MP IRESs *in vivo*

Experiment done using bicistronic constructs in the WGE system demonstrated that the HVS IRESs were active *in vitro*. The sequences responsible for IRES activity within CP IRES were identified. The next set of experiments were designed to test if the HVS CP and MP IRESs could function *in vivo*. This would indirectly indicate whether these IRESs may have an active role in the viral infection cycle.

The *in vivo* constructs were similar to *in vitro* constructs in architecture except that transcription was driven by the 35S promoter. All the deletion clones of CP IRES were created whereas for the MP IRESs, the 365 nt and 165 nt upstream sequences alone were tested. The kenaf protoplast system was used to test these pFF19G based 35S constructs. Owing to poor GUS activity in this system the IRESs were cloned into pCAMBIA 1300 so that they could be tested in *N.*

benthamiana by agro-inoculation. This construct was modified by cloning the castor bean catalase intron within the GUS ORF, to avoid any background activity observed due to the expression of GUS in Agrobacterium (Ohta et al., 1990). Initial experiments resulted in very low levels of GUS activity making it difficult to assay the efficiency of IRESs. It has been reported previously that T-DNA sequences by themselves could trigger RNA silencing which is directed against transcribed regions within the construct (Canto et al., 2002). This could be a reason for low levels of GUS expression observed in leaves infiltrated with the 35S IRES constructs. Co-infiltration of test constructs with a viral suppressor of gene silencing prevents the onset of PTGS (Voinnet et al., 2003). Constructs harbouring the HCRSV CP which is a gene silencing suppressor (unpublished) was co-infiltrated with the 35S IRES constructs. There was an increase in observed GUS activity which in turn facilitated meaningful comparison of efficiencies of different IRES constructs *in vivo*.

In vivo assay of HVS CP IRES constructs showed that the full-length IRES HVS CP-134 could express GUS at high levels as compared to its deletion mutants (Fig. 5.4b). Construct 35SGFP-HVS CPMID-GUS and 35SGFP-HVS PPT II del-GUS showed very low levels of GUS activity (Fig. 5.4b HVS CP-MID & HVS PPT II del). Deletion of PPT I (35SGFP-HVS PPTI Del-GUS) showed a reduction in GUS activity as compared to HVS CP-134 (Fig. 5.4c). The reduction in GUS activity was not significant and is not in agreement with *in vitro* tests (Fig. 5.2c, HVS PPT I del).

The MP IRES tests revealed that both MP 365 and MP 165 were active *in vivo* (Fig. 5.4b HVS MP 365 &HVS MP 165). The GUS activities of constructs

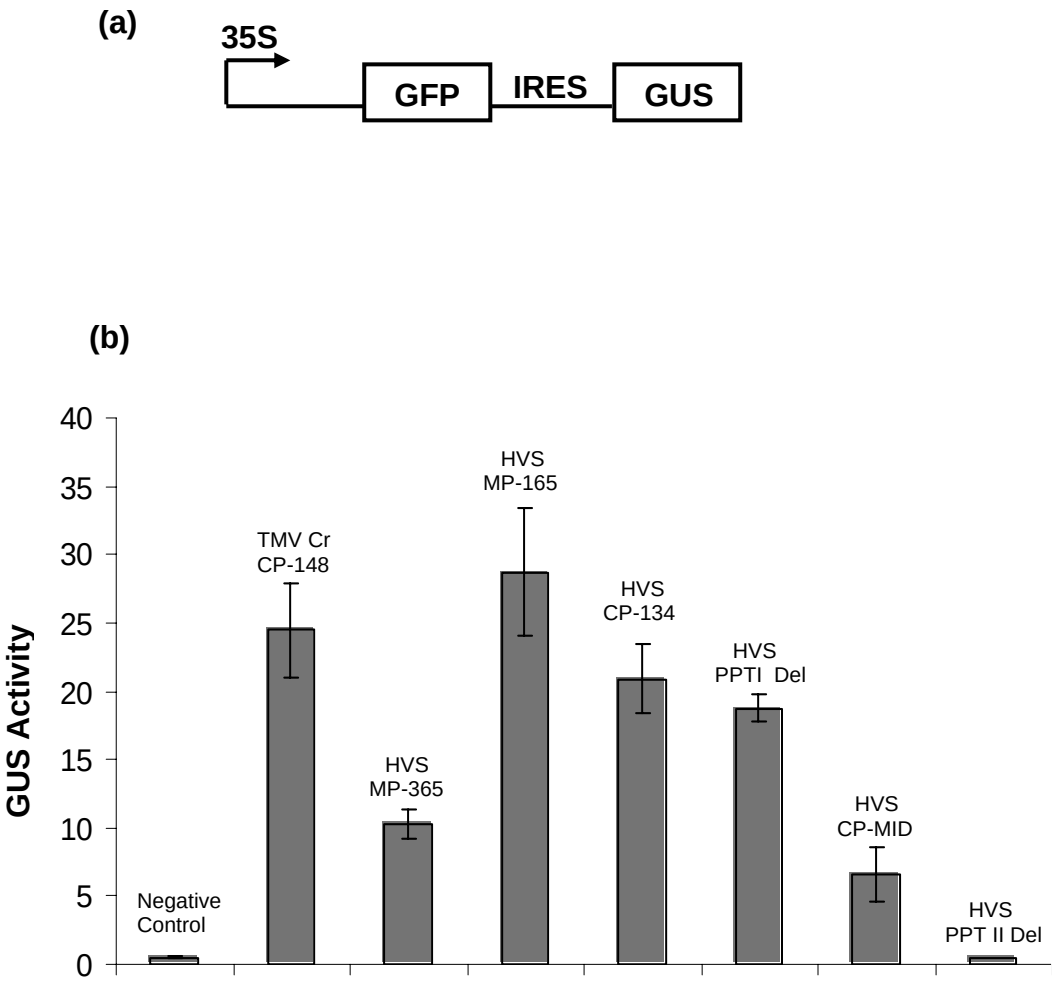


Fig. 5.4 (a) Schematic of 35S bicistronic construct. **(b)** *In vivo* GUS activity of HVS MP and CP IRES bicistronic deletion constructs under the control of 35S promoter. GUS activity was assayed by fluorimetric method and the values were taken for plotting the histogram. Mean values of five individual GUS fluorimetric reactions were taken and plotted with standard error bars. TMV-Cr CP 148 was used as a positive control. 35S GFP- Spacer-GUS served as negative control.

harbouring TMV-Cr CP-148 , HVS CP-134, MP-165 IRESs were similar. Promoter-less IRES-GUS constructs of MP365, MP 165 and CP 134 do not show any GUS activity (data not shown). These results further confirm that HVS sequences function as IRES elements rather than as cryptic pol II promoters *in vivo*.

5.7 Discussion

Viruses employ different mechanisms in order to compete with host for the recruiting cellular machinery. Use of alternative mechanisms for translation of viral gene products confers a distinct advantage for the viruses over the hosts. Several plant viruses have been reported to contain sequences that are capable of expressing 3' proximal ORFs by cap independent translation (Skulachev et al., 1999; Koh et al., 2003). Viruses lacking a 5' cap have compensatory mechanisms such as IRES sequences which helps in efficient translation of their gene products (Levis and Astier-Manifacier, 1993).

In HVS, there are two IRES sequences that may operate by entirely different mechanisms. The IRES element present upstream of CP gene has two polypurine tracts PPT I & PPT II, that are responsible for its IRES activity. Deletion of the PPT regions abolishes IRES activity *in vitro* (Fig. 5.2c). Activity of HVS CP-134 was comparable to TMV-Cr CP148. The activity of both the IRESs could be attributed to polypurine rich regions present within the IRESs.

Preliminary investigations suggest that secondary structures within the MP IRES are responsible for its IRES activity. Four stem loop structures were located within the 365 bp MP IRES region. Two of these were present within the 5' 200 nt region

(MP Upstream). Deletion of this region resulted in a stimulatory activity that correlated with enhanced GUS activity (Fig. 5.3b, HVS MP 365 Vs HVS MP 165). The upstream stem loops may have a negative regulatory influence on MP 365. The presence of AUG codons or secondary structures within this region hampering ribosome accessibility could lead to reduction of IRES activity. Stem loop structures located within MP 165 could individually drive the expression of the second ORF in bicistronic constructs, though the levels of expression were approximately 50% lower than for MP 165 (Fig. 5.3b, HVS MP 5'Del & HVS MP 3'Del). Both these stem loops may contribute equally towards the activity of MP 165.

The HVS CP and MP IRESs were found to be active *in vivo* in whole plant assays. Deletion of PPT I of CP IRES did not show a significant decrease in IRES activity. This is in contrast to results obtained *in vitro* in the WGE system. Generation of monocistronic RNA by a splicing mechanism may account for the increase in activity of this deletion construct *in vivo*.

IRESs vary to a great extent in secondary structures and their requirement for factors necessary to perform their function (Pestova and Hellen, 2003; Ali et al., 2001). HVS CP 134 contains polypurine tracts which is responsible for its activity. Artificially created polypurine rich sequences were found to promote internal initiation *in vitro* in the WGE and *in vivo* (tobacco protoplasts and HeLa cells). Such sequences were considered to be archetypal in nature and can function in almost all cell types (Dorokhov et al., 2002). Mechanisms underlying the ability to promote internal initiation by such sequences are not understood.

Similarly this study implicates two stem loop structures present within MP165 for the IRES like activity. Sequences within IRES regions that are complementary to 18S rRNA have been proven to be determinants of IRES activity (Koh et al., 2003; Zhou et al 2003). The mechanism by which MP 165 brings about internal initiation needs to be understood.

It is still questionable whether the 3' proximal genes in HVS could be expressed simultaneously from a sgRNA by cap dependent translation and from the gRNA by internal initiation. Like picornaviruses, it is possible that the plant viruses may employ IRES dependent translation under conditions of cellular stresses. The functional role of IRES dependent mechanism in viral infection cycle needs to be investigated.

CHAPTER 6

FUTURE DIRECTIONS

6.1 Genome analysis of HVS

The present study was aimed to characterize HVS and to construct biologically active cDNA clones of HVS. Based on nucleotide and amino acid identities, HVS was identified to be a new member of the tobamovirus genus. Although HVS shares very low sequence homology with members of the tobamoviruses, it is more closely related to subgroup II viruses. HVS, in spite of sharing significantly high sequence similarities with FHV, is distinctly different from this hibiscus infecting tobamovirus. Several interesting features within this tobamovirus are being reported for the first time in the present study.

In general, the tobamoviral 3' UTR consists of a tRNA-like structure at the 3' terminus and a pseudoknot domain. The tRNA-like structure is essential for minus-strand synthesis (Osman et al., 2000) and the pseudoknot domain acts as a translational enhancer (Gallie et al., 1991). The presence of a unique 3' UTR region in HVS makes it very different among tobamoviruses. The 3' UTR of HVS could be folded into a tRNA-like domain. However the pseudoknot domain upstream of the tRNA-like structure present in all other tobamoviruses was replaced by a poly(A) tract (77-96 bp) in HVS. This is a feature that HVS shares with BSMV (Agranovsky et al., 1978) although HVS did not cluster with BSMV when included in phylogenetic analysis. Functionally the poly(A) tract may substitute for the pseudoknot domain as it may act as a translational enhancer. It was observed that a poly(A) tract 15-43 nucleotides in length is sufficient for

associating itself with PABP, subsequently resulting in translation of proteins as a circular complex by the interaction of PABPs with the 5'-cap (Dreher, 1999).

Sequencing the 3'-terminus of several clones generated by conventional cDNA synthesis and cloning methods revealed that the poly(A) tract found within the 3' UTR region began at the same position i.e. immediately after the CP stop codon. This suggests that the poly(A) tract may be synthesized due to the slippage of viral polymerase and that it was not an artifact. A putative polyadenylation signal AAUAUA was found 105 nt upstream of poly(A) tract in HVS. In potexviruses such signals are present immediately preceding the poly(A) tail (Morozov et al., 1983) or 120 nt upstream of it (Abouhaidar, 1988). The poly(A) tail at the 3' end of viral sequences could be added by poly-A polymerase present in the cytoplasm (Jupin et al., 1990) or by slippage of viral polymerases. Sequences found within vesicular stomatitis virus (VSV) has been shown to influence the backward slippage of polymerase which results in elongation of the A-tract (Barr et al., 1997). The mechanism by which poly(A) tract is added in HVS remains obscure. The presence of an internal poly(A) tract strongly favours template-dependent polymerase slippage rather than a terminal transferase type of enzymatic activity. With the construction of infectious cDNA clones, the effect of mutating the polyadenylation signal on infectivity could be understood. To ascertain whether the poly(A) tract is indeed a replacement for pseudoknots, TMV clones with a poly(A) tract of similar length, in place of pseudoknots while retaining the TMV tRNA-like structure could be constructed and tested for infectivity. Double-helical segment I upstream of the tRNA-like structure is indispensable for TMV replication

(Takamatsu et al., 1990). Other domains within the pseudoknots could be removed and replaced with the poly(A) tract with marginal effects on infectivity.

6.2 Attempts towards generating infectious cDNA clones of HVS

Construction of infectious cDNA clone is essential to understand the functional aspects of the structures present within the viral genome and viral gene products. Several attempts were made to construct biologically active cDNA clones of HVS. Initially due to the lack of information on the 3' UTR apart from the poly(A) tract, near full-length clones of HVS without the 3' UTR or with a poly$(A)_{40}$ region immediately downstream of the CP gene were constructed. *In vitro* transcripts of these constructs were unable to replicate in kenaf protoplasts and were not infectious in *C. quinoa* and *N. benthamiana* whole plants. HVS clones with 3' UTRs of other tobamoviruses were constructed and tested for infectivity in kenaf protoplasts, *C. quinoa* and *N. benthamiana* whole plants. *In vitro* transcripts of these clones were unable to replicate in any of these systems. Assuming that the HVS replicase is unable to replicate the hybrid transcript, the replicase component was provided from wild type viruses from which the 3' UTRs were derived from by co-inoculation. Such experiments did not yield positive results. Finally, HVS with its authentic 3' UTR was tested for infectivity in all the above mentioned systems and it was found to be non-infectious. HVS clones were constructed by amplifying a 6.3 kb sequence using PCR. It is possible that these fragments may carry mutations that are lethal due to misincorporation of nucleotides during PCR amplification. It has been shown that the Taq error-rate estimate to be 0.27×10^{-4} misincorporation per bp per cycle (Bracho et al., 1998). Viral RdRp cannot

perform a proof reading function and hence have very high error rates of 7 X 10^{-4} to 5.4 X 10^{-3} (Ward et al., 1988). Under such circumstances in a given host, RNA of a virus may be present as a population containing viable as well as non-viable mutant molecules. This could influence the choice of template that is being amplified and cloned while constructing infectious clones. Population cloning strategy has been shown to be very useful in the construction of infectious cDNA clone of CymMV (Yu and Wong, 1998). This method is based on utilizing a pool of clones in each cloning step to generate a population of intermediate clones. Using the population of intermediate clones a large population of independent full-length clones could be constructed and tested. The efficiency of this method had been reported to be approximately 50%; i.e., the chances of obtaining an infectious clone is quite high. The advantage in this method is that slight inaccuracies in PCR could be tolerated. Moreover, there is a high chance for such mutant fragments being removed during subsequent digestion and cloning steps. A similar strategy could be employed for constructing infectious cDNA clones of HVS.

6.3 Identification of IRES-like elements within HVS genome

Similar to TMV-Cr, HVS has IRES regions present upstream of the MP and CP ORFs. A polypurine tract present within the CP IRES was responsible for its activity. It is not known as to how the polypurine tracts promote internal initiation. Secondary structures present within the MP IRES region are thought to be involved in conferring internal initiation. These structures could individually drive the expression of second ORF (GUS) in bicistronic constructs. This suggests that

the MP IRES is modular in nature which is a property observed in cellular IRESs (Chappell et al., 2000; Chappell and Mauro, 2003). The ability of the elements within the MP IRES to promote internal initiation needs to be investigated. Nevertheless, the current study shows that these IRESs are active *in vitro* in the WGE system and *in vivo* in whole plant assays.

Interaction studies leading to the an understanding of translation factors that can bind to the HVS IRESs would provide insights into the requirements of this plant viral IRESs to be functional. The activity of picornaviral IRES has been shown to be stimulated by its interaction with the poly(A) tail region. Interaction of the poly(A) tail with the IRES is achieved by association of eIF4G-PABP (Michel et al., 2001). Similarly, interaction of the HVS IRES with the poly(A) tract could be studied in depleted rabbit reticulocyte (RRL) systems or HeLa extracts. HeLa extracts show lower translational activity and are thus considered less efficient than RRL extracts. Depletion of RRL involves removal of ribosomal subunits, their associated translation factors and certain general RNA-binding proteins using ultracentrifugation. Under such conditions cap-poly(A) synergy is more pronounced (Michel et al., 2000).

With the construction of infectious cDNA clones of HVS, it would become possible to understand the influence of the poly(A) tract on the infectivity of HVS. Further insights into the ability of the poly(A) tract to affect the stability of RNA, its role in translation enhancement and the influence of a putative polyadenylation signal within HVS CP ORF in the generation of internal poly(A) tract could be gained.

REFERENCES

Abel, P.P., R.S. Nelson, B. De, N. Hoffmann, S.G. Rogers, R.T. Fraley, and R.N. Beachy. (1986). Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. Science **232**:738-743.

Abouhaidar, M. G. (1988). Nucleotide sequence of the capsid protein gene and 3' non-coding region of papaya mosaic virus RNA. J. Gen.Virol. **69:**219-26.

Adkins, S., R.W. Siegel, J.H. Sun, and C.C. Kao. (1997). Minimal templates directing accurate initiation of subgenomic RNA synthesis *in vitro* by the brome mosaic virus RNA-dependent RNA polymerase. RNA **3**:634-647.

Adkins, S., I. Kamenova, D. Achor, and D.J. Lewandowski. (2003). Biological and molecular characterization of a novel tobamovirus with a unique host range. Plant Dis. **87**:1190-1196.

Agranovsky, A.A., V.V. Dolja, V.M. Kavsan, and J.G. Atabekov. (1978). Detection of polyadenylate sequences in RNA components of barley stripe mosaic virus. Virol. **91**:95-105.

Ahlquist, P., V. Luckow, and P. Kaesberg. (1981). Complete nucleotide sequence of brome mosaic virus RNA3. J. Mol. Biol. **153**:23-38.

Aldaoud, R., W.O. Dawson, and G.E. Jones. (1989). Rapid, random evolution of the genetic structure of replicating tobacco mosaic virus populations. Intervirol. **30**:227-233.

Ali, I.K., L. McKendrick, S.J. Morley, and R.J. Jackson. (2001). Activity of the hepatitis A virus IRES requires association between the cap-binding translation initiation factor (eIF4E) and eIF4G. J. Virol. **75**:7854-7863.

Alonso, E., I. Garcia-Luque, A. de la Cruz, B. Wicke, M.J. Avila-Rincon, M.T. Serra, C. Castresana, and J.R. Diaz-Ruiz. (1991). Nucleotide sequence of the genomic RNA of pepper mild mottle virus, a resistance-breaking tobamovirus in pepper. J. Gen. Virol. **72**:2875-2884.

Altschuh, D., A.M. Lesk, A.C. Bloomer, and A. Klug. (1987). Correlation of co-ordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. J. Mol. Biol. **193**:693-707.

Atabekov ,J.G., and V.V. Dolja. (1986) Hordeiviruses: Structure and Replication In:The Plant Viruses. **2**. The rod shaped plant viruses (eds) Van Regenmortel MHV and Fraenkel-Conrat, pp 397-420. Plenum Press, New York.

Banerjee, N., J.Y.Wang, and M.Zaitlin. (1995). A single nucleotide change in the coat protein gene of tobacco mosaic virus is involved in the induction of severe chlorosis. Virol. **207**:234-239.

Barr, J.N., S.P. Whelan, and G.W. Wertz. (1997). cis-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation. J. Virol. **71:**8718-25.

Beachy, R.N., and M. Zaitlin. (1977). Characterization and *in vitro* translation of the RNAs from less-than-full-length, virus-related, nucleoprotein rods present in tobacco mosaic virus preparations. Virol. **81**:160-169.

Beier, H., M. Barciszewska, G. Krupp, R. Mitnacht, and H.J.Gross. (1984). UAG readthrough during TMV RNA translation: Isolation and sequence of two $tRNAs^{Tyr}$ with suppressor activity from tobacco plants. EMBO J. **3**:351-356.

Bendahmane, M., J.H. Fitchen, G. Zhang, and R.N. Beachy. (1997). Studies of coat protein-mediated resistance to tobacco mosaic tobamovirus: correlation between assembly of mutant coat proteins and resistance. J. Virol. **71**:7942-7950.

Bendahmane, M., J. Szecsi, I. Chen, R.H. Berg, and R.N. Beachy. (2002). Characterization of mutant tobacco mosaic virus coat protein that interferes with virus cell-to-cell movement. Proc. Natl. Acad. Sci. USA. **99**:3645-3650.

Bernstein, P., S.W. Peltz, and J. Ross. (1989). The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability *in vitro*. Mol. Cell. Biol. **9**:659-670.

Berzal-Herranz, A., A. de la Cruz, F. Tenllado, J.R. Diaz-Ruiz, L. Lopez, A.I. Sanz, C. Vaquero, M.T. Serra, and I. Garcia-Luque. (1995). The Capsicum L3 gene-mediated resistance against the tobamoviruses is elicited by the coat protein.Virol. **209**:498-505.

Bloomer, A.C., J.N. Champness, G. Bricogne, R. Staden, and A. Klug. (1978). Protein disk of tobacco mosaic virus at 2.8 A resolution showing interactions within and between subunits. Nature **276**:362-368

Bloomer, A.C., and P.J.G. Butler. (1986) In: The Plant Viruses **2**. The Rod-Shaped Plant Viruses. Van Regenmortel MHV and Fraenkel-Conrat H (eds), pp 19-57. Plenum Press, New York.

Bol, J.F., L. van Vloten-Doting, and E.M. Jaspars. (1971). A functional equivalence of top component a RNA and coat protein in the initiation of infection by alfalfa mosaic virus. Virol. **46**:73-85.

Bracho, M.A., A. Moya, and E. Barrio. (1998). Contribution of Taq polymerase-induced errors to the estimation of RNA virus diversity. J. Gen. Virol. **79:**2921-2928.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248-254.

Brierley, I., P. Digard, and S.C. Inglis. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell **57:**537-547.

Brunt, A.A., K. Crabtree, M.J. Dallwitz, A.J. Gibbs, L. Watson, and E.J. Zurcher. (1996 onwards). `Plant Viruses Online: Descriptions and Lists from the VIDE Database. Version: 20th August 1996.' URL http://biology.anu.edu.au/Groups/MES/vide.

Butler, P.J., and A. Klug. (1971). Assembly of the particle of tobacco mosaic virus from RNA and disks of protein. Nat. New. Biol. **229**:47-50.

Bustamante, P.I., and Hull R. (1998). Plant virus gene expression strategies. Electronic Journal of Biotechnology [online]. **1**:Available from: http://www.ejbiotechnology.info/content/vol1/issue2/full/3/.

Canto, T., F. Cillo, and P. Palukaitis. (2002). Generation of siRNAs by T-DNA sequences does not require active transcription or homology to sequences in the plant. Mol. Plant Microbe Interact. **15:**1137-1146.

Champness, J.N., A.C. Bloomer, G. Bricogne, P.G. Butler, and A. Klug. (1976). The structure of the protein disk of tobacco mosaic virus to 5A resolution.
Nature **259**:20-24.

Chappell, S.A., G.M. Edelman, and V.P. Mauro. (2000). A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. Proc. Natl. Acad. Sci. USA. **97**:1536-1541.

Chappell, S.A., G.C. Owens, and V.P. Mauro. (2001). A 5' Leader of Rbm3, a Cold Stress-induced mRNA, Mediates Internal Initiation of Translation with Increased Efficiency under Conditions of Mild Hypothermia. J. Biol. Chem. **276**:36917-36922.

Chappell, S.A., and V.P. Mauro. (2003). The internal ribosome entry site (IRES) contained within the RNA-binding motif protein 3 (Rbm3) mRNA is composed of functionally distinct elements. J. Biol. Chem. **278**:33793-33800.

Chen, M.H., J. Sheng, G. Hind, A.K. Handa, and V. Citovsky. (2000). Interaction between the tobacco mosaic virus movement protein and host cell pectin methylesterases is required for viral cell-to-cell movement. EMBO J. **19**:913-920.

Chen M.H., and V. Citovsky. (2003). Systemic movement of a tobamovirus requires host cell pectin methylesterase. Plant J. **35**:386-392.

Chng, C.G., S.M. Wong, P.H. Mahtani, C.S. Loh, C.J. Goh, M.C. Kao, M.C. Chung, and Y. Watanabe. (1996). The complete sequence of a Singapore isolate of odontoglossum ringspot virus and comparison with other tobamoviruses. Gene **171**:155-161.

Citovsky, V., D. Knorr, G. Schuster, and P. Zambryski. (1990). The P30 movement protein of tobacco mosaic virus is a single-strand nucleic acid binding protein. Cell **60**:637-647.

Citovsky, V., M.L. Wong, A.L. Shaw, B.V. Prasad, and P. Zambryski. (1992). Visualization and characterization of tobacco mosaic virus movement protein binding to single-stranded nucleic acids. Plant Cell **4**:397-411.

Citovsky, V., B.G. McLean, J.R. Zupan, and P. Zambryski. (1993). Phosphorylation of tobacco mosaic virus cell-to-cell movement protein by a developmentally regulated plant cell wall-associated protein kinase. Genes Dev. **7**:904-910.

Cooper, B., I. Schmitz, A.L. Rao, R.N. Beachy, and J.A. Dodds. (1996). Cell-to-cell transport of movement-defective cucumber mosaic and tobacco mosaic viruses in transgenic plants expressing heterologous movement protein genes. Virol. **216**:208-213.

Culver, J.N., K. Lehto, S.M. Close, M.E. Hilf, and W.O. Dawson. (1993).Genomic position affects the expression of tobacco mosaic virus movement and coat protein genes. Proc. Natl. Acad. Sci. USA. **90**:2055-2059.

Dam, E., K. Pleij, and D. Draper. (1992). Structural and functional aspects of RNA pseudoknots. Biochem. **31**:11665-11676.

Dawson, W.O., D.L. Beck, D.A. Knorr, and G.L. Grantham. (1986). cDNA cloning of the complete genome of tobacco mosaic virus and production of infectious transcripts. Proc. Natl. Acad. Sci. USA **83**:1832-1836.

Dawson, W.O., P. Bubrick, and G.L. Grantham. (1988). Modification of the tobacco mosaic virus coat protein gene affecting replication, movement and symptomatology. Phytopathol. **78**:783-789.

de la Cruz, A., L. Lopez, F. Tenllado, J.R. Diaz-Ruiz, A.I. Sanz, C. Vaquero, M.T. Serra, and I. Garcia-Luque. (1997). The coat protein is required for the elicitation

of the Capsicum L2 gene-mediated resistance against the tobamoviruses. Mol. Plant. Microbe Interact. **10**:107-113.

Deom, C.M., S. Wolf, C.A. Holt, W.J. Lucas, and R.N. Beachy. (1991). Altered function of the tobacco mosaic virus movement protein in a hypersensitive host. Virol. **180**:251-256.

Deom, C.M., X.Z. He, R.N. Beachy, and A.K. Weissinger. (1994). Influence of heterologous tobamovirus movement protein and chimeric-movement protein genes on cell-to-cell and long-distance movement. Virol. **205**:198-209.

Devereux, J., P. Haeberli, and O. Smithies. (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387-395.

Dorokhov, Y.L., K. Makinen, O.Y. Frolova, A. Merits, J. Saarinen, N. Kalkkinen, J.G. Atabekov , and M. Saarma. (1999). A novel function for a ubiquitous plant enzyme pectin methylesterase: the host-cell receptor for the tobacco mosaic virus movement protein. FEBS Lett. **461**:223-228.

Dorokhov, Y.L., M.V. Skulachev, P.A. Ivanov, S.D. Zvereva, L.G. Tjulkina, A.Merits, Y.Y. Gleba, T. Hohn, and J.G. Atabekov. (2002). Polypurine (A)-rich sequences promote cross-kingdom conservation of internal ribosome entry. Proc. Natl. Acad. Sci. USA. **99**:5301-5306.

Dreher, T.W. (1999). Functions of the 3'- untranslated regions of positive strand RNA viral genomes. Ann. Rev. Phytopathol. **37**:151-174.

Dunigan, D.D., and M. Zaitlin. (1990). Capping of tobacco mosaic virus RNA. Analysis of viral-coded guanylyltransferase-like activity. J. Biol. Chem. **265**:7779-7786.

Erickson, F.L., S. Holzberg, A. Calderon-Urrea, V. Handley, M. Axtell, C. Corr, and B. Baker. (1999). The helicase domain of the TMV replicase proteins induces the N-mediated defence response in tobacco. Plant J. **18**:67-75.

Fauquet, C.M. (1999) Taxonomy,classification and nomenclature of viruses. In : A.Granoffand R.G. Webster (eds) Encyclopedia of Virol., 2nd edn, pp. 1730-1756. Academic Press, New York.

Fedorkin, O.N., O.N. Denisenko, A.S. Sitikov, D.A. Zelenina, L.I. Lukasheva, S.Yu. Morozov, and J.G. Atabekov. (1995). A protein product of the tobamovirus open translation frame forms a stable complex with translation elongation factor eEF-1alpha. Dokl. Akad. Nauk. **343**:703-704.

Felden, B., C. Florentz, A. McPherson, and R. Giege. (1994). A histidine accepting tRNA-like fold at the 3'-end of satellite tobacco mosaic virus RNA. Nucleic Acids Res. **22**:2882-2886.

Fenczik, C.A., H.S. Padgett, C.A. Holt, S.J. Casper, and R.N. Beachy. (1995). Mutational analysis of the movement protein of odontoglossum ringspot virus to identify a host-range determinant. Mol. Plant. Microbe Interact. **8**:666-673.

Florentz, C., J.P. Briand, and R.Giege. (1984). Possible functional role of viral tRNA-like structures. FEBS Letters **176**:295-300.

Fraenkel-Conrat. (1957). Degradation of tobacco mosaic virus with acetic acid.Virol. **4**:1-4.

Fukuda, M., T. Meshi, Y. Okada, Y. Otsuki, and I. Takebe. (1981). Correlation between particle multiplicity and location on virion RNA of the assembly initiation site for viruses of the tobacco mosaic virus group. Proc. Natl. Acad. Sci. USA, **78**: 4231-4235.

Furuichi, Y., S. Muthukrishnan, J. Tomasz, and A.J. Shatkin. (1976). Mechanism of formation of reovirus mRNA 5'-terminal blocked and methylated sequence, m7GpppGmpC. J. Biol. Chem. **251**:5043-5053.

Futterer, J., Z. Kiss-Laszlo, and T. Hohn. (1993). Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. Cell **73**:789-802.

Futterer, J., H.M. Rothnie, T. Hohn, and I. Potrykus. (1997). Rice tungro bacilliform virus open reading frames II and III are translated from polycistronic pregenomic RNA by leaky scanning. J. Virol. **71**:7984-7989.

Gaines, P., and D.M. Wojchowski. (1999). pIRES-CD4t, a dicistronic expression vector for MACS- or FACS-based selection of transfected cells. Biotechniques **26**:683-688.

Gallie, D.R., D.E. Sleat, J.W. Watts, P.C. Turner, and T.M. Wilson. (1987). The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and in vivo. Nucleic Acids Res. **15**:3257-3273.

Gallie, D.R., D.E. Sleat, J.W. Watts, P.C. Turner, and T.M. Wilson. (1988). Mutational analysis of the tobacco mosaic virus 5'-leader for altered ability to enhance translation. Nucleic Acids Res. **16**:883-893.

Gallie DR, and V. Walbot. (1990). RNA pseudoknot domain of tobacco mosaic virus can functionally substitute for a poly(A) tail in plant and animal cells. Genes Dev. **4**:1149-1157.

Gallie, D.R., J.N. Feder R.T. Schimke and V. Walbot. (1991). Functional analysis of the tobacco mosaic virus tRNA-like structure in cytoplasmic gene regulation. Nucleic Acids Res. **19**:5031-5036.

Gallie, D.R., and V. Walbot. (1992). Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation. Nucleic Acids Res. **20**:4631-4638.

Gallie, D.R., and M. Kobayashi. (1994). The role of the 3'-untranslated region of non-polyadenylated plant viral mRNAs in regulating translational efficiency. Gene **142**:159-165.

Gal-On, A., Y. Antignus, A. Rosner, and B. Raccah. (1992). A zucchini yellow mosaic virus coat protein gene mutation restores aphid transmissibility but has no effect on multiplication. J. Gen. Virol. **73**:2183-2187.

Gan, W., and R.E. Rhoads. (1996). Internal initiation of translation directed by the 5'-untranslated region of the mRNA for eIF4G, a factor involved in the picornavirus-induced switch from cap-dependent to internal initiation. J. Biol. Chem. **271**:623-626.

Garcia-Arenal, F. (1988). Sequence and structure at the genome 3' end of the U2-strain of tobacco mosaic virus, a histidine-accepting tobamovirus. Virol. **167**:201-206.

Gibbs, A. (1986). Tobamovirus classification. In: The Plant Viruses.**2**. The rod shaped plant viruses (eds) Van Regenmortel MHV and Fraenkel-Conrat, pp 167-180. Plenum Press, New York.

Gibbs, A.J. (1999). Evolution and origins of tobamoviruses. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **354**:593-602.

Giesman-Cookmeyer, D., S. Silver, A.A. Vaewhongs, S.A. Lommel, and C.M. Deom. (1995). Tobamovirus and dianthovirus movement proteins are functionally homologous. Virol. **213**:38-45.

Goelet, P., G.P. Lomonossoff, P.J. Butler, M.E. Akam, M.J. Gait, and J. Karn.(1982). Nucleotide sequence of tobacco mosaic virus RNA. Proc. Natl. Acad. Sci. USA. **79**: 5818-5822.

Goldbach R, and J. Wellink. (1988). Evolution of plus-strand RNA viruses. Intervirol. **29**:260-267.

Goregaoker, S.P., and J.N. Culver. (2003). Oligomerization and activity of the helicase domain of the tobacco mosaic virus 126- and 183-kilodalton replicase proteins. J. Virol. **77**:3549-3556.

Gradi, A., Y.V. Svitkin, H. Imataka, and N. Sonenberg. (1998). Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. Proc. Natl. Acad. Sci. USA **95**:11089-11094.

Grdzelishvili, V.Z., S.N. Chapman, W.O.Dawson, and D.J. Lewandowski. (2000). Mapping of the Tobacco mosaic virus movement protein and coat protein subgenomic RNA promoters in vivo. Virol. **275**:177-192.

Gubler, U., and B.J. Hoffman. (1983).A simple and very efficient method for generating cDNA libraries. Gene **25:**263-9.

Guilley, H., G. Jonard, B. Kukla, and K.E. Richards. (1979). Sequence of 1000 nucleotides at the 3' end of tobacco mosaic virus RNA. Nucleic Acids Res. **6**:1287-1308.

Haenni, A.L., S. Joshi, and F. Chapeville. (1982). tRNA-like structures in the genomes of RNA viruses. Prog. Nucleic. Acid. Res. Mol. Biol. **27**:85-104.

Hamamoto, H., Y. Sugiyama, N. Nakagawa, E. Hashida, Y. Matsunaga, S.Takemoto, Y. Watanabe, and Y. Okada. (1993). A new tobacco mosaic virus vector and its use for the systemic production of angiotensin-I-converting enzyme inhibitor in transgenic tobacco and tomato. Biotechnology **11**:930-932.

Heinlein, M., B.L. Epel, H.S. Padgett, and R.N. Beachy. (1995). Interaction of tobamovirus movement proteins with the plant cytoskeleton. Science **270**:1983-1985.

Herman, R.C., S. Adler, R.A. Lazzarini, R.J. Colonno, A.K. Banerjee, and H. Westphal . (1978). Intervening polyadenylate sequences in RNA transcripts of vesicular stomatitis virus. Cell. **15**:587-596.

Higgins, T.J.V., P.B. Goodwin, and P.R. Whitfield. (1976). Occurence of short particles in beans infested with cowpea strain of TMV. II. Evidence that short particles contain the cistron for coat-protein. Virol. **71**:486-497.

Hilf, M.E., and W.O. Dawson. (1993).The tobamovirus capsid protein functions as a host-specific determinant of long-distance movement. Virol.**193**:106-114.

Hirashima, K., and Y. Watanabe. (2001). Tobamovirus replicase coding region is involved in cell-to-cell movement. J. Virol. **75**:8831-8836.

Hirashima K, and Y. Watanabe. (2003). RNA helicase domain of tobamovirus replicase executes cell-to-cell movement possibly through collaboration with its non-conserved region. J. Virol. **77**:12357-12362.

Hoffman, M.A., and A.C. Palmenberg. (1996). Revertant analysis of J-K mutations in the encephalomyocarditis virus internal ribosomal entry site detects an altered leader protein. J. Virol. **70**:6425-6430.

Hood, E.E., S.B. Gelvin, S. Melchers, and A. Hoekema (1993) New Agrobacterium helper plasmids for gene transfer to plants (EHA105). Trans. Res. **2**:208-218.

Houwing, C.J., and E.M. Jaspars. (2000). Activation of the alfalfa mosaic virus genome by viral coat protein in non-transgenic plants and protoplasts. The protection model biochemically tested. Arch. Virol. **145**:13-35.

Hu, M.C., P. Tranque, G.M. Edelman, and V.P. Mauro. (1999). rRNA-complementarity in the 5' untranslated region of mRNA specifying the Gtx homeodomain protein: vidence that base- pairing to 18S rRNA affects translational efficiency. Proc. Natl. Acad. Sci. USA. **96**:1339-1344.

Hunter, T.R., T. Hunt, J. Knowland, and D. Zimmern. (1976). Messenger RNA for the coat protein of tobacco mosaic virus. Nature **260**:759-764.

Hunter, T., R. Jackson, and D. Zimmern. (1983). Multiple proteins and subgenomic mRNAs may be derived from a single open reading frame on tobacco mosaic virus RNA. Nucleic Acids Res. **11:**801-821.

Hwang, D.J., I.M. Roberts, and T.M. Wilson. (1994). Expression of tobacco mosaic virus coat protein and assembly of pseudovirus particles in Escherichia coli. Proc. Natl. Acad. Sci. USA. **91**:9067-9071.

Ikeda, R., E. Watanabe, Y. Watanabe, and Y. Okada. (1993). Nucleotide sequence of tobamovirus Ob which can spread systemically in N gene tobacco. J. Gen. Virol. **74**:1939-1944.

Ikegami, M., Y. Isomura, Y. Matumoto, M. Chatani, and N. Inouye. (1995). The complete nucleotide sequence of odontoglossum ringspot virus (Cy-1 strain) genomic RNA. Microbiol. Immunol. **39**:995-1001.

Ishikawa, M., T. Meshi, F. Motoyoshi, N. Takamatsu, and Y. Okada. (1986). In vitro mutagenesis of the putative replicase genes of tobacco mosaic virus. Nucleic Acids Res. **14:**8291-8305.

Ishikawa, M., T. Meshi, T. Ohno, and Y. Okada. (1991). Specific cessation of minus-strand RNA accumulation at an early stage of tobacco mosaic virus infection. J. Virol. **65**:861-868.

Isomura, Y., Y. Matumoto, A. Murayama, M. Chatani , N. Inouye, and M. Ikegami. (1991). Molecular cloning, sequencing and expression in Escherichia coli of the odontoglossum ringspot virus coat protein gene. J. Gen. Virol. **72**:2247-2249.

Ivanov, P.A., O.V. Karpova, M.V. Skulachev, O.L. Tomashevskaya, N.P. Rodionova, Y.L. Dorokhov, and J.G. Atabekov. (1997). A tobamovirus genome that contains an internal ribosome entry site functional *in vitro*. Virol. **232**:32-43.

Jackson, A.O., D.M. Mitchell, and A. Siegel. (1971). Replication of tobacco mosaic virus. I. Isolation and characterization of double-stranded forms of ribonucleic acid. Virol. **45**:182-191.

Jang, S.K., H. G. Krausslich, M.J.H. Nicklin, G.M. Duke, A.C. Palmenberg, and E. Wimmer. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus-RNA directs internal entry of ribosomes during invitro translation. J. Virol. **62**:2636-2643.

Jang, S.K., and E. Wimmer. (1990). Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. Genes Dev. **4**:1560-1572.

Jaspars, E.M., and C.J. Houwing. (2002). A genome-activating N-terminal coat protein peptide of alfalfa mosaic virus is able to activate infection by RNAs 1, 2 and 3 but not by RNAs 1 and 2. Further support for the messenger release hypothesis. Arch.Virol.**147**:857-863.

Jefferson, R.A., T.A. Kavanagh, and M.W. Bevan. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6**:3901-3907.

Johannes, G., M.S. Carter, M.B. Eisen, P.O. Brown, and P. Sarnow. (1999). Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. Proc Natl Acad Sci USA. **96**:13118-13123.

Jupin, I., S. Bouzoubaa, K. Richards, G. Jonard, and H. Guilley. (1990). Multiplication of beet necrotic yellow vein virus RNA 3 lacking a 3' poly(A) tail is accompanied by reappearance of the poly(A) tail and a novel short U-rich tract preceding it. Virol. **178:**281-284.

Karger, E.M., O.Y. Frolova, N.V. Fedorova, L.A. Baratova, T.V. Ovchinnikova, P. Susi, K. Makinen, L. Ronnstrand, Y.L. Dorokhov , and J.G. Atabekov. (2003).Dysfunctionality of a tobacco mosaic virus movement protein mutant mimicking threonine 104 phosphorylation. J.Gen.Virol. **84**:727-732.

Kashiwazaki, T., S.Yamashita, and Y. Doi. (1982). Ann. Phytopath.Soc.Japan **48**:395.

Kawakami, S., H.S. Padgett, D. Hosokawa, Y. Okada, R.N. Beachy, and Y.Watanabe. (1999). Phosphorylation and/or presence of serine 37 in the movement protein of tomato mosaic tobamovirus is essential for intracellular localization and stability in vivo. J. Virol. **73**:6831-6840.

Khan, A.H., K. Morita, C. Parquet Md Mdel, F. Hasebe, E.G. Mathenge, and A. Igarashi. (2002) Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. J. Gen. Virol. **83**: 3075-3084.

Kim, K.H., and C. Hemenway. (1996). The 5' nontranslated region of potato virus X RNA affects both genomic and subgenomic RNA synthesis. J. Virol. **70**:5533-5540.

Kim, K.H., and C.L. Hemenway. (1999). Long-distance RNA-RNA interactions and conserved sequence elements affect potato virus X plus-strand RNA accumulation. RNA **5**:636-645.

Knapp, E., W.O. Dawson, and D.J. Lewandowski. (2001). Conundrum of the lack of defective RNAs (dRNAs) associated with tobamovirus Infections: dRNAs that can move are not replicated by the wild-type virus; dRNAs that are replicated by the wild-type virus do not move. J. Virol. **75**:5518-5525.

Koh, D.C., S.M. Wong, and D.X. Liu. (2003). Synergism of the 3'-untranslated region and an internal ribosome entry site differentially enhances the translation of a plant virus coat protein. J. Biol. Chem. **278**:20565-20573.

Kolupaeva, V.G., C.U. Hellen, and I.N. Shatsky. (1996). Structural analysis of the interaction of the pyrimidine tract-binding protein with the internal ribosomal entry site of encephalomyocarditis virus and foot-and-mouth disease virus RNAs. RNA **2**:1199-1212.

Kolupaeva, V.G., T.V. Pestova, C.U. Hellen, and I.N. Shatsky. (1998). Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. J. Biol. Chem. **273**:18599-18604.

Koonin, E.V., A.R. Mushegian, E.V. Ryabov, and V.V. Dolja. (1991). Diverse groups of plant RNA and DNA viruses share related movement proteins that may possess chaperone-like activity. J. Gen. Virol. **72**:2895-2903.

Koonin, E.V., and V.V. Dolja. (1993). Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit. Rev. Biochem. Mol. Biol. **28**:375-430.

Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1-45.

Kozak, M. (1986). Regulation of protein synthesis in virus-infected animal cells. Adv. Virus. Res. **31**:229-292.

Kozak, M. (1989). The scanning model for translation: an update. J. Cell. Biol. **108**: 229-241.

Kozak, M. (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. **266**: 867-870.

Kozak, M. (2001)**.** Constraints on reinitiation of translation in mammals. Nucleic. Acids. Res. **29**:5226-5232.

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. Gene **299**:1-34.

Kubota, K., S. Tsuda, A. Tamai, and T. Meshi. (2003).Tomato mosaic virus replication protein suppresses virus-targeted posttranscriptional gene silencing. J. Virol. **77**:11016-11026.

Laemmli, U.K. (1970) Cleavage of structural proteins during assembly of head bacteriophage T4. Nature **227**:680-685.

Lartey, R.T., T.C. Voss, and U.Melcher. (1995).Completion of a cDNA sequence from a tobamovirus pathogenic to crucifers. Gene**166**:331-332.

Lartey, R.T., T.C. Voss, and U. Melcher.(1996). Tobamovirus evolution: gene overlaps, recombination, and taxonomic implications. Mol. Biol. Evol.**13**:1327-1338.

Le Quesne, J.P., M. Stoneley, G.A. Fraser, and A.E. Willis. (2001). Derivation of a structural model for the c-myc IRES. J. Mol. Biol. **310**:111-126.

Leathers, V., R.Tanguay, M. Kobayashi, and D.R.Gallie. (1993). A phylogenetically conserved sequence within viral 3' untranslated RNA pseudoknots regulates translation. Mol. Cell Biol. **13**:5331-5347.

Lebeurier, G., A. Nicolaieff, and K.E. Richards. (1977). Inside-out model for self-assembly of tobacco mosaic virus. Proc. Natl. Acad. Sci. USA. **74**:149-153.

Lehrach, H., D. Diamond, J.M. Wozney, and H. Boedtker. (1977). RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical re-examination. Biochem. **16**:4743-4751.

Lehto, K., G.L. Grantham, and W.O. Dawson. (1990a).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. Virol. **174**:145-157.

Lehto, K., P. Bubrick, and W.O. Dawson. (1990b). Time course of TMV 30K protein accumulation in intact leaves.Virol. **174**:290-293.

Levis, C., and S. Astier-Manifacier. (1993). The 5' untranslated region of PVY RNA, even located in an internal position, enables initiation of translation.Virus Genes **7**:367-379.

Lewandowski, D.J., and W.O. Dawson. (2000). Functions of the 126- and 183-kDa proteins of tobacco mosaic virus. Virol. **271**:90-98.

Liang, X.Z., S.W. Ding, and S.M. Wong. (2002) Development of a kenaf (*Hibiscus cannabinus L.*) protoplast system for replication study of hibiscus chlorotic ringspot virus. Plant Cell Rep. **20**:982-986.

Lim, A.A., S. Tachibana, Y. Watanabe, and S.M. Wong. (2002) Expression and purification of a neuropeptide nocistatin using two related plant viral vectors. Gene **289**:69-79.

Linder, P., P.F. Lasko, M. Ashburner, P. Leroy, P.J. Nielsen, K. Nishi, J. Schnier, and P.P. Slonimski . (1989). Birth of the D-E-A-D box. Nature **337**:121-122.

Liu, D.X., and S.C. Inglis. (1992). Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus. J. Virol. **66**:6143-6154.

Lopez de Quinto, S., and E. Martinez-Salas. (1997). Conserved structural motifs located in distal loops of aphthovirus internal ribosome entry site domain 3 are required for internal initiation of translation. J. Virol. **71**:4171-4175.

Lopez de Quinto, S., and E. Martinez-Salas. (2000). Interaction of the eIF4G initiation factor with the aphthovirus IRES is essential for internal translation initiation in vivo. RNA **10**:1380-1392.

Lopez de Quinto, S., M. Saiz, D. de la Morena, F. Sobrino, and E. Martinez-Salas. (2002). IRES-driven translation is stimulated separately by the FMDV 3'-NCR and poly(A) sequences. Nucleic Acids Res. **30**:4398-4405.

Luukkonen, B.G., W. Tan, and S. Schwartz. (1995). Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. J. Virol. **69**:4086-4094.

Malpica, J.M., A. Fraile, I. Moreno, C.I.Obies, J.W. Drake, and F. Garcia-Arenal (2002). The rate and character of spontaneous mutation in an RNA virus. Genetics **162**:1505-1511.

Mans, RM, C.W. Pleij, and L. Bosch. (1991). tRNA-like structures. Structure, function and evolutionary significance.Eur. J. Biochem. **201**:303-324.

McLean, B.G., J. Zupan, and P.C. Zambryski. (1995). Tobacco mosaic virus movement protein associates with the cytoskeleton in tobacco cells. Plant Cell **7**:2101-14.

Merits, A., R. Kettunen, K. Makinen, A. Lampio, P. Auvinen, L. Kaariainen, and T. Ahola. (1999). Virus-specific capping of tobacco mosaic virus RNA: methylation of GTP prior to formation of covalent complex p126-m7GMP. FEBS Lett. **455**:45-48.

Meshi, T., T. Ohno, H. Iba , and Y. Okada. (1981). Nucleotide sequence of a cloned cDNA copy of TMV (cowpea strain) RNA, including the assembly origin, the coat protein cistron, and the 3' non-coding region. Mol. Gen. Genet. **184**:20-25.

Meshi, T., F. Motoyoshi, T. Maeda, S. Yoshiwoka, H. Watanabe, and Y. Okada. (1989). Mutations in the tobacco mosaic virus 30-kD protein gene overcome Tm-2 resistance in tomato. Plant Cell **1**:515-522.

Michel, Y.M., D. Poncet, M. Piron, K.M. Kean, and A.M. Borman. (2000). Cap-Poly(A) synergy in mammalian cell-free extracts. Investigation of the requirements for poly(A)-mediated stimulation of translation initiation.J. Biol. Chem. **275:**32268-32276.

Michel, Y.M., A.M. Borman, S. Paulous, and K.M. Kean. (2001). Eukaryotic initiation factor 4G-poly(A) binding protein interaction is required for poly(A) tail-mediated stimulation of picornavirus internal ribosome entry segment-driven translation but not for X-mediated stimulation of hepatitis C virus translation. Mol. Cell Biol. **21:**4097-109.

Miller, W.A., T.W. Dreher, and T.C.Hall. (1985). Synthesis of brome mosaic virus subgenomic RNA *in vitro* by internal initiation on (-) sense genomic RNA. Nature **313**:68-70.

Miller, W.A., and G. Koev. (2000). Synthesis of subgenomic RNAs by positive-strand RNA viruses. Virol. **273**:1-8.

Morozov, S.Yu., V.M. Zakchariev, B.K. Chernov, V.S. Prasolov, Yu. V. Kozlov, J.G. Atabekov, and K.G. Skryabin. (1983). The analysis of the primary structure and localization of the coat protein gene of the genomic RNA of potato virus X. Dok. Akad. Nauk SSSR. **271**:211-215

Morozov, S.Yu., O.N. Denisenko, D.A. Zelenina, O.N. Fedorkin, A.G. Solovyev, E. Maiss, R. Casper, and J.G. Atabekov. (1993). A novel open reading frame in tobacco mosaic virus genome coding for a putative small, positively charged protein. Biochimie **75**:659-665.

Mundry, K.W., P.A. Watkins, T. Ashfield, K.A. Plaskitt, S. Eisele-Walter, and T.M. Wilson. (1991). Complete uncoating of the 5' leader sequence of tobacco mosaic virus RNA occurs rapidly and is required to initiate cotranslational virus disassembly *in vitro*. J. Gen. Virol. **72** :769-777.

Murashige, T., and F. Skoog. (1962) A revised medium for rapid growth and bioassay with tobacco tissue cultures. Physiol. Plant. **15**:473-497.

Murray, K.E., A.W. Roberts, and D.J. Barton. (2001). Poly(rC) binding proteins mediate poliovirus mRNA stability. RNA **7**:1126-1141.

Namba, K., and G. Stubbs. (1986). Structure of tobacco mosaic virus at 3.6 A resolution: implications for assembly. Science **231**:1401-1406.

Namba, K., R. Pattanayek, and G. Stubbs. (1989). Visualization of protein-nucleic acid interactions in a virus. Refined structure of intact tobacco mosaic virus at 2.9 A resolution by X-ray fiber diffraction. J. Mol. Biol. **208**:307-325.

Oberg, B., and L. Philipson. (1972). Binding of histidine to tobacco mosaic virus RNA. Biochem. Biophys. Res. Commun. **48**:927-932.

Ohlmann, T., M. Rau, V.M. Pain, and S.J. Morley. (1996). The C-terminal domain of eukaryotic protein synthesis initiation factor (eIF) 4G is sufficient to support cap-independent translation in the absence of eIF4E. EMBO J. **15**:1371-82.

Ohno, T., M. Aoyagi, Y. Yamanashi, H. Saito, S. Ikawa, T. Meshi, and Y. Okada. (1984). Nucleotide sequence of the tobacco mosaic virus (tomato strain) genome and comparison with the common strain genome. J. Biochem. (Tokyo). **96**:1915-1923.

Ohta, S., S. Mita, T. Hattori, and K. Nakamura. (1990) Construction and expression in tobacco of a b-glucuronidase (GUS) reporter gene containing an intron within the coding sequence. Plant Cell Physiol. **31**:805-814.

Okada, Y., and T. Ohno. (1972). Assembly mechanism of tobacco mosaic virus from its ribonucleic acid and protein. Mol. Gen. Genet. **114**:205-213.

Ooshika, I., Y.Watanabe, T. Meshi, Y.Okada, K. Igano, K. Inouye, and N.Yoshida.(1984). Identification of the 30K protein of TMV by imunoprecipitation with antibodies directed against a synthetic peptide. Virol. **132**:71-78.

Osman, T.A., and K.W. Buck. (1996). Complete replication in vitro of tobacco mosaic virus RNA by a template-dependent, membrane-bound RNA polymerase. J. Virol. **70:**6227-6234.

Osman, T.A., C.L. Hemenway, and K.W. Buck. (2000). Role of the 3' tRNA-like structure in tobacco mosaic virus minus-strand RNA synthesis by the viral RNA-dependent RNA polymerase *In vitro*. J. Virol. **74**:11671-11680.

Padgett, H.S., and R.N. Beachy. (1993). Analysis of a tobacco mosaic virus strain capable of overcoming N gene-mediated resistance. Plant Cell **5**:577-586.

Page, R.D.M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. Comp. Applc. Biosci. **12**: 357-358.

Palukaitis, P., and M. Zaitlin. (1986). Tobacco mosaic virus. Infectivity and replication. In: The Plant Viruses. **2**. The rod shaped plant viruses (eds) Van Regenmortel MHV and Fraenkel-Conrat, pp 105-131. Plenum Press, New York.

Pelham, H.R. (1978). Leaky UAG termination codon in tobacco mosaic virus RNA. Nature **272**:469-471.

Pelletier, J., and N. Sonenberg. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature **334**: 320-325.

Pestova, T.V., I.N. Shatsky, S.P. Fletcher, R.J. Jackson, and C.U. Hellen. (1998). A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes. Dev. **12**:67-83.

Pestova, T.V., and C.U. Hellen. (2003). Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. Genes. Dev. **17**:181-186.

Pilipenko, E.V., T.V. Pestova, V.G. Kolupaeva, E.V. Khitrina, A.N. Poperechnaya, V.I. Agol, and C.U. Hellen. (2000). A cell cycle-dependent protein serves as a template-specific translation initiation factor. Genes. Dev. **14**:2028-2045.

Pleij, C.W.A., J.P. Abrahams, A. van Belkum, K. Rietveld, and L. Bosch. (1987). The spatial folding of the 3' non coding region of aminoacylatable plant viral RNAs. In: Positive strand RNA Viruses. Brinton M and Rueckert R (eds) pp 299-316. Alan R. Liss, New York.

Poch, O., I. Sauvaget, M. Delarue, and N. Tordo. (1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J. **8** :3867-3874.

Poole, T. L., C. Wang, R.A. Popp, L.N.D. Potgieter, A. Siddiqui, and M.S. Collett.(1995) Pestivirus translation initiation occurs by internal ribosome entry. Virol. **206**:750-754.

Quadt, R., and E.M.J. Jaspars. (1989). RNA polymerases of plus-strand RNA viruses of plants. Mol. Plant Microbe Interact. **2**:219-223.

Ruiz, M.T., O. Voinnet, and D.C. Baulcombe. (1998). Initiation and maintenance of virus-induced gene silencing. Plant Cell **10**:937-946.

Ryabov, E.V., D.J. Robinson, and M.E. Taliansky. (1999). A plant virus-encoded protein facilitates long-distance movement of heterologous viral RNA. Proc. Natl. Acad. Sci. USA. **96**:1212-1217.

Ryabov, E.V., D.J. Robinson, and M. E. Taliansky. (2001). Umbravirus-encoded proteins both stabilize heterologous viral RNA and mediate its systemic movement in some plant species. Virol. **288**:391-400.

Saito, T., Y. Imai, T. Meshi, and Y. Okada . (1988). Interviral homologies of the 30K proteins of tobamoviruses. Virol. **167**:653-656.

Saito, T., K.Yamanaka, Y. Watanabe, N. Takamatsu, T. Meshi, and Y. Okada. (1989). Mutational analysis of the coat protein gene of tobacco mosaic virus in relation to hypersensitive response in tobacco plants with the N' gene. Virol. **173**:11-20.

Sambrook, J., E.F. Fritsch and T. Maniatis. (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York.

Sasaki, J., and N. Nakashima. (2000). Methionine-independent initiation of translation in the capsid protein of an insect RNA virus. Proc. Natl. Acad. Sci. USA. **97**:1512-1515.

Sawicki, S.G., and D.L. Sawicki. (1990). Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J. Virol. **64**:1050-1056.

Schmitz, J., D. Prufer, W. Rohde, and E.Tacke. (1996). Non-canonical translation mechanisms in plants: efficient *in vitro* and in planta initiation at AUU codons of the tobacco mosaic virus enhancer sequence. Nucleic Acids Res. **24**:257-263.

Shatkin, A.J. (1976) Capping of eukaryotic mRNAs. Cell **9**: 645-653.

Shivprasad, S., G.P. Pogue, D.J. Lewandowski, J. Hidalgo, J. Donson, L.K.Grill, and W.O. Dawson. (1999). Heterologous sequences greatly affect foreign gene expression in tobacco mosaic virus-based vectors. Virol. **255**:312-323.

Siegel, A., V. Hari, I. Montgomery, and K. Kolacz. (1976). A messenger RNA for capsid protein isolated from tobacco mosaic virus-infected tissue. Virol. **73**:363-371.

Siegel, A., V. Hari, and K. Kolacz. (1978). The effect of tobacco mosaic virus infection on host and virus-specific protein synthesis in protoplasts. Virol. **85**:494-503.

Silver, S., Quan, S., and C.M. Deom. (1996). Completion of the nucleotide sequence of sunn-hemp mosaic virus: a tobamovirus pathogenic to legumes.Virus Genes **13**:83-85.

Simon-Buela, L., and F. Garcia-Arenal. (1999). Virus particles of cucumber green mottle mosaic tobamovirus move systemically in the phloem of infected cucumber plants. Mol. Plant Microbe Interact. **12**:112-118.

Sit, T.L., A.A. Vaewhongs, and S.A. Lommel. (1998). RNA-mediated trans-activation of transcription from a viral RNA. Science **281**:829-832.

Skulachev, M.V., P.A. Ivanov, O.V. Karpova, T. Korpela, N.P. Rodionova, Y.L. Dorokhov, and J.G. Atabekov. (1999). Internal initiation of translation directed by the 5'-untranslated region of the tobamovirus subgenomic RNA $I_{(2)}$.Virol. **263**:139-154.

Skuzeski, J.M., L.M. Nichols, R.F. Gesteland, and J.F. Atkins. (1991). The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. J. Mol. Biol. **218**:365-373.

Sleat, D.E., P.C. Turner, J.T. Finch, P.J.G. Butler, and T.M.A. Wilson.(1986). Packaging of recombinant RNA molecules into pseudovirus particles directed by the origin of assembly sequence from tobacco mosaic virus RNA.Virol. **155**:299-308.

Sleat, D.E., D.R. Gallie, J.W. Watts, C.M. Deom, P.C. Turner, R.N. Beachy, and T.M. Wilson. (1988). Selective recovery of foreign gene transcripts as virus-like particles in TMV-infected transgenic tobaccos. Nucleic Acids Res. **16**:3127-3140.

Solis, I., and F. Garcia-Arenal. (1990). The complete nucleotide sequence of the genomic RNA of the tobamovirus tobacco mild green mosaic virus.Virol.**177**:553-558.

Solovyev, A.G., D.A. Zelenina, E.I. Savenkov, V.Z. Grdzelishvili, S.Y. Morozov, D.E. Lesemann, E. Maiss, R. Casper, and J.G. Atabekov. (1996). Movement of a barley stripe mosaic virus chimera with a tobacco mosaic virus movement protein.Virol. **217**:435-441.

Spitsin, S., K. Steplewski, N. Fleysh, H. Belanger, T. Mikheeva, S. Shivprasad, W .O. Dawson, H. Koprowski, and V. Yusibov. (1999). Expression of alfalfa mosaic virus coat protein in tobacco mosaic virus (TMV) deficient in the production

of its native coat protein supports long-distance movement of a chimeric TMV. Proc. Natl. Acad. Sci. USA. **96** :2549-2553.

Srinivasan, K.G., R. Narendrakumar, and S.M. Wong. (2002). Hibiscus virus S is a new subgroup II tobamovirus: evidence from its unique coat protein and movement protein sequences. Arch. Virol. **147**:1585-1598.

Stoneley, M., F.E. Paulin, J.P. Le Quesne, S.A. Chappell, and A.E. Willis. (1998). C-Myc 5' untranslated region contains an internal ribosome entry segment. Oncogene **16**:423-428.

Strauss, E.G., J.H. Strauss, and A.J. Levine. (1996) Virus evolution. In: Fundamental Virology, Third Edition. Fields, B.N., Knipe, P.M., and Howley, P.M. (eds.). pp 141-159. Lippincott - Raven, Philadelphia.

Sulzinski, M.A., K.A. Gabard, P. Palukaitis, and M. Zaitlin. (1985) Replication of tobacco mosaic virus. VIII. Characterisation of a third subgenomic TMV RNA. Virol. **145**:132-140.

Sutic, D.D., R.E. Ford, and M.T. Tosic. (1999) Handbook of plant virus diseases. CRC press LLC, Florida, USA.

Takamatsu, N., Y. Watanabe, T. Meshi, and Y. Okada. (1990). Mutational analysis of the pseudoknot region in the 3' noncoding region of tobacco mosaic virus RNA. J. Virol. **64**:3686-3693.

Takamatsu, N., Y., Watanabe, T., Iwasaki, T., Shiba, T., Meshi, and Y. Okada. (1991). Deletion analysis of the 5' untranslated leader sequence of tobacco mosaic virus RNA. J. Virol. **65**:1619-1622.

Taliansky, M., I.M. Roberts, N. Kalinina, E.V. Ryabov, S.K. Raj, D.J. Robinson, and K.J. Oparka. (2003). An umbraviral protein, involved in long-distance RNA movement, binds viral RNA and forms unique, protective ribonucleoprotein complexes. J. Virol. **77**:3031-3040.

Tan, S.H., M. Nishiguchi, M. Murata, and F. Motoyoshi. (2000).The genome structure of kyuri green mottle mosaic tobamovirus and its comparison with that of cucumber green mottle mosaic tobamovirus. Arch. Virol. **145**:1067-1079.

Tan, S.W., S.M. Wong, and R.M. Kini. (2000). Rapid simultaneous detection of two orchid viruses using LC- and/or MALDI-mass spectrometry. J.Virol. Methods **85**:93-99.

Tang, C.K., and D.E. Draper. (1989). Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. Cell **57**:531-536.

Thomas, A.A., E. ter Haar, J. Wellink, and H.O. Voorma. (1991). Cowpea mosaic virus middle component RNA contains a sequence that allows internal binding of ribosomes and that requires eukaryotic initiation factor 4F for optimal translation. J Virol. **65**:2953-2959.

Thomas, J.J., B. Falk, C. Fenselau, J. Jackman, and J. Ezzell. (1998). Viral characterization by direct analysis of capsid proteins. Anal. Chem. **70**:3863-3867.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G.Higgins. (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**: 4876-4882.

Thompson, S.R., and P. Sarnow. (2000). Regulation of host cell translation by viruses and effects on cell function. Curr. Opin. Microbiol. **3**:366-370.

Timmermans, M.C., P. Maliga, J. Vieira, and J. Messing. (1990) The pFF plasmids: cassettes utilising CaMV sequences for expression of foreign genes in plants. J. Biotechnol. **14**:333-344.

Todd, S., J.S. Towner, D.M. Brown, and B.L. Semler. (1997).Replication-competent picornaviruses with complete genomic RNA 3' noncoding region deletions. J. Virol. **71**:8868-8874.

Tranque, P., M.C. Hu, G.M. Edelman, and V.P. Mauro. (1998). rRNA complementarity within mRNAs: a possible basis for mRNA-ribosome interactions and translational control. Proc. Natl. Acad. Sci. USA. **95**:12238-12243.

Tsukiyama-Kohara, K., N. Iizuka, M. Kohara, and A. Nomoto. (1992). Internal ribosome entry site within hepatitis C virus RNA. J. Virol. **66**:1476-1483.

Turner, D.R., and P.J. Butler. (1986).Essential features of the assembly origin of tobacco mosaic virus RNA as studied by directed mutagenesis. Nucleic Acids Res. **14**:9229-9242.

Turner, D.R., L.E. Joyce, and P.J. Butler. (1988). The tobacco mosaic virus assembly origin RNA. Functional characteristics defined by directed mutagenesis. J. Mol. Biol. **203**:531-547.

Tyc K, M. Konarska, H.J. Gross, and W. Filipowicz. (1984). Multiple ribosome binding to the 5'-terminal leader sequence of tobacco mosaic virus RNA. Assembly of an 80S ribosome X mRNA complex at the AUU codon. Eur. J. Biochem.**140**:503-511.

Ugaki, M., M. Tomiyama, T. Kakutani, S. Hidaka, T. Kiguchi, R. Nagata, T. Sato, F. Motoyoshi, and M. Nishiguchi. (1991). The complete nucleotide sequence of cucumber green mottle mosaic virus (SH strain) genomic RNA. J. Gen. Virol. **72**:1487-1495.

Vagner, S., M.C. Gensac, A. Maret, F. Bayard, F. Amalric, H. Prats, H., and A.C. Prats. (1995a). Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes. Mol. Cell. Biol. **15**:35-44.

Vagner, S., A. Waysbort, M. Marenda, M.C. Gensac, F. Amalric, and A.C. Prats. (1995b). Alternative translation initiation of the Moloney murine leukemia virus mRNA controlled by internal ribosome entry involving the p57/PTB splicing factor. J. Biol. Chem. **270**:20376-20383.

van Belkum A, J.P. Abrahams, C.W. Pleij, and L. Bosch. (1985). Five pseudoknots are present at the 204 nucleotides long 3' noncoding region of tobacco mosaic virus RNA. Nucleic Acids Res. **13**:7673-7686.

Van Regenmortel, M.H.V. (1975) Antigenic relationships between strains of tobacco mosaic virus. Virol. **64**: 415-420.

Varma, A. (1986) Sunn-hemp mosaic virus. In: Van Regenmortel MHV, Heinz Frankel Conrat (eds) The Rod Shaped Plant Viruses **2**. pp 249-266. Plenum press, New York.

Voinnet, O., S. Rivas, P. Mestre, and D. Baulcombe. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J. **33:**949-56.

Waigmann, E., M.H. Chen, R. Bachmaier, S. Ghoshroy, and V. Citovsky. (2000). Regulation of plasmodesmal transport by phosphorylation of tobacco mosaic virus cell-to-cell movement protein. EMBO J. **19**:4875-4884.

Ward, C.D., M.A. Stokes, and J.B. Flanegan. (1988). Direct measurement of the poliovirus RNA polymerase error frequency in vitro. J. Virol. **62:**558-62.

Watanabe, Y., T. Meshi, and Y. Okada. (1984a). The initiation site for transcription of the TMV 30-kDa protein messenger RNA. FEBS Lett. **173**:247-250.

Watanabe, Y., Y. Emori, I. Ooshika, T. Meshi, T. Ohno, and Y.Okada. (1984b). Synthesis of TMV-Specific RNAs and proteins at the early stage of infection in tobacco protoplasts: transient expression of the 30K protein and its mRNA. Virol. **133**:18-24.

Watanabe, T., A. Honda, A. Iwata, S. Ueda, T. Hibi, and A. Ishihama. (1999). Isolation from tobacco mosaic virus-infected tobacco of a solubilized template-specific RNA-dependent RNA polymerase containing a 126K/183K protein heterodimer. J. Virol. **73**:2633-2640.

Waterworth, H.E., and A. Hadidi. (1998). In: Economic losses due to plant viruses. A.Hadidi, R.K.Khetarpal and H.Koganezawa (eds) Plant Virus Disease Control, pp.1-13 APS press,St Paul, MN.

Weber, H., S. Schultze, and A.J. Pfitzner. (1993). Two amino acid substitutions in the tomato mosaic virus 30-kilodalton movement protein confer the ability to overcome the Tm-2(2) resistance gene in the tomato. J. Virol. **67:**6432-6438.

Wen, X.Y., S. Mandelbaum, Z.H. Li, M. Hitt, F.L.Graham, T.S.Hawley, R.G. Hawley, and A.K. Stewart. (2001). Tricistronic viral vectors co-expressing interleukin-12 (1L-12) and CD80 (B7-1) for the immunotherapy of cancer: preclinical studies in myeloma. Cancer Gene Ther. **8**:361-370.

Whitfield, P.R., and T.J.V. Higgins. (1976).Occurence of short particles in beans infected with cowpea strain of TMV. Virol. **71**: 471- 485.

Widholm, J.M. (1972) The use of fluorescein diacetate and phenosafranine for determining viability of cultured plant cells. Stain Tech. **47**:189-194.

Wittman-Liebold, B., and H.G. Wittman. (1967). Coat proteins of strains of two RNA viruses, comparison of their amino acid sequences. Mol. Gen. Genet. **100**: 358-363.

Wolf, S., C.M. Deom, R.N. Beachy, and W.J. Lucas. (1989). Movement protein of tobacco mosaic virus modifies plasmodesmatal size exclusion limit. Science **246**: 377-379.

Wu, L., L. Jiang, Z. Zhou, J. Fan, Q. Zhang, H. Zhu, Q. Han, and Z. Xu. (2003). Expression of foot-and-mouth disease virus epitopes in tobacco by a tobacco mosaic virus-based vector. Vaccine **21**:4390-4398.

Yoon, J.Y., B.E. Min, J.K. Choi, and K.H. Ryu. (2002). Genome structure and production of biologically active *in vitro* transcripts of cucurbit-infecting Zucchini green mottle mosaic virus. **92:** 156-163

Yu, H.H., and S.M. Wong. (1998).Synthesis of biologically active cDNA clones of cymbidium mosaic potexvirus using a population cloning strategy. Arch. Virol. **143**:1617-1620.

Yueh, A., and R.J. Schneider. (2000). Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to 18S rRNA. Genes Dev. **14:** 414-421.

Zhang, X., and M.M. Lai. (1994). Unusual heterogeneity of leader mRNA fusion in a murine coronavirus: Implications for the mechanism of RNA transcription and recombination. J Virol. **68**:6626-6633.

Zhou, W., G.M. Edelman, and V.P. Mauro. (2003). Isolation and identification of short nucleotide sequences that affect translation initiation in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA. **100**:4457-4462.

Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. Science **244**:48-52.