2014

RNA SEQUENCE DETERMINANTS OF A COUPLED TERMINATION-REINITIATION STRATEGY FOR TRANSLATION OF DOWNSTREAM ORF IN HELMINTHOSPORIUM VICTORIAE VIRUS 190S AND OTHER VICTORIVIRUSES (FAMILY *TOTIVIRIDAE*)

Hua Li
University of Kentucky, mayawww@gmail.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Hua Li, Student

Dr. Said A. Ghabrial, Major Professor

Dr. Lisa J. Vaillancourt, Director of Graduate Studies

RNA SEQUENCE DETERMINANTS OF A COUPLED TERMINATION-REINITIATION STRATEGY FOR TRANSLATION OF DOWNSTREAM ORF IN HELMINTHOSPORIUM VICTORIAE VIRUS 190S AND OTHER VICTORIVIRUSES (FAMILY *TOTIVIRIDAE*)

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Agriculture at the University of Kentucky

By

HUA LI

Lexington, Kentucky

Director: Dr. Said A. Ghabrial, Professor of Plant Pathology

Lexington, Kentucky

2014

ABSTRCT OF DISSERTATION

RNA SEQUENCE DETERMINANTS OF A COUPLED TERMINATION-REINITIATION STRATEGY FOR TRANSLATION OF DOWNSTREAM ORF IN HELMINTHOSPORIUM VICTORIAE VIRUS 190S AND OTHER VICTORIVIRUSES (FAMILY *TOTIVIRIDAE*)

Double-stranded RNA fungal virus Helminthosporium victoriae virus 190S (genus *Victorivirus*, family *Totiviridae*) contains two large open reading frames (ORFs) that overlap in the tetranucleotide AUGA. Translation of the downstream ORF, which encodes the RNA-dependent RNA polymerase (RdRp), was previously proposed to depend on ribosomal reinitiation following termination of the upstream ORF, which encodes the capsid protein. In this study, I provided evidence to confirm that coupled termination-reinitiation (stop-restart) is indeed used. A dual-fluorescence method was established to define the RNA sequence determinants for RdRp translation. Stop-restart depends on a 32-nt stretch of RNA sequence immediately upstream of the AUGA motif, including a predicted pseudoknot structure. The presence of similar sequence motifs and predicted RNA structures in other victoriviruses suggest that they all share a related stop–restart strategy for RdRp translation. The close proximity of the secondary structure to the AUGA motif appears to be especially important for promoting translation of the downstream ORF. Normal strong preferences for AUG start codons and canonical sequence context for translation initiation of the downstream ORF appear somewhat relaxed. With dual-fluorescence system, reinitiation efficiency of the downstream ORF was determined to be ~ 3.9%. Pseudoknot swapping between the one in HvV190S and those predicted from other victoriviruses showed that reinitiation from the downstream ORF of HvV190S is quite tolerant to varying primary sequences of the various pseudoknots. Mutational analysis by introducing different combinations of nucleotide mutations into pseudoknot stems reproducibly confirmed the determinant role of pseudoknot on reinitiation using two different experimental systems. Together, these results provide the first example of coupled termination-reinitiation regulated by a simple pseudoknot structure. These data expanded the understanding of coupled termination-reinitiation mechanism employed by RNA viruses and refined a new model for genus *victorivirus*, the largest genus in the family *Totiviridae*. The dual fluorescence system used in this study represented the first application of an efficient *in vivo* assay for recording low-frequency events in filamentous fungi.

Key words: Helminthosporium victoriae virus 190S, translation reinitiation, stop-restart, RNA-dependent RNA polymerase, pseudoknot structure

HUA LI

January 25, 2014

RNA SEQUENCE DETERMINANTS OF A COUPLED TERMINATION-REINITIATION STRATEGY FOR TRANSLATION OF DOWNSTREAM ORF IN HELMINTHOSPORIUM VICTORIAE VIRUS 190S AND OTHER VICTORIVIRUSES (FAMILY *TOTIVIRIDAE*)

By

HUA LI

Said A. Ghabrial
Director of Dissertation

Lisa J. Vaillancourt
Director of Graduate Studies

January 25, 2014

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my mentor, Dr. Said A. Ghabrial. I want to thank him to accept me to be his last graduate student although he is close to his retirement and to help me accomplish my Ph.D. I want to thank him for his excellent instructions, great patience and encouragement, tremendous support and help that he have given to me in the past many years. Without him, I couldn't have been possible to present the work here. I am very lucky to have Dr. Ghabrial as a mentor as he always updated me with the latest knowledge in the field of fungal virology. His continuous enthusiasm and serious attitude throughout his research life have inspired me and will benefit me for my entire life.

I sincerely thank all my committee members, Dr. Brian C. Rymond, Dr. Mark L. Farman, Dr. Aardra P. Kachroo and outside examiner Dr. Rebecca Dutch for their valuable suggestions during committee meetings, great encouragement during qualification exams and critical questions on my dissertation. I also thank our lab collaborator, Dr. Max L. Nibert from Harvard University for giving great suggestions and editing on my dissertation project and manuscript preparation.

Many thanks to the departmental DGS, Dr. Lisa J. Vaillancourt for giving great instructions during the entire graduate study. I also appreciate a lot for the department chair, Dr. Chris L. Schardl for serving as a DGS during my qualification exams and the long graduation process when Dr. Vaillancourt is on her sabbatical leave.

I also sincerely thank the Ghabrial lab manager, Wendy M. Havens for her constant support and especially her amazing personality. Without her, my work in the lab couldn't

have been so convenient and efficient and happy too. I appreciate all the help from other members in the Ghabrial lab including Suryadevara S. Rao and Mohamed H. El-Habbak and previous postdoc fellows Ajay Singh and Jiatao Xie, also for their friendship to make me feel very relaxed and always be able to focus on the work.

I would also like to thank all the departmental faculties for bringing me great teachings and giving great talks. Thanks to all the department personnel for ever giving me any type of help in the past years. Thanks to all the lab managers for sharing your lab instruments and welcoming me to use them for many times.

My academic accomplishment could not have been achieved without the love and the unconditional support from my parents and my sister. Thanks for their enormous support, their understanding and their hard work through the past many years that have helped me step so far. My beloved, passed, great farther has given me the greatest love in the world and taught me to be a hardworking, responsible and morally upright person with himself as a great role model. I would also like to thank my husband Congqing and my precious daughter Mei, the two incredible persons in my life. My daughter, although bringing me sleepless nights after the birth, is continuously creating great and proud happiness into my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 CHAPTER ONE: INTRODUCTION

1.1 The discovery of dsRNA mycovirus HvV190S and its genomic characterization

Filamentous fungus*, Helminthosporim victoriae (Teleomorph: Cochliobolus victoriae*), was first described in 1946 as the causal agent of Victoria blight of oats, named after the parent cultivar, Victoria (1). The disease rose to epidemic magnitude and resulted in serious yield losses in most oat-growing regions of the United States from 1945 to 1948 (2). A disease of *H. victoriae* was observed by Lindberg in 1959 based on cultural abnormalities of some isolates (3). Diseased isolates were characterized by reduced rates of growth, excessive aerial mycelium, excessive sectoring, reduced spore production as shown in Fig. 1.1 and reduced virulence (4). Lindberg referred to these abnormalities as a disease of the fungus and showed that this disease was transmitted to normal isolates by hyphal anastomosis. He suggested a viral role in the disease of *H. victoriae*, however, at that time, viruses were not known to infect fungi (5). The year 1962 marks the first discovery of a fungal virus in cultivated mushroom (6), which suggests that infection by fungal viruses may be common. In 1978, Ghabrial and his group identified for the first time two distinct viruses infecting *H. victoriae* and designated them as HvV190S and HvV145S according to their sedimentation values (7). HvV190S consists of a nonsegmented RNA and is the prototype of the genus *Victorivirus*, family of *Totiviridae* whereas HvV145S genome comprises four RNA fragments leading to the creation of the Family *Chrysoviridae* (8, 9). Although diseased isolates of *H. victoriae* are doubly infected with those two viruses, mixed infection, however, is not required for development of the disease phenotype. Transfection assays with purified virions suggested that HvV190S is the major cause of the disease of *H. victoriae* and that HvV145S, like other chrysoviruses, does not appear to affect colony morphology (8). HvV190S infected *H. victoriae* exhibits symptoms typical of a disease phenotype, which is unusual for dsRNA fungal viruses, in that most of them do not cause symptoms in their respective hosts (10). Thus, HvV190S may provide a useful model system for studying fungal viruses that have debilitating or hypovirulent effects in their hosts, which may contribute to biological control of economically important plant pathogens (11). Molecular studies suggested that HvV190S possesses an undivided double-stranded RNA genome (dsRNA) packaged in isometric particles (12, 13). The ~5200-bp dsRNA

genome of HvV190S is nonsegmented and comprises two large overlapping ORFs, both of which are contained within the same genome-length mRNA (Fig. 1.2A). The upstream ORF (ORF1) encodes the capsid protein (CP), whereas the downstream ORF (ORF2), which is in the −1 frame relative to ORF1, encodes the RNA-dependent RNA polymerase (RdRp) (14). Several lines of evidence indicate that ORF1 translation initiates at the AUG codon at positions 290–292 (14, 15), even though this codon (underlined below) resides in an unfavorable context (16) for translation initiation (UCCAUGU). Other findings have verified that translation of HvV190S ORF1 terminates at the UGA codon at positions 2606–2608 (14, 15).

Although a single ORF encodes CP, the particles of HvV190S contain three related forms of CP: p88, p83, and p78 (Fig. 1.2B). Both i*n vivo* and *in vitro* studies have indicated that p88 and p83 are phosphorylated, whereas p78 is not (17). Results of *in vitro* translation studies with either the denatured HvV190S dsRNA or the full-length *in vitro* transcript have furthermore indicated that p88 is the primary translation product (17, 18). Expression studies in bacterial and eukaryotic systems have confirmed this finding and have also shown that p88 undergoes cleavages near its C-terminus to generate p83 and p78 (14, 15). Based on these findings, it has been proposed that the sites for phosphorylation reside within the C-terminal proximal region that is shared by p88 and p83, but is removed from p78 by limited proteolysis. The 5′ end of the genomic plus strand is uncapped, the mRNA contains a long (289-nt) 5′ untranslated region (UTR), and the 5′-proximal region is predicted to be highly structured, suggesting that ORF1 translation is likely to involve a cap-independent internal-initiation mechanism (14, 15).

HvV190S was initially classified within the genus *Totivirus* under the family *Totiviridae* (19). Like members of genus *Totivirus*, the 5179-bp HvV190S dsRNA genome comprises two large overlapping ORFs, both of which are contained within the same genome-length mRNA (Fig. 1.2A). Unlike members in genus *Totivirus* and numerous viruses from other genera in the family *Totiviridae*, which express the RdRp as a CP/RdRp fusion protein consequent to ribosomal frameshifting (20-24), HvV190S RdRp is expressed as a separate, nonfused protein (Fig. 1.2B) (25). Thus, in 2009, HvV190S and 10 other, related viruses from filamentous fungi have been recognized and grouped together as members of the new genus *Victorivirus* with HvV190S as the prototype strain (19). The RdRp-encoding ORF2 of HvV190S has its first in-frame AUG codon overlapping the CP stop

codon in the tetranucleotide AUGA at positions 2605–2608 (14, 25). This suggests that the expression of RdRp occurs by a mechanism different from that utilized by most totiviruses to express their RdRps (19).

1.2 Classic cap-dependent scanning translation mechanism and alternative IRES strategy in eukaryotic cells

Translation is a process that decodes mRNA into proteins in the factory of ribosomes with assisting factors, which involves three phases: initiation, elongation and termination. Protein biosynthesis is principally regulated at the initiation phase, other than during elongation and termination (26, 27). Viruses depend solely on their hosts to provide the ribosomal components necessary to synthesize their proteins, and thus any translation mechanism utilized by virus must be compatible with the host system (28). Most eukaryotic transcripts contain a single open reading frame and initiate translation by the scanning mechanism (29). The scanning model mainly postulates a mechanism by which eukaryotic ribosomes select initiation codon in an mRNA as simplified in Fig. 1.3A. It states that 40S ribosome subunit with eukaryotic initiation factors, GTP and Met-tRNAi initially binds to a cap structure on the 5’ end of mRNA and then migrates and scans downstream along the 5’-untranslated region (5’UTR), and finally stops at the proper AUG codon for initiation of translation. Such initiation of translation preferentially uses the 5’-proximal AUG triplets as start sites and an insertion of any upstream ORF inhibits initiation at the downstream coding sequences. Seminal studies have been documented primarily in mammalian cells and continuously evidenced in budding yeast (29-31). Features that contribute to proper selection of initiation codon depends mainly on the context around first AUG. A consensus sequence surrounding the first AUG was indicated by Kozak for vertebrate mRNAs and the strong sequence bias is limited to only the first 5 nt upstream of the AUG, 5’–GCC(A/G)CCAUGG-3’ (32-34). There is an overpresence of A or G at position -3, C at positions -1, -2, -4 and -5 and a dominance of G at position +4. In contrast, the 5’ UTR is highly biased for A with also a particularly strong preference for A at position -3 compared to what occurs in mammals (34, 35). Studies on filamentous fungi showed that they require a consensus sequence like 5’-CAXXAUGYC-3’, in which X prefer A or C whereas Y can be any nucleotides (36). Internal ribosomal entry site (IRES)-mediated translation is an alternative mechanism for translation initiation other than the classical ribosomal scanning strategy. Eukaryotic

cells have been confirmed to employed IRES (37). It is predicted that up to 10% of cellular mRNAs initiate translation by this mechanism. IRES regions are known to be located in the 5’ UTR of the mRNA and are capable of recruiting eukaryotic ribosomal translation initiation complex to a site that is a considerable distance from the cap structure (Fig 1.3B). A common feature of viral IRES is the long and highly structured 5’UTR, which bind ribosomal initiation complex and catalyze the formation of a functional ribosome. IRES is mainly employed by mammalian viruses such as hepatitis C virus, picornavirus and cricket paralysis virus. In contrast, sites for cellular IRES do not have much in common in terms of their sequence or secondary structure (26).

1.3 Noncanonical translation mechanisms

The cap-dependent ribosomal scanning mechanism presents a problem to many RNA viruses with either polycistronic or cap-deficient monocistronic genomes. Thus, a number of novel translation strategies have been developed to facilitate access of ribosomes to their unique open reading frames. The most common and earliest discovered strategies are polyprotein synthesis and translation from subgenomic cistrons. In addition, other strategies include: leaky scanning past the start codon of the first ORF (38), ribosomal shunting to a downstream start codon (39), programmed ribosomal frameshifting (40), read-through of stop codon (28), and more recently elucidated, coupled termination-reinitiation (also referred to as translational coupling or stop-restart) (41).

1.4 Polyprotein translation, stop codon read-through and ribosomal frameshifting

Among these mechanisms, polyprotein translation, stop codon read-through and ribosomal frameshifting involve a fusion protein synthesis for the downstream cistronic genes with regard to upstream genes. In polyprotein synthesis strategy (Fig 1.3C), the viral genome contains a long ORF, which translates a large polyprotein. The precursor protein is immediately cleaved into multiple mature proteins by virally-encode proteinases, which are usually small and functional for virus activity (42). Polyprotein translation was firstly reported on potyviruses. This group contains at least 180

designated members, which are very similar, in terms of their genomic structure and gene expression (42, 43).

Stop codon read-through (Fig 1.3D) describes “leaky” or unfavorable termination codon in the first cistron of a genomic viral RNA, which permits some of the ribosomes to read through into a downstream cistron as a result, giving rise to a second longer functional polypeptide (44). Many viruses employ in-frame read-through of stop codons to express low level of 3’extended proteins with regard to 5’ translated products. These include tobamoviruses, tobraviruses, tombusviruses, carmoviruses and luteoviruses.

Ribosomal frameshifting (Fig 1.3E) is usually involved in the translation of gag and pol gene expression. In most cases, viruses, in which gag and pol proteins are arranged in two different reading frames, are translated into a single protein via ribosomal frameshifting (21). Ribosomal frameshifting was recognized as a translation misreading in the early 1970s in bacteria *Escherichia coli* at an occasional frequency of $10^{-3\sim-4}$. The emerging viral cases started from late 1980S in retrovirus. Later it was found in eukaryotic positive RNA, dsRNA, plant RNA viruses and some bacteriophages (45). The achievement of reading of two different ORFs is determined mostly by two features within the RNA sequence: 1) a heptanucleotide slippery stretch 5’-XXXYYYN-3’ located at the frameshift site, where the ribosome switches the reading frame; 2) a stem-loop structure, usually a pseudoknot, which is located a few nucleotides downstream the heptanulceotide sequence (21, 40, 46). Primary work has been documented on the bias for sequence bias in the fungal yeast virus ScV-L-A and retroviruses such as Rous sarcoma virus (RSV) (47, 48). In the consensus sequence of XXXYYYN, the first triplet represents the same nucleotide from A, U, G, or C in the heptanucleotide region; the second triplet represents also the same nucleotide that was found naturally to be A or U for the slippery site (48); Z represents a single nucleotide in the last position that was shown to be mostly A, U or C, but can be replaced with G artificially (47).

1.5 Subgenomic translation, leaky scanning, ribosomal shunting and stop-restart

The expression of some internal genes such as coat proteins of RNA viruses is frequently mediated via subgenomic RNAs (Fig 1.3F), mostly mRNAs, in which the

subgenomic (+) RNA can be synthesized via internal transcription based on (-) RNA of genome length as template (49). The representative RNA viruses are tobacco mosaic virus (TMV) (50), cucumber mosaic virus (CMV) (51), brome mosaic virus (BMV) (52) and barley yellow dwarf virus (BYDV) (53).

Leaky scanning (Fig 1.3G) is a translation strategy employed by viruses to express two different proteins from a single RNA via ribosome skipping of the first AUG, which is present in sub-optimal context and continuing translation of next AUG (38). As mentioned for scanning mechanism, in most cases, position -3 has a strong preference for A or G in the following consensus sequence, 5'-(A/G)XXAUGY, in which X and Y have a different bias for nucleotides among different organism (16). However, suboptimal initiation sites in a weak context and a non-AUG codon, such as CUG, ACG or GUG, in a good surrounding context can cause fractional initiation with a low level of ribosomal scanning (28). This observation is due to the presence of weak basepairing between AUG codon and Met-tRNA (54). When the initiation sites are in different reading frames, two proteins of different sequences are made.

Ribosome shunting (Fig 1.3H) describes a pathway of translation initiation in which ribosomes bind to the mRNA in a cap-dependent manner, then jump over a large region of the mRNA containing RNA secondary structure, upstream AUGs and short ORFs to "land" at or immediately upstream of the major coding ORF AUG (55, 56). This pathway was first described for plant Cauliflower mosaic virus (CaMV) (57). Research suggested that 5' cap and the presence of special elements in the leader were required for shunting. The elements include a giant stem-loop structure, in which the base of the stem-loop and a small ORF (sORF) a few nucleotides upstream of the stem-loop are crucial, while the central part is dispensable. The scanning translation of sORF will pause at the donor site when encountering the stem-loop and bypass the secondary structure part to land at the corresponding acceptor site and translate the next downstream AUG (56).

For coupled termination-reinitiation mechanism (Fig 1.3I), ribosomes translate the upstream ORF of an mRNA, then following termination, a proportion of the 40S subunits remain bound to the mRNA and go on to reinitiate at the start codon of the downstream ORF (58). So far, only a few RNA viruses have been confirmed to use coupled

termination-reinitiation to express their downstream ORFs. The negative strand influenza B virus BM2 protein was the first example described to use such mechanism in 1990 (59). After more than a decade, this mechanism has been found for the expression of the M2-2 protein of respiratory syncytial virus (RSV) and pneumovirus of mice (PVM) (60) and the downstream protein VP10 and VP2 of the rabbit calicivirus RHDV and the feline calicivirus FCV (61, 62). Interestingly in 2009, the hypovirus CHV1 has also been discovered to utilize translational coupling (63). Research on these viruses suggested they shared some common characteristics. These features are: 1) An initiation codon is necessary for the translation of corresponding downstream genes, 2) The termination codon of the upstream ORF is critical for the initiation of the downstream ORF, 3) These two ORFs must overlap or are tightly spaced and 4) Sequences in the upstream ORF affect reinitiation of the downstream ORF.

1.6 Objectives

(i) Examine the RdRp translation mechanism from the downstream ORF of HvV190S;
(ii) Identify required RNA sequence elements for corresponding mechanism utilized by RdRp in HvV190S.

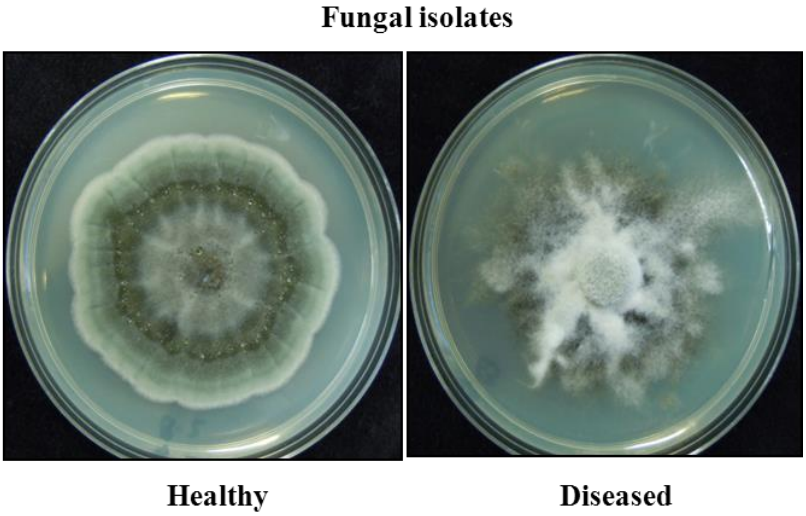


Figure 1.1 Colony morphology of healthy (left) and diseased (right) isolates of *H. victoriae* growing on PDAY medium.

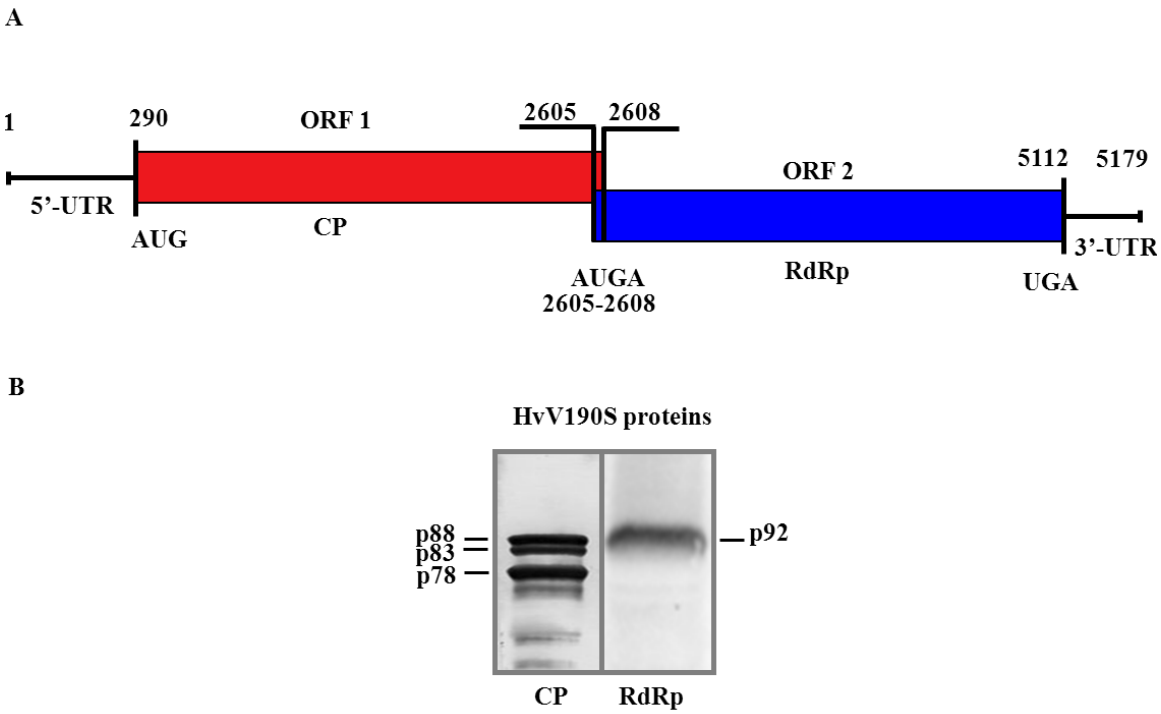
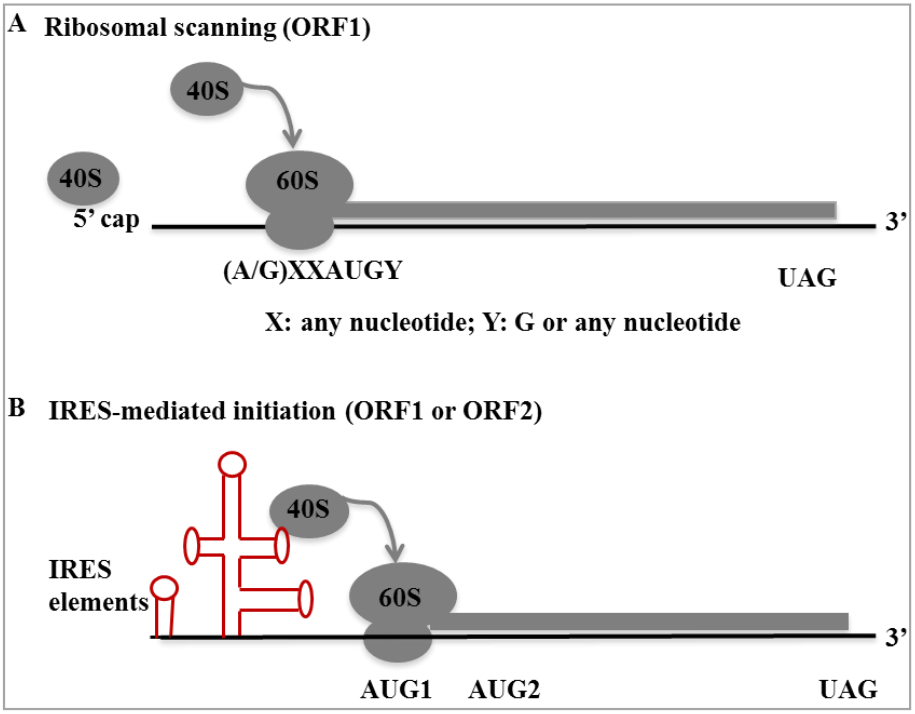
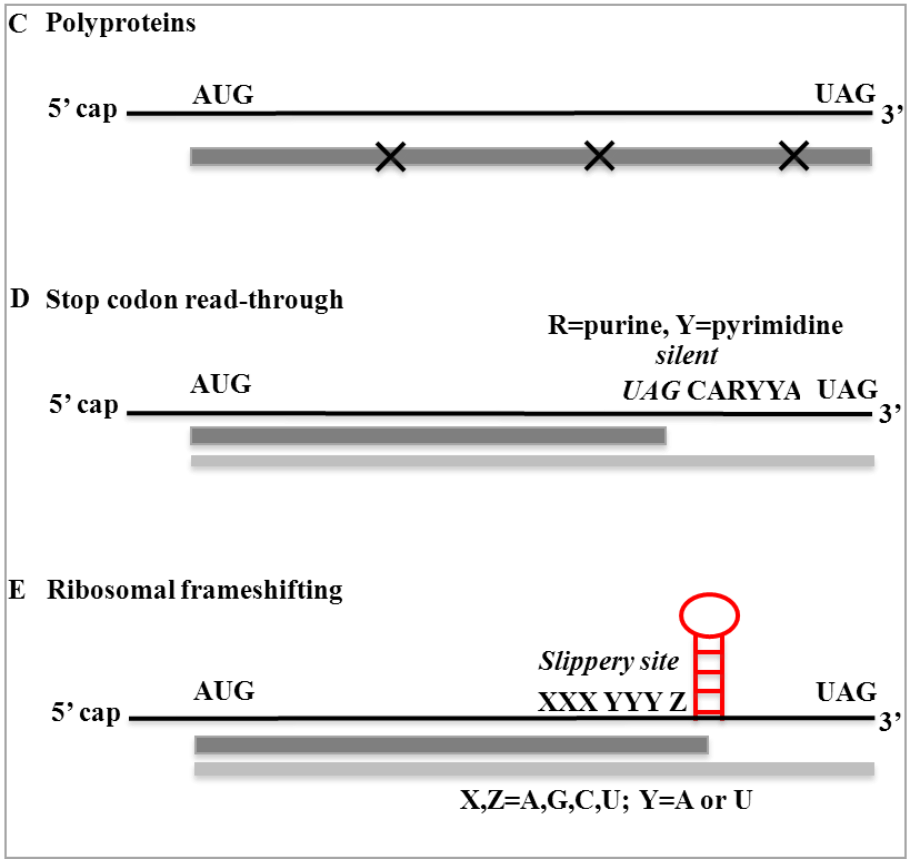


Figure 1.2 Schematic representation of genomic organization of HvV190S and analysis of expressed viral proteins.

(A)The 5179 nucleotide-long genomic plus strand of HvV190S contains two large ORFs, with the 5'-proximal ORF encoding the CP and the 3'-proximal ORF encoding the RdRp (not drawn to scale). The start codon of ORF2 overlaps the stop codon of ORF1 in the tetranucleotide AUGA at position nt 2605- nt 2608. Start codon for ORF1 and stop codon for ORF2 and their locations on the genome are indicated. The genome also contains 5'- and 3'-untranslated regions (UTRs). (B) Immunoblot analysis of viral proteins of HvV190S obtained from purified viral particles. Three types of CP are detected with the CP-specific antiserum whereas a single separate protein is detected with the RdRp-specific antiserum.


A Ribosomal scanning (ORF1)
40S
40S
60S
5' cap
3'
(A/G)XXAUGY
UAG
X: any nucleotide; Y: G or any nucleotide
B IRES-mediated initiation (ORF1 or ORF2)
40S
IRES
elements
60S
3'
AUG1
AUG2
UAG


C Polyproteins
5' cap
AUG
UAG
3'
D Stop codon read-through
R=purine, Y=pyrimidine
silent
5' cap
AUG
UAG CARYYA UAG
3'
E Ribosomal frameshifting
Slippery site
5' cap
AUG
XXX YYY Z
UAG
3'
X,Z=A,G,C,U; Y=A or U

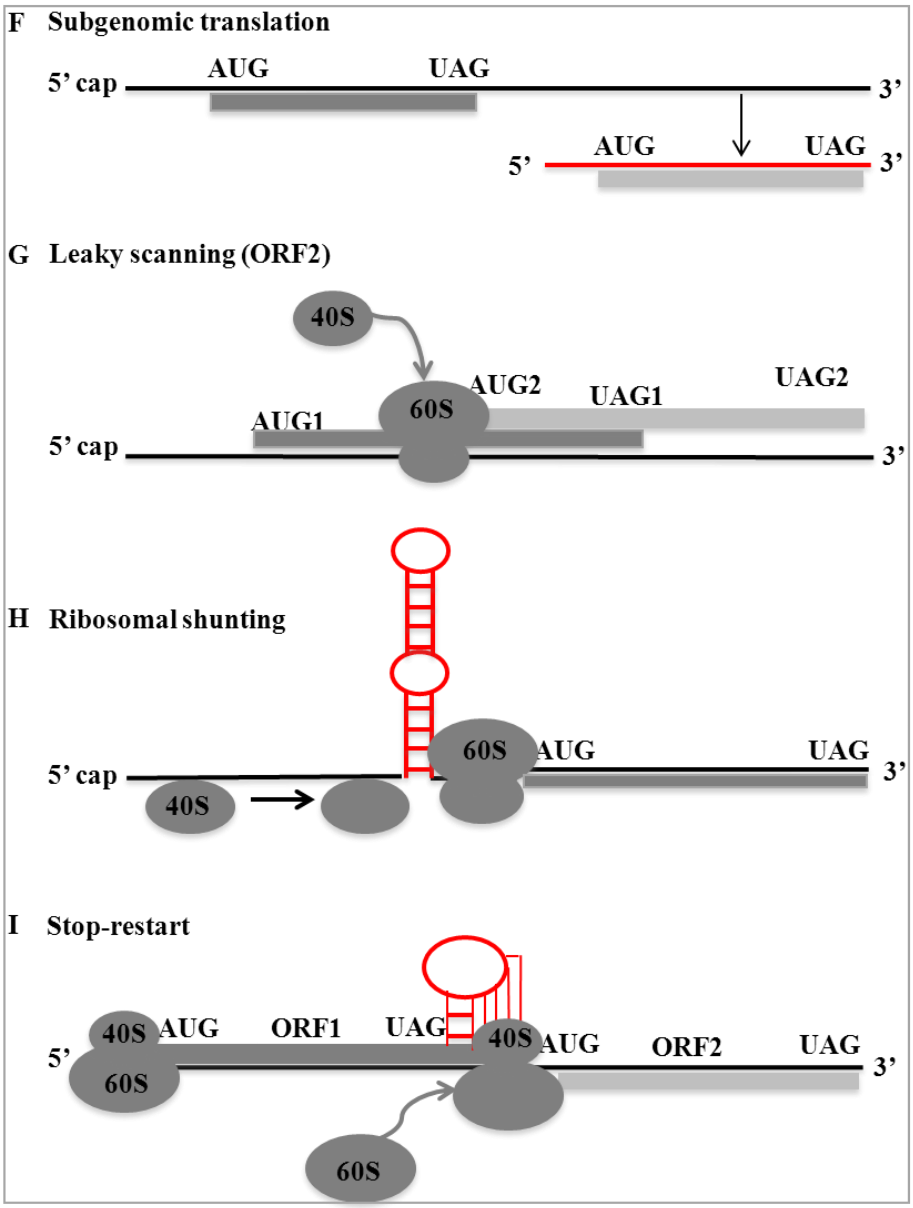


Figure 1.3 Translational mechanisms for RNA viruses.

(A) Kozak's ribosomal scanning translation for the majority of mRNAs in eukaryotic cells. 40S subunit (indicated as small oval) recognizes the 5' cap structure and migrates linearly along 5'UTR of a standard mRNA (in black line), then 40S recognizes and pauses at first AUG followed with 60S large subunit (indicated as large oval) joining of 40S subunit and finally the translation of the first ORF (in gray rectangle) is initiated. The proper translation initiation depends on two features, one is the 7-methylguanylate cap linked to the 5' end of an mRNA and one is stringent context around first AUG. Unless specifically mentioned, representative drawings for mRNA, 40S and 60S ribosomal

subunits, first ORF in dark gray rectangle are the same for all figures. (B) Internal ribosomal entry site (IRES)-mediated translation initiation. Such mechanism in RNA viruses involves a long and highly structured IRES element (in red lines) that is located at the 5' UTR of viral sequences. IRES binds ribosomal initiation complex and catalyzes the formation of a functional 40S ribosomal subunits and then directs them to the proper initiation codon followed 60S subunits joining and initiation of translation. (C) A large number of viruses utilize polyprotein strategy to translate their genomes. This polyprotein strategy will synthesize a single large polyprotein from the entire genomic RNA followed by proteolytic processing at different sites (indicated with symbol "X") that cleaves the large polyperotein into several small but functional proteins. (D) In some viral RNAs, a suppressible termination codon is embedded at the 3' end of upstream ORF. In conjunction with the downstream recoding signal CARYYA (R=purine; Y=pyrimidine), a small proportion of ribosomes skip termination codon, permitting the synthesis of an elongated product (in light gray rectangle) of the upstream protein. (E) Frameshifting occurs in ±1, ±2 frame of the downstream ORF relative to the upstream ORF and typically in -1 frame. It is promoted by a pair of recoding signals: a slippery heptanucleotide (XXX YYY Z, where X is any nucleotide; Y is A or U; Z is A, U or C) and a pseudoknot (in red lines) that is thought to induce ribosome pausing. The encoded products include a product from the entire upstream ORF and an elongated product from a switched downstream frame. (F) Subgenomic translation involves the presence of a subgenome RNA which is transcribed from the genomic RNA. Thus, in this strategy the subgenomic RNA will be translated independently from the genomic RNA into a separate protein. (G) In many viral RNAs, more than one initiation site can be used. This can occur when the 5' proximal initiation codon is weakly recognized by ribosomes in a weak context or at a non-AUG codon. The translational products may be two proteins of totally unrelated sequences or one protein with its N-terminal sequence identical to that of the upstream one. (H) Ribosomal shunting involves a highly structured sequence (giant stem-loops in red lines) located upstream of the proper initiation site. The base parts of the stem-loop are essential for scanning translation, which consist of a donor sites and an acceptor site. Ribosomes pause at the donor site when encountering the stem-loop and bypass the middle part of it and land at the downstream acceptor site and translate the nearby AUG. (I) In stop-restart, ribosomes translate the upstream ORF of an mRNA, then following termination, a proportion of the 40S subunits remain bound to the mRNA and go on to reinitiate at the start codon of the downstream ORF. In

caliciviruses such as RHDV, a binding between 18S rRNA in 40S subunits and viral sequences is thought to promote the maintaining of 40S subunits onto mRNA for the continuous translation of the downstream ORF.

2 CHAPTER TWO: MATERIALS AND METHODS

2.1 Fungal and bacterial strains

Helminthosporium victoriae strain B-2ss (ATCC 42020) is a virus-free fungal isolate and used for the preparation of protoplasts and fungal transformation experiments with recombinant plasmids. Another strain A-9 (ATCC 42018), a diseased *H. victoriae* isolate that contains *victorivirus* HvV190S and *chrysovirus* HvV145S, was used as a source of HvV190S virions and HvV190S proteins CP and RdRp. All fungal cultures were grown on potato dextrose agar supplemented with 0.5% (wt/vol) yeast extract (PDAY) at 20°C, and stored at 4°C on the same medium. *E. coli* strain DH5α was used to propagate all the recombinant plasmids constructed in this study.

2.2 Construction of transformation vector and recombinant plasmids used in experiments with virus-like particles

A transformation/expression vector for *H. victoriae* was previously constructed by the Ghabrial lab (9, 64). The vector, designated p190S, comprises a full-length cDNA clone of the HvV190S genome inserted under the control of a *gpd1* promoter from *Cochlioblus heterostrophus* (65) and a *trp*C terminator from *Aspergillus nidulans* (66). The vector also contains a hygromycin (*HygB*) selectable marker gene (67).

To introduce site-directed, insertion, or deletion mutations into the region flanking the AUGA stop–restart motif in the HvV190S cDNA, an adaptation of the overlap-extension protocol using PCR was used (68, 69). Two unique restriction-enzyme sites, for cleavage by *Fse*I and *Stu*I at nt positions 2171 (434 nt upstream of the AUGA) and 2990 (385 nt downstream of the AUGA) in the HvV190S cDNA, respectively, were selected to border the region for mutagenesis (Fig. 2.1). A three-step PCR amplification procedure was used to generate mutated fragments (AB, CD, and AD) with two common primers (primers a and d; *Fse*I-a-F and *Stu*I-d-R, respectively; Table 2.1) and a series of mutagenic primers (primer types b and c; Table 2.1). As these mutagenic primers have terminal complementarity, two overlapping fragments can be fused together in a subsequent extension reaction that includes the outside primers a and d (see illustration

in Fig. 2.1). Finally, PCR product AD was digested with *Fse*I and *Stu*I and inserted into similarly digested HvV190S in vector p190S. Nucleotide sequences related to PCR amplification were verified with the Big Dye terminator cycle sequencing kit and an ABI DNA sequencer for constructs mentioned in this chapter and elsewhere in the dissertation.

2.3 Construction of transformation vector and recombinant plasmids for dual-reporter assay

Fungal *trpC* terminator was from cosmid vector pMLF2 provided by Dr. Mark L. Farman. Plasmid pBG containing fungal promoter *gpd1* is provided by Dr. Olen C. Yoder. Plasmid pRL-CMV containing Renilla luciferase (Rluc) was purchased from Invitrogen, and pFR-LUC containing Firefly luciferase (Fluc) (Acc.No. AF058756) was donated by Dr. Ajay Singh. eGFP and RFP genes were from vectors pSAN6-eGFP/RFP-N1 that were gifts from Dr. Michael Goodin in the department. pGEM-T easy vector was purchased from Invitrogen. Vector pET-21a(+) was purchased from Invitrogen. Vector p190S described in previous section was used as the start vector to generate constructs for dual-reporter assay. Primers used for the construction of all plasmids used in dual-reporter assay were listed in Table 2.2.

The *gpd1* promoter in p190S (nt 1 to nt 662) was originally from plasmid pBG, which contains *gpd1* sequence nt 1 to nt 677. Compared to the wild type g*pd1* promoter from *C. heterostrophus* (Acc.No. X63516), thirty more nucleotides (*ACTCTCTCAAAGCATCACTCTCAGTTCAAC*) were added to the *gpd1* sequence of pGB during the construction of plasmid Rluc-his. For this purpose, four primers (a, b, c and d) and a PCR-based overlap-extension protocol as described in the previous section was used. Forward primer, *Nsi*I-a-F, and reverse primer, *Xho*I-d-R, indicated by a and d, were used as the outside primers. Reverse primer Rluc-b-R and forward primer Rluc-c-F, indicated by b and c, were used as the inside primers for PCR amplification. *Nsi*I-a-F contains the unique *Nsi*I enzyme locating at nt position 198 on the *gpd1* promoter sequence, and *Xho*I-d-R contains partial 3' sequence of Rluc followed by a 6X His tag sequence immediately upstream the stop codon TAA and the unique *Xho*I enzyme. Rluc-c-F contains an inserted unique enzyme site *Fse*I for the convenience of subsequent cloning of other reporter gene constructs. Plasmid pBG and pRL-CMV were

used as templates to achieve PCR fragment AB and fragment CD, respectively. Finally, fragment AD was digested with unique enzymes *Nsi*I and *Xho*I and cloned into vector p190S. The resultant construct, Rluc-His, contains the wild type *gpd1* promoter sequence, indicated by *gpd1',* followed by the open reading frame (ORF) of Rluc gene. The construction of three other reporter genes into vector Rluc-His involved a PCR amplification of specific ORF (Fluc, *e*GFP*,* RFP) plus a 6XHis sequence with specific primers and two unique enzymes *Fse*I and *Xho*I were introduced to ends of each fragment (see Table. 2.2). Finally, PCR fragments containing specific reporter gene were digested with *Fse*I/ *Xho*I and cloned separately into similarly digested Rluc-his vector, and the resultant constructs were named Fluc-His*, e*GFP-His and RFP-His respectively. Then eGFP was fused into HvV190S coat protein open reading frame (CP ORF: nt 290-nt 2608 in the full-length cDNA of HvV190S) sequence, and RFP was fused to RNA-dependent-RNA polymerase ORF (RdRp ORF: nt 2605-nt 5112). For this purpose, two constructs were made, GFP-CP and RdRp-RFP-His. The primer pair *Fse*I-GFP-FP and *Xho*I-GFP-RP2 was used to amplify GFP fragment without its stop codon (with unique enzymes *Fse*I and *Xho*I at ends), and primer pair *Xho*ICP*Apa*IFP and *Xho*ICP*Apa*IRP was used to amplify CP fragment (with unique enzymes *Xho*I and *Apa*I at ends). Finally *Fse*I/*Xho*I digested GFP and *Xho*I/*Apa*I digested CP were ligated into *Fse*I/*Apa*I digested plasmid eGFP-His*. Apa*I-digested T*trpC* terminator was firstly digested off eGFP-His and then ligated back into GFP-CP. Primer pair *Fse*IRdRp*Cla*IFP and *Fse*IRdRp*Cla*IRP was used to amplify RdRp fragment removing its stop codon (unique enzymes *Fse*I and *Cla*I at ends). Primer pair *Cla*IRFP*Xho*IFP and *Xho*I-RFP-RP was used to amplify fragment RFP (unique enzymes *Cla*I and *Xho*I at ends). Finally *Fse*I/*Cla*I digested RdRp fragment and *Cla*I/*Xho*I digested RFP was cloned into *Fse*I/*Xho*I-digested RFP-His, and the new construct was designated as RdRp-RFP-His*.* These two constructs were used to create the dual-reporter plasmid GFP-CP:RFP and the dual-reporter-CP plasmid GFP-CPt-RFP. Primer pair *Fse*I-GFP-FP and *Xho*ICP2608*Cla*IRP was used to amplify the full-length fused ORF (GFP-CP) with *Fse*I and *Cla*I at ends. Primers *Fse*I-GFP-FP and *Fse*IGFPCPtr1711*Cla*IRP amplified the truncated fused GFP-CPt ORF, in which CP ORF was truncated from nt position 1711 to nt 2608 (including the stop codon). The two large fragments were digested with *Fse*I/*Cla*I and ligated separately into similarly digested RdRp-RFP-His.

Dual-reporter plasmid GFP-CP:RFP was used as the new cloning vector for constructs related to dual-fluorescence assay, which include CP5' deletion constructs (except construct CP5'_dall), CP3' deletion constructs, PK7-PK12, PK_Loop1, PK_Loop2, PK_Flank1 and PK_Flank2. PCR amplification was performed to obtain a specific fragment from nt 290 to nt 2608 of CP sequence with designed base mutations or deletions for each construct. Different combination of primers that contain specific mutations and one of the pair of forward primer, *Xho*ICP*Cla*IFP and reverse primer, *Xho*ICP*Cla*IRP were used to amplify the specific fragment (see primer list in Table. 2.2). Part of CP3' deletion constructs (CP3'_d3-d6) used the primer pair, *Xho*ICP*Cla*IFP and *Xho*ICP*Cla*IRP for amplification. Restriction sites of two unique enzymes *Xho*I and *Cla*I were added at ends of PCR products during amplification. Each particular fragment was digested with *Xho*I/*Cla*I, and cloned into *Fse*I/*Cla*I-cut GFP-CP:RFP, together with previously amplified *Fse*I/*Xho*I-digested-GFP fragment. The construction of CP5'_dall was different from above because the majority of CP sequence was deleted. The primer pair *Fse*I-GFP-FP and CP5'_d6-RP was used to amplify a fragment that contains GFP sequence and the two remaining nucleotides of CP after deletion. This fragment was digested with *Fse*I/*Cla*I and cloned into *Fse*I/*Cla*I –cut GFP-CP:RFP.

For the construction of plasmids PK_BfTV1, PK_GaRV-L1, PK_SsRV1, PK_MoV1, PK_TcV1, PK_CeRV1, mutations with a relatively long exogenous oligonucleotide were introduced into the terminal region of CP ORF, thus a PCR-based overlap-extension protocol, as described previously, was used. Two primers: *Xho*ICP*Cla*IFP (previously mentioned) and *Stu*I-d-R, were designated as a and d, were the common primers. These two outside primers together with each inside primer pairs (designated as b and c, see Table 2.2) amplified each new fragment spanning nt 290-nt 2990 on the HvV190S genome with *Xho*I at the 5' end. With these fragments as templates, primer pair *Xho*ICP*Cla*IFP and the c primer for each construct amplified a fragment from nt 290-nt 2608 of CP ORF with swapped pseudoknot sequences. Another unique enzyme *Cla*I was introduced into each fragment by c primer. Subsequent cloning of the PCR products into vector GFP-CP:RFP was as described in the last paragraph.

2.4 Construction of plasmids for *in vitro* translation and bacterial expression

For the construction of plasmids used *in vitro* transcription and translation experiments, the pGEM-T easy vector containing T7 promoter was modified. Six oligonucleotides GCCACC were introduced after *Apa*I restriction site for high efficient recognition of first AUG. Plasmid T7-CP5'_d5 was the first to be made. Primer pairs *Apa*IGFP*Xho*IFP and *Cla*IRFP*Xba*IRP (Table. 2.2) were used to amplify the whole large fragment including GFP, truncated CP ORF and RFP using plasmid CP5'_d5 as amplification template. This large fragment was digested with *Apa*I/*Xba*I and cloned into similarly digested pGEM-T vector. All other plasmids including T7-eGFP-CP:RFP, T7-eGFP-CPt-RFP, T7-CP5'_d1-CP5'_d4 and T7-CP no-stop were generated by digesting corresponding plasmids used in dual- fluorescence assay with *Xho*I/*Cla*I and ligating the resultant fragments into *Xho*I/*Cla*I-digested T7-CP5'_d5.

For the construction of plasmids expressed in bacterial strain BL21 DE3 for Rluc and Fluc, pET-21a(+) was used as the cloning vector. Primer pair *BamHI*RlucFP and *XhoI*RlucRP amplified a fragment including the entire ORF of Rluc from plasmid pRL-CMV whereas primer pair *BamHI*FlucFP and *XhoI*FlucRP amplified the entire ORF of Fluc from plasmid pFR-luc. The stop codon for each ORF was removed during amplification. Restriction sites for two unique enzymes, *BamHI* and *XhoI*, were introduced into amplified fragments, which were subjected to digestion with these two enzymes and cloned into vector pET-21a(+). Resultant constructs were designated as pET21a-Rluc and pET21a-Fluc respectively. The 6XHis tag was fused to the 3' end of these two luciferase genes.

2.5 Bacterial transformation and plasmid extraction

The heat shock method was used for *E. coli* transformation (Molecular cloning 2nd, Cold Spring Harbor Laboratory Press). *E. coli* DH5α competent cells were kindly provided by Wendy M. Havens, lab manager of the Ghabrial lab. Approximately 20 μl of cells were mixed with 5~ 10 μl (50~100 ng) of plasmid DNA and after 30 min incubation on ice, the cells were subjected to heat shock at 42°C for 40 sec followed by a 2 min incubation on ice. Approximately 400 μl of LB broth was mixed with the transformed cells followed by 1

hr incubation at 220 rpm at 37°C. One hundred to two hundred microliter aliquots were plated onto LB agar medium containing appropriate antibiotic.

2.6 Bacterial expression of Fluc and Rluc

Plasmids were introduced into bacterial strain BL21 DE3 as described under bacterial transformation. Thirty μl transformed bacterial cells with desired constructs were added to 3 ml LB broth medium followed with shaking at 220 rpm overnight at 37°C. Thirty microliter of overnight culture was transferred again to 3 ml LB broth and incubated for 2.5-3 hr at same conditions followed by determination of OD600. After the optimal concentration (0.5-0.8) was reached, Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added into culture to a final concentration of 1 mM and culture was further incubated for another 3-4 hr. The extraction of both soluble and insoluble proteins from BL21 DE3 cells was according to the manual protocol of Bugbuster® Extract Reagent (Novagen). Finally, 100 μl and 50 μl supernatant were achieved from soluble and insoluble extraction respectively and saved for further Western blotting analysis.

2.7 Preparation of fungal protoplasts and fungal transformation

Protoplasts were prepared from mycelia of *H. victoriae* virus-free strain B-2ss as described previously (70). Briefly, Mycelia were cultured on PDAY for several days followed with continuous culturing in potato dextrose broth supplemented with yeast extract (PDBY) at 20°C with a shaking at 125 rpm. Following the homogenization of mycelia after 72 hr growth, two milliliters mycelia were transferred to new PDBY media. After 24 hr growth, mycelia were centrifuged at 3000 rpm and incubated in filter-sterilized enzyme osmoticum (0.7 M NaCl, 100 μg/ml chitinase, 10 mg/ml driselase, 10 mg/ml lysing enzyme). After 3 hr incubation at 30°C with a shaking at 75 rpm, digested mycelia-enzyme mixture passed through and washed with 0.7 M NaCl once and STC buffer (1.2 M sorbitol, 10 mM Tris pH 7.5, 10 mM $CaCl_2$) three times and finally dissolved in STC. The concentration of protoplasts were determined with hemocytometer and adjusted to 10^{-8} with STC buffer. Protoplasts were distributed into small aliquots and stored at -80°C for later transformation.

Polyethylene glycol (PEG)-mediated DNA uptake method (71) was used to transform fungal protoplasts with different constructs. Briefly, 150 μl fungal protoplasts were mixed with 10-20 μl plasmid DNA and incubated for 30 min at room temperature followed with a gradually addition of 200 μl, 200 μl and 800 μl 60% PEG 4000 solution containing 50 mM $CaCl_2$ and 10 mM Tris-HCl pH7.5. After incubation for another 25 min at room temperature, protoplast-DNA mixture were washed with STC buffer once and finally dissolved in 500 μl STC. These protoplast-DNA mixture in STC were continuously mixed with 15 ml regeneration media (RM) (0.1% casein hydrolysate, 0.1% yeast extract, 1 M sucrose, 2.5% agar) and poured onto sterile plate followed with overnight growth at 30°C. All the plasmid constructs contained a hygromycin B phosphotransferase gene (*hyg*) as a selection marker. Finally, from overlayed RM plate with 1% agar-H_2O media containing hygromycin at 50 μg/ml, transformants were selected for further studies as representatives of each construct.

2.8 Extraction of fungal total protein

For extracting total protein, mycelia of *H. victoriae* transformants containing different construct were cultured on PDAY media covered with a layer of sterile cellophane. Mycelia from 10-day old cultures were collected and ground in phosphate buffered saline (PBS) buffer, PH 7.4, supplemented with 1 mM protease inhibitor phenylmethylsulfonyl fluoride (PMSF). The homogenate was centrifuged for 10 min at 10,000 rpm at room temperature. The supernatant was collected for further testing with the fluorescence assay or Western blotting.

2.9 Preparation of viral particles

For the preparation of virus-like particles from *H. victoriae* virus-free isolate that was transformed with p190S or p190S-derived constructs made in this study, mycelia from those cultures were grown on PDAY media for 14 days in a stationary condition. Purification of VLPs was performed as described previously (72). Briefly, mycelia from 14 day-growth were collected and homogenized 1 min for three times in a Waring blender in 0.1 M sodium phosphate buffer (pH7.4) containing 0.2 M KCl and 0.5% β-mercaptomethanol at a rate of 3 ml/g of wet mycelium followed with twice 1 min

homogenization in an equal volume of chloroform. The homogenate was centrifuged for 20 min at 12,000 rpm at 4°C and supernatant were carefully collected and followed with first high speed centrifugation for 2.5 hr at 27,000 rpm at 4°C. The pellets were suspended in 0.1 M sodium phosphate buffer (pH 7.0) with shaking overnight at 4°C and followed with second high speed centrifugation for 1 hr at 40,000 rpm. The resultant pellets were suspended in phosphate buffer and layered onto 10-40% gradient sucrose and centrifuged 2.5 hr at 24,000 rpm. Gradient fractions containing VLPs were collected from the sucrose density gradients and pelleted by overnight centrifugation at 40,000 rpm. The resultant pellets were suspended in 50 mM Tris-HCl buffer (pH 8.0) and stored at -20°C for Western blotting.

2.10 Western blotting

Western blotting analysis was performed as described previously (73). Briefly, samples of fungal total protein extracts were concentrated 10 times in a Savant Speed Vac Concentrator. These samples together with VLP preparations were loaded onto 10 % polyacrylamide gels (10 µl loading for detecting RFP or RdRp, 1 µl loading for detection of GFP or CP) and fractionated by SDS-PAGE. After blotting to a polyvinylidene fluoride membrane, the blots were probed overnight with antisera to CP or RdRp (analysis of VLPs), or with antisera to GFP, RFP or His-tag (analysis of total proteins). Subsequent to washing, membranes were incubated with appropriate second antiserum conjugated with alkaline phosphatase. Then following several washes, membranes were subjected to color development using BCIP/NBT reagents (5-Bromo-4-chloro-3-indolyl phosphate/Nitro blue tetrazolium).

2.11 RNA isolation and Northern blotting

Total RNA was extracted from virus-infected or vector-transformed fungal mycelia using the TRIzol reagent (Invitrogen) according to the manufacturer's protocol. Full-length plus-stranded RNA transcript of HvV190S (~ 5.2 kb) was synthesized in vitro using T7 RNA polymerase and plasmid pT7-HvV190S containing full-length HvV190S cDNA as template (74). RNA samples (15 µg) were then fractionated in a denaturing agarose gel. Blotting to Hybond-N+ membranes and hybridization were carried out as described previously (70).

PCR fragments corresponding to the CP ORF or the RdRp ORF of HvV190S cDNA were amplified and randomly radiolabeled using [α-^{32}P] dCTP for probing the blots.

2.12 RNA structure predictions

Prediction of H-type RNA pseudoknots was performed using the program HPKNOTTER (75) and Dotknot (Version 1.2) (76). These two programs are implemented at http://genome.cs.nthu.edu.tw/HPKNOTTER/ and http://dotknot.csse.uwa.edu.au/, respectively. The program Hpknotter was run using the website's default settings for "Class: General" and using pknotsRG (version 1.2) (77) as the prediction kernel. The program Dotknot was run at the default settings on the web site. The long stem–loop structure in SsRV2 was predicted using the program Mfold with the default settings (78) as implemented at http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form. Mfold was also used for the prediction of secondary structure for immediately downstream sequence of AUGA region of HvV190S. The free energy values for each structure were calculated by the respective program.

2.13 *In vitro* transcription and translation

Unless otherwise stated, the *in vitro* coupled transcription/translation was carried out in TnT® T7 Quick Coupled Transcription/Translation Systems (Promega) according to manual instructions. Typical reactions were 25 µl and composed of 50 % wheat germ extracts, 1 µl reaction buffer, 0.5 µl T7 RNA polymerase, 0.5 µl amino acid mixture minus methionine, 0.5 ul RNasin ribonuclease inhibitor (40 U/µl) and 1-2 µl [^{35}S]-methionine (10 mCi/ml) plus 1 µg of plasmid DNAs or luciferase control DNA. Reactions were incubated for 2 h at 30 °C. Samples were prepared for gel electrophoresis by adding 4 µl 4 × sample buffer into 10 µl of reaction products, boiled for 3 minutes and resolved by SDS-PAGE on a 10% polyacryamide gel. The detection of products on the gels was determined by direct measurement of [^{35}S]-methionine incorporation using a typhoon imager. All *in vitro* reactions were repeated twice independently.

2.14 Dual fluorescence assay

One hundred and fifty microliter samples of fungal total protein extracts were loaded into 96-well opaque plate (greiner bio-one) and fluorescence was measured using a SpectraMax M2 microplate reader according to manufacturer's manual (Molecular devices). EGFP fluorescence was detected using 485-nm excitation/ 520-nm emission spectra, while RFP (DsRed) was detected using 543-nm excitation/ 590-nm emission spectra. Total protein from each transformant was independently extracted. Fluorescence measurement for each extract was repeated at least three times using representative samples to obtain a mean value.

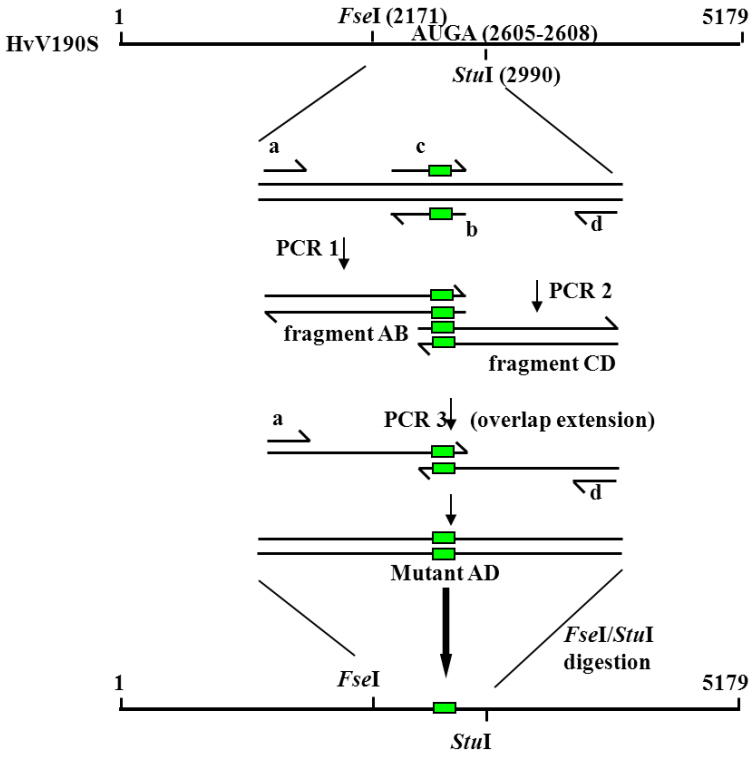


Figure 2.1 Schematic representation of the overlap-extension protocol used for mutagenesis.

To introduce site-directed, insertion, or deletion mutations into the region flanking the AUGA stop–restart motif in the HvV190S cDNA, two unique restriction enzymes sites, *Fse*I and *Stu*I at nt positions 2171 and 2990, respectively, were selected to border the region for site-directed mutagenesis. A three-step PCR amplification procedure was used to generate mutated fragments AB, CD, and AD. Two common primers (primers a and d; *Fse*I-a-F and *Stu*I-d-R, respectively; Table 2.1) and a series of mutagenic primers containing nucleotide mismatches represented by green rectangles (primer types b and c; Table 2.1) were designed to introduce a variety of mutations into the region flanking the AUGA overlap. For each mutant construct, fragment AB was synthesized using the common forward primer *Fse*I-a-F and the pertinent mutagenic type-b primer for reverse priming; fragment CD was synthesized using the pertinent mutagenic type-c primer for forward priming and the common reverse primer *Stu*I-d-R; and finally, fragment AD was synthesized using the amplified fragments AB and CD as templates and the common primer pair *Fse*I-a-F and *Stu*I-d-R. All AD fragments were doubly digested with *Fse*I and *Stu*I and cloned into similarly digested vector p190S.

Table 2.1 List of primers used in experiments involving purified virus-like particles.

Primers	Sequence (5'-3')	Template
*Fse*I-a-F	GTCTTT***GGCCGGCC***AGATGTCGGT (*Fse*I in bold italic)	
*Stu*I-d-R	TTCGCGAA***AGGCCT***TGTCGTTG (*Stu*I in bold italic)	
pStartM1-b-R	TGAGGATCACTCA**C**TGTCCC	p190S
pStartM1-c-F	GACA**G**TGAGTGATCCTCAGG	p190S
pStartM2-b-R	CGTAACCAGCAAGCA**C**ATGG	p190S
pStartM2-c-F	CCAT**G**TGCTTGCTGGTTACG	p190S
pStartM3-b-R	TGAGGATCACTCA**C**TGTCCC	pStartM2
pStartM3-c-F	GACA**G**TGAGTGATCCTCAGG	pStartM2
pStartM4-b-R	CTGAGGATCAC**CT**ATTGT	p190S
pStartM4-c-F	ACAAT**AG**GTGATCCTCAG	p190S
pStartM5-b-R	Same as pStartM2-b-R	pStartM4
pStartM5-c-F	Same as pStartM2-c-F	pStartM4
pStartM6-b-R	AAG**T**A**T**ATGGGAATTGGC	pStartM1
pStartM6-c-F	AT**A**T**A**CTTGCTGGTTAC	pStartM1
pStop+20-b-R	GATCAC**C**CATTGTCCCTCGGC	p190S
pStop+20-c-F	CGGGCCGAGGGACAATG**G**GTGATC	p190S
pStop+8-b-R	CCGTAGGCT**A**TGGAACGT	pStop+20
pStop+8-c-F	ACGTTCCA**T**AGCCTACGG	pStop+20
pStop+2-b-R	GTTCCTGAG**TCA**CACCCATTGTCC	pStop+20
pStop+2-c-F	CGAGGGACAATGGGTG**TGA**CTC	pStop+20
pStop-2-b-R	CACCCATTGTC**A**CTCGGC	pStop+20
pStop-2-c-F	GCCGAG**T**GACAATGGGTG	pStop+20
pStop-8-b-R	ATCAGCCT**A**GGCAGCGG	pStop+20
pStop-8-c-F	TGCC**T**AGGCTGATCGG	pStop+20
pStop-12-b-R	GCAGCGGC**TCA**GGGTGCGTG	pStop+20
pStop-12-c-F	ACGCACCC**TGA**GCCGCTGCCC	pStop+20
PK0-b-R	CCTCGG**T**C**T**GATC**T**G**T**CTG**T**G**TT**GC**TTT**GGGGGGTGCGTGGATG GCAGTG	p190S
PK0-c-F	ACGCACCCCCC**AAA**GC**AA**C**A**CAG**A**C**A**GATC**A**G**A**CCGAGGGAC	p190S
Restore1-b-R	**GCCTGGGCAGCGGCGGGCT**ATTGTCCCTCGGTCTGATC	PK0
Restore1-c-F	**GCTGCCCAGGCTGATCGGGCCGAGGGACAATGA**GTG	PK0
Restore2-b-R	CGGG**CC**A**C**TGTCCCTCGGTCTGATC	Restore1

Table 2.1 Continued

Primers	Sequence (5'-3')	Template
Restore2-c-F	GGACA**G**T**GG**CCCGCCGCTGCCC	Restore1
Spacer6-1-b-R	GCCCGATCAGCCTGGGCAGCG	p190S
Spacer6-1-c-F	ATCGGGCCGGACAATGAGTGATCCTC	p190S
Spacer6-2-b-R	TCATTGCTCGGCCCGATCAG	p190S
Spacer6-2-c-F	CCGAGCAATGAGTGATCCTC	p190S
Spacer3-b-R	ACTCATTGGGCCCGATCAGCCTGG	p190S
Spacer3-c-F	CAATGAGTGATCCTCAGGAACG	p190S
Spacer18-b-R	TTGTCCCTCGGCCCGATCAGCCTGG	p190S
Spacer18-c-F	GAGGGACAAGAGGGACAATGAGTGATC	p190S
Spacer27-b-R	TCCCTCGGCCCGATCAGCCTGG	Spacer18
Spacer27-c-F	GGGCCGAGGGACAAGAGGGACAAGAGGGACAATGA	Spacer18
RdRp5'_d1-b-R	TTTGGAATCACTCATTGTCCCTCGGCCCG	p190S
RdRp5'_d1-c-F	ATGAGTGAT TCCAAAGCCTACGGA	p190S
RdRp5'_d2-b-R	ACGCTACTTCCCCGAGAAGT CCGTAGGC	p190S
RdRp5'_d2-c-F	GGGGAAGTAGCGTCAGCCAATTCCC	p190S
p190S-CP3'd1-b-R	TCATTGATCAGCCTGGGCAGC	p190S
p190S-CP3'd1-c-F	TGAT CAATGAGTGATCCTCAGG	p190S
p190S-CP3'd2-c-F	CGCT CGGGCCGAGGGACAATGAG	p190S
p190S-CP3'd2-c-F	CGCT CGGGCCGAGGGACAATGAG	p190S
p190S-CP3'd3-b-R	TGGGCTGCGTGGATGGCAGTGG	p190S
p190S-CP3'd3-c-F	ACGCAGCCCAGGCTGATCGG	p190S
p190S-CP3'd4-b-R	GGGGAGTGGGGACGGGGGCAGC	p190S
p190S-CP3'd4-c -F	CCCACTCCCCCCGCCGCTGCCC	p190S
p190S-CP3'd5-b-R	GATGGCGGCAGCAGGGTTGG	p190S
p190S-CP3'd5-c-F	TGCCGCCATCCACGCACC	p190S
p190S-CP3'd6-b-R	ACGGGCACGTTGCCGCCGCCACCG	p190S
p190S-CP3'd6-c-F	AACGTGCCCGTCCCCACTGCCATCC	p190S
PK1-b-R	AGC**CCG**GGGGGGTGCGTGGATGGC	p190S
PK1-c-F	CCC**CGG**GCTGCCCAGGCTGATC	p190S
PK2-b-R	ATCA**CGG**TGGGCAGCGGCG	p190S
PK2-c-F	CCA**CCG**TGATCGGGCCGAGG	p190S

Table 2.1 Continued

Primers	Sequence (5'-3')	Template
PK3-b-R	A**CGG**TGGGCAGC**CCG**GGGGGGTGCGTG	p190S
PK3-c-F	**CGG**GCTGCCCA**CCG**TGATCGGGCCG	p190S
PK4-b-R	TCAGCCT**CCCG**AGCGGCGG	p190S
PK4-c-F	CCGCT**CGGG**AGGCTGATCG	p190S
PK5-b-R	TCCCTCG**CGGG**GATCAGCC	p190S
PK5-c-F	TGATC**CCCG**CGAGGGACAATG	p190S
PK6-b-R	same as PK5-b-R	PK4
PK6-c-F	same as PK5-c-F	PK4
PK2-1-b-R	ATC**T**CGGTGGGCAGCGGCGG	p190S
PK2-1-c-F	CCACCG**A**GATCGGGCCGAGG	p190S
PK3-1-b-R	GGGTGCGTGGTAGGCAGTG	p190S
PK3-1-c-F	ATCCACGCACCCCCC**TAA**GCTGCCCA**TTA**TGATCGGGCCGAGGGAC	p190S
PK3-2-b-R	AT**CATAAT**GGGCAG**ATTATG**GGGGTGCGTGGATGGC	p190S
PK3-2-c-F	CC**ATTATG**ATCGGGCCGAGGGACAATG	p190S
PK6-1-b-R	ATCAGCCT**TTTA**AGCGGCGG	p190S
PK6-1-c-F	**AA**AGGCTGATC**TTTA**CGAGGGACAATG	p190S
PK6-2-b-R	ATCAGCCT**TTTAAG**CGGCGGGG	p190S
PK6-2-c-F	**AA**AGGCTGATC**TTTAAG**AGGGACAATGAGTG	p190S
18S-20M-b-R	CC**CTAGT**GCCTGGGCAGCGG	p190S
18S-20M-c-F	AGGC**ACTAG**GGGCCGAGGGACAATG	p190S
18S-40M-b-R	GGT**CGCACCT**TGGCAGTGGGGAC	p190S
18S-40M-c-F	A**AGGTGCG**ACCCCCCGCCGCTGC	p190S
18S-65M-b-R	A**GCCCCCGTCGTCC**GTTGGAAGATCCATCGTCG	p190S
18S-65M-c-F	**ACGACGGGGGC**TCCCCACTGCCATCCAC	p190S
18S-90M-b-R	TCG**AGCCTTCC**AGGTGGGCCGGCAGG	p190S
18S-90M-c-F	T**GGAAGGCT**CGATGGATCTTCCAAC	p190S

I. Primers designated “a” or “d” are common outside primers; Primers designated “b” or “c” are mutagenizing primers that were designed to introduce mutations.
II. Mutations introduced are bolded; Unique enzymes inserted are in italic.

Table 2.2 List of primers used in experiments for dual-fluorescence assay.

Primers	Sequence (5'-3')	Template
Rluc-His		
NsiI-a-F	CATATGCATGAGGCTGTAAG	
XhoI-d-R	***CTCGAGTTAGTGGTGGTGGTGGTGTTGTTCATTTTTGAGAACTCGCTCAACG***	
Rluc-b-R	***CCAGTGATGCTTTGAGAGAGTTAATAGACTTGAATTGAGGAATGG***	pGB
Rluc-c-F	**AAGCATCACTGGCCGGCCAACATGACTTCGAAAGTTTATGATCC**	pRL-CMV
Fluc-His		
FseI-Fluc-FP	***GGCCGGCCAACATGGAAGACGCCAAAAAC***	pFR-Luc
XhoI-Fluc-RP	***CTCGAGTTAGTGGTGGTGGTGGTGGTGCAATTTGGACTTTCCGCCCTTCTTGGCC***	pFR-Luc
eGFP-His		
FseI-GFP-FP	***GGCCGGCCAACATGGTGAGCAAGGGCG***	pSAN6-eGFP-N1
XhoI-GFP-RP	***CTCGAGTTAGTGGTGGTGGTGGTGGTGCTTGTACAGCTCGTCCATGCCGAGAGTG***	pSAN6-eGFP-N1
RFP-his		
FseI-RFP-FP	***GGCCGGCCAACATGGCCTCCTCCGAG***	pSAN6-RFP-N1
XhoI-RFP-RP	***CTCGAGTTAGTGGTGGTGGTGGTGGTGGGCGCCGGTGGAGTGGCGGCCCTCG***	pSAN6-RFP-N1
eGFP-CP		
XhoI-GFP-RP2	***CTCGAGCTTGTACAGCTCGTCCATG***	pSAN6-eGFP-N1
XhoICPApaIFP	***CTCGAGATGTCTCACACCACGATC***	p190S
XhoICPApaIRP	***GGGCCCTCAT*** TGTCCCTCG	p190S
RdRp-RFP-His		
FseIRdRpClaIFP	***GGCCGGCCAACATGAGTGATCCTCAG***	p190S
FseIRdRpClaIRP	***ATCGATAGAACGGATATGATATG***	p190S
ClaIRFPXhoIFP	***ATCGATATGGCCTCCT*** CCGAGGAC	pSAN6-RFP-N1
eGFP-CPt-RFP		
FseIeGFPCP1711ClaIRP	***ATCGATCTTCATACGTATGTGATAC***	eGFP-CP
eGFP-CP:RFP		
XhoICP2608ClaIRP	***ATCGATACTCATTGTCCCTCGGCCCGATC***	eGFP-CP
CP 5' and 3' deletions		
XhoICPApaIFP	***CTCGAGATGTCTCACACCACGATC***	
Non-stop-RP	***ATCGATACCCATTGTCCCTCGGCCCGATC***	eGFP-CP:RFP
CP5'_d1-FP	***CTCGAGACTCATGCCGACACTATGC***	eGFP-CP:RFP
CP5'_d2-FP	***CTCGAGAACAGGTCCGCCGCGCCTGG***	eGFP-CP:RFP
CP5'_d3-FP	***CTCGAGAACGGTCCCCCTGCCGGC***	eGFP-CP:RFP

Table 2.2 Continued

Primers	Sequence (5'-3')	Template
CP5'_d4-FP	***CTCGAG***GCTGCCCCCGTCCCCAC	eGFP-CP:RFP
CP5'_d5-FP	***CTCGAG***CCCCCCGCCGCTG	eGFP-CP:RFP
*Xho*ICP2608*Cla*IRP	***ATCGAT***ACTCATTGTCCCTCGGCCCGATC	
*Fse*I-GFP-FP	***GGCCGGCC*AAC**ATGGTGAGCAAGGGCG	pSAN6-eGFP-N1
CP5'd6-RP	***ATCGAT***ACTCATTGCTTGTACAGCTCGTCCATG	pSAN6-eGFP-N1
CP3'd1-RP	***ATCGAT***ACTCATTGATCAGCCTGGGCAGCGG	eGFP-CP:RFP
CP3'd2-RP	***ATCGAT***ACTCATTGTCCCTCGGCC	eGFP-CP:RFP
CP3'd3-RP	*Xho*ICP*Apa*IFP/*Xho*ICP2608*Cla*IRP	p190S-CP3'd3
CP3'd4-RP	*Xho*ICP*Apa*IFP/*Xho*ICP2608*Cla*IRP	p190S-CP3'd4
CP3'd5-RP	*Xho*ICP*Apa*IFP/*Xho*ICP2608*Cla*IRP	p190S-CP3'd5
CP3'd6-RP	*Xho*ICP*Apa*IFP/*Xho*ICP2608*Cla*IRP	p190S-CP3'd6
Pseudoknot swapping		
*Xho*ICP*Apa*IFP	***CTCGAG***ATGTCTCACACCACGATC	
*Stu*I-d-R	TTCGCGAA***AGGCCT***TGTCGTTG (*Stu*I in bold italic)	
PK_BfTV1-b-R	**CGTTACCTGGGGCAGCAGG**GGGGGGTGCGTGGATGGCAGTG	eGFP-CP:RFP
PK_ BfTV1-c-F	**CCCAGGTAACGGGCCGGACACGATATGAGT*ATCGAT***GATCCTCAGGAAC	eGFP-CP:RFP
PK_GaRV-L1-b-R	**GCTTCAGCAGCAGTAGCTTC**GGGGGGTGCGTGGATGGCAGTG	eGFP-CP:RFP
PK_GaRV-L1-c-F	**GCTGCTGAAGCAGTGCCCGCTCAATGAGT*ATCGAT***GATCCTCAGGAACG	eGFP-CP:RFP
PK_SsRV1-b-R	**CCGTCAGGCTGGGCGGGGCC**GGGGGGTGCGTGGATGGCAGTG	eGFP-CP:RFP
PK_SsRV1-c-F	**AGCCTGACGGGCCCGCCAATGAATAAGAGT*ATCGAT***GATCCTCAGGAAC	eGFP-CP:RFP
PK_MoV1-b-R	**CAGGCTAGGCGCCGCGCC**GGGGGGTGCGTGGATGGCAGTG	eGFP-CP:RFP
PK_MoV1-c-F	**CGCCTAGCCTGCACGAATAGATATGAGT*ATCGAT***GATCCTCAGGAACG	eGFP-CP:RFP
PK_TcV1-b-R	**CCACCATAGCCTGGGGGTCGGC**GGGGGGTGCGTGGATGGCAGTG	eGFP-CP:RFP
PK_TcV1-c-F	**GCTATGGTGGGCGACCAGATCTAAATGAGT*ATCGAT***GATCCTCAGGAAC	eGFP-CP:RFP
PK_CeRV1-b-R	**GGCCCTCATTAGTGGGGGCAGCCAC**GGGGGGTGCGTGGATGGCAGTGGG	eGFP-CP:RFP
PK_CeRV1-c-F	**CCACTAATGAGGGCCGAAACGATGTCTAGA*ATCGAT***GATCCTCAGGAAC	eGFP-CP:RFP

Table 2.2 Continued

Primers	Sequence (5'-3')	Template
Mutations in pseudoknot and flanking regions		
PK7-R	***ATCGAT***ACTCATTGTCCCTCGGCCCGATCAGCCTGGGCAGC**GCC**GG	eGFP-CP:RFP
PK8-R	***ATCGAT***ACTCATTGTCCCTCGGCCCGATCA**GGC**TGG	eGFP-CP:RFP
PK9-R	***ATCGAT***ACTCATTGTCCCTCGGCCCGATCA**GGC**TGGGCAGC**GCC**GG	eGFP-CP:RFP
PK10-R	***ATCGAT***ACTCATTGTCCCTCGGCCCGATCAGCCT**GCGC**AGCGG	eGFP-CP:RFP
PK11-R	***ATCGAT***ACTCATTGTCCCTCG**GCGC**GATCAG	eGFP-CP:RFP
PK12-R	***ATCGAT***ACTCATTGTCCCTCG**GCGC**GATCAGCCT**GCGC**AGCGG	eGFP-CP:RFP
PK_Loop1-R	***ATCGAT***ACTCATTGTCCCTCGGCCCGATCAGCC**A**GGGC**TCGGG**CGGG	eGFP-CP:RFP
PK_Loop2-R	***ATCGAT***ACTCATTGTCCCTCGGCCC**CTAGT**GCCTGGGCAGC	eGFP-CP:RFP
PK_Flank1-R	***ATCGAT***ACTCATTGTCCCTCGGCCCGATCAGCCTGGGCAGCGGC**CCCCCC**TGCGTG	eGFP-CP:RFP
PK_Flank2-	***ATCGAT***ACTCAT**ACAGGGAGC**GCCCGATCAGCC	eGFP-CP:RFP
In vitro translation		
*Apa*IGFP*Xho*IFP	***GGGCCC*GCCACC**ATGGTGAGCAAGGGCG	pSAN6-eGFP-N1
*Cla*IRFP*Xba*I	***TCTAGA***TTAGGCGCCGGTGGAGTG	pSAN6-RFP-N1
*Xho*ICP*Cla*Id	***ATCGAT***ACTCATTGATCAGCCTGGGCAGCGG	p190S
*Xho*ICP*Cla*Id2RP	***ATCGAT***ACTCATTGTCCCTCGGCC	p190S-CP3’d2
Bacterial expression of Rluc and Fluc		
*BamHI*RlucFP	***GGATCC***ATGACTTCGAAAGTTTATG	pRL-CMV
*Xho*IRlucRP	***CTCGAG***TTGTTCATTTTTGAGAAC	pRL-CMV
*BamHI*FlucF	***GGATCC***ATGGAAGACGCCAAAAAC	pFR-Luc
*Xho*IFlucRP	***CTCGAG***CAATTTGGACTTTCC	pFR-Luc

(i) Primers designated “a” or “d” are common outside primers; Primers designated “b” or “c” are mutagenizing primers that were designed to introduce mutations.

(ii) Mutations introduced are bolded; Unique enzymes inserted are in italic.

3 CHAPTER THREE: HvV190S utilizes a coupled termination-reinitiation strategy to translate the downstream RNA dependent RNA polymerase (RdRp) gene from its bicistronic genome

Part of this work was published as a regular paper in the Journal of Virology, 2011

3.1 Introduction

Unlike viruses in the genus *Totivirus*, which express their RdRp only as CP-RdRp fusion proteins, HvV190S, the prototype strain of genus *Victorivirus*, express its RdRp as a separate, nonfused protein. Thus ribosomal frameshifting and in-frame read-through of stop codon can be ruled out as the mechanism utilized by HvV190S for expressing its RdRp because both of these strategies generate fusion proteins (14) Then expression of the downstream RdRp ORF from the HvV190S bicistronic genome may occur by one or a combination of the following strategies: leaky scanning, ribosomal shunting, IRES and coupled termination-reinitiation, polyprotein synthesis/processing and subgenomic RNA translation. Leaky scanning seems unlikely because of the very long distance (2316 nt) separating the two start codons and the presence of several AUGs in a more favorable context (AXXAUGG). The longest distance between two start codons from other examples of translations by leaky scanning is usually less than 150 nt (57). It is also unlikely that RdRp ORF is expressed via ribosomal shunting or IRES mediated internal initiation. These two mechanisms involve a highly structured untranslated region upstream the start codon (37, 56). On the contrary, instead of the short inter-region between the two ORFs, the 5' UTR of the upstream CP ORF was found to be highly structured which implies IRES or ribosomal shunting for translation for the upstream ORF (14). In the mechanism of polyprotein translation, initiation codons and termination codons are deficient for internal ORFs. However, independent initiation and termination codons are predicted from HvV190S genome for CP ORF and RdRp ORF repectively although not tested (14, 15). RdRp may use subgenomic strategy to expression a separate protein from a subgenomic RNA although no subgenomic promoter sequence upstream the RdRp ORF has been identified. When RdRp ORF is expressed in *Escherichia coli* from a monocistronic construct employing its predicted initiation codon, the translation product is indistinguisheable in both size and serological reactivity from the HvV190S particle-associated RdRp (14). Since polyprotein processing or translation

via subgenomic mRNA are unlikely, HvV190S most likely utilizes a coupled termination-reinitiation strategy to translate its downstream RdRp ORF based on the fact that the termination codon of the CP ORF overlaps the start codon of the downstream RdRp ORF in a tetranucleotide AUGA, which meets the most common feature for this mechanism.

In this chapter, an *in vivo* transformation/expression system was established in *H. victoriae* to investigate the translation mechanism for the downstream RdRp gene from bicistronic HvV190S genome. Crucial experiments were performed to rule out some translational possibilities and provide evidence that a coupled termination-reinitiation was truly utilized by HvV190S to express its downstream RdRp ORF.

3.2 Results

3.2.1 Efficient expression of HvV190S RdRp from transformation vector in the native host

The RdRp ORF of HvV190S has been previously cloned into monocistronic or bicistronic vectors and transformed into yeast cells *Schizosaccharomyces pombe* and bacterial cells *E. coli* to investigate its expression (25). Although RdRp expression from dicistronic vectors was unsuccessful in bacteria cells, it was expressed at low levels in yeast cells, suggesting that some eukaryotic host factors may be involved for its translation. Therefore, to investigate the translation mechanism used by the RdRp ORF, full-length HvV190S cDNA was first cloned into a fungal transformation/expression vector under the control of a fungal *gpd1* promoter and a fungal *trpC* terminator (Fig. 3.1A). This recombinant plasmid, designated p190S, was then transformed into a virus-free host isolate of *H. victoriae*. Following transformation, the viral genome was integrated into the host DNA. In addition, transcription of the bicistronic mRNA followed by translation into CP and RdRp were respectively demonstrated by Northern and Western blotting analysis (see below). Viral dsRNA replication was not launched in these transformants, however, and only empty virus-like particles (VLPs) accumulated in the transformed hyphae, albeit to substantially higher levels than in natural infections (64). As in natural infections, the primary translation product of ORF1, p88, was phosphorylated and proteolytically processed to generate phosphorylated p83 and nonphosphorylated p78.

Western blotting of empty virus-like particles (VLPs) from different *H. victoriae* transformants detected both CP and RdRp, in similar sizes and amounts as those proteins present in viral particles purified from naturally infected *H. victoriae* (Fig. 3.1B). The successful expression of both CP and RdRp from plasmid p190S in *H. victoriae* suggested that this bicistronic cDNA construct, as well as the corresponding transformation and expression system, could be used to explore the mechanism of RdRp translation from the downstream ORF. Notably, no CP/RdRp fusion products larger than the nonfused RdRp (92 kDa) were detected by Western blotting using either CP- or RdRp-specific antiserum.

To verify that RdRp was expressed from genome-length HvV190S mRNA in *H. victoriae*, Northern blotting was performed. Total RNA was prepared from virus-infected *H. victoriae*,

p190S transformants, and empty-vector transformants, and then hybridized to CP- or RdRp-specific gene probes. A strong hybridization signal to genome-length mRNA (~5.2 kb) was detected for the infected strain as expected, as well as for all of the p190S transformants (Fig. 3.1C). In contrast, no band consistent with a subgenomic mRNA that might serve as an alternative or preferred template for RdRp translation was detected.

3.2.2 Initiation codons for HvV190S RdRp translation from bicistronic mRNA

Previous studies have shown that when the RdRp ORF is expressed in *E. coli* from a monocistronic construct employing its predicted start codon (HvV190S nt positions 2605–2607), the translation product is indistinguishable in both size and serological reactivity from the particle-associated RdRp of HvV190S (14). In addition to the start codon at 2605–2607 (AUG1), however, ORF2 contains two other in-frame AUG codons that are fairly close to its 5´ end: AUG2 at nt positions 2686–2688 and AUG3 at nt positions 2803–2805. Moreover, no out-of-frame AUG codons are found over this interval. To verify that HvV190S RdRp translation initiates from AUG1 in its native host, mutations of these three AUG codons were introduced into plasmid p190S by site-directed mutagenesis. After transformation of the mutated plasmids into virus-free *H. victoriae*, the level of RdRp expression and incorporation into VLPs was examined using RdRp-specific antiserum. The intensity of CP bands expressed from the upstream ORF, examined using CP-specific antiserum, served as an internal standard.

Mutation of AUG1 to GUG in construct pStartM1 resulted in an RdRp smaller in size than that expressed from the wild type construct (Fig. 3.2A). Mutation of AUG2 to GUG in construct pStartM2, on the other hand, did not influence RdRp size (Fig. 3.2A). The yield of RdRp from wild type, pStartM1, and pStartM2 constructs were all similar. Mutation of both AUG1 and AUG2 to GUG in construct pStartM3, however, severely reduced RdRp expression (Fig. 3.2A). These results suggest that GUG cannot serve as an effective reinitiation codon at either the AUG1 or the AUG2 position of the downstream ORF and also that AUG2, but not AUG3, can serve as an effective reinitiation codon when AUG1 is mutated, probably due to a requirement for proximity of the stop and restart signals as characterized further below. These results indicated that RdRp translation initiates from AUG1, which is located in the AUGA stop–restart motif (nt positions 2605–2607).

3.2.3 Effect of noncanonical start codons and context on HvV190S RdRp translation from dicistronic mRNA

In light of recent evidence for effective use of noncanonical start codons in another system utilizing coupled termination–reinitiation for downstream ORF translation (58), it was interesting to study this phenomenon in the HvV190S system. When mutating AUG1 to AUA in construct pStartM4, it was found that the expressed RdRp was indistinguishable in both size and yield from wild type protein (Fig. 3.2B). This suggested that translation from the downstream ORF could effectively initiate from the noncanonical start codon AUA. To verify that with pStartM4 RdRp translation was indeed effectively initiating from AUA at the AUG1 position, rather than from AUG2, construct pStartM5 was generated, in which AUG2 was mutated to the ineffective start codon GUG (see Fig. 3.2A for preceding evidence that GUG is ineffective). The RdRp expressed from pStartM5 was again indistinguishable in both size and yield from wild type (Fig. 3.2B). As a final test of AUA use as an effective start codon in this system, an additional construct pStartM6 was generated, in which AUG1 was mutated to the ineffective start codon GUG and AUG2 was mutated to AUA. In this case, a smaller RdRp was expressed, though in similar yield to wild type (Fig. 3.2B), consistent with AUA use as an effective start codon at the AUG2 position.

In the report by Powell et al. (58), the authors also determined that the sequence context immediately surrounding the reinitiation codon has little effect on the efficiency of downstream ORF translation. To address this issue in HvV190S, the relatively good context in the wild type construct, GGGACAAUGA (purine at position –3, pyrimidine at position –2; initiator AUG underlined) was changed into poorer contexts in constructs Kozak-3, GGG**C**CAAUGA (pyrimidine at –3; mutation in bold), and Kozak-2, GGGA**A**AAUGA (purine at –2; mutation in bold). Yields of RdRp expression were similar to wild type with both these mutants (Fig. 3.2C), suggesting that the sequence context immediately surrounding the reinitiation codon has little effect on the efficiency of downstream ORF translation in HvV190S.

3.2.4 Termination of CP translation is essential for reinitiation of RdRp translation

In all of the mutant constructs in Fig. 2, the stop codon of ORF1 remains in place, albeit changed from UGA to UAG in constructs pStartM4 and pStartM5 (see Fig. 3.2B). To assess the influence of uncoupling termination of CP translation from reinitiation for RdRp translation, the authentic UGA stop codon of ORF1 (at nt positions 2606–2608 in the AUGA stop–restart motif) was mutated to UGG. As a result, ORF1 is expected to terminate at the next in-frame stop codon, which is UAG at nt positions 2666–2668, 20 codons downstream of the native ORF1 stop (Fig. 3.3, construct pStop+20). As expected, the resulting full-length CP product (~97 kDa) was larger than the primary translation product of wild type ORF1, p88, though it was still processed to p83 and p78, consistent with the fact that the proteolysis to generate these smaller forms occurs near the CP C-terminus (15). Notably, no RdRp was detected with construct pStop+20 (Fig. 3.3), suggesting that uncoupling the stop–restart signals in the AUGA motif severely reduces expression of the downstream ORF.

To study further the effects of the position of the CP stop codon relative to the RdRp reinitiation codon, a series of mutants were generated, in which a stop codon in frame with ORF1 was introduced at various positions upstream or downstream of the ORF2 start codon, with the authentic UGA stop codon of ORF1 mutated to UGG in all cases (Fig. 3.3). In constructs pStop+8 and pStop+2, the inserted stop codon was respectively positioned 8 or 2 codons downstream of the ORF2 start codon, whereas in constructs pStop-2, pStop-8, and pStop-12, the inserted stop codon was respectively inserted 2, 8, or 12 codons upstream of the ORF2 start codon. The results showed that proximity of the CP stop codon to its wild type position is required for reinitiation for RdRp translation, since little or no RdRp was detected with constructs pStop+8, pStop-8, and pStop-12. One caveat is that in the case of mutant pStop-8, relocation of the stop codon by 8 codons upstream of the RdRp start codon disrupts a predicted pseudoknot structure in this region, which was shown to be important for reinitiation (see following sections). Notwithstanding this caveat, these results imply that the relative location of the CP stop codon is important for the efficiency of reinitiation for RdRp translation, consistent with a coupled termination–reinitiation mechanism.

3.2.5 A pseudoknot structure is predicted upstream of the AUGA motif and appears important for reinitiation for RdRp translation

RNA structure predictions using the program HPKNOTTER (75) suggested that an H-type pseudoknot may closely precede the AUGA stop–restart motif in HvV190S (Fig. 3.4). As for other pseudoknots, this one is predicted to be formed when nucleotides in a hairpin loop base-pair with nucleotides in a single-stranded region outside the hairpin to form a second stem adjacent to the hairpin stem. To address whether this predicted pseudoknot plays a role in reinitiation for RdRp translation, multiple mutations were introduced in an effort to disrupt the structure (construct PK0; Fig. 3.5A). Indeed, no pseudoknot was predicted by the program HPKNOTTER for the mutated sequence in PK0. Upon analysis with three different transformants of construct PK0, little or no RdRp was detected (Fig. 3.5A), suggesting the importance of the predicted pseudoknot sequence for ORF2 reinitiation.

Next it was of interest to determine whether RdRp expression could be restored if a wild type version of the predicted pseudoknot sequence were reintroduced into construct PK0 downstream of the mutated one. Two constructs were made (Restore1 and Restore2), both having a new, wild type pseudoknot-to-AUGA cassette inserted downstream of the mutated sequence and the native AUGA position. In Restore1, the native AUGA motif was mutated to AUAG, which nonetheless retains the CP stop codon at its authentic position; as a result, the inserted wild type version of the predicted pseudoknot sequence is located downstream of the CP stop codon (Fig. 3.5B). Only limited restoration of RdRp expression was seen with this mutant, which also showed CP expression of the expected, native size. In Restore2, on the other hand, the native AUGA motif was mutated to GUGG; as a result, the CP stop codon is shifted to the inserted, downstream AUGA motif and the inserted wild type version of the predicted pseudoknot sequence is now upstream of this stop codon, mimicking the native organization (Fig. 3.5B). Much stronger restoration of RdRp expression was seen with this mutant, which also showed CP expression of the expected, larger size. From these results, it can be concluded that the predicted pseudoknot sequence is an important determinant for promoting translation of downstream ORF2 and that it must be located upstream of the CP stop codon for optimal activity.

3.2.6 Effect of the spacer length between pseudoknot and AUGA motif on reinitiation for RdRp translation

The length of spacer sequence between the predicted pseudoknot and the AUGA motif in HvV190S is 9 nt. To address whether spacers smaller or larger than 9 nt would impact RdRp expression, a set of deletion and insertion mutants were constructed and analyzed (Fig. 3.6). Deletion of 3 nt (leaving a 6-nt spacer) had little effect on RdRp expression from constructs Spacer6-1 and Spacer6-2, whereas deletion of 6 nt (leaving only a 3-nt spacer) essentially abolished RdRp expression from construct Spacer3. Reciprocally, insertion of one duplicate spacer region (giving an 18-nt spacer) had little effect on RdRp expression from construct Spacer18, whereas insertion of two duplicate spacer regions (giving a 27-nt spacer) essentially abolished RdRp expression from construct Spacer27. Based on these results, it can be concluded that the length of spacer sequence between the predicted pseudoknot and the AUGA motif is important, with spacers smaller than 6 nt or larger than 18 nt not supporting RdRp expression. Whether the sensitivity to spacer length is determined by the distance between predicted pseudoknot and CP stop codon and/or between predicted pseudoknot and RdRp start codon is not addressed by this experiment, but based on preceding findings (e.g., effective initiation of RdRp from AUG2 in Fig. 3.2), it seems that when the spacer length is larger than that in wild type HvV190S, the spacing between predicted pseudoknot and CP stop codon is the more relevant determinant.

3.2.7 Other victoriviruses contain similar sequence motifs for coupled termination–reinitiation for RdRp translation

Ten other viruses were assigned to the genus *Victorivirus* in 2009 with HvV190S as the prototype strain of the genus (19). The number of victoriviruses is still growing. BbRV1 from entomopathogenic fungus *Beauveria bassiana* (79), TcV1 from *Tolypocladium cylindrosporum* (80), and AfSV1, RnVV1 and UvRV1 from pathogenic fungus *Aspergillus foetidus* (81),*Ustilaginoidea virens* (82) and *Rosellinia necatrix* (83) are discovered as possible *victoriviruses* based on similarly predicted secondary structure in their RNA sequences. Fourteen of 16 members have an H-type pseudoknot predicted by HPKNOTTER (75) at nearly the same position as in HvV190S (Fig. 3.7). The defined Class 1 in H-type pseudoknot represents that the nt length of stem 1 is less than that of stem 2. Two exceptions are Sphaeropsis sapiens virus 2 (SsRV2) and Ustilaginoidea

virens RNA virus 1 (UvRV1), which contain a predicted long stem–loop structure instead of a predicted pseudoknot at nearly the same position. In all 14 victoriviruses with a predicted pseudoknot, including HvV190S, the hairpin stem (stem 1; cyan in Fig. 3.4) is more conserved in number of base pairs involved in stem formation (3 or 4) than is stem 2 (orange in Fig. 3.4), which comprises from 3 to 6 bp. Both stems, however, show a preponderance of G:C base pairs. Compared to stems, hairpin loops (both loop1 and loop 2) are less conserved in nucleotide numbers, loop 1 possesses 1 or 4 nucleotides and loop 2 with 2 or 6 nucleotides. Loop 2 consists of more AU than loop 1 in their sequences. The spacer length between the predicted pseudoknot and stop/restart sites vary from a minimum of 6 nt in Sphaeropsis sapiens virus 1 (SsRV1) to a maximum of 11 nt in Beauveria bassiana virus 1(BbRV1) and UvRV1. The spacer length between the predicted pseudoknots and the CP stop codons in all victoriviruses are very similar, ranging from a minimum of 8 nt in Magnaportha oryzae virus 1 (MoV1) to a maximum of 12 nt in Chalara elegans RNA virus 1 (CeRV1). The spacer lengths between the predicted pseudoknots and the RdRp start codons in all of the victoriviruses are also similar, ranging from a minimum of 6 nt in SsRV1 to a maximum of 13 nt in MoV1, TcV1(Tolypocladium cylindrosporum virus 1), BbRV1 and UvRV1. The start/stop region can be distributed into three types. Seven of the 16 victoriviruses contain the AUGA stop–restart motif, the remaining 8—CeRV1, MoV1, SsRV1, and SsRV2—show a small difference: the stop and start codons are close together, but do not overlap (2 nt spacers in each). Moreover, in 3 of these viruses—CeRV1, SsRV1, and SsRV2—the start codon comes first, whereas the remaining 5— BbRV1, MoV1, RnVV1 (Rosellinia necatrix victorivirus 1), TcV1 and UvRV1—the stop codon comes first. Victoriviruses show a varying scale for minimum free energy to form a stable pseudoknot, with the lowest value of -16 kcal/mol in MoV2 and highest value of -5.8 kcal/mol in MoV1. These observations suggest that each victorivirus, similar to the prototype HvV190S, expresses its RdRp from the downstream ORF via a coupled termination–reinitiation mechanism, even though some variation in the relevant signals, overlapping vs. nonoverlapping stop and restart codons is tolerated.

3.2.8 Effect of deletions of downstream sequences of AUGA motif on RdRp reinitiation

Although no example suggested an involvement for the downstream RNA sequence on

reinitiaton in RNA viruses, it did occur for reinitiation mechanism on study of SART1. A retrotransposon SART1 showed for a first time that a downstream RNA secondary structure contributes to ORF2 translation while other features are similar to those of the reinitiation mechanism for RNA viruses (84). Recently, research on RnVV1 proposed a secondary structure in the downstream ORF that might be involved in reinitiation (83). Coincidentally, HvV190S was predicted to form a stable secondary structure (two stem-loops, RdRP stem 1 and stem 2) in the downstream sequences nearby the overlapped AUGA motif (Fig. 3.8 Left panel). To test if the secondary structure contributes to reinitiation of RdRp gene, 12 nt-deletions were made to those two stems separately. After analyzing the purified VLPs, mutants containing deletions expressed RdRp protein as efficiently as wild type (Fig. 3.8 Right panel), which suggested that the downstream RNA sequences are not involved in reinitiation of the downstream gene, same as other RNA viruses did on reinitiation. However, the whole-length of the downstream RNA sequences was not yet tested.

3.2.9 Effect of deletions of upstream sequences of AUGA motif on RdRp reinitiation

In Fig.3.5, a sequence element was identified to be required for reinitiation of translation of RdRp from the downstream ORF of the bicistronic genome of HvV190S. This comprised a 32-nucleotide-stretch of RNA sequence immediately upstream of the AUGA motif. In a continuation of this study, the aim was to define the required sequence elements within the entire upstream CP ORF. Signals for termination-reinitiation among caliciviruses have been located within ~ 80 nucleotide at the 3' end of the upstream ORF (62, 85). A series of deletions to the 3' end of HvV190S CP sequence spanning ~170 nucleotides were made (Fig. 3.9 upper panel). After constructs containing deletions were introduced into virus-free *H. victoriae* isolate, virus-like particles were purified and subjected to immunoblot analysis using antisera specific for HvV190s capsid protein and RdRp polymerase. No RdRp expression was detected with the RdRp-specific antiserum from any of those constructs (Fig. 3.9 bottom panel). However, conclusion cannot be made that all upstream ~170 nucleotides are required for RdRp reinitiation bacause the inability to detect RdRp might be due to unsuccessful packaging the synthesized RdRp into VLPs. Thus, defining the RNA sequence boundary required for RdRp reinitiation via analyzing purified VLPs would not be useful for this purpose. Instead, a dual-reporter

system was established that is not only able to define the sequence boundary, but also to generate quantitative data on expression on RdRp expression level.

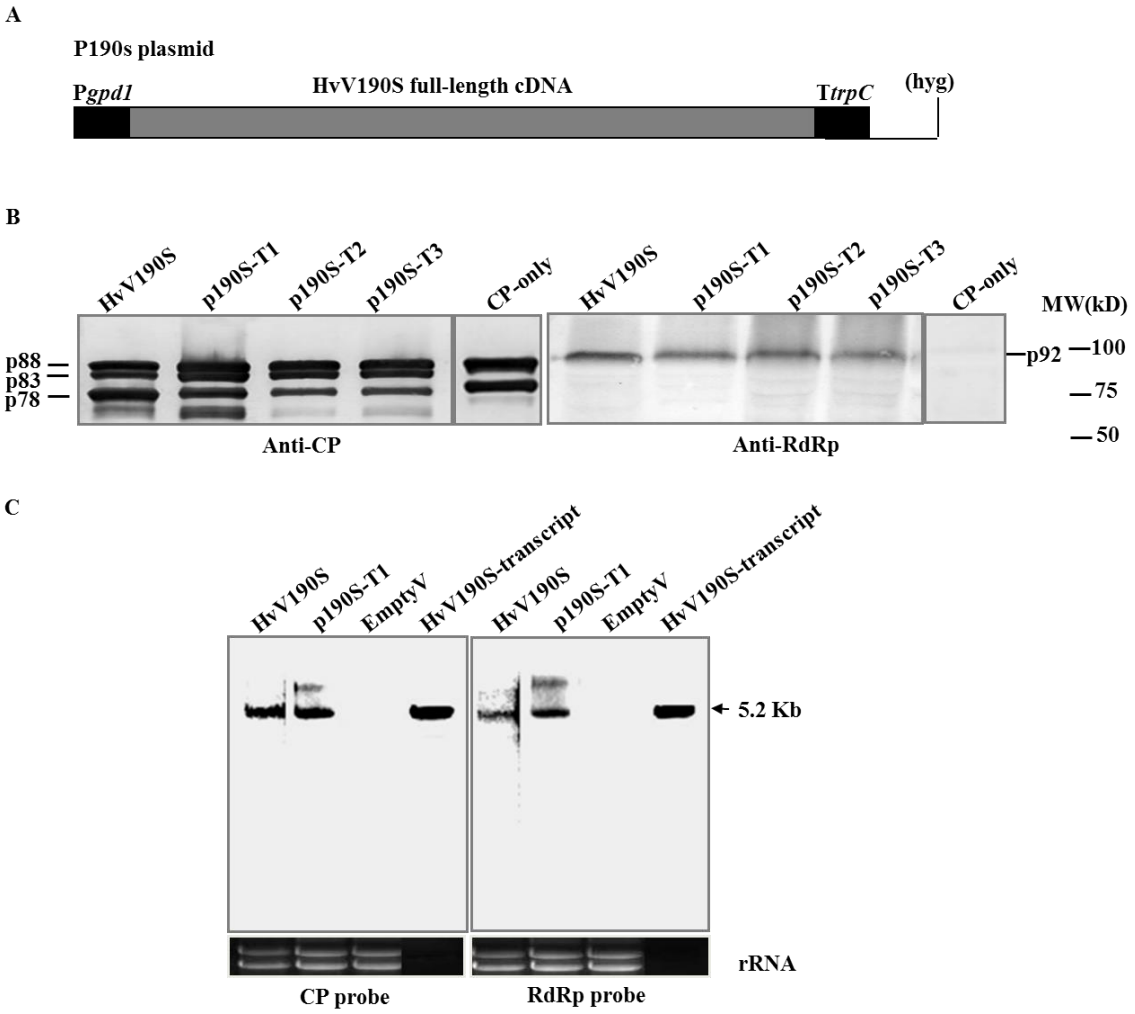
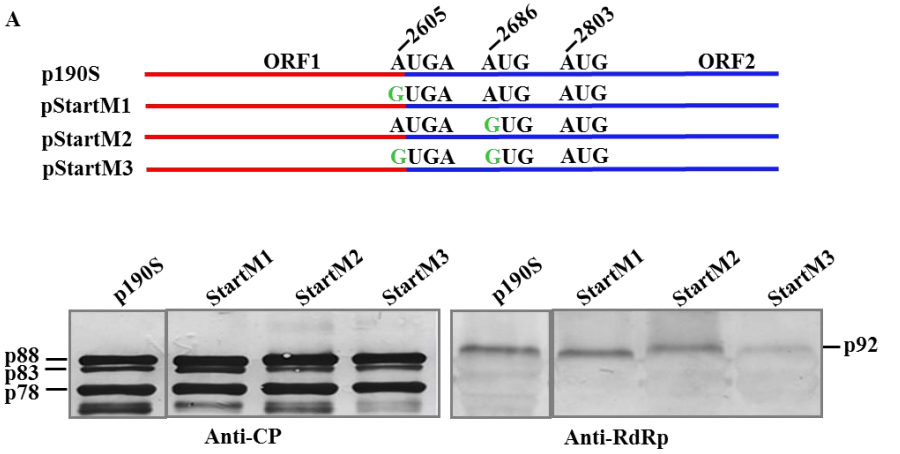
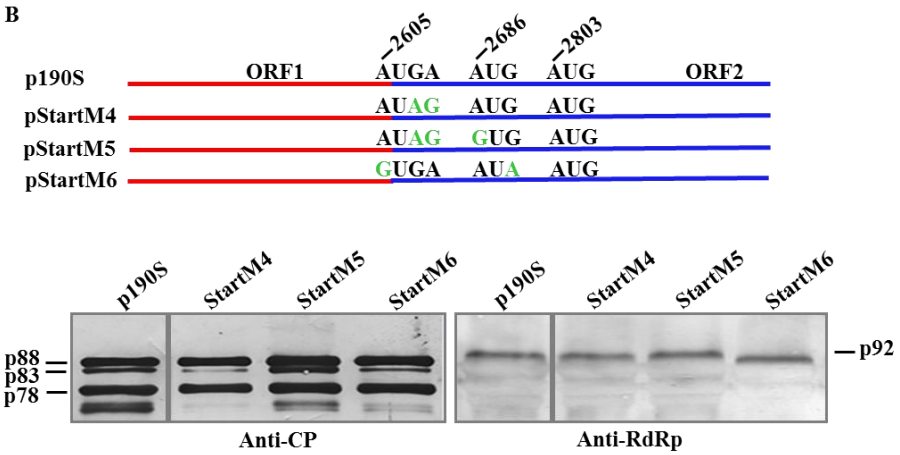


Figure 3.1 Construction of the fungal transformation/expression vector p190S and expression of CP and RdRp from plasmid p190S transformed into *H. victoriae.*

(A) Schematic representation of the structure of the fungal transformation/expression vector p190S containing full-length cDNA of HvV190S. Pgpd1, gpd1 promoter (nt 1-nt 662) from *Chochliobuls heterostrophus*; TtrpC, trpC terminator from *Aspergillus nidulans*; hyg, hygromycin B resistance gene. (B) Western blot analysis of viral proteins expressed from plasmid p190S in a virus-free *H. victoriae* isolate using CP- or RdRp-specific antiserum. Virus particles were purified from the wild type strain and p190S transformants, and subjected to sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) and Western blotting. Because of the preponderance of CP relative to RdRp in the particle samples, ten times more sample was loaded into each gel lane for detecting

RdRp. The designations of different proteins are given, and sizes of protein standards are indicated at right. HvV190S, viral proteins prepared from wild type strain A-9; p190S-T1, -T2, and -T3, viral proteins prepared from three different p190S transformants. CP-only, which solely contains the CP ORF sequence of HvV190S genome and used as a negative control for expression of RdRp, was adapted from previously obtained transformants described by Dunn et al. (73). (C) Northern blot analysis of total RNA isolated from the wild type strain and p190S transformants. Hybridization was carried out using CP- or RdRp-specific DNA probes corresponding to ORF1 and ORF2, respectively. Empty vector lacking the HvV190S cDNA sequence served as a negative control (EmptyV). Full-length HvV190S transcripts synthesized *in vitro* by purified virions served as a positive control (HvV190ST).





C

	-3	+4
p190S	.CGAGGGACA	**AUG**AGUGA.
	-3	
Kozak-3	.CGAGGGCCA	**AUG**AGUGA.
	-2	
Kozak-2	.CGAGGGAAA	**AUG**AGUGA.

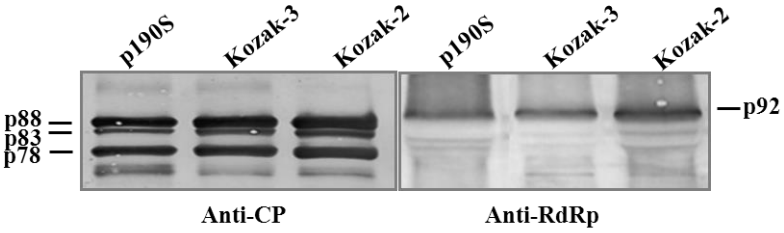


Figure 3.2 Identification of the reinitiation codon of the downstream RdRp ORF and effects of using noncanonical start codons and suboptimal AUG context.

(A, B) Top panels, CP and RdRp coding sequences are shown as red and blue lines, respectively. The first three AUGs in frame within ORF2 are indicated, and their genomic nt positions are shown above each. Specific mutations that were introduced in these potential start codons are shown in green. Bottom panels, Western blot analysis of viral translational products in VLPs purified from transformants of p190S and mutated HvV190S constructs. Viral translational products are indicated at left and right, and were detected with CP- or RdRp-specific antiserum as indicated at bottom. Ten times more sample was loaded into each SDS-PAGE gel lane for detecting RdRp. (C) Top, sequences surrounding the RdRp start codon (underlined) are shown. Specific mutations that were introduced to alter the Kozak context of this start codon are shown in green and labeled with their relative positions. Bottom, Western blot analysis comparable to those in A and B.

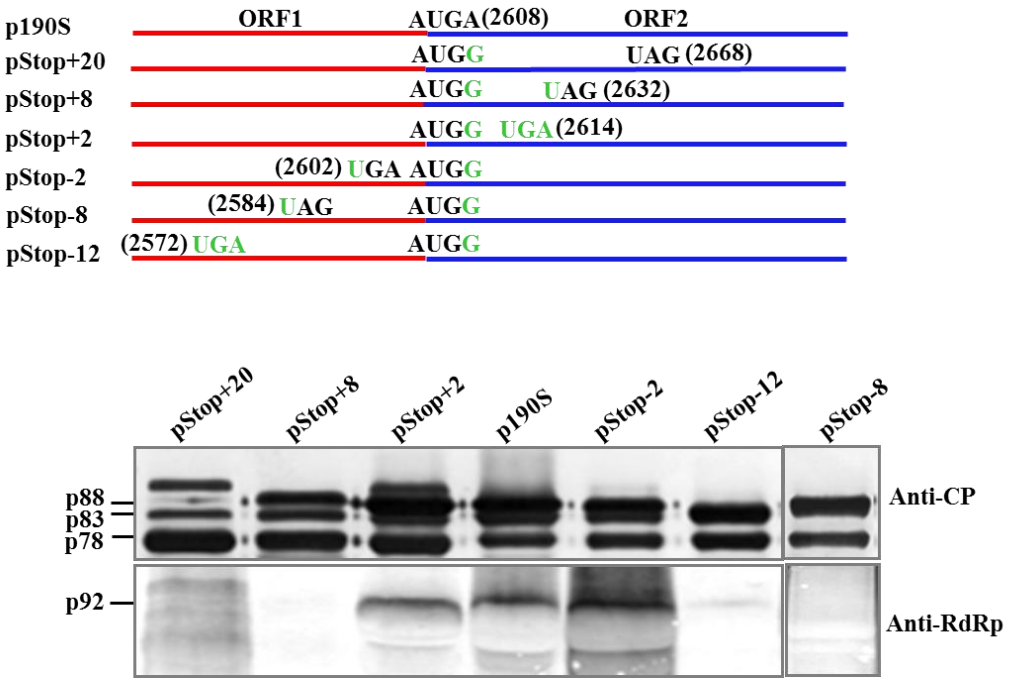


Figure 3.3 Influence of uncoupling the termination-reinitiation signals in the tetranucleotide AUGA overlap on expression of the downstream RdRp ORF. Top, schematic representation of the different constructs in which the position of the CP stop codon was varied relative to that of wild type construct p190S. Numbers in parentheses refer to the genomic positions of the last nt in each stop codon. Construct designations refer to the number of codons separating the newly engineered in-frame stop codon from the position of the authentic stop codon in wild type construct p190S, with "+" and "–" signs indicating downstream or upstream positions, respectively. Mutations introduced into the wild type sequence are shown in green. Bottom, Western blot analysis comparable to those in Fig. 3.2. Viral translational products are indicated at the left and were detected with CP- or RdRp-specific antiserum as indicated at the right.

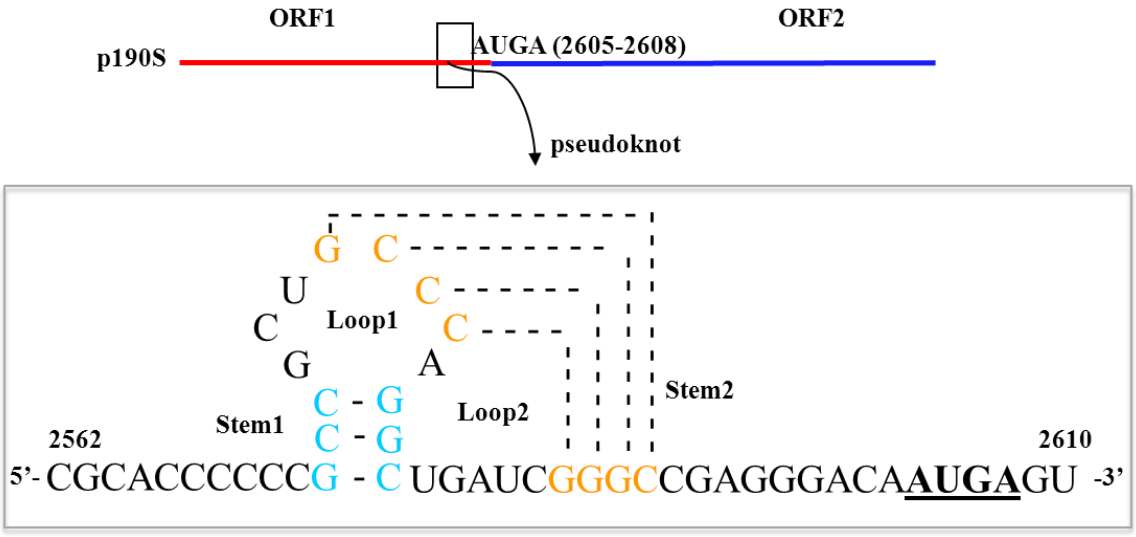


Figure 3.4 H-type pseudoknot predicted for sequences upstream of stop–restart region of HvV190S.

A predicted RNA pseudoknot structure is encoded near the 3´ end of ORF1 preceding the stop-restart site. Nucleotide positions are indicated and the stop-restart site is italicized and underlined. Dashed lines indicate base-pairs predicted to form stem1 in cyan and stem 2 in orange. Two loops are formed as the result of stem formation. Loop 1 includes nucleotides GCU plus unpaired A whereas loop 2 consists of UGAUC.

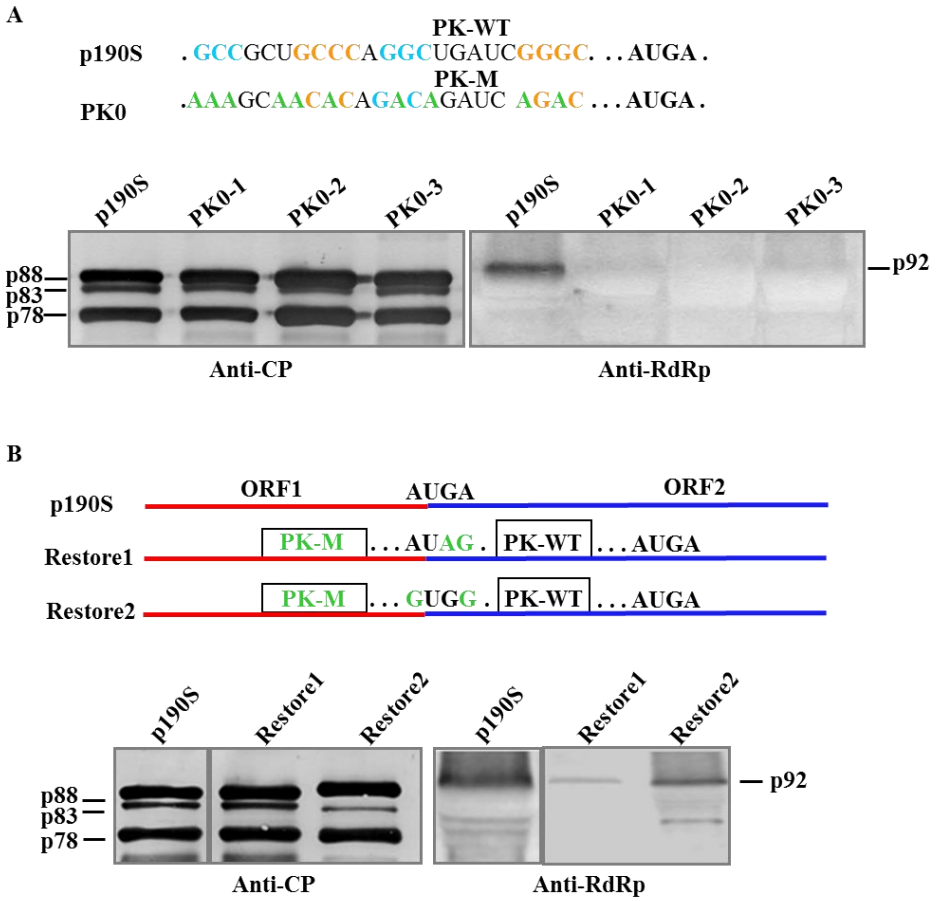


Figure 3.5 Effect of disrupting or restoring the predicted pseudoknot sequence on RdRp expression.

(A) Top, primary sequences of the predicted pseudoknot in the wild type construct p190S (PK-WT) and the mutated version (PK-M) in construct PK0 are shown. Stem 1 (cyan) and stem 2 (orange) are color-coded to match Fig. 3.4. Specific mutations are shown in green. The spacer sequence CGAGGGACA between predicted pseudoknot and AUGA motif is indicated by dots. Bottom, Western blot analysis comparable to those in Fig. 3.2. Viral translational products are indicated at left and right, and were detected with antisera to CP or RdRp. (B) Top, schematic representation of the different constructs in which a wild type pseudoknot-to-AUGA cassette (green) is inserted downstream of PK-M and the authentic AUGA position. Mutations made in the original AUGA motif are also shown in green. Bottom, Western blot analysis comparable to that in A. Viral translational products are indicated at left and right, and were detected with CP- or RdRp-specific antiserum as indicated at bottom.

p190S CCCCCCGCCGCUGCCCAGGCUGAUCGGGCCGAGGGACAAUGAGU

Spacer3 CCCCCCGCCGCUGCCCAGGCUGAUCGGGCC.....CAAUGAGU

Spacer6-1 CCCCCCGCCGCUGCCCAGGCUGAUCGGGCC...GGACAAUGAGU

Spacer6-2 CCCCCCGCCGCUGCCCAGGCUGAUCGGGCCGAG...CAAUGAGU

Spacer18 CCCCCCGCCGCUGCCCAGGCUGAUCGGGC(CGAGGGACA)2AUGAGU

Spacer27 CCCCCCGCCGCUGCCCAGGCUGAUCGGGC(CGAGGGACA)3AUGAGU

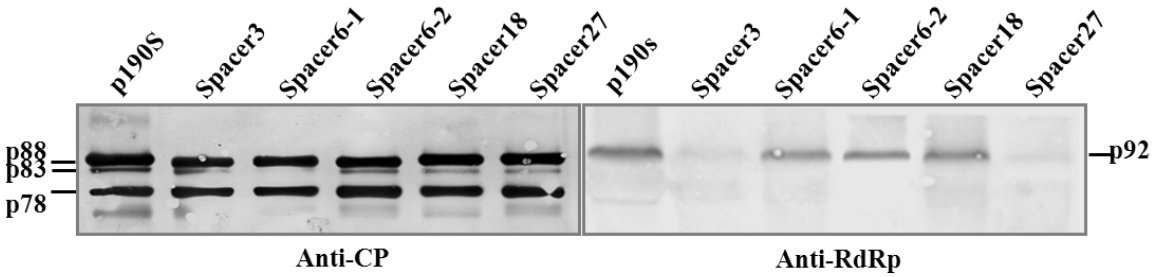


Figure 3.6 Length requirements for the spacer region between predicted pseudoknot and AUGA motif.

Top, schematic representations of deletion or insertion mutations in the spacer region (9 nt) of wild type construct p190S. Stem 1 (cyan) and stem 2 (orange) are color-coded to match Figs. 3.4. Specific mutations are shown in green. Spacer regions in constructs Spacer3 and Spacer6 contain 3 and 6 nt, respectively; spacer regions in constructs Spacer18 and Spacer27 are respectively duplicated (18 nt) and triplicated (27 nt) relative to wild type. Bottom, Western blot analysis comparable to those in Fig. 3.2. Viral translational products are indicated at left and right, and were detected with CP- or RdRp-specific antiserum as indicated at bottom.

Virus	Sequence	Class	MFE	Stem 1	Stem 2	Loop 1	Loop 2	nt to stop/re start
HvV190S	CCCCCCGCCGCUGCCCAGGCUGAUCGGGCCGAGGGACA**AUGA**	1	-8.1	3	4	4	5	9
BbRV1	UUGCCGGUGCUGCCCCACCCGGAGGGGCCGAACCCCGAG**UAAUG**	1	-8.3	4	4	2	4	11
BfTV1	CGUUCCCCUGCUGCCCCAGGUAACGGGGCCGGACACGAU**AUGA**	1	-7.9	4	4	2	4	10
CeRV1	UUCUUCUGUGGCUGCCCCCACUAAUGAGGGGCCGAAACG**AUG**UC**UAG**	1	-8.3	4	4	2	6	7
CmRV	GUUCACGGUGCCGCCGGAGCCUUACCGGCUCCACAAGU**AUGA**	1	-8.6	3	5	4	3	9
EfV1	CCUGAGGCUGCCACCGGUGCCGAACCGGUGCCCCAAGC**AUGA**	1	-12	3	6	4	3	9
GaRV-L1	CCAGCCGAAGCUACUGCUGCUGAAGCAGUGCCCGCUCA**AUGA**	1	-12	3	6	1	2	9
HmTV-17	AUGCAACAGGACGCAGCCCUGCAAGCUGCGGGGGCUCA**AUGA**	1	-11	4	6	1	3	8
MoV1	CCCCCGAACCCCACGGCGCGGCGCCUAGCCUGCACGAA**UAG**AU**AUG**	1	-5.8	3	3	2	2	8
MoV2	AACGAGCAGGCCGCCGGCGCCCAGCCGGCCCCCGCUCA**AUGA**	1	-16	3	6	1	2	9
RnVV1	CGGUUGCCGCUGCCCCGGCUGAACGGGCCGCCGAACAG**UAAUG**	1	-9.8	4	4	2	5	10
SsRV1	UGUAGCACCCGGCCCCGCCCAGCCUGACGGGCCCGCCA**AUG**AA**UAA**	1	-7.8	3	4	4	4	6
TcV1	CUGUUGCCGACCCCCAGGCUAUGGUGGGCGACCAGAUC**UAAAUG**	1	-9.2	3	4	4	5	10
UvRV1	CACCUAGCGCUGCCCCGCAGAGUGGGCCGACCCCCGAG**UAAUG**	1	-7.5	3	4	2	5	11
SsRV2	GAAGGCGCCGACGGUAAAGUACCGCCGGCGCCCCGCA**AUG**AA**UAA**		-22					
AfSV1	CCUGCAGCUGAUGAGGAUGAUAACCAACCCCCTCAACCAGUUGCGAU CGGUAACGGAAU**UAATG**		-16					
Consensus for pseudoknot structure		1	≤5.8	3–4	3–6	1–4	2–6	6–11

Figure 3.7 Sequences of HvV190S-like fungal viruses (victoriviruses) predicted to form a pseudoknot (or a long stem-loop) structure upstream of the stop-restart motif.

Program HPknotter (two versions 1.04 and 1.2) was employed for pesudoknot prediction. Thirty and fifty nucleotides upstream of stop-restart region for each virus were used for prediction. Stem 1 (cyan) and stem 2 (orange) are color-coded to match Figs. 3.4–3.6. Stop-restart sites (s/s) are bolded and underlined. Type of pseudoknot and minimum free energy (MFE) for this structure are listed for each virus together with the length of each stem and each loop sequences. Numbers of nucleotides between each predicted structure and stop-restart site are also indicated.

Virus names and abbreviations: BbRV1, Beauveria bassiana RNA virus 1; BfTV1, Botryotinia fuckeliana totivirus 1; CeRV1, Chalara elegans RNA virus 1; CmRV, Coniothyrium minitans RNA virus; EfV1, Epichloe festucae virus 1; GaRV-L1, Gremmeniella abietina RNA virus L1; HmTV1-17, Helicobasidium mompa totivirus 1-17; HvV190S, Helminthosporium victoriae virus 190S; MoV1 and MoV2, Magnaporthe oryzae viruses 1 and 2; RnVV1, Rosellinia necatrix victorivirus 1; SsRV1 and SsRV2, Sphaeropsis sapnea RNA viruses 1 and 2; TcV1, Tolypocladium cylindrosporum virus 1; UvRV1, Ustilaginoidea virens RNA virus 1; AfSV1, Aspergillus foetidus slow virus 1. MFE, minimum free energy value for structure prediction at 37˚C. SsRV2 and AfSV1 were predicted to possess a long stem-loop in this region with program Mfold and their MFE were indicated in red.

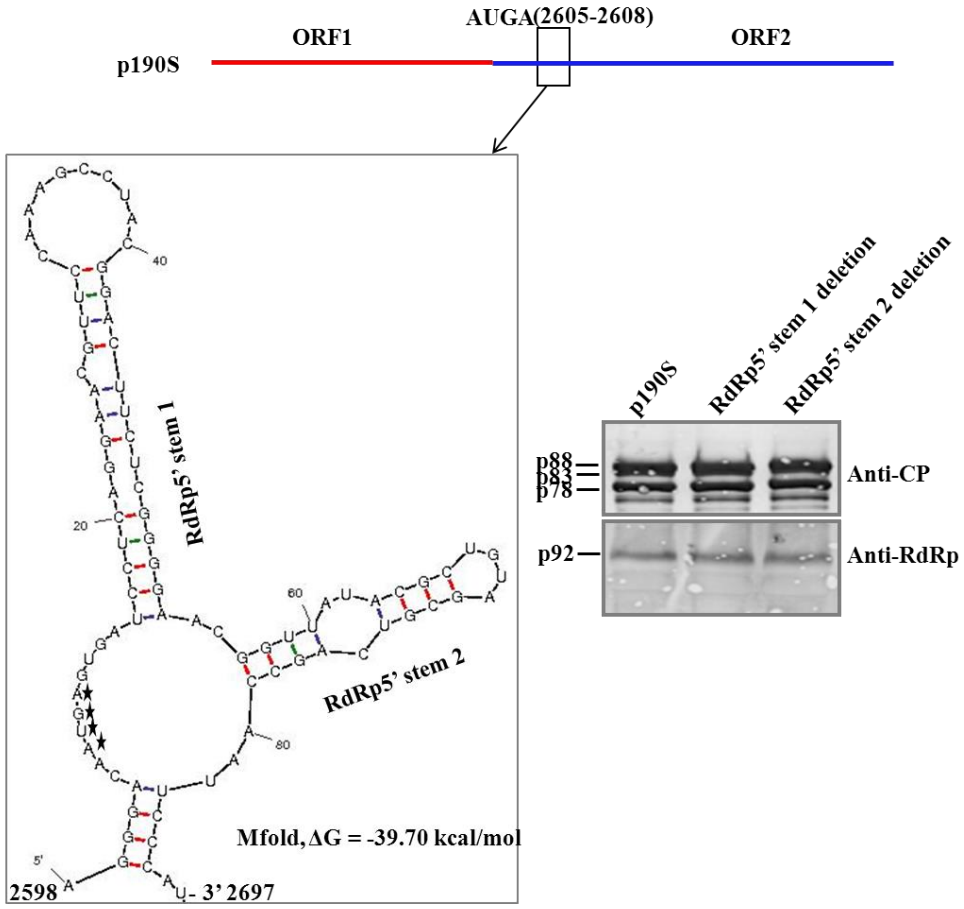


Figure 3.8 Effect of disruption of predicted secondary structure at the 5' end of RdRp ORF on RdRp expression.

Left, a stable secondary structure was predicted with Mfold downstream of AUGA motif (indicated with stars) in the 5' end of RdRp ORF consisting of two stem-loops. Stem 1 includes nucleotides 2613-2651, and stem 2 includes nucleotides 2654 to 2676. Right, 12 nucleotide deletions are introduced into stem 1 and stem 2 separately and mutant constructs were transformed into virus-free *H. victoriae*. Subsequent to transformation, VLPs were purified and subjected to Western blotting with CP- and RdRp-specific antiserum.

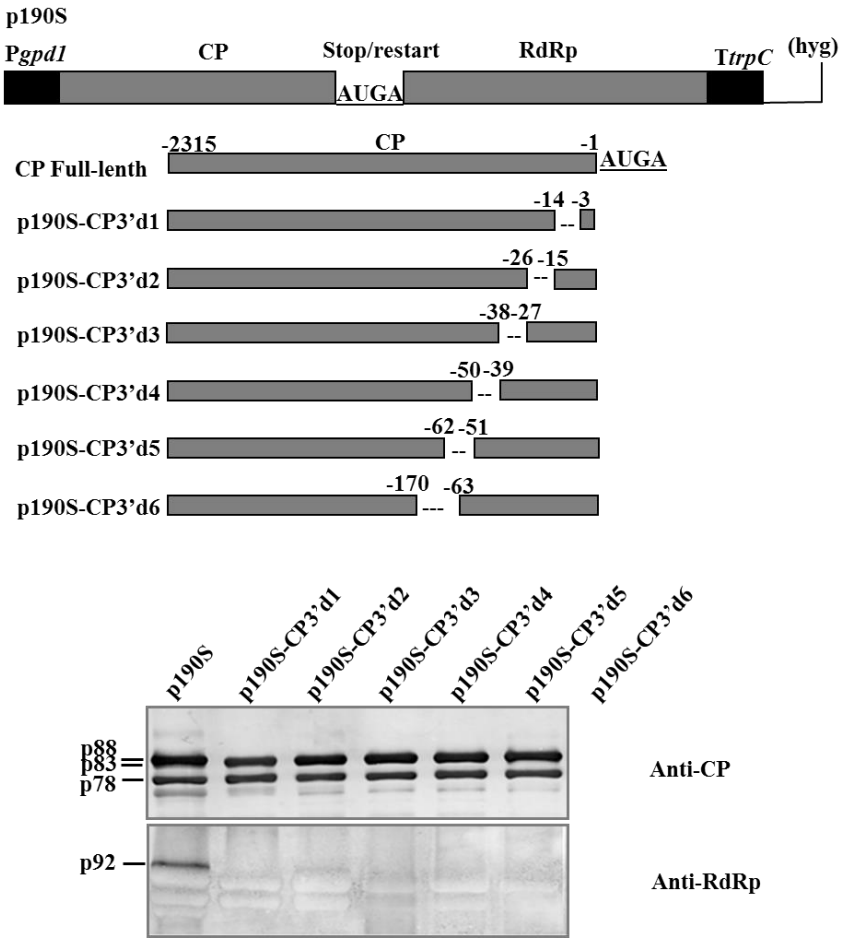


Figure 3.9 Effect of nucleotide deletions at the 3'end of CP ORF on RdRp expression.

Top, Schematic representation of the basic structure of plasmid p190S as described in Fig. 3.1A. The stop/restart site is bolded and underlined. Below, different nucleotide deletions (dashed) were introduced into the 3' end of CP ORF and positions for the deleted part of each construct was indicated with numbers at both ends relative to the first A in tetranucleotides AUGA. The full-length CP ORF is drawn at top of mutant diagrams and its start codon is located at -2315 upstream of the start/stop site. Bottom, Western blot analysis of viral proteins expressed from different constructs using CP- or RdRp-specific antisera. VLPs were purified from transformants with original plasmid p190S and p190S-derived mutant constructs. The positions of different proteins from VLPs are shown on the left.

4 CHAPTER FOUR: An H-type pseudoknot structure is essential for coupled termination-reinitiation translation of the downstream ORF of HvV190S

4.1 Introduction

In Chapter 3 HvV190S was confirmed to utilize a termination-dependent reinitiation to express its RdRp gene from the downstream ORF2. A sequence element of 32nt-strentch containing a predicted pseudoknot structure immediately upstream of ORF2 has been verified to be important for the reinitiation of RdRp. However, further determination of other sequence elements spanning the entire CP sequence was not yet completed. Also the detailed explanation on the role of pseudoknot in reinitiation has not performed.

The mechanism of translational coupling has been best studied with certain caliciviruses, feline calicivirus and rabbit haemorrhagic disease virus. In these viruses, 84–87 nt of RNA sequence immediately upstream of the overlapping stop–restart codons, termed the "termination upstream ribosome binding site" (TURBS) are required for efficient reinitiation (62, 86). Two distinct regions within the TURBS have been shown to play significant roles. Motif 2 is not conserved among different caliciviruses, and its function remains unknown. Motif 1, on the other hand, is conserved and is complementary to a small single-stranded region at the tip of helix 26 of 18S rRNA, which is juxtaposed to mRNA in the translating ribosome (87). This complementary region is thought to tether the mRNA to the 40S subunit, allowing time for the ribosome to acquire the initiation factors necessary for translation of the downstream ORF (88). There is also evidence that TURBS is involved in recruiting eukaryotic initiation factor 3 (eIF3) and eIF3/40S complexes (89). The multi-subunit eIF3 complex plays multiple roles in translation initiation including dissociating the 60S and 40S ribosomal subunits after termination (90), and it is therefore possible that the TURBS is involved in both ribosome dissociation/recycling and 40S tethering (58). A mechanism for stop–restart similar to that of calciviruses has been proposed for influenza B viruses (85). Stop–restart in pneumoviruses, however, appears to occur by a mechanism distinct from that of caliciviruses. Present evidence suggests that the presence of a large secondary structure in a region upstream of stop and restart codons is required for reinitiation of translation in pneumoviruses (60, 91), but the nature of this structure and how it works to promote reinitiation remain unknown.

The rates of translational reinitiation for downstream ORFs in RNA viruses are mostly below 10% relative to upstream ORF translation (63, 92). Accuracy of these results depends on the quality of the measuring systems. In the past, monitoring such rates have mainly relied on detection of enzymatic activities of different combinations of two (upstream and downstream) enzymes including β-galactosidase, chloramphenicol acetyltransferase (CAT), and luciferase (60, 63, 93), which have often involved two independent, unrelated, and thus less accurate measuring methods (91, 93). More recently, researchers have used autoradiography to quantitate translation levels for the target ORFs by measuring radioactivity incorporated into the translation products (58, 61, 62, 92). Discovery of the dual-luciferase system has permitted researchers to detect both luciferase activities at the same time in a single sample and have also increased accuracy because of high sensitivity of the activity measurements (94-97). These and other attributes make the dual-luciferase system nearly ideal for further applications in a broad range of fields except for the relatively high cost of reagents. In recent years, fluorescent proteins have offered excellent alternatives to luciferases based on their high sensitivities and absence of expensive disposable reagents in their assays. Fluorescent proteins have gained broad use as a visualization tool for gene expression and gene localization *in vivo* (98-101). In addition, they can be used as quantitative reporters for gene expression in that measured fluorescence intensity units can accurately reflect fluorophore mRNA and protein levels *in vivo* (102, 103). Dual-fluorescence reporter assay systems have been successfully established in mammalian cell lines (104) as well as in yeast (105). The application of dual-fluorescence assays for quantitation in other organisms including filamentous fungi has seldom been reported, although reliable expression of fluorescent proteins has been reported in many cases, though mostly for visualization rather than quantitation (106, 107).

In this Chapter, a dual-fluorescence measurement system was established to analyze the translational stop-restart mechanism of HvV190S. The boundaries of the required upstream sequence element were defined, and the role of a predicted pseudoknot structure within this region was evaluated in detail. Whether host 18S rRNA is involved in translational termination–reinitiation in this system was also investigated. Together, the new results and conclusions expand the understanding of the mechanism of stop-restart employed by RNA viruses.

4.2 Results

4.2.1 Expression of reporter genes in the fungus, *H. victoriae*

To establish a dual-reporter gene system, four reporter genes were tested for expression in *H. victoriae*, host of HvV190S. Namely, these are the genes coding for Renilla luciferase (Rluc) and firefly luciferase (Fluc), which are commonly paired in dual-luciferase assays (94, 95), plus enhanced green fluorescent protein (eGFP) and the DsRed variant of red fluorescent protein (RFP) (101). A modified version of the fungal transformation/expression vector p190S (108) was used (Fig. 4.1A; also see Methods). A 6xHis tag sequence was added to the 3' end of each reporter gene for the detection of its expression using a His-tag specific antiserum. Following transformation of protoplasts prepared from a virus-free *H. victroriae* strain, transfromants with each reporter construct were randomly selected and total protein was extracted and subjected to Western blot analysis. Surprisingly, the widely used luciferases for studying recording signals in other organisms were not expressed in *H. victoriae* (Fig.4.1B and 4.1C) although successfully expressed in *E. coli* and *in vitro* translation experiments. On the other hand, the pair of fluorescent proteins, eGFP and RFP were expressed at high levels in *H. victoriae* and fluorescence was observed with the naked eye in the medium on which the fungal cultures were grown or in total protein extracts (Fig. 4.1D for *eGFP* transformants, Fig. 4.1E for *RFP* transformants). Both green and red fluorescence were strongly detected under epifluorescence microscopy and no background signals were detected in non-transformed cultures. Western blot analysis of randomly selected transformants using a His-tag antiserum showed that the different transformants expressed varying levels of GFP or RFP (Figs. 4.1D and 4.1E).

4.2.2 Dual-fluorescence reporter system for quantitation of downstream ORF expression

In order to construct a bicistronic vector with two genes coding for fluorescent proteins, I tested *apriori* whether eGFP and RFP can be expressed stably in *H. victoriae* when fused to HvV190S proteins. Two constructs were made with eGFP coding sequence fused to the the N-terminus of the CP ORF and with RFP fused to the C terminus of the

RdRp ORF, and the resultant constructs were used to transform protoplasts of a virus-free *H. victoriae* isolate. Western blot analysis of selected transformants with these two constructs suggested that eGFP fusion construct successfully expressed eGFP-CP fusion with an expected size of ~114 kDa (see Fig. 4.2B.) whereas RFP fusion construct somehow did not express any protein (data not shown). Thus in further attempts to construct a fluorescent bicistronic vector, eGFP-CP:RFP-His, eGFP-CP fusion ORF was set as the upstream ORF whereas RFP alone was arranged as the downstream ORF. Thus, this dual-fluorescence vector possesses the following two main features: 1, It contains the full-length sequence of CP ORF from HvV190S genome; 2, The stop codon of its upstream ORF overlapping the start codon of the downstream ORF in the tetranucleotides AUGA, which mimics the wild type genomic organization of HvV190S except that the RdRp ORF was replaced by RFP ORF followed with a 6×His tag sequence (Fig.4.2A.).

The construct was then introduced into a virus-free *H. victoriae* isolate and total protein was extracted. Translational products from the upstream and downstream ORFs were analyzed using Western blotting with specific antiserums. Results showed a few proteins bands from upstream ORF were detected with an antiserum to GFP with the largest band sizing ~114 KDa (Fig. 4.2B.). It was reported previously that there are three types of related CP proteins expressed from p190S vector (Fig. 3.1B in Chapter 3), with the primary translational product p88 is co- or posttranslationally proteolyticly processed into two related shorter proteins, p83 and p78. In this study, three major types of proteins were similarly detected indicating translational products of the upstream fusion protein encountered similar posttranslational processing. When the protein loading amount for Western blot analysis of the downstream translational product was increased 10 times relative to that used for the detection of the upstream product, a strong band was shown with a similar size to the RFP protein (~27 KDa) using a specific antiserum to the downstream His-tag sequence. This suggested that translation of RFP was successfully reinitiated from the downstream ORF of the recombinant bicistronic vector although it contained no RdRp sequence.

To assess the effectiveness of this system as a quantitative tool to identify the changes in RdRp reinitiation, random transformants from this construct were selected and respective green and red fluorescence was measured. Results showed that red

fluorescence displayed a parallel change with the green fluorescence although expressed at much lower levels (Fig. 4.2C.). The measured activity of red fluorescence was able to scatter at a linearization regression line relative to the measured activity of green fluorescence, suggesting that the activity measured for each fluorescent protein represented well the concentration of RFP and GFP proteins. Thus, for further study, I used this system with GFP as the internal control to quantitate the level of RFP expression.

To present evidence that RFP expression is due to reinitiation of translation from this dual reporter vector which differs from the original genome of HvV190S, it was therefore necessary to verify that RFP expression from this vector indeed results from reinitiation of translation following termination of translation of the upstream ORF. To achieve this, a construct CP no-stop was made, in which the stop codon (UGA) of the CP ORF was mutated to UGG in the AUGA motif (Fig. 4.3A.). It has been shown before that such construct will no longer promote downstream RdRp translation as verified by absence of RdRp from purified virus-like particles (Fig. 3.3 in Chapter 3), Therefore, it was expected that RFP reinitiation would also be abolished from this construct. As expected, anti-His tag antiserum did not detect RFP expression from this construct (Fig. 4.3B.). Thus, in this dual-fluorescence system, initiation of translation of the downstream ORF depends indeed on termination of translation of the upstream ORF same as was previously shown for HvV190S genome.

4.2.3 Reinitiation efficiency is lower than 5%

The efficiency of reinitiation was found to below for a number of viruses employing a coupled termination-reinitiation strategy (less than 20%) and some are even lower than 5%, which was determined by upstream RNA sequence elements. To determine the efficiency in this case, a new construct with two reporters was made based on the dual-fluorescence system. Approximately 900 nucleotides were deleted from the 3’ end of the CP ORF to create a truncated fusion GFP-CP ORF (eGFP-CPt), then fused again to the RFP ORF (Fig. 4.4A). The construct, designated as eGFP-CPt-RFP, thus was expected to express a large single fusion protein with each end of CP protein containing a reporter gene. Compared to construct eGFP-CP:RFP, it will permit us to estimate the reinitiation efficiency when using GFP as an internal control for normalization of differential

expression caused by insertion numbers of vectors integrated into fungal chromosome or due to different mycelium amounts used for protein extraction.

The construct was transformed into a virus-free *H. victoriae* and total protein was extracted from different transformants. Western blotting was performed using a GFP, CP or RFP-specific antiserum. Results indicated two small proteins, instead of a large fusion protein, were independently detected with GFP or RFP-specific antiserum (Fig. 4.4B.) while no proteins were detected with the CP- specific antiserum (data not shown). Surprisingly, the small proteins detected by GFP or RFP- specific antiserum respectively appeared to be the same size as GFP or RFP protein expressed using a monocistronic construct (Fig. 4.4B.), which implied these two small proteins might be from the proteolytic process of the primary large fusion product. To test this idea, an *in vitro* translation vector was constructed copying exactly the whole sequence from plasmid eGFP-CPt-RFP used for the *in vivo* experiment and was translated in wheat germ extracts under the control of a T7 polymerase promoter. In vitro translation experiment was also performed for wild type vector eGFP-CP:RFP, serving as a control of translational products. Translational products were labeled with ^{35}S-methionine and separated by SDS-PAGE on a 10% gel. The autoradiograph showed there were no small proteins expressed with a size similar to GFP or RFP for eGFP-CPt-RFP, instead, two large proteins were expressed with the larger protein close to ~77 KDa as expected from the large fusion frame (Fig.4.4D). Although there was a slightly smaller protein also expressed from this construct, it was probably translated from the next initiation AUG in the GFP ORF, which is 78 codons downstream. For wild type construct T7- eGFP-CP:RFP, there were also several large proteins with expected sizes expressed from the first or next initiation codons of the upstream fused frame. In this construct, RFP was successfully expressed from the downstream ORF via reinitiation translation *in vitro*. That the small proteins were generated via limited proteolysis rather than sequence errors or other translation mechanism was further supported by measuring the fluorescence. Both green and red fluorescence were measured using the same droplets of total protein extracts. Red fluorescence gave a much higher readouts relative to the green one (Fig. 4.4C.) compared to those from vector eGFP-CP:RFP that showed a much smaller relative value (Fig. 4.2C.), which suggested both GFP and RFP were expressed at a high efficiency in vector eGFP-CPt-RFP, probably from the same ORF.

The linearization regression test confirmed the postulate since RFP showed an obviously perfect linearization correlation to GFP (Fig. 4.4C right panel).
Six transformants each from the constructs, eGFP-CP:RFP and eGFP-CPt-RFP were used to extract total protein and each experiment was repeated three times for fluorescence measurement. The mean ratio of the RFP to GFP from four extractions of six transformants of eGFP-CP:RFP was divided by the mean ratio of the RFP to GFP from four extractions of six transformants of eGFP-CPt-RFP. And the value was 3.9% (Fig. 4.5), suggesting only 3.9% of RFP was reinitiated from downstream ORF.

4.2.4 Deletion from 5' end of CP ORF showed that only a 38-nucleotide stretch is required for RFP reinitiation

Based on the dual-fluorescence system, it was able to define the required RNA sequence in the upstream ORF for reinitiation of RFP translation. By generating a series of nucleotide deletions from the 5' of CP ORF, 1002, 2001, 2196, 2247, 2277 and 2313 nts out of the total 2315 nucleotides upstream of the AUGA motif were removed, resulting 1313bp, 314bp, 68bp, 38bp and 2bp remaining in the upstream CP ORF for constructs d1, d2, d3, d4, d5 and dall, respectively (Fig. 4.6A.). All constructs were introduced into virus-free protoplasts of *H. victoriae*. Transformants from each construct were analyzed with Western blotting (Fig. 4.6B.). Different levels of RFP expression were detected for constructs CP5'_d1 to CP5'_d5, at either high or low levels. Only construct CP5'_dall totally abolished the RFP expression. By comparing the sequences deleted, results suggested that the remainder 38 nucleotides in CP5'_d5 were important for RFP reinitiation, these were removed in CP5'_dall to abolish RFP expression. That was supported by fluorescence results. Five or six transformants from each construct were used to measure their GFP or RFP fluorescence and the mean ratio of RFP to GFP for each construct was compared to the one from construct eGFP-CP:RFP containing full-length of CP ORF. When converted to percentage, the CP5'_d2 to CP5'_d5 gave comparable values that were not significantly different from the full-length construct (Fig. 4.6C.). CP5'_d1 gave a significantly higher expression of RFP. However, CP5'_dall showed an extremely low rate (<5%) resembling the value from CP no-stop in which RFP reinitiation was abolished when the stop codon of the upstream ORF was removed. All the results in this study suggest that the terminal 38 nucleotides upstream of the AUGA motif were essential and sufficient for reinitiation of translation from the

downstream RFP ORF. Reinitation of RFP was totally abolished by deleting these 38 nucleotides.

Initially it was expected to detect a varying size of proteins will be expressed from the upstream GFP- truncated CP ORF for deletion constructs respectively at 75KDa, 36KDa, 30KDa, 28KDa, 27KDa, 26KDa. However, GFP-specific antiserum did not detect varying sizes of upstream products. Instead, it detected the same size protein expressed from upstream for all constructs except the one with full-length CP inserted, in which expected fusion proteins were showing on the desired upper position except the similar small protein at lower position. The size of this small protein again was similar to GFP protein as the previous construct eGFP-CPt-RFP in Fig. 4.4B, suggesting the fused translational products from the upstream ORF may undergo the same proteolytic process when translated in vivo. The postulate was confirmed by *in vitro* transcription and translation experiments (Fig. 4.6D) in the same way as in Fig. 4.4D. The whole construct sequences starting from the first AUG of GFP-CP ORF to the stop codon of RFP ORF, for all mutant constructs were copied by high-fidelity polymerase PCR amplification and cloned into a T7 *in vitro* expression vector and translated in wheat germ extracts supplied with [^{35}S]-methionine. The wild type vector eGFP-CP:RFP and CP no-stop were made here as controls for translation from both upstream and downstream. The expected size for the upstream products for these two constructs are 114 KDa and 116 KDa respectively. The downstream RFP ORF is expected to be translated for eGFP-CP:RFP but not for CP no-stop. Autoradiography of the radiolabeled translational products showed that predicted differently sized upstream proteins were translated successfully together with a consistent RFP protein from downstream ORF except constructs T7-CP5'_d5 and T7-CP5'_dall. In T7-CP5'_d5 and T7-CP5'_dall, there were only one band detected although it was expected two separate bands. Since GFP was fused with a highly truncated CP ORF for these two constructs, one with 39 nucleotides and one with just one nucleotide left resulting in 1 kDa differential at the molecular weight for translated proteins, the fused upstream products were hardly separated in repeated experiments on 10% polyacrylamide gel.

4.2.5 Results of deletions from the 3' end of CP ORF are consistent with those from deletions from the 5' end

Following the experiment in Fig. 4.6, it was of interest to test in detail the effect of deletions within the 3' terminal 38-nucleotides on RFP reinitiation (Fig. 4.7A.). Constructs CP3'_d1 to CP3'_d3, involves sets of successive 12 nt-deletions within the target 38-nucleotides. It also included three more constructs that included deletions upstream the 38-nt region, extending the sequences tested from -2 to -170 nt upstream of the AUGA motif. Subsequent to transforming fungal protoplasts with the above described constructs, five or six transformants were selected for further analysis by Western blotting and fluorescence measurement. The first three constructs containing successive deletions from nucleotide -2 to -38 abolished RFP reinitiation, whereas the last three deletion constructs showed efficient RFP expression when compared to the wild type construct (Fig. 4.7B.). These results were confirmed by fluorescence measurement. Compared to the wild type construct, constructs CP3'_d1 to CP3'_d3 showed significantly much lower of RFP expression while CP3'_d4 to CP3'_ d6 showed comparable expression levels (Fig. 4.7C.). Once again, this suggests that the terminal 38 nucleotides of CP ORF are crucial for downstream reinitiation.

4.2.6 Swapping the pseudoknot of HvV190S with those predicted for HvV190S-like victoriviruses showed reinitiation from the downstream ORF of HvV190S was quite tolerant to various pseudoknots

In previous work, it predicted an H-type pseudoknot structure closely preceding the stop/restart AUGA motif in HvV190S in Fig. 3.5 in Chapter 3 (108). In this pseudoknot, nucleotides in a hairpin loop form base pairs with nucleotides in a single-stranded region downstream the hairpin to form a second stem adjacent to the hairpin stem. This secondary structure seems to occur among other 11 members in the same genus of *Victorivirus* as HvV190S (8). And the member of victoriviruses is still growing based on this predicted secondary structure such as recently discovered RnVV1 from pathogenic fungus *Rosellinia necatrix* (83). All predicted pseudoknots resemble in their short stems and the limited distance between pseudoknots and stop/restart sites, which suggest a similar role on translation of downstream genes and thus a similar reinitiation mechanism for downstream translation.

Pseudoknot in HvV190S is located within the terminal 38 nucleotides identified as crucial sequence elements for RFP translation, from -10 nt to -32 nt (Fig. 4.8A). To test if it is the pseudoknot structure, in another words, the stem-loops, plays a role on RFP reinitiation other than its primary sequences, it was swapped with those predicted for different members in the genus of *Victorivirus* since different pseudoknots involve different primary sequences. Six pseudoknots were selected based on different features in stop-restart sites and they were from BfTV1, GaRV-L1, SsRV1, MoV1, TcV1 and CeRV1 (Fig. 4.8B). Among them, BfTV1 and GaRV-L1 were mostly similar to HvV190S, possessing the tetranucleotides AUGA, whereas the others were different, SsRV1 and CeRV1 possessed start codon coming before stop codon, and the left two MoV1 and TcV1 on the contrary. The replaced region includes the whole sequences spanning from each pseudoknot to each stop/restart site. Replaced pseudoknots were successfully re-predicted for all viruses with program Hpknotter after introduced into HvV190S. The reinitiation efficiency for various pseudoknots was achieved with fluorescence assay, in which it suggested that pseudoknots from BfTV1 and CeRV1 gave even higher level of RFP expression than HvV190S pseudoknot whereas other three from GaRV-L1, MoV1 and TcV1 gave a lower level (almost 50%). Although the reinitiation of RFP translation is quite tolerated to various pseudoknots, SsRV1 (11%) somehow failed to translate RFP efficiently (Fig. 4.8C).

In Fig.3.7 of chapter 3, SsRV2 was also classified into the genus of victorivirus based on the prediction of secondary structure, which was a long stem-loop structure, but formed at the similar position as a pseudoknot. So it was of interest to test if the long stem-loop may function similar to the pseudoknot in reinitiation of translation. A fragment containing SsRV2 long stem-loop primary sequence was inserted in frame into plasmid p190S at similar position to HvV190S pseudoknot (Fig. 4.9 top panel). After analyzing VLPs from transformants containing this swapped structure, Western blotting showed that the long stem-loop structure did not launch reintiation of RdRp translation (Fig. 4.9 bottom right) although it was still formed in HvV190S after swapping (Fig. 4.9 bottom left).

4.2.7 Confirmation of the role of pseudoknot structure on reinitiation of translation using two strategies

Although the sequence boundary was defined to be the upstream ORF 3' terminal 38

nucleotides that were predicted to include a similar secondary structure among victoriviruses, it might also involve other structures. To test if it is indeed the predicted pseudoknot structure in HvV190S that plays a critical role in RdRp reinitiation, two sets of experiments were carried out to test additional structural features, one set of experiments involved purified VLPs and the other was based on the fluorescence assay. Construction of all vectors used for purification of VLPs was based on previous vector p190S. Mutation was made at one strand of one stem at a time and also restored the structure by mutating two strands at another time. Constructs PK1 and PK2 involved mutating each strand of stem 1, and PK4 and PK5 involved mutating each strand of stem 2, thus the nts in stem 1 and stem 2 were unpaired in theses constructs (Fig. 4.10A top.). Stems in PK3 and PK6 were base-paired again by mutating both strands of each stem. The mutations were made to be complementary to the original sequences. All constructs were introduced into virus-free *H. victoriae* and transformants were subjected to VLP purification and Western blot analysis, as performed in Chapter 3. CP or RdRp-specific antisera were used for detection of the upstream and downstream translation products. It was expected that PK1, PK2, PK4 and PK5 do not express RdRp whereas PK3 and PK6 do if the pseudoknot structure is actually necessary for RdRp reinitiation. Results showed that only PK2 and PK3 gave an efficient expression of RdRp whereas others (PK1, PK4, PK5, PK6) expressed little or no RdRp, which contradicted expectation at first glance (Fig. 4.10A, bottom panel).

To determine whether pseudoknot structures in mutant constructs were actually disrupted or restored as expected, Hpknotter was employed to re-predict secondary structure. No pseudoknot structures were predicted for PK1, PK5 and PK6. On the contrary, a pseudoknot structure was predicted for construct PK2-4. Compared to results of Western blot analysis, only PK4 showed inconsistency between pseudoknot prediction and Western blot analysis. Compared to wild type pseudoknot in HvV190S, reformed pseudoknot predicted for PK2 and PK4 showed different nt composition for stem 1 (PK2) or stem 2 (PK4). To test whether a predicted alternative pseudoknot was indeed formed with construct PK2, a single nucleotide mutation was introduced into the second strand of the alternative stem 1 (U in GUG changed into A), thus disrupting base pairing in stem 1. Results of Western blot analysis using construct PK2-1 once again were consistent with structure prediction since no RdRp expression was detected using an RdRp specific antibody (Fig. 4.10B).

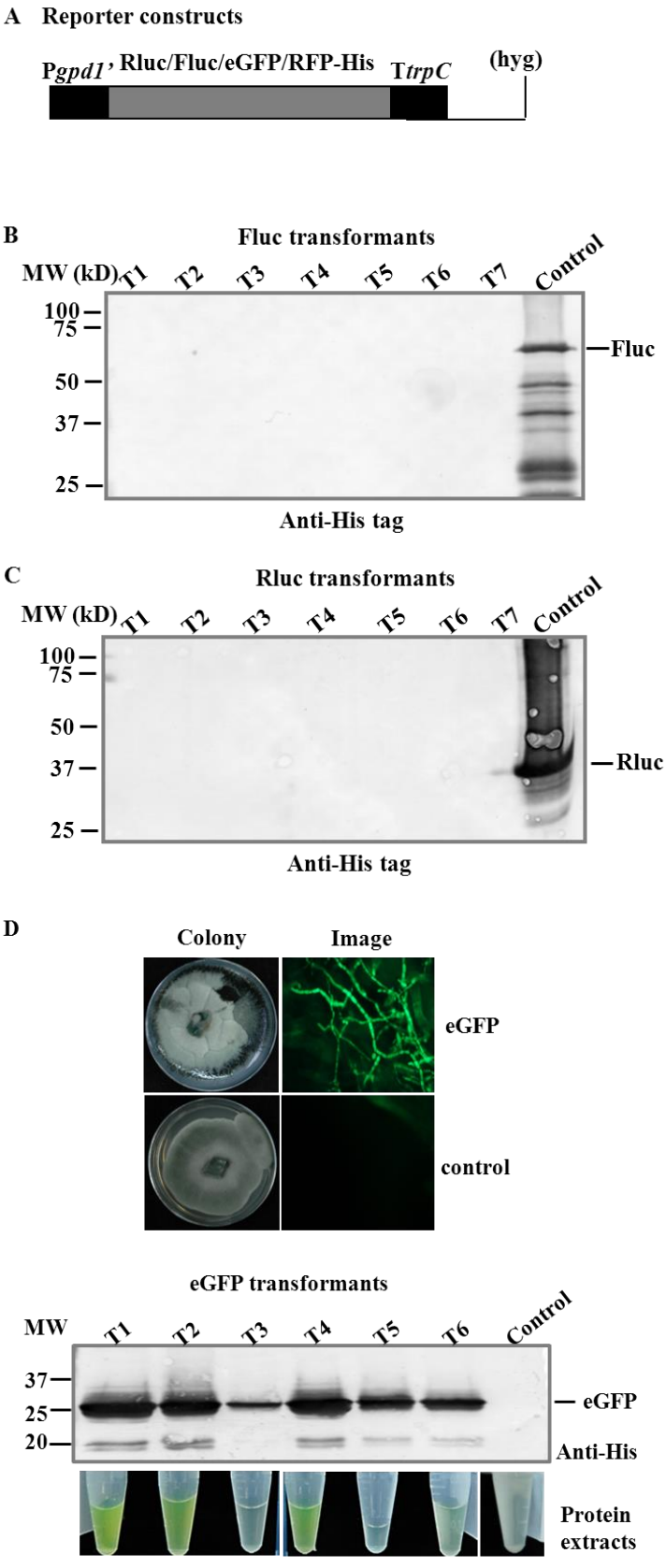
To reproduce the correlation between structure prediction and detected RdRp, translation subsequent to nucleotide mutation and also to justify the accuracy of the method employed for analyzing the RdRp expression from purified VLPs, a different set of mutations were made to both stem 1 and stem 2 in the same way but performed in the dual-fluorescence system. Mutations were made as shown in Fig. 4.10C and results for each construct from prediction are listed in the right column. Mutations in PK7-PK9 were related to stem 1 whereas PK10-PK12 on stem 2. PK7, PK8, PK10 and PK11 were mutated in one strand of each stem whereas PK9 and PK12 involved re-pairing the two strands of each stem. Results from fluorescence measurement suggested that PK9 and PK12 expressed RFP as efficiently as the wild type construct. PK 7-8 and PK 10-11 with a much higher MFE than wild type for formation of a pseudoknot structure did not express RFP efficiently.

4.2.8 Role of pseudoknot loops and flanking regions in RFP expression

In Fig. 4.6 and Fig. 4.7, the terminal thirty-eight nucleotides containing a predicted pseudoknot structure were determined to be crucial for RFP expression from the bicistronic genome. In Fig. 4.10, the detailed mutations introduced into stems of the pseudoknot suggested the determinant role of the predicted pseudoknot on RFP reinitiation. As I mentioned earlier the formation of pseudoknot comprises two stems, stem 1 and stem 2, and also two loops, loop 1 and loop 2. To investigate whether RNA sequences in both loops and sequences flanking both stem-loops play any role in RFP expression, four constructs were made to address this question. In construct PK_Loop1 and PK_Loop2, nucleotides were made complementary to the original nucleotides in loop 1 and loop 2 respectively. In PK_Flank1, six Cs flanking stem 1 were mutated to complementary Gs whereas nine nucleotides flanking stem 2 were mutated to their complementary ones. After mutations introduced into related sequences, pseudoknot was predicted and prediction results were listed on the right column with corresponding programs (Fig. 4.11 top panel). All constructs were introduced into dual-fluorescence system. Result showed that mutations made to sequences in loop 2, flanking sequence 1 and flanking sequence 2 did not affect the downstream RFP expression whereas mutation in loop 1 did (Fig. 4.11 bottom panel). Structure prediction suggest that when mutations were introduced into loop 1 sequence, they altered the formation of pseudoknot structure and thus abolished reinitiation of RFP translation.

4.2.9 Binding of viral sequences to complementary sequences in 18S rRNA

The binding between viral sequence and 18S rRNA sequence was tested and confirmed to play a role for downstream reinitiation in a few examples, including caliciviruses and influenza B virus (88, 92). The binding was found to occur within 80 nucleotides preceding the downstream ORF of those viruses. Although it was already identified that a pseudoknot play a definitive positive role in RFP reinitiation instead of binding to 18S rRNA, the possibility that the latter might also be involved in the reinitiation was not excluded (108). Approximately 100 nucleotides upstream of the AUGA motif of HvV190S were examined for potential base pairing with 18S rRNA from the fungus *C. heterostrophus* (Fig. 4.12A top panel). Since the sequence of the host fungus, *H. victoriae* is unavailable; the closely related *C. heterostrophus* was selected for the test. The 18S rRNA from *C. heterostrophus* was found to base pair with HvV190S sequences at four sites, binding to sequences around -20 nt, -40 nt, -65 nt and -90 nt upstream of the stop/restart site. Constructs were generated with mutations introduced that are complementary to original viral sequences at the corresponding sites and were transformed into fungal protoplasts for purification of VLPs. Western blot analysis of viral proteins from VLPs with antisera specific to CP and RdRp showed that only p18S-40M construct failed to express RdRp (Fig. 4.12A bottom panel). The expression of this construct was repeated using the dual-fluorescence system, again the 18S_40M gave a much lower RFP expression from the downstream ORF (Fig. 4.12B). Results from deletion experiments related to the 5' and 3' ends of CP sequence in Figs. 4 and 5 already showed that upstream sequences beyond -38 nucleotides did not alter reinitiation of translation of the downstream gene, ruling out that binding may occur close to -40 nt, -65 nt and -90 nt. Therefore, the result obtained with the 18S_40M following nucleotide mutations suggest that the unsuccessful expression of RdRp was due to the change in pseudoknot structure rather than any other possibility.


A Reporter constructs
Pgpd1, Rluc/Fluc/eGFP/RFP-His
TtrpC
(hyg)
B
Fluc transformants
MW (kD)
T1
T2
T3
T4
T5
T6
T7
Control
100
75
50
37
25
Fluc
Anti-His tag
C
Rluc transformants
MW (kD)
T1
T2
T3
T4
T5
T6
T7
Control
100
75
50
37
25
Rluc
Anti-His tag
D
Colony
Image
eGFP
control
eGFP transformants
MW
T1
T2
T3
T4
T5
T6
Control
37
25
20
eGFP
Anti-His
Protein extracts

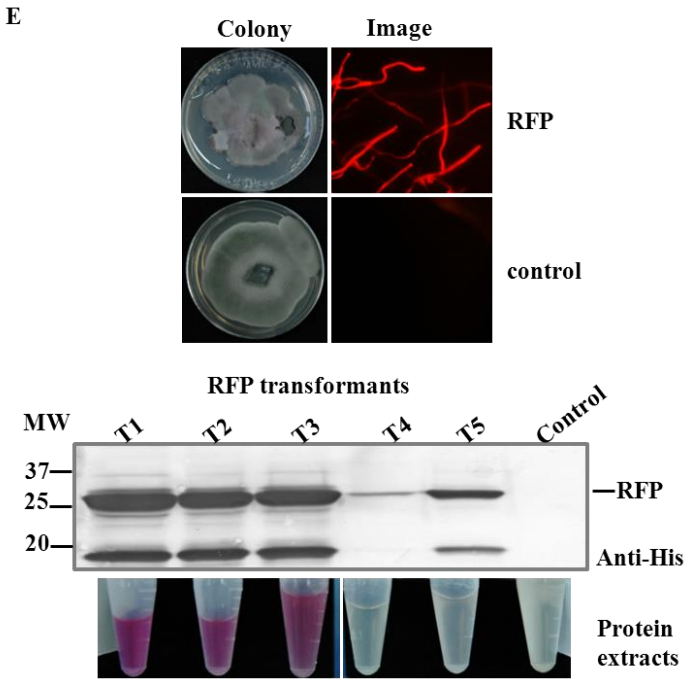
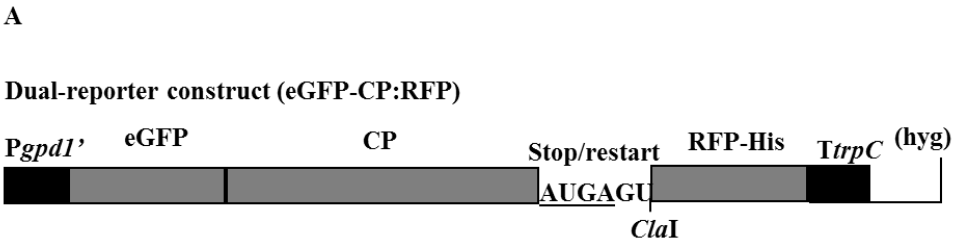


Figure 4.1 Expression of reporter genes in *H. victoriae.*

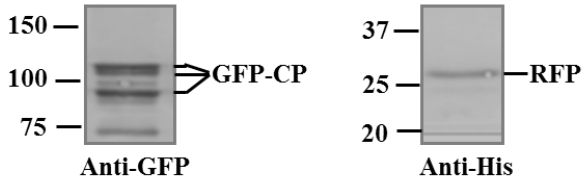
(A) Schematic representation of monocistronic reporter constructs to express two luciferase genes and two fluorescent protein genes in fungus *H. victoriae*. *Fluc*, Firefly luciferase; *Rluc*, Renilla luciferase; *eGFP*, enhanced green fluorescence gene; *RFP*: DsRed fluorescence gene. Reporter genes were inserted separately between fungal promoter and fungal terminator sequences, and a 6x His-tag sequence was fused to the 3'-terminus of each gene. Pgpd1', *gpd*1 promoter (nt 1-nt 707) from *Cochliobolus heterostrophus* (Acc.No. X63516); T*trpC*, His and *hyg* was similarly represented as in Fig. 3.1A. (B) and (C) Western blot analysis of total protein extracts from random-selected transformants (T1-T7) containing Fluc-His construct (B) and Rluc-His construct (C). Specific His-tag antiserum was used for detecting translational products. Controls for *in vivo* Fluc and Rluc expression are from bacterially expressed proteins in *E. coli* BL21 DE3 cells using plasmid pET-21a-Fluc and pET-21a-Rluc. Fluc was expressed from bacterial cells with an expected size of ~66 kDa whereas Rluc with an expected size of ~36 kDa. (D) Top, epifluorescence images of green fluorescence from the hyphae of transformants obtained from the transformation of fungal protoplasts with the eGFP construct. Virus-free *H. victoriae* isolate B-2 was used for preparing protoplasts for all transformation experiments. Empty vector transformants serve as negative controls

(Control). Bottom, Western blot analysis of total protein extracts from randomly-selected *eGFP* transformants (T1-T6). The eGFP protein was detected with a His-tag specific antiserum. The positions of protein size marker bands are indicated on the left. Photographs of protein extracts from different transformants were taken under laboratory light conditions and listed below. Western blot analysis was carried out by SDS-PAGE on 10% polyacrylamide gels. (E) Similar experiments were carried out as in (B), but with the *RFP* construct. Red fluorescence and RFP protein from randomly selected transformants, T1-T5, were examined.

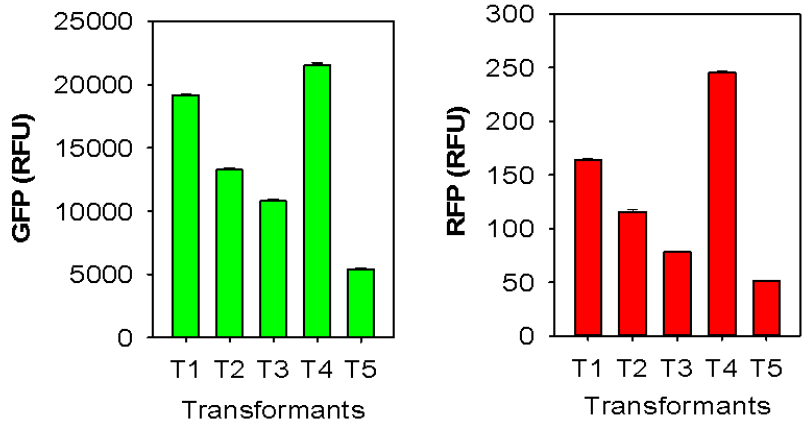
A


Dual-reporter construct (eGFP-CP:RFP)
Pgpd1'
eGFP
CP
Stop/restart
AUGAGU
ClaI
RFP-His
TtrpC
(hyg)

B


150
100
75
GFP-CP
Anti-GFP
37
25
20
RFP
Anti-His

C


GFP (RFU)
25000
20000
15000
10000
5000
0
T1 T2 T3 T4 T5
Transformants
RFP (RFU)
300
250
200
150
100
50
0
T1 T2 T3 T4 T5
Transformants

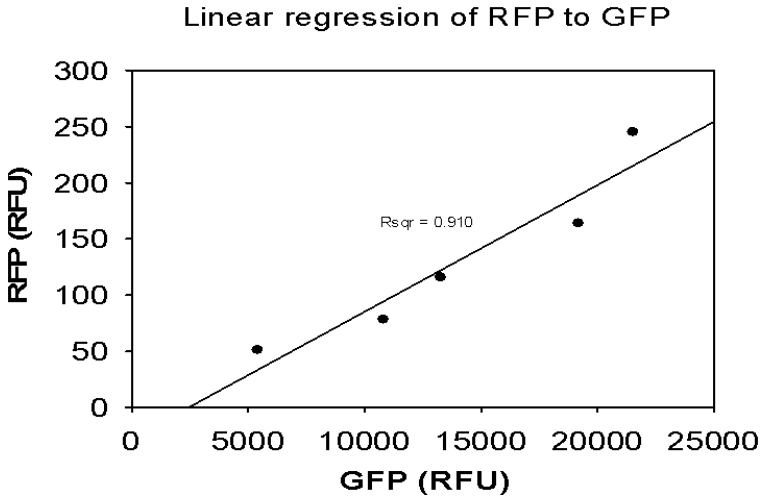

Linear regression of RFP to GFP
RFP (RFU)
300
250
200
150
100
50
0
Rsqr = 0.910
0
5000
10000
15000
20000
25000
GFP (RFU)

Figure 4.2 Establishment of a dual-fluorescence system and expression of reporter genes from the dual-reporter construct.

(A) Schematic representation of the dual-reporter construct eGFP-CP:RFP. This bicistronic plasmid contains two ORFs, with the 5'-proximal ORF encoding eGFP fused to full-length CP ORF of HvV190S (eGFP-CP) and the 3'-proximal ORF encoding RFP ORF with a His-tag sequence fused to its 3' end (RFP-His). The start codon of RFP-His ORF overlaps the stop codon of eGFP-CP ORF in the stop/restart site AUGA, indicated in boldface and underlined. In RFP ORF, the first two codons AUGAGU are derived from RdRp ORF followed by the unique *Cla*I restriction site sequence and complete RFP ORF. (B) Western blot analysis of transformants obtained from fungal transformation with plasmid eGFP-CP:RFP. Antiserum to eGFP and to His-tag were used to detect the upstream eGFP-CP fusion protein and the downstream RFP protein. The positions of eGFP and RFP proteins are indicated on the right, and those of the protein size markers on the left. Ten times more sample was loaded into SDS-PAGE gel for detecting RFP than eGFP-CP, and this is applicable to all Western blot analysis pertinent to bicistronic constructs. (C) Fluorescence characterization of eGFP and RFP expressed from the bicistronic plasmid eGFP-CP:RFP. Total protein extracts from five transformants (T1-T5) were subjected to fluorescence measurement in a microplate reader. Green fluorophore readouts (Left, upper panel) and red fluorophore readouts (right, upper panel) from each individual transfromants were given as relative fluorescence units (RFU). Error bars indicate standard error. The linearization relationship of each RFP readout to the corresponding GFP readout from five transfromants was analyzed using the scatter plot (Bottom).

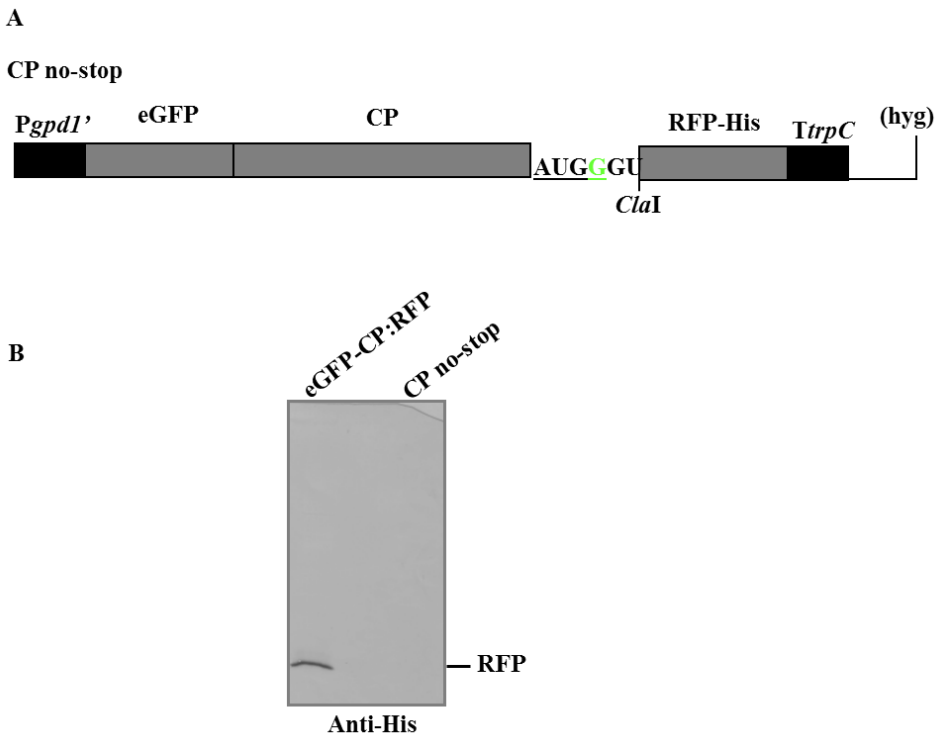
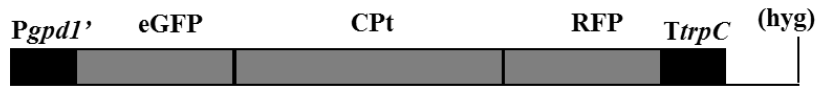


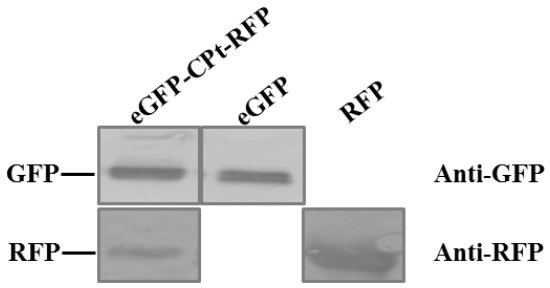
Figure 4.3 Effect of mutating CP stop codon in AUGA motif on expression of the downstream gene.

(A) Schematic representation of the plasmid "CP no-stop" in which the stop codon of fused eGFP-CP ORF was mutated from UGA to UGG (mutated nucleotide in green). In this construct, the stop/restart signal was uncoupled. After transformed into *H. victoriae*, the translational product from the downstream ORF was analyzed using Western blotting and a His-tag specific antiserum (B).

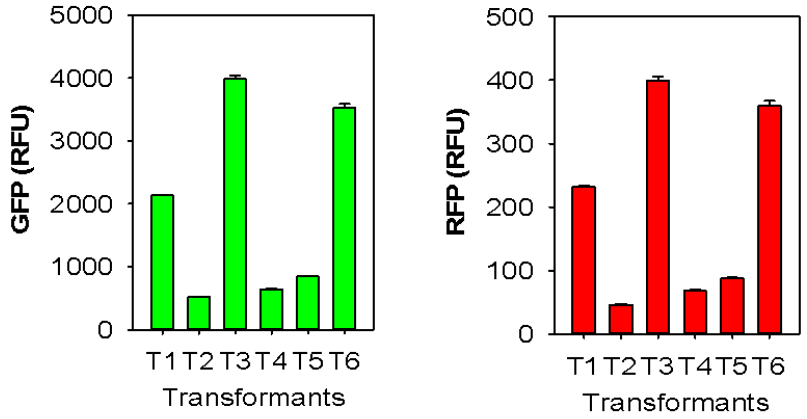
A

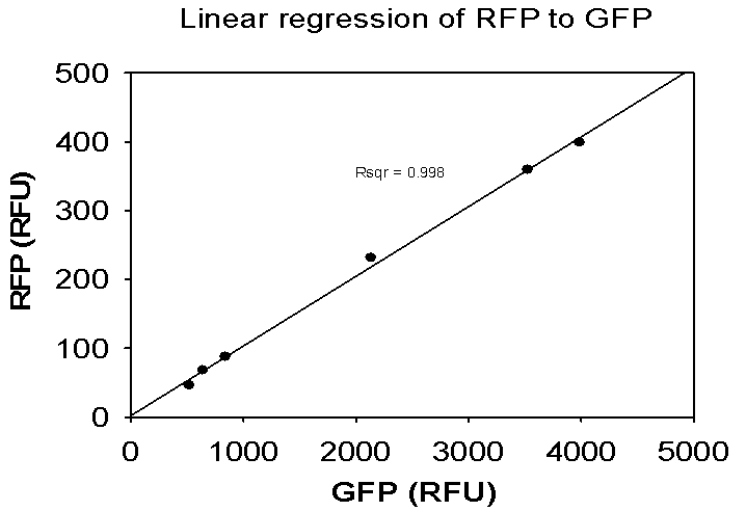

Dual-reporter-CP fusion construct (eGFP-CPt-RFP)
Pgpd1'
eGFP
CPt
RFP
TtrpC
(hyg)

B


eGFP-CPt-RFP
eGFP
RFP
GFP
RFP
Anti-GFP
Anti-RFP

C


GFP (RFU)
5000
4000
3000
2000
1000
0
T1 T2 T3 T4 T5 T6
Transformants
RFP (RFU)
500
400
300
200
100
0
T1 T2 T3 T4 T5 T6
Transformants


Linear regression of RFP to GFP
Rsqr = 0.998
RFP (RFU)
500
400
300
200
100
0
0
1000
2000
3000
4000
5000
GFP (RFU)

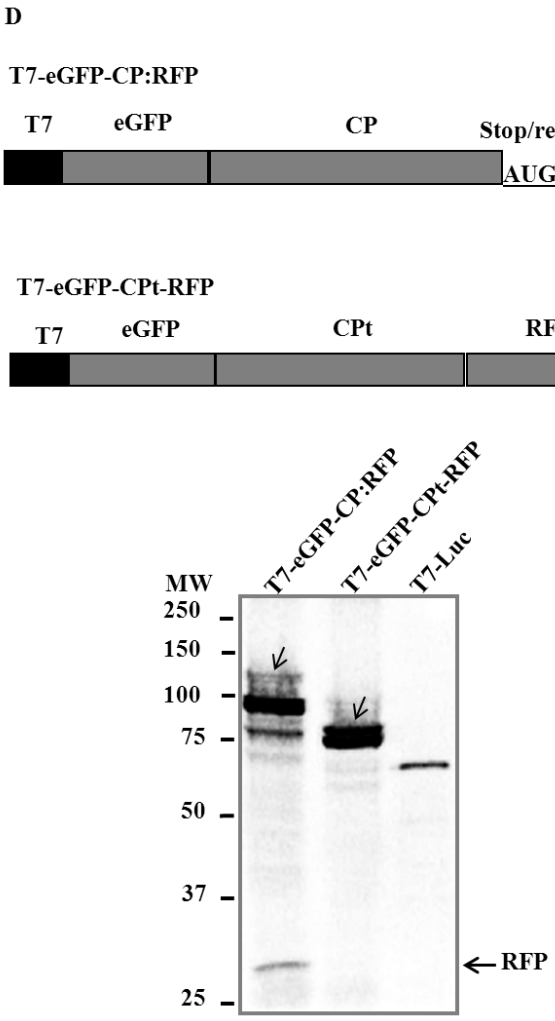


Figure 4.4 Construction of a monocistronic plasmid containing two reporter genes and its expression in *H. victoriae*.

(A) Schematic diagram representing the structure of eGFP-CPt-RFP plasmid, in which both eGFP and RFP reporters were fused to the CP ORF at its 5'- or 3'- termini, respectively, thus generating a monocistronic vector of fused three genes (fused 3-gene construct). The C-terminal region (~ 900 nt) of CP ORF was truncated to remove its putative proteolytic sites and designated as CPt. (B) Western blot analysis of transformants with the monocistronic plasmid shown in A. Specific antisera to GFP and RFP were used to detect the translational products. Transformants from the monocistronic constructs of a single reporter genes, as shown in Fig. 4.1, served as positive controls for eGFP and RFP, respectively. (C) Fluorescence measurement of eGFP and RFP expressed from the fused 3-gene plasmid. Experiments were carried out

similarly to that in Fig. 4.2C. Six transformants (T1-T6) obtained were used to measure their green and red fluorescence and to plot the linear regression graph. Error bars indicate standard error. (D) *In vitro* translation of reporter genes from bi- and mono-cistronic recombinant plasmids. Top, Schematic representation of the basic structures of two plasmids for *in vitro* translation, T7-eGFP-CP:RFP and T7-eGFP-CPt-RFP. The construction of the two plasmids is based on pGEM-T easy vector from Invitrogen. Their genomic organizations are same as those in Fig.4.2A and Fig. 4.4A except that different promoters and backbone vectors were used. T7, T7 polymerase promoter sequence. Bottom, The autoradiograph of *in vitro* translational products labeled with ^{35}S-methionine and fractionated on 10% polyacrylamide gel using SDS-PAGE. Luc gene (luciferase, 66 kDa) supplied with the kit serves as a control for translational product. The positions of different proteins expressed from *in vitro* translation for each construct were indicated with arrows. The positions of protein size marker bands are indicated on the left.

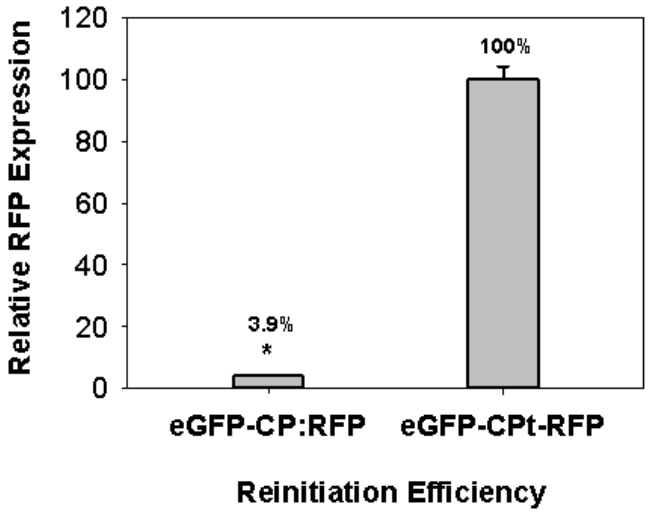
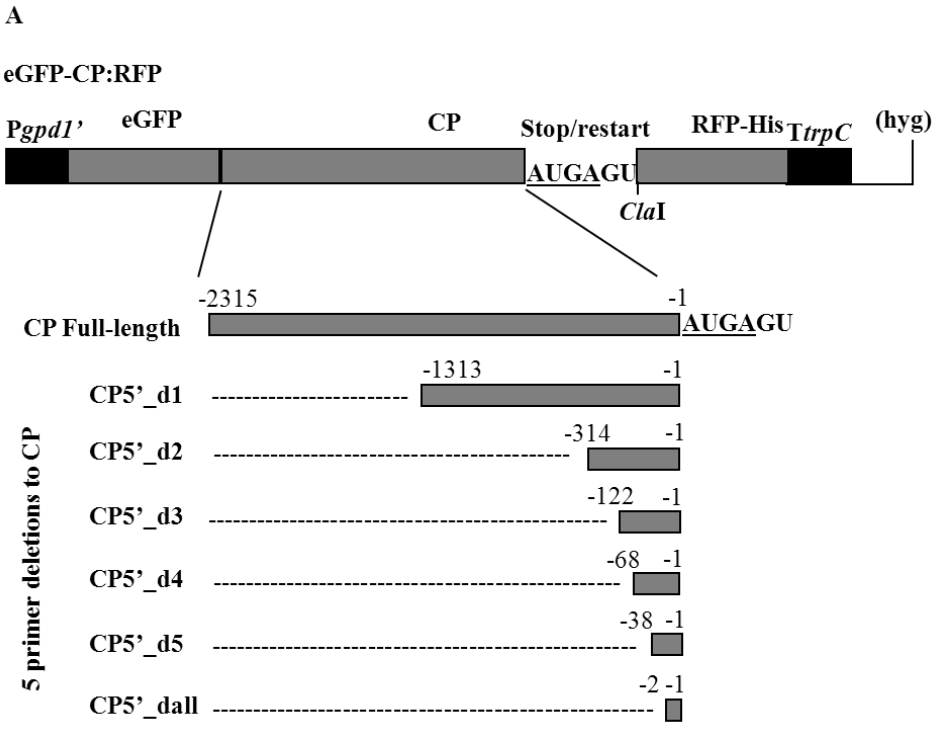


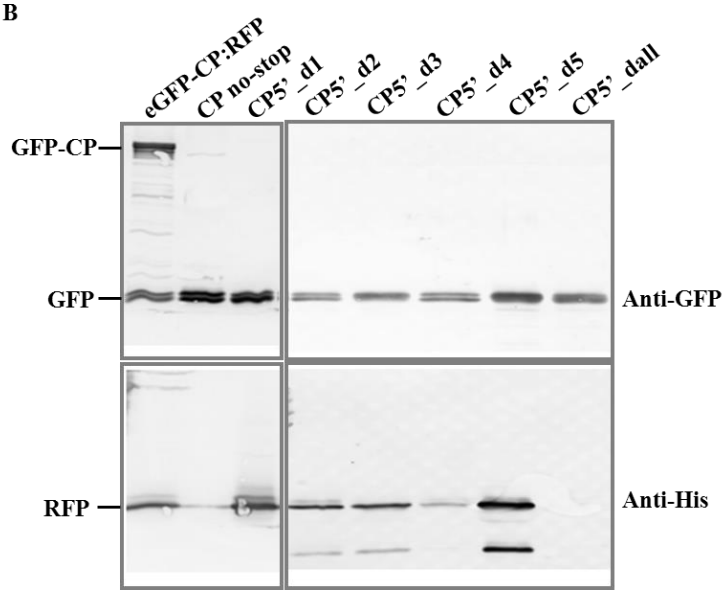
Figure 4.5 Determination of reinitiation efficiency of the downstream ORF in HvV190S.

Determination of reinitiation efficiency of RFP translation by comparing the monocistronic plasmid, eGFP-CPt-RFP (without stop/restart motif) and the bicistronic plasmid, eGFP-CP:RFP (with stop/restart signal). The analysis was based on the fluorescence readouts from five transformants used in Fig. 4.2C and six transformants used in Fig. 4.4C. The mean ratio of RFP to GFP from eGFP-CP:RFP was calculated and compared to the mean ratio of RFP to GFP of eGFP-CPt-RFP. Results show the mean and standard deviations for four replicates of independent protein extracts for each transformant.

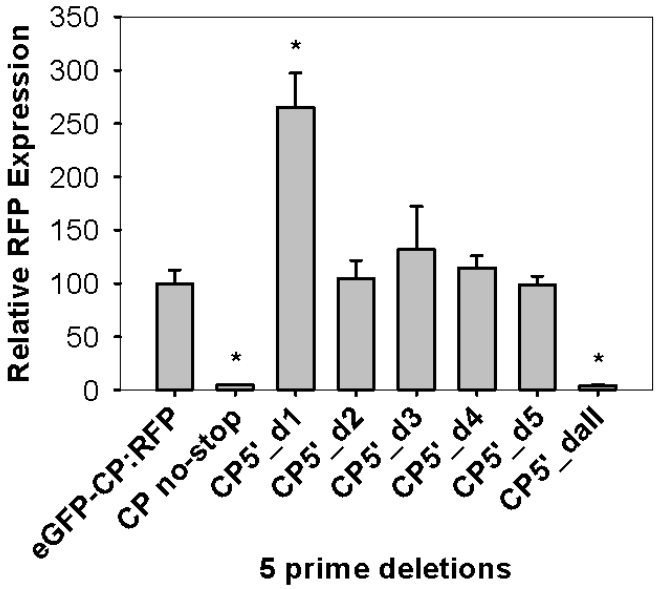
A


eGFP-CP:RFP
Pgpd1'
eGFP
CP
Stop/restart
RFP-His
TtrpC
(hyg)
AUGAGU
ClaI
-2315
-1
CP Full-length
AUGAGU
5 primer deletions to CP
-1313
CP5'_d1
-314
CP5'_d2
-122
CP5'_d3
-68
CP5'_d4
-38
CP5'_d5
-2
CP5'_dall

B


eGFP-CP:RFP
CP no-stop
CP5'_d1
CP5'_d2
CP5'_d3
CP5'_d4
CP5'_d5
CP5'_dall
GFP-CP
GFP
Anti-GFP
RFP
Anti-His

C



* p < 0.01 vs eGFP-CP:RFP

D

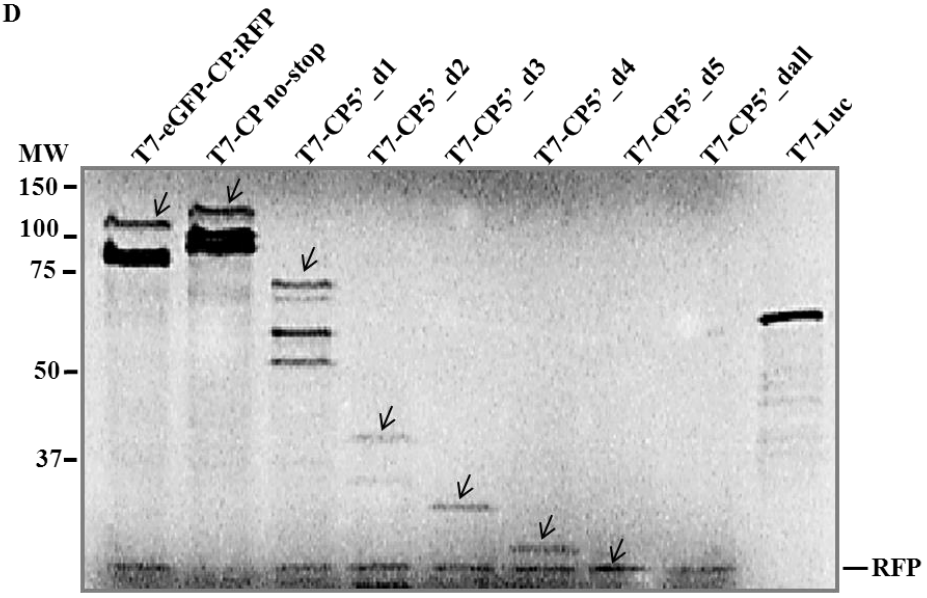
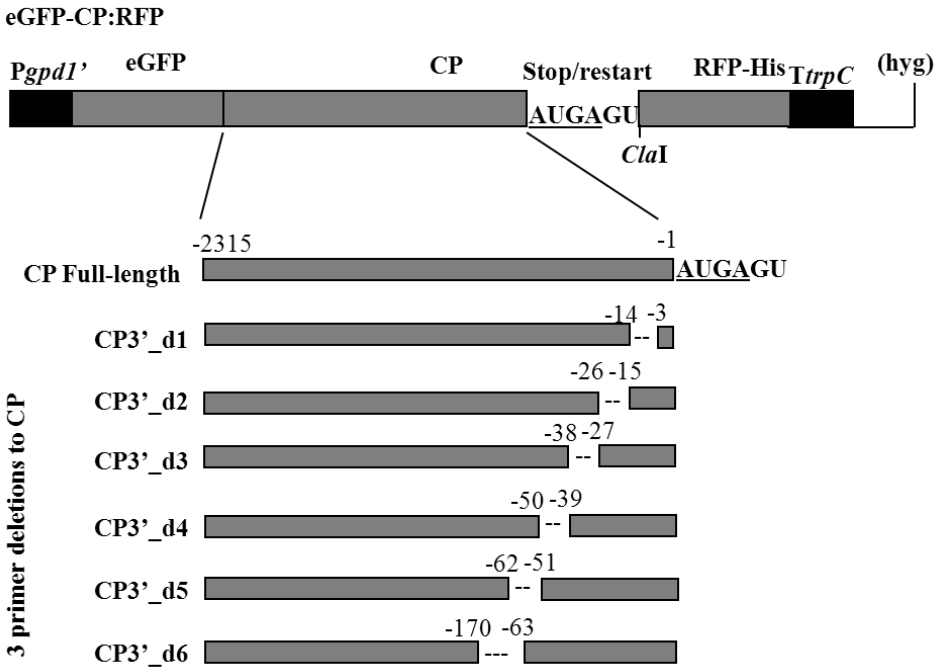


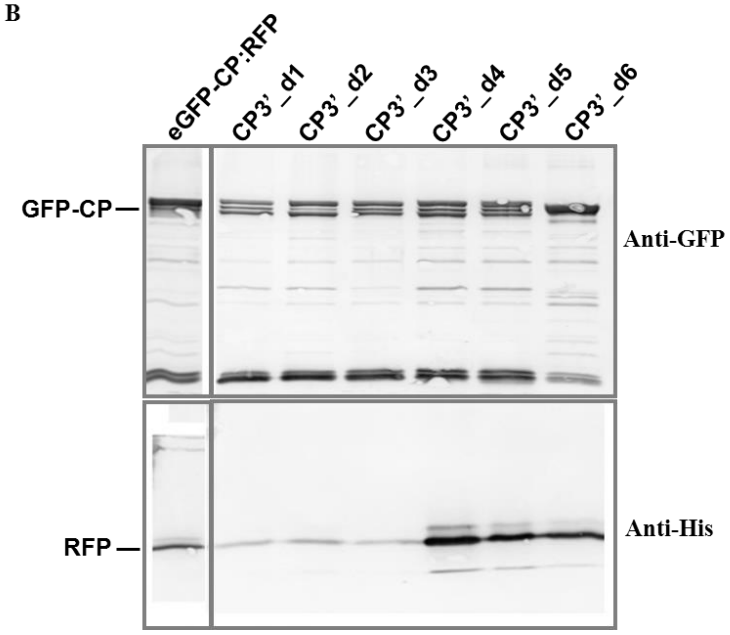
Figure 4.6 Definition of the upstream sequence boundary for RFP reinitiation by introducing a series of truncation into to the 5' end of the CP ORF.

(A) Schematic diagram representing the different truncations introduced into the upstream CP ORF to generate six constructs based on the dual-reporter plasmid eGFP-CP:RFP. The truncated region was indicated by a dashed line for each construct and the remainder CP ORF was indicated with numbers at ends relative to the first A in the stop/restart site, AUGA. The full-length CP ORF is drawn at top of mutant diagrams and its start codon is located at -2315 upstream of the start/stop site. (B) Western blot analysis to demonstrate the influence of truncations introduced into the 5' end of CP ORF on RFP expression. The positions of different products detected with antisera to GFP or His-tag are indicated on the left. CP no-stop construct was included as a negative control for RFP expression. (C) At least five transformants from each construct were measured for green and red fluorescence. The mean ratio of RFP to GFP from different transformants for each construct was divided by that from the full-length construct. The rate was converted into percentage for this experiment. Asterisks indicate a statistically significant difference between the input two constructs (original construct and truncated construct) based on t-test ($P < 0.01$). (D) *In vitro* translation of mutant bicistronic plasmids with introduced truncation to the 5' of the CP ORF. The organization of mutant constructs for *in vitro* translation were the same as in Fig. 4.6A. The different constructs made and analysis of their translational products were similar to those in Fig. 4.4D. *Luc* gene (luciferase, 66 kDa) was supplied with the kit and served as a control for translational product. The positions of different proteins from *in vitro* translation are shown with arrows. The positions of protein size marker bands are indicated on the left.

A


eGFP-CP:RFP
Pgpd1'
eGFP
CP
Stop/restart
RFP-His
TtrpC
(hyg)
AUGAGU
ClaI
-2315
-1
CP Full-length
AUGAGU
3 primer deletions to CP
CP3'_d1
-14 -3
CP3'_d2
-26 -15
CP3'_d3
-38 -27
CP3'_d4
-50 -39
CP3'_d5
-62 -51
CP3'_d6
-170 -63

B


eGFP-CP:RFP
CP3'_d1
CP3'_d2
CP3'_d3
CP3'_d4
CP3'_d5
CP3'_d6
GFP-CP
Anti-GFP
RFP
Anti-His

C

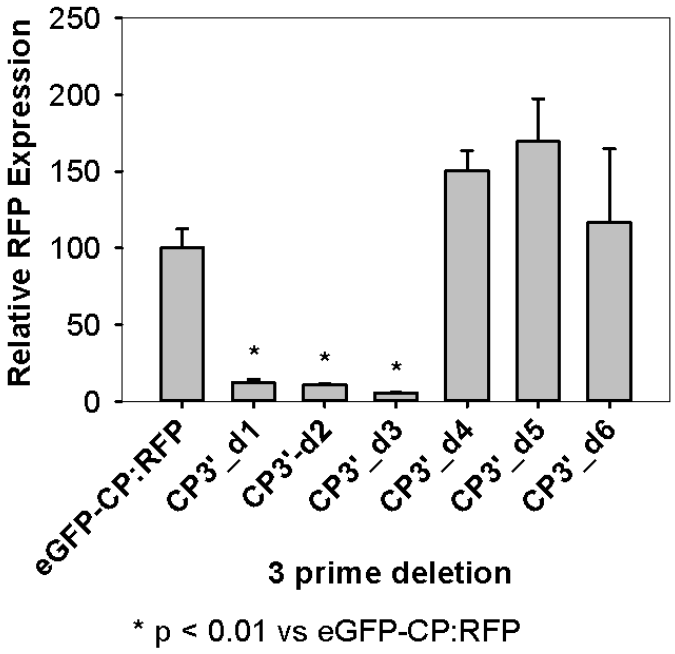
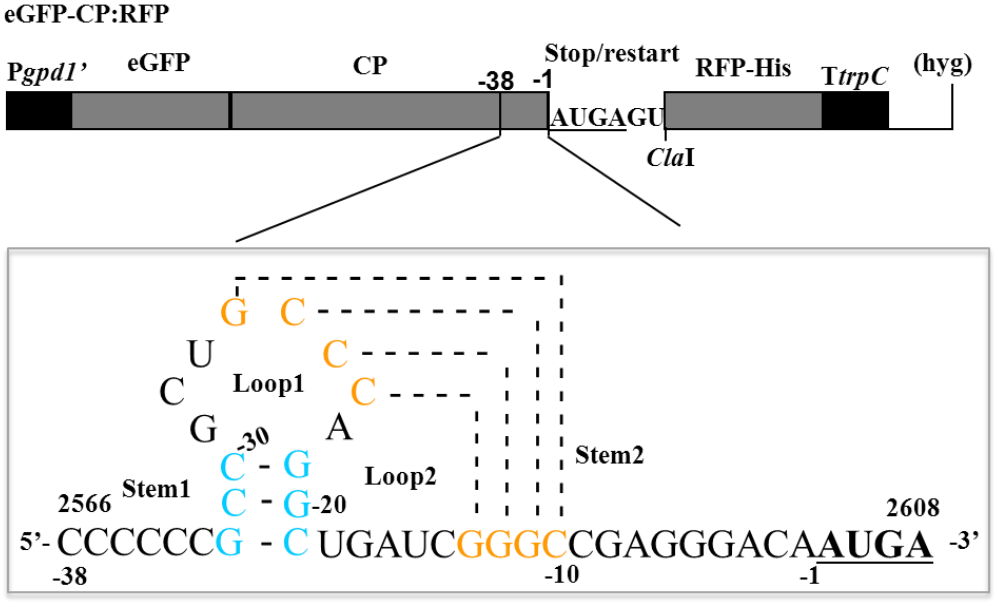


Figure 4.7 Further confirmation of upstream sequence boundary for RFP reinitiation by introducing nucleotide deletions to the 3' end of CP ORF.

(A) Schematic diagram represents 12 nt-deletions (except CP3'_d6 with a larger deletion) were introduced into the 3' end of CP ORF at different positions relative to the stop/restart site based on the dual-reporter vector eGFP-CP:RFP. The deleted parts for each construct were indicated with dashed lines with both ends indicated with numbers relative to the first A in stop/restart site AUGA. The full-length CP ORF is shown at the top of the representation of mutant constructs similar to Fig. 4.6A. (B) Western blot analysis was performed similarly to Fig. 4.6B. (C) Fluorescence analysis was done in the same manner as in Fig. 4.6C.

A



B

Virus	Sequence	Class	MFE
HvV190S	GCCGCUGCCCAGGCUGAUCGGGCCGAGGGACA**AUGA**GU	1	-8.1
BfTV1	CCUGCUGCCCCAGGUAACGGGCCGGACACGAU**AUGA**GU	1	-7.9
CeRV1	GUGGCUGCCCCCACUAAUGAGGGCCGAAACG**AUG**UC**UAG**A	1	-8.3
GaRV-L1	GAAGCUACUGCUGCUGAAGCAGUGCCCGCUCA**AUGA**GU	1	-11.5
MoV1	GGCGCGGCGCCUAGCCUGCACGAA**UAG**AU**AUG**AGU	1	-5.8
SsRV1	GGCCCCGCCCAGCCUGACGGGCCGCCA**AUG**AA**UAA**GAGU	1	-7.8
TcV1	GCCGACCCCCAGGCUAUGGUGGGCGACCAGAUC**UAAAUG**AGU	1	-9.2

C

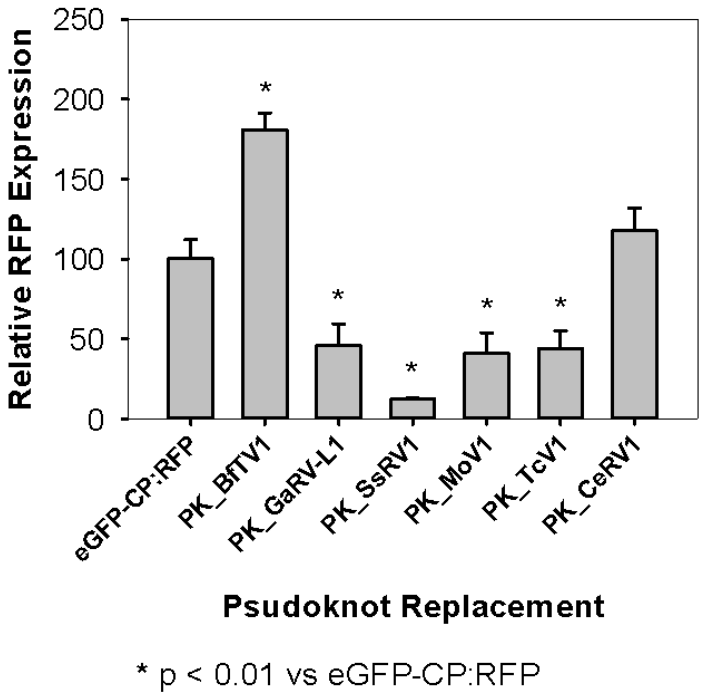


Figure 4.8 Effect of swapping HvV190S pseudoknot with that from other victoriviruses on RFP expression from the downstream ORF.

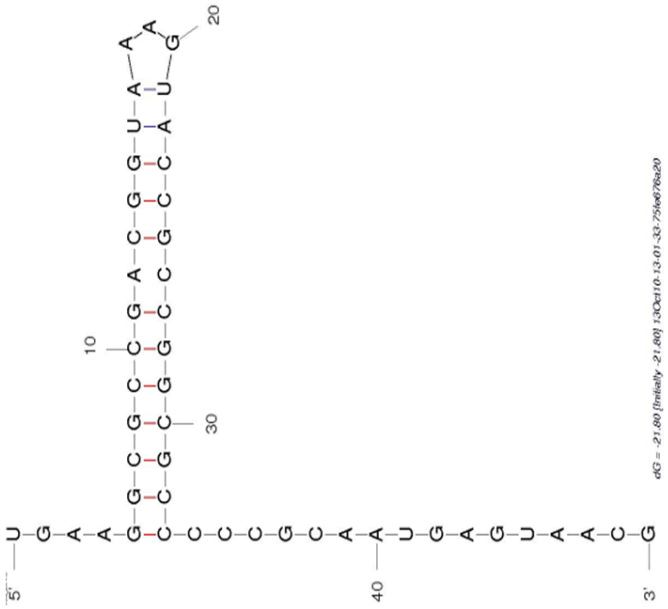
(A) Diagram representing a predicted H-type RNA pseudoknot structure within the terminal 38 nucleotides of CP ORF. This pseudoknot structure consists of stem 1 (cyan) and stem 2 (orange) (indicated with dashes lines) and two resultant loops (loop 1 and loop 2). The nucleotide positions relative to stop/restart sites are indicated. (B) List of victoriviruses selected for the pseudoknot swapping with that predicted in HvV190S. Pseudoknot swapping was based on the dual-fluorescence plasmid eGFP-CP:RFP. The swapped part includes the whole sequence spanning each pseudoknot to the stop/restart site (the entire sequences listed here for each virus). The second codon AGU from RdRp ORF (in green) was added downstream the stop/restart sites if it is lacking in the sequence of viruses followed by the *Cla*I restriction site sequence and RFP ORF. Nucleotides involved in each stem of the pseudoknots are indicated in the same way as for HvV190S for the color-coded stem 1 and stem 2. Stop/restart site for each is in boldface and underlined. Abbreviations of virus names are listed in Fig. 3.7 of Chapter 3. Pseudoknot swapped into HvV190S was predicted with HPknotter and results were shown in the right column together with corresponding MFE. (C) Five

transformants from each construct were selected to measure the green and red fluorescence. The mean ratio of RFP to GFP from different transformants for each construct was compared to that from wild type pseudoknot-containing construct.

A

HvV190S CCCCCC (GCCGCUGCCCAGGCUGAUCGGGC) CGAGGGACA**AUGA**GU

SsRV2 GAA (GGCGCCGACGGUAAAGUACCGCCGGCGCC) CCCGCA**AUG**AG**UAA**C



Long stem-loop predicted for HvV190S after swapping

B

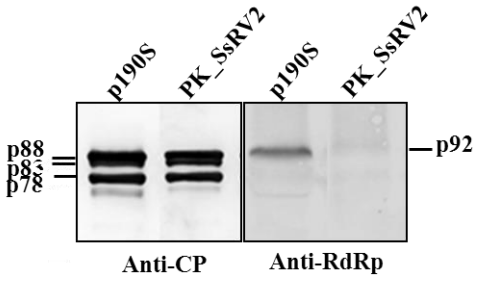
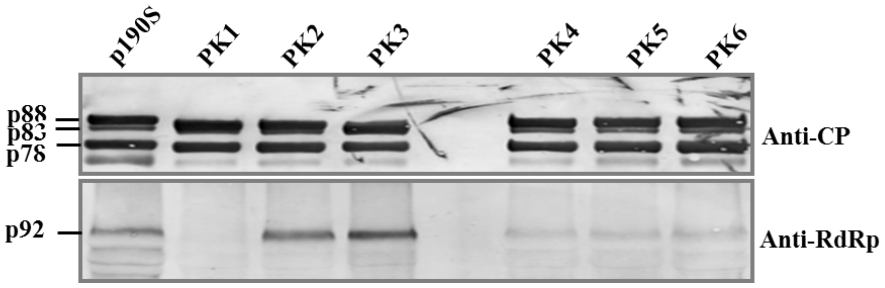


Figure 4.9 Effect of replacing the predicted pseudoknot in HvV190S with the long stem-loop predicted in SsRV2 on RdRp reinitiation.

Long stem-loop in SsRV2 alone (in bracket) was inserted into the similar position for HvV190S pseudoknot. This long stem-loop was also indicated with color for its nucleotides similarly as HvV190S pseudoknot as in Fig. 4.8A, in which nucleotides in cyan represent the lower part of the long stem-loop whereas those in orange represent the upper part. The re-predicted stem-loop after swapped into HvV190S was indicated on the bottom. (B) Western blot analysis of VLPs containing construct with swapped stem-loop structure in HvV190S with CP- and RdRp- antiserum.

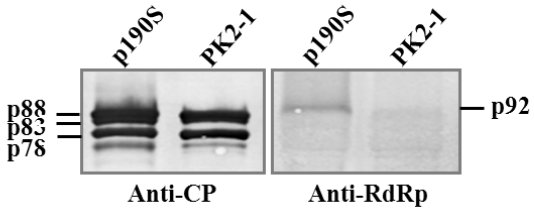
A

Construct	Sequence	Class	MFE
p190S	CCCCCCGCCGCUGCCCAGGCUGAUCGGGCCGAGGGACAAUGAGU	1	-8.1
PK1	CCCCCCCGGGCUGCCCAGGCUGAUCGGGCCGAGGGACAAUGAGU	NO	NO
PK2	CCCCCCGCCGCUGCCCACCGUGAUCGGGCCGAGGGACAAUGAGU	1	-7.4
PK3	CCCCCCCGGGCUGCCCACCGUGAUCGGGCCGAGGGACAAUGAGU	1	-9.5
PK4	CCCCCCGCCGCUCGGGAGGCUGAUCGGGCCGAGGGACAAUGAGU	1	-9.2
PK5	CCCCCCGCCGCUGCCCAGGCUGAUCCCCGCGAGGGACAAUGAGU	NO	NO
PK6	CCCCCCGCCGCUCGGGAGGCUGAUCCCCGCGAGGGACAAUGAGU	NO	NO



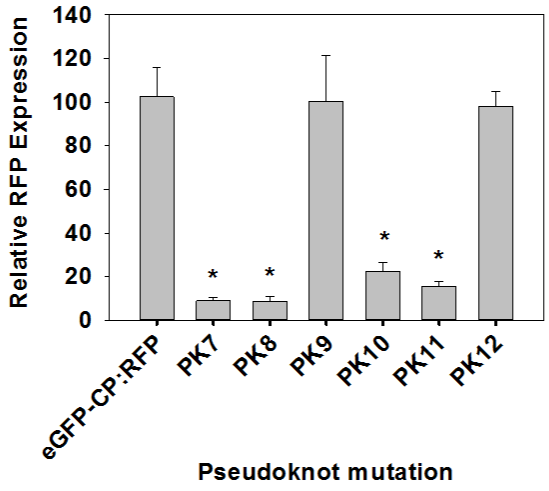
B

Construct	Sequence	Class	MFE
PK2	CCCCCCGCCGCUGCCCACCGUGAUCGGGCCGAGGGACAAUGAGU	1	-7.4
PK2-1	CCCCCCGCCGCUGCCCACCGAGAUCGGGCCGAGGGACAAUGAGU	NO	NO



C

Construct	Sequence	Class	MFE
eGFP-CP:RFP	CCCCCCGCCGCUGCCCAGGCUGAUCGGGCCGAGGGACAAUGAGU	1	-8.1
PK7	CCCCCCGGCGCUGCCCAGGCUGAUCGGGCCGAGGGACAAUGAGU	1	-4.6
PK8	CCCCCCGCCGCUGCCCAGCCUGAUCGGGCCGAGGGACAAUGAGU	1	-4.9
PK9	CCCCCCGGCGCUGCCCAGCCUGAUCGGGCCGAGGGACAAUGAGU	1	-10.6
PK10	CCCCCCGCCGCUGCGCAGGCUGAUCGGGCCGAGGGACAAUGAGU	1	-3.9
PK11	CCCCCCGCCGCUGCCCAGGCUGAUCGCGCCGAGGGACAAUGAGU	NO	NO
PK12	CCCCCCGCCGCUGCGCAGGCUGAUCGCGCCGAGGGACAAUGAGU	1	-9.8



* p < 0.01 vs eGFP-CP:RFP

D

Construct	Sequence													
WT	P	P	A	A	A	Q	A	D	R	A	E	G	Q	*
PK1	P	P	R	A	A	Q	A	D	R	A	E	G	Q	*
PK2	P	P	A	A	A	H	R	D	R	A	E	G	Q	*
PK2-1	P	P	A	A	A	H	R	D	R	A	E	G	Q	*
PK3	P	P	R	A	A	H	R	D	R	A	E	G	Q	*
PK4	P	P	A	A	R	E	A	D	R	A	E	G	Q	*
PK5	P	P	A	A	A	Q	A	D	P	R	E	G	Q	*
PK6	P	P	A	A	R	E	A	D	P	R	E	G	Q	*
PK7	P	P	G	A	A	Q	A	D	R	A	E	G	Q	*
PK8	P	P	A	A	A	Q	P	D	R	A	E	G	Q	*
PK9	P	P	G	A	A	Q	P	D	R	A	E	G	Q	*
PK10	P	P	A	A	A	Q	A	D	R	A	E	G	Q	*
PK11	P	P	A	A	A	Q	A	D	R	A	E	G	Q	*
PK12	P	P	A	A	A	Q	A	D	R	A	E	G	Q	*

Figure 4.10 Influence of nucleotide mutations in stem 1 and stem 2 of predicted pseudoknot on RFP expression.

(A) Top panel, List of mutagenizing constructs containing nucleotide mutations complementary to the original sequences of both stems. Mutations (in green colored) were introduced to one strand (for PK1, PK2, PK4 and PK5) or to both strands (for PK3 and PK6). Wild type pseudoknot in HvV190S is listed at the top of mutant constructs and colored similarly as previously described. Sequences were predicted for pseudoknot formation subsequent to mutation using HPknotter and the results are listed in the right column together with corresponding MFE. The reformed pseudoknot in mutant constructs were underlined. The construction of all mutant plasmids was based on plasmid p190S in Chapter 3. Bottom panel, Western blot analysis of viral proteins expressed from different constructs with CP- or RdRp-specific antisera. Virus-like particles (VLPs) were purified from transformants with the original plasmid p190S and p190S-derived mutant constructs. The positions of different proteins from VLPs are shown on the left. (B) Top panel, Sequence of a construct with a nt mutation (in underlined green) introduced into the alternative stem1 formed in PK2 in (A). Bottom panel, Western blot analysis of transformants containing PK2-1 construct. (C) Top panel, similar to (A), different combination of mutations were introduced into two stems of predicted pseudoknot structure, but construction of all mutant plasmids was based on

the dual-reporter plasmid eGFP-CP:RFP. Bottom panel, fluorescence intensity measurement with transformants containing constructs shown at the top. (D) Amino acid analysis in CP reading frame for sequences immediately upstream of stop/restart sites for all constructs in (A), (B) and (C). "*****" represented stop codon signal for translation of CP ORF. Amino acids were colored similarly to the primary sequences for stem1 (cyan) and stem2 (orange) of the wild type. Mutated amino acids encoded by mutant constructs were indicated in green.

Construct	sequence	Class	MFE
eGFP-CP:RFP	CCCCCCGCCGCUGCCCAGGCUGAUCGGGCCGAGGGACA**AUGA**GU	1	-8.1
PK_Loop1	CCCCCCGCCCGAGCCCUGGCUGAUCGGGCCGAGGGACA**AUGA**GU	**NO**	**NO**
PK_Loop2	CCCCCCGCCGCUGCCCAGGCACUAGGGGCCGAGGGACA**AUGA**GU	1	-10.8
PK_Flank1	GGGGGGGCCGCUGCCCAGGC**U**GAUCGGGCCGAGGGACA**AUGA**GU	1	-12.5
PK_Flank2	CCCCCCGCCGCUGCCCAGGCUGAUCGGGCGCUCCCUGU**AUGA**GU	1	-12.4

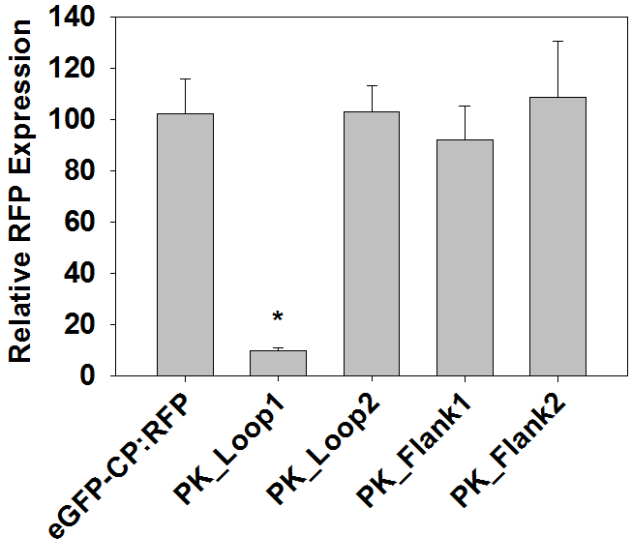
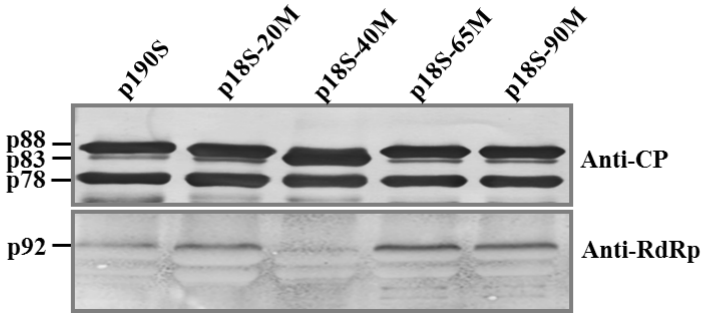


Figure 4.11 Role of pseudoknot loops and flanking region sequences on RFP expression.

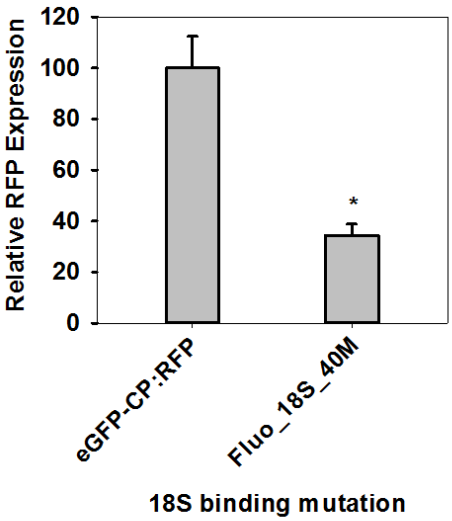
Top, Sequences of a series of mutagenizing constructs that introduce mutations into sequences in predicted loops and flanking regions relevant to pseudoknot structure in HvV190S are listed. Original sequence of construct eGFP-CP:RFP is listed at top. Mutations were introduced separately to all nucleotides in loops1 and 2 and also those comprising two regions bordering the predicted pseudoknot shown in Fig.4.8A. Mutated nucleotides are complementary to the original ones and are printed in green. Pseudoknots were predicted for mutant constructs and the results are shown in the right column with corresponding MFE. Bottom, fluorescence intensity measurement with transformants containing constructs shown at the top of the figure.

A

	Location in 18S rRNA	18S rRNA sequence (5'→3')	Corresponding viral sequence (5'→3')	Location relative to stop/restart	Class	MFE
18S_20M	977-983	CGAUCAG	CUGAUCG	-19 to -13	1	-10.8
18S_40M	1146-1152	GCGUGGA	UCCACGC	-46 to -40	NO	NO
18S_65M	854-865	CGGGGGCAUCAG	CUGCUGCCCCCG	-70 to -59	1	-10.6
18S_90M	171-178	UCGGAAGG	CCUUCCGA	-95 to -88	1	-10.6



B



* p < 0.01 vs eGFP-CP:RFP

Figure 4.12 Potential binding of HvV190S sequences to complementary sequences in 18S rRNA.

(A) Top, List of possible binding sites found between 18S rRNA of *C. heterostrophus* And ~100 nucleotides stretch upstream of stop/restart site of HvV190S. The 18S rRNA relevant nucleotides involved in putative binding to HvV190s sequence and their positions in 18S rRNA are listed. The corresponding viral sequences for binding and their positions are also indicated, which is relative to the stop/restart site with a "-"sign indicating positions upstream of this signal site. Predicted viral sequences are completely or partially mutated to their complementary sequences. Results from pseudoknot prediction for each mutant construct was shown at right with corresponding MFE. GenBank accession number for 18S rRNA is AY544727.1. Bottom, Western blot analysis of translational products for each construct with antisera specific for CP and RdRp. (B) Confirmation of low expression level of the downstream RFP ORF from plasmid p18S_40M using the dual-fluorescence assay. Same mutations as introduced in the upstream viral sequence in p18S_40M were introduced into eGFP-CP:RFP to generate plasmid Fluo_18S_40M. Five Fluo_18S_40M transformants were selected at random to measure green and red fluorescence. The mean ratio of RFP to GFP from different transformants was compared to that from the original construct.

5 CHAPTER FIVE: Discussion

Several lines of evidence have previously been presented in support of the conclusion that HvV190S expresses its RdRp from the downstream ORF (ORF2) of its genome-length, bicistronic mRNA via an internal-initiation mechanism, proposed to be coupled termination–reinitiation (14, 15, 25). Moreover, as inferred from shared AUGA or similar such sequence motifs involving ORF1 (CP) stop codon and ORF2 start codon, RdRp expression strategy is thought to be a distinguishing characteristic of at least 14 other viruses closely related to HvV190S, which constitute the genus *Victorivirus* in family *Totiviridae* (8, 19). Despite the previous evidence, the occurrence of translational coupling in these viruses has not been widely appreciated.

In this dissertation, two RNA sequence determinants were identified for coupled translation of RdRp from the downstream ORF of HvV190S: 1, Translation coupling requires that the CP stop and RdRp start codons remain relatively closely spaced (Fig. 3.3); 2, A predicted pseudoknot structure is present within close proximity upstream of the CP stop codon (Fig. 3.5 and Fig. 3.6). Translation of the upstream (CP) ORF and termination at its stop codon are absolute requirements for reinitiation at the RdRp start codon, distinguishing this process from IRES-mediated initiation (37), as well as from leaky scanning (38) and ribosomal shunting (39, 109). Leaky scanning can be further ruled out since there are >20 AUG codons in reasonably favorable context upstream of the RdRp start codon. The fact that RdRp is expressed from its downstream start codon as a separate, nonfused protein also distinguishes this process from ribosomal frameshifting (40) and in-frame read-through of termination codons (28), which generate fusion proteins.

The identification of a predicted RNA structure element (pseudoknot) as a key determinant of the stop–restart mechanism in HvV190S and other victoriviruses is an exciting aspect of this study. Results indicate that the predicted structure element must be located closely upstream of the CP stop codon, presumably so that it can tether the terminating ribosome or components thereof, which can then reinitiate with finite efficiency at the next downstream start codon (58). Even though the restart codon is very close to the predicted structure element (only 6–13 nt downstream) and CP stop codon (within 2 nt) in all of wild-type victoriviruses, the fact that HvV190S RdRp can initiate at the AUG2 position, which is 90 nt downstream of the predicted pseudoknot (80 nt downstream the CP stop codon)

when the AUG1 codon is mutated to GUG, suggests that the very close juxtaposition seen in the wild-type viruses is not strictly required, though it may be favored for other reasons that have yet to identified. There does appear to be a limit to how far downstream the restart codon can be located, however, in that reinitiation does not occur effectively at the AUG3 position, which is 207 nt downstream of the putative pseudoknot in HvV190S, when both AUG1 and AUG2 are mutated to GUG.

In this study, a dual fluorescence assay was developed in *H. victoriae* that can be adapted for studies on low frequency events in fungi. This is the first report for using such dual fluorescence assay in a filamentous fungus. The use of fluorescent reporters not only circumvents the extra steps needed for an enzyme assay or expensive substrates for luciferase reporters or unavoidable exposure to isotopes, but also circumvents the need for considering gene insertion copy number and requiring equal amount of material used each time whenever they are subjected for quantitative analysis. This method also provides an effective alternative to dual luciferase assays for research on recording events in those fungi in which luciferases are not expressed well.

The dual-fluorescence vector, eGFP-CP:RFP, which was developed for this study, maintains the wild-type organization of the HvV190S genome including the AUGA motif and the regulatory sequences within the CP ORF. Translation of the RFP-His coding sequence, which replaced the RdRp ORF in the HvV190S genome, is effectively initiated from the downstream ORF of this construct. The findings that parallel changes in red and green fluorescence were manifest in the different eGFP-CP:RFP transformants suggested that fluorescence intensities correspond well with the expression levels of the two fluorescent proteins. The high consistency of this system, which was substantiated by strong correlation ($R^2 = 0.910$) in linear-regression analysis, allowed us to use it as a sensitive and reliable quantitative tool for measuring the rate of re-initiation from the downstream ORF of a bicistronic construct. For determining the efficiency of reinitiation in the dual-fluorescence system, this construct, eGFP-CPt-RFP, was established to express a single protein with each end of truncated CP fused to a reporter. The importance of this construct is that it allows us to quantitate the relative fluorescence intensities of eGFP and RFP when translated from the same ORF, and thus in approximately equal amounts, in *H. victoriae*, thereby providing a means for normalizing relative levels of eGFP and RFP translation when using instead the eGFP-

CP:RFP construct. The extremely high correlation value (R^2 = 0.998) for RFP to GFP in vector eGFP-CPt-RFP suggested both RFP and GFP are translated from the same monocistronic ORF. By comparing the bicistronic vector and the monocistronic vector with GFP as a control for two vectors, RdRp reinitiation efficiency for HvV190S downstream ORF was determined to be ~ 3.9%. This rate suggests that 3.9% of ribosomes that terminate at the upstream ORF will continue to translate RdRp ORF, in other words, it suggests every 100 moles of CP molecules are synthesized, there will be appropriate 4 moles of RdRp molecules are made. The value is similar to the observed molar ratio in HvV190S viral particles of 1~2 molecules of polymerase to 120 molecules of capsid proteins (12, 73). With this assay, the entire upstream sequence of CP ORF was examined for the definition of RNA elements required for reinitiaton of RdRp translation. RNA sequence elements were defined to comprise the terminal 38 nucleotides at the 3' end of the CP ORF (Fig. 4.6). Although the limited proteolytic processing was observed for both above two vectors on their translational fusion proteins, the processed fused proteins due to the N- or C- terminal modification did not seem to affect the value of measured fluorescence since results were reproduced in all experiments described in Fig. 4.7. However, experiments have not yet set up to confirm if it is true that all the constructs have the same rate for the degradation on fusion proteins. It will be interesting to check if the RFP to GFP ratio changes when transformants are challenged with protease inhibitors that inhibit the limited proteolytic processing or reduced the natural degradation by comparing to those from non-challenged.

I have previously attempted to define the required RNA sequence boundary at the 3' end of CP ORF for RdRp translation in experiments involving purified VLPs (See Fig.3.9). However, none of these constructs showed RdRp expression, based on detection of packaged RdRp in purified VLPs. No conclusions were drawn from these experiments since it was not possible to distinguish between lack of RdRp synthesis and unsuccessful RdRp packaging. However, based on other results presented in this study, it can now be clarified that only the terminal 38 nucleotides out of 170 nucleotides contribute to reinitiation of translation of the downstream ORF. Upstream sequences beyond those 38 nucleotides might be involved in the mediation of RdRp packaging into VLPs.

Pseudoknot in SsRV1 is most similar to that of HvV190S since they share the same number of nucleotides in the stem-loops and have similar GC content; they differ the most in the length of the inter-region and organization of stop-restart codons (Fig. 4.8B). It was previously confirmed that reinitiation of downstream RdRp is tolerated by varying stop/restart sites, either overlapped or separated stop/restart codons (Fig. 3.3) provided they are closely located to each other. Such flexible requirement for stop-restart codons is also true for caliciviruses (62). Thus the disparity in expression levels of RFP from various pseudoknots should not be due to variations in stop/restart sites. Then next consideration is the spacer-region between pseudoknots and stop/restart sites. The existence of an authentic pseudoknot in SsRV1 was supported by efficient RFP reintiation from mutant PK9 in Fig. 4.10C, in which mutated pseudoknot of HvV190S was changed to that of SsRV1. The difference between PK9 and PK_SsRV1 is that PK9 possesses wild type spacer-region as in HvV190S whereas PK_SsRV1 contains its own short spacer-region. This suggested that the pseudoknot predicted for SsRV1 indeed reinitiated downstream gene translation in HvV190S and should play a similar role in SsRV1. Thus the short inter-region provides the most likely explanation of the low level of RFP reinitiation. The length of spacer region known to launch reinitiation has a minimum of 6 nucleotides to a maximum of 13 nucleotides for HvV190S, which was experimentally tested (Fig. 3.6). SsRV1 represents the minimum one with 6 nucleotides, BfTV1 and TcV1 represent the largest two with 10 nucleotides and CeRV1, MoV1 and GaRV-L1 represent the middle three with 7, 8 and 9 nucleotides, respectively. Since the authentic reinitiation efficiency for RdRp of SsRV1 in its original host was not yet reported, it is thus unclear whether it is really due to the short inter-region. The long stem-loop in SsRV2, when inserted into HvV190S in a similar position did not launch reinitiation of RdRp translation from the downstream ORF of HvV190S genome, which suggest that the long stem-loop structure might not perform a similar function as pseudoknot regarding translation reinitiation. The unsuccessful reinitiation of RdRp expression from the downstream ORF in the recombinant SsRV2-HvV190S construct (see Fig. 4.8C) suggests that SsRV2 might employ a different strategy to translate its downstream ORF. Alternatively, there might be sequence errors in the region upstream the stop-restart motif that affect pseudoknot prediction.

Direct examination for a correlation between the existence of a predicted pseudoknot in HvV190S and reinitiation of translation was performed in experiments described in Fig.

4.10. Results of Western blot analysis of RdRp expression and packaging within VLPs and those of fluorescence intensity measurements matched very well the results of pseudoknot prediction using program HPknotter (with one exception; PK4 in Fig. 4.10A). These findings strongly suggest that there is a positive correlation between predicted secondary structures and authentic expression of the downstream ORF. Although PK4 did not express RdRp as expected, this might be explained based on its suboptimal spacer length of 5 nts between the reformed pseudoknot and stop/restart. A minimum of 6 nts was found for naturally occurring victoriviruses (see Fig. 3.7) and experimentally tested efficient spacer length (see Fig. 3.6). However, this construct raises the question that changes at the amino acid level instead of nt level (pseudoknot structure) might affect CP-RdRp interaction, thus packaging of RdRp into VLPs.

To address the argument that amino acid level changes instead of changes in nt sequences (subsequently pseudoknot structure) or a combination of both control reinitiation of the downstream ORF and packaging in VLPs, deduced amino acid sequences immediately upstream of stop/restart site for all constructs used in Fig. 4.10A-C were determined and listed in Fig. 4.10D. PK1 contains codon changes due to nucleotide mutations introduced into the primary sequence of one strand of stem 1 whereas PK2 contains codon changes due to nucleotide mutations introduced into the other strand of stem 1. Different results from PK1 and PK2 (PK2 expressed RdRp while PK1 did not) may suggest that amino acid changes due to mutations in the first strand of stem 1 probably controls RdRp packaging. However, PK3 rules out this possibility since it contains codon changes that are also contained in PK1 and PK2, but it expressed RdRp, suggesting that lack of detection of RdRp expression is not necessarily due to amino acids changes and consequently absence of packaging. To confirm that it is due to the function of an alternative pseudoknot existing in PK2, construct PK2-1 was made (Fig. 4.10B), in which a single nucleotide A was introduced that did not change amino acid sequence but did disrupt the formation of stem 1. Comparing results of VLPs from PK2 and PK2-1, we can state that base pairing in stem 1 instead of changes in amino acids encoded by the primary sequences of stem 1 affects RdRp reinitiation. It is difficult, however, to differentiate the effect of the primary sequences of stem 2 and its encoded amino acids from base pairing role of stem 2 on RdRp reintiation/packaging since all three constructs with mutations in stem 2 (PK4-6) did not package RdRp or reinitiate RdRp translation. However, experiments with constructs PK10-12 in Fig. 4.10C,

perfectly confirmed that reinitiation of translation from the downstream ORF was controlled authentically by a simple pseudoknot rather than at the amino acid level or primary nt sequences. Since experiments in Fig.4.10 involved two different methods, analyzing the expression of RdRp through purified VLPs was proven indirectly correct since it was reproduced well by dual-fluorescence assay. That is, nucleotide mutations in pseudoknot structure did not affect the packaging of RdRp protein into VLPs. The existence of an authentic pseudoknot structure and its critical role in reinitiaton of translation has thus been determined.

It needs to be pointed out that in this dual-fluorescence assay, the 5' UTR and the majority of RdRp sequence in the genome of HvV190S was lacking, thus it may be argued that the identified sequence elements might not be sufficient since the missing sequences might be playing a role. Recently, reinitiation of translation has been verified for a protein kinase gene, ORF36 of Kaposi's sarcoma-associated herpesvirus (KSHV, dsDNA virus). A short upstream ORF (uORF) located at its 5' UTR seems critical in ORF36 reinitiation (110). However, this example resembles the translation mechanism used by eukaryotic genes, such as GCN4 gene (111). Study of SART1 (a retrotransposon) however, deserves attention because its major features are reminiscent of the reinitiation mechanism in RNA viruses except that two downstream RNA secondary structures were identified to contribute to reinititiation (84). A recent study on victorivirus RnVV1 also proposed that a secondary structure in the downstream ORF might be involved in reinitiation of translation (83). Because HvV190S and some other victoriviruses are predicted to form a comparable secondary structure downstream the AUGA motif, I carried out mutational analysis to disrupt secondary structure of two stem loops predicted downstream the AUGA motif of HvV190S. Results showed all mutants expressed RdRp efficiently (Fig. 3.8). Thus results indicate that downstream sequences are dispensable. Furthermore, when RdRp sequence was replaced by RFP for the downstream ORF together with the 5' UTR was completely removed from the bicistronic constructs, efficient reinitiation of RFP translation was obtained provided the required upstream pseudoknot and AUGA motif were maintained. This finding confirms that both downstream RdRp sequence and 5'UTR, although possible, are not essential.
In summary, this study presents compelling evidence that victoriviruses employ a coupled translation (stop-restart) strategy to express their RdRp from their downstream ORF. The stop-restart mechanism is utilized mostly by viruses with ssRNA genomes

(caliciviruess, pneumoviruses and hypoviruses) and victoriviruses represent the only known example of dsRNA viruses that use this strategy. Reinitiation depends on a pseudoknot structure, 32 nt upstream AUGA motif and pseudoknot structure needs to be located in a limited spacer distance to the AUGA motif which requires a minimum of 6 nts and a maximum of 18 nts. The determinant role of mRNA-rRNA interactions in reinitiation on caliciviral and influenza B viral RNAs, probably used to tether 40S ribosomal subunits to the mRNA after termination in time for initiation factors to be recruited to the AUG of the downstream ORF, however, was not critical in reinitiation on HvV190S. Initiation factor eIF3 was shown to be involved in termination-reinitiation process for both caliciviral and influenza B viral RNAs. Supplementary eIF3 has been shown to stimulate reinitiation at the wild-type FCV TURBS and BM2 ORF of influzenza B by associating with TURBS that binds 40S subunits which are destined to reinitiate translation (89, 92). eIF3 has also been shown to play an important role in reinitiation on eukaryotic GCN4 gene. eIF3 is retained on ribosomes throughout uORF1 translation and, upon termination, interacts with its 5' enhancer to stabilize mRNA association with post-termination 40S subunits and enable resumption of scanning for reinitiation downstream (111, 112) At present, HvV190S represents a novel mechanism for RNA virus termination-reinitiation. Wether or not eIF3 plays a similar role in reinitiation on HvV190S remains unknown. It will be very interesting to identify such initiation factors and to explore the roles they have played in reinitiation on HvV190S, which will uncover this novel mechanism possessed by victoriviruses more in detail.

6 APPENDIX

LIST OF ABBREVIATIONS

Abbreviation	Expansion
CP	Coat protein
RdRp	RNA-dependent RNA polymerase
ORF	Open reading frame
UTR	Untranslated region
dsRNA	Double-stranded RNA
GTP	Guanosine-5'-triphosphate
mRNA	Message RNA
IRES	Internal ribosome entry site
RNA	Ribonucleic acid
TMV	Tobacco mosaic virus
CMV	Cucumber mosaic virus
BMV	Brome mosaic virus
BYDV	Barley yellow dwarf virus
CaMV	Cauliflower mosaic virus
hyg	Hygromycin B phosphotransferase
RSV	Respiratory syncytial virus
RSV	Rous sarcoma virus
PVM	Pneumovirus of mice
RHDV	Rabbit haemorrhagic disease virus
FCV	Feline calicivirus
CHV1	Cryponectria hypovirus 1
HvV190S	*Helminthosporium victoriae* virus 190S
HvV145S	*Helminthosporium victoriae* virus 145S
gpd1	Glycerol-3-phosphate dehydrogenase 1
cDNA	Complementary DNA
eGFP	Enhanced green fluorescence protein
RFP	Red fluorescence protein
Rluc	Renilla luciferase
Fluc	Firefly luciferase
CAT	*Chloramphenicol acetyltransferase*
BfTV1	*Botryotinia fuckeliana* totivirus 1
GaRV-L1	*Gremmeniella abietina* RNA virus L1
SsRV1	*Sphaeropsis sapnea* RNA viruses 1
MoV1	*Magnaporthe oryzae* viruses 1
TcV1	*Tolypocladium cylindrosporum* virus 1
CeRV1	*Chalara elegans* RNA virus 1

BbRV1	*Beauveria bassiana* RNA virus 1
CmRV	*Coniothyrium minitans* RNA virus
EfV1	*Epichloe festucae* virus 1
HmTV-17	*Helicobasidium mompa* totivirus 1-17
MoV2	*Magnaporthe oryzae* viruses 2
RnVV1	*Rosellinia necatrix* victorivirus 1
SsRV2	*Sphaeropsis sapnea* RNA viruses 1
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Liquid broth
DNA	Deoxyribonucleic acid
PEG	Polyethylene glycol
PBS	Phosphate buffered saline
PMSF	Phenylmethylsulfonyl fluoride
PDA	Potato dextrose agar
VLP	Virus-like particle
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
NBT	Nitro blue tetrazolium
dCTP	Deoxycytidine triphosphate
°C	Degree centigrade
PCR	Polymerase chain reaction
g/µg	Gram/microgram
l/µl	Liter/microliter
h	hour
EDTA	Ethylene diamine tetra acetic acid
DTT	Dithiothreitol
EtBr	Ethidium bromide
M/Mm/uM	Molar/millimolar/micromolar
MOPS	3-(N-morpholino) propanesulfonic acid

REFERENCES:

1. Meehan FaM, H.C. (1946) A new Helminthosporium blight of oats. Science 104:413-414.

2. Litzenberger SC (1949) Nature of Susceptibility to Helminthosporium Victoriae and Resistance to Puccinia Coronata in Victoria Oats. Phytopathology 39(4):300-318.

3. Lindberg GD (1959) A Transmissible Disease of Helminthosporium-Victoriae. Phytopathology 49(1):29-32.

4. Psarros EE & Lindberg GD (1962) Morphology and Respiration of Diseased and Normal Helminthosporium-Victoriae. Phytopathology 52(7):693-699.

5. Lindberg GD (1960) Reduction in Pathogenicity and Toxin Production in Diseased Helminthosporium Victoriae. Phytopathology 50(6):457-460.

6. Hollings M (1962) Viruses Associated with a Die-Back Disease of Cultivated Mushroom. Nature 196(4858):962-&.

7. Sanderlin RS & Ghabrial SA (1978) Physicochemical Properties of 2 Distinct Types of Virus-Like Particles from Helminthosporium-Victoriae. Virology 87(1):142-151.

8. Ghabrial SA, Dunn SE, Li H, Xie JT, & Baker TS (2013) Viruses of Helminthosporium (Cochlioblus) victoriae. Adv Virus Res 86:289-325.

9. Ghabrial SA, Soldevila AI, & Havens WM (2002) Molecular genetics of the viruses infecting the plant pathogenic fungus Helminthosporium victoriae In Molecular biology of double-stranded RNA: Concepts and applications in agriculture, forestry and medicine. S.M. Tavantzis, ed. CRC Press, Boca Ratoon.:213-236.

10. Ghabrial SA (1998) Origin, adaptation and evolutionary pathways of fungal viruses. Virus Genes 16(1):119-131.

11. Ghabrial SA & Suzuki N (2009) Viruses of Plant Pathogenic Fungi. Annu Rev Phytopathol 47:353-384.

12. Caston JR, et al. (2006) Three-dimensional structure and stoichiometry of Helmintosporium victoriae190S totivirus. Virology 347(2):323-332.

13. Soldevila AI, Havens WM, & Ghabrial SA (2000) A cellular protein with an RNA-binding activity co-purifies with viral dsRNA from mycovirus-infected Helminthosporium victoriae. Virology 272(1):183-190.

14. Huang SH & Ghabrial SA (1996) Organization and expression of the double-stranded RNA genome of Helminthosporium victoriae 190S virus, a totivirus infecting a plant pathogenic filamentous fungus. P Natl Acad Sci USA 93(22):12541-12546.

15. Huang SH, Soldevila AI, Webb BA, & Ghabrial SA (1997) Expression, assembly, and proteolytic processing of Helminthosporium victoriae 190S totivirus capsid protein in insect cells. Virology 234(1):130-137.

16. Kozak M (1989) Context Effects and Inefficient Initiation at Non-Aug Codons in Eukaryotic Cell-Free Translation Systems. Mol Cell Biol 9(11):5073-5080.

17. Ghabrial SA & Havens WM (1992) The Helminthosporium-Victoriae 190s Mycovirus Has 2 Forms Distinguishable by Capsid Protein-Composition and Phosphorylation State. Virology 188(2):657-665.

18. Ghabrial SA, Bibb JA, Price KH, Havens WM, & Lesnaw JA (1987) The Capsid Polypeptides of the 190s Virus of Helminthosporium-Victoriae. J Gen Virol 68:1791-1800.

19. Ghabrial SA & Nibert ML (2009) Victorivirus, a new genus of fungal viruses in the family Totiviridae. Arch Virol 154(2):373-379.

20. Icho T & Wickner RB (1989) The Double-Stranded-Rna Genome of Yeast Virus L-a Encodes Its Own Putative Rna-Polymerase by Fusing 2 Open Reading Frames. J Biol Chem 264(12):6716-6723.

21. Wickner RB (1996) Double-stranded RNA viruses of Saccharomyces cerevisiae. Microbiol Rev 60(1):250-&.

22. Wang AL, Yang HM, Shen KA, & Wang CC (1993) Giardiavirus Double-Stranded-Rna Genome Encodes a Capsid Polypeptide and a Gag Pol-Like Fusion Protein by a Translation Frameshift. P Natl Acad Sci USA 90(18):8595-8599.

23. Scheffter SM, Ro YT, Chung IK, & Patterson JL (1995) The Complete Sequence of Leishmania Rna Virus Lrv2-1, a Virus of an Old-World Parasite Strain. Virology 212(1):84-90.

24. Lee SE, Suh JM, Scheffter S, Patterson JL, & Chung IK (1996) Identification of a ribosomal frameshift in Leishmania RNA virus 1-4. J Biochem-Tokyo 120(1):22-25.

25. Soldevila AI & Ghabrial SA (2000) Expression of the totivirus Helminthosporium victoriae 190S virus RNA-dependent RNA polymerase from its downstream open reading frame in dicistronic constructs. J Virol 74(2):997-1003.

26. Jackson RJ, Hellen CUT, & Pestova TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Bio 11(2):113-127.

27. Hinnebusch AG (2011) Molecular Mechanism of Scanning and Start Codon Selection in Eukaryotes. Microbiol Mol Biol R 75(3):434-467.

28. Dreher TW & Miller WA (2006) Translational control in positive strand RNA plant viruses. Virology 344(1):185-197.

29. Kozak M (1978) How Do Eukaryotic Ribosomes Select Initiation Regions in Messenger-Rna. Cell 15(4):1109-1123.

30. Kozak M (1984) Selection of initiation sites by eucaryotic ribosomes: effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. Nucleic Acids Res 12(9):3873-3893.

31. Sherman F, Stewart JW, & Schweingruber AM (1980) Mutants of yeast initiating translation of iso-1-cytochrome c within a region spanning 37 nucleotides. Cell 20(1):215-222.

32. Kozak M (1987) An Analysis of 5'-Noncoding Sequences from 699 Vertebrate Messenger-Rnas. Nucleic Acids Res 15(20):8125-8148.

33. Kozak M (1986) Point Mutations Define a Sequence Flanking the Aug Initiator Codon That Modulates Translation by Eukaryotic Ribosomes. Cell 44(2):283-292.

34. Kozak M (1989) The Scanning Model for Translation - an Update. J Cell Biol 108(2):229-241.

35. Shabalina SA, Ogurtsov AY, Rogozin IB, Koonin EV, & Lipman DJ (2004) Comparative analysis of orthologous eukaryotic mRNAs: potential hidden functional signals. Nucleic Acids Res 32(5):1774-1782.

36. Edelmann SE & Staben C (1994) A statistical analysis of sequence features within genes from Neurospora crassa. Exp Mycol 18:70-81.

37. Hellen CUT & Sarnow P (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. Gene Dev 15(13):1593-1612.

38. Kozak M (2002) Pushing the limits of the scanning mechanism for initiation of translation. Gene 299(1-2):1-34.

39. Ryabova LA & Hohn T (2000) Ribosome shunting in the cauliflower mosaic virus 35S RNA leader is a special case of reinitiation of translation functioning in plant and animal systems. Gene Dev 14(7):817-829.

40. Dinman JD, Icho T, & Wickner RB (1991) A -1 Ribosomal Frameshift in a Double-Stranded-Rna Virus of Yeast Forms a Gag Pol Fusion Protein. P Natl Acad Sci USA 88(1):174-178.

41. Firth AE & Brierley I (2012) Non-canonical translation in RNA viruses. J Gen Virol 93:1385-1409.

42. Goldbach R (1987) Genome Similarities between Plant and Animal Rna Viruses. Microbiol Sci 4(7):197-202.

43. Wellink J, Rezelman G, Goldbach R, & Beyreuther K (1986) Determination of the Proteolytic Processing Sites in the Polyprotein Encoded by the Bottom-Component Rna of Cowpea Mosaic-Virus. J Virol 59(1):50-58.

44. Zaccomer B, Haenni AL, & Macaya G (1995) The Remarkable Variety of Plant Rna Virus Genomes. J Gen Virol 76:231-247.

45. Brierley I (1995) Ribosomal Frameshifting on Viral Rnas. J Gen Virol 76:1885-1892.

46. Brierley I, Digard P, & Inglis SC (1989) Characterization of an Efficient Coronavirus Ribosomal Frameshifting Signal - Requirement for an Rna Pseudoknot. Cell 57(4):537-547.

47. Jacks T, Madhani HD, Masiarz FR, & Varmus HE (1988) Signals for Ribosomal Frameshifting in the Rous-Sarcoma Virus Gag-Pol Region. Cell 55(3):447-458.

48. Dinman JD & Wickner RB (1992) Ribosomal Frameshifting Efficiency and Gag Gag-Pol Ratio Are Critical for Yeast M(1) Double-Stranded-Rna Virus Propagation. J Virol 66(6):3669-3676.

49. Bustamante PIaH, R. (1998) Plant virus gene expression strategies. Electronic J Biotech 1(2):65-82.

50. Palukaitis P, Garciaarenal F, Sulzinski MA, & Zaitlin M (1983) Replication of Tobacco Mosaic-Virus .7. Further Characterization of Single-Stranded and Double-Stranded Virus-Related Rnas from Tmv-Infected Plants. Virology 131(2):533-545.

51. Jaspars EMJ, Gill DS, & Symons RH (1985) Viral-Rna Synthesis by a Particulate Fraction from Cucumber Seedlings Infected with Cucumber Mosaic-Virus. Virology 144(2):410-425.

52. Miller WA, Dreher TW, & Hall TC (1985) Synthesis of Brome Mosaic-Virus Subgenomic Rna Invitro by Internal Initiation on (-)-Sense Genomic Rna. Nature 313(5997):68-70.

53. McGrath PF, Lister RM, & Hunter BG (1996) A domain of the readthrough protein of barley yellow dwarf virus (NY-RPV isolate) is essential for aphid transmission. Eur J Plant Pathol 102(7):671-679.

54. Kozak M (1999) Initiation of translation in prokaryotes and eukaryotes. Gene 234(2):187-208.

55. Dominguez DI, et al. (1998) Ribosome shunting in cauliflower mosaic virus - Identification of an essential, and sufficient structural element. J Biol Chem 273(6):3669-3678.

56. Hohn T, et al. (2001) Shunting is a translation strategy used by plant pararetroviruses (Caulimoviridae). Micron 32(1):51-57.

57. Futterer J, Kisslaszlo Z, & Hohn T (1993) Nonlinear Ribosome Migration on Cauliflower Mosaic Virus-35s Rna. Cell 73(4):789-802.

58. Powell ML, Brown TDK, & Brierley I (2008) Translational termination-re-initiation in viral systems. Biochem Soc T 36:717-722.

59. Horvath CM, Williams MA, & Lamb RA (1990) Eukaryotic Coupled Translation of Tandem Cistrons - Identification of the Influenza-B Virus Bm2 Polypeptide. Embo J 9(8):2639-2647.

60. Gould PS & Easton AJ (2007) Coupled translation of the second open reading frame of M2 mRNA is sequence dependent and differs significantly within the subfamily Pneumovirinae. J Virol 81(16):8488-8496.

61. Meyers G (2003) Translation of the minor capsid protein of a calicivirus is initiated by a novel termination-dependent reinitiation mechanism. J Biol Chem 278(36):34051-34060.

62. Meyers G (2007) Characterization of the sequence element directing translation reinitiation in RNA of the calicivirus, rabbit hemorrhagic disease virus. J Virol 81(18):9623-9632.

63. Guo LH, Sun LY, Chiba S, Araki H, & Suzuki N (2009) Coupled termination/reinitiation for translation of the downstream open reading frame B of the prototypic hypovirus CHV1-EP713. Nucleic Acids Res 37(11):3645-3659.

64. Ghabrial SA (2008) Totiviruses. In B.W.J. Mahy and M. H. V. Van Regenmortel (ed.), Encyclopedia of Virology, 3rd ed. Elsevier, Oxford, United Kingdom.

65. Vanwert SL & Yoder OC (1992) Structure of the Cochliobolus-Heterostrophus Glyceraldehyde-3-Phosphate Dehydrogenase Gene. Curr Genet 22(1):29-35.

66. Cullen D, Leong SA, Wilson LJ, & Henner DJ (1987) Transformation of Aspergillus-Nidulans with the Hygromycin-Resistance Gene, Hph. Gene 57(1):21-26.

67. Staben C, et al. (1989) Use of a bacterial Hygromycin B resistance gene as a dominant selectable marker in Neurospora crassa transformation. Fungal Genetics Newsl 36:79-81.

68. Ho SN, Hunt HD, Horton RM, Pullen JK, & Pease LR (1989) Site-Directed Mutagenesis by Overlap Extension Using the Polymerase Chain-Reaction. Gene 77(1):51-59.

69. Horton RM, Cai ZL, Ho SN, & Pease LR (1990) Gene-Splicing by Overlap Extension - Tailor-Made Genes Using the Polymerase Chain-Reaction. Biotechniques 8(5):528-&.

70. Zhao TY, Havens WM, & Ghabrial SA (2006) Disease phenotype of virus-infected Helminthosporium victoriae is independent of overexpression of the cellular alcohol oxidase/RNA-binding protein Hv-p68. Phytopathology 96(3):326-332.

71. Yoder OC (1988) Cochliobolus heterostrophus, cause of Southern corn leaf blight. Advances in Plant Pathology 6:19.

72. Ghabrial SA & Havens WM (1989) Conservative Transcription of Helminthosporium-Victoriae 190s Virus Double-Stranded-Rna Invitro. J Gen Virol 70:1025-1035.

73. Dunn SE, et al. (2013) Three-dimensional Structure of Victorivirus HvV190S Suggests Coat Proteins in Most Totiviruses Share a Conserved Core. Plos Pathog 9(3).

74. Soldevila AI & Ghabrial SA (2001) A novel alcohol oxidase/RNA-binding protein with affinity for mycovirus double-stranded RNA from the filamentous fungus Helminthosporium (Cochliobolus) victoriae - Molecular and functional characterization. J Biol Chem 276(7):4652-4661.

75. Huang CH, Lu CL, & Chiu HT (2005) A heuristic approach for detecting RNA H-type pseudoknots. Bioinformatics 21(17):3501-3508.

76. Sperschneider J, Datta A, & Wise MJ (2011) Heuristic RNA pseudoknot prediction including intramolecular kissing hairpins. Rna 17(1):27-38.

77. Reeder J, Steffen P, & Giegerich R (2007) pknotsRG: RNA pseudoknot folding including near-optimal structures and sliding windows. Nucleic Acids Res 35:W320-W324.

78. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31(13):3406-3415.

79. Herrero N, Duenas E, Quesada-Moraga E, & Zabalgogeazcoa I (2012) Prevalence and Diversity of Viruses in the Entomopathogenic Fungus Beauveria bassiana. Appl Environ Microb 78(24):8523-8530.

80. Herrero N & Zabalgogeazcoa I (2011) Mycoviruses infecting the endophytic and entomopathogenic fungus Tolypocladium cylindrosporum. Virus Res 160(1-2):409-413.

81. Kozlakidis Z, Herrero N, & Coutts RHA (2013) The complete nucleotide sequence of a totivirus from Aspergillus foetidus (vol 158, pg 263, 2013). Arch Virol 158(2):519-519.

82. Zhang TT, Jiang YH, Huang JB, & Dong WB (2013) Complete genome sequence of a putative novel victorivirus from Ustilaginoidea virens. Arch Virol 158(6):1403-1406.

83. Chiba S, Lin YH, Kondo H, Kanematsu S, & Suzuki N (2013) A Novel Victorivirus from a Phytopathogenic Fungus, Rosellinia necatrix, Is Infectious as Particles and Targeted by RNA Silencing. J Virol 87(12):6727-6738.

84. Kojima KK, Matsumoto T, & Fujiwara H (2005) Eukaryotic translational coupling in UAAUG stop-start codons for the bicistronic RNA translation of the non-long terminal repeat retrotransposon SART1. Mol Cell Biol 25(17):7675-7686.

85. Powell ML, Napthine S, Jackson RJ, Brierley I, & Brown TDK (2008) Characterization of the termination-reinitiation strategy employed in the expression of influenza B virus BM2 protein. Rna 14(11):2394-2406.

86. Luttermann C & Meyers G (2007) A bipartite sequence motif induces translation reinitiation in feline calicivirus RNA. J Biol Chem 282(10):7056-7065.

87. Matassova NB, Venjaminova AG, & Karpova GG (1998) Nucleotides of 18S rRNA surrounding mRNA at the decoding site of translating human ribosome as revealed from the cross-linking data. Bba-Gene Struct Expr 1397(2):231-239.

88. Luttermann C & Meyers G (2009) The importance of inter- and intramolecular base pairing for translation reinitiation on a eukaryotic bicistronic mRNA. Gene Dev 23(3):331-344.

89. Poyry TAA, Kaminski A, Connell EJ, Fraser CS, & Jackson RJ (2007) The mechanism of an exceptional case of reinitiation after translation of a long ORF reveals why such events do not generally occur in mammalian mRNA translation. Gene Dev 21(23):3149-3162.

90. Pisarev AV, Hellen CUT, & Pestova TV (2007) Recycling of eukaryotic posttermination ribosomal complexes. Cell 131(2):286-299.

91. Gould PS & Easton AJ (2005) Coupled translation of the respiratory syncytial virus M2 open reading frames requires upstream sequences. J Biol Chem 280(23):21972-21980.

92. Powell ML, et al. (2011) Further Characterisation of the Translational Termination-Reinitiation Signal of the Influenza B Virus Segment 7 RNA. Plos One 6(2).

93. Kim YC, Su L, Maas S, O'Neill A, & Rich A (1999) Specific mutations in a viral RNA pseudoknot drastically change ribosomal frameshifting efficiency. P Natl Acad Sci USA 96(25):14234-14239.

94. Grentzmann G, Ingram JA, Kelly PJ, Gesteland RF, & Atkins JF (1998) A dual-luciferase reporter system for studying recoding signals. Rna 4(4):479-486.

95. Harger JW & Dinman JD (2003) An in vivo dual-luciferase assay system for studying translational recoding in the yeast Saccharomyces cerevisiae. Rna 9(8):1019-1024.

96. McNabb DS, Reed R, & Marciniak RA (2005) Dual luciferase assay system for rapid assessment of gene expression in Saccharomyces cerevisiae. Eukaryot Cell 4(9):1539-1549.

97. McCormick CJ, Salim O, Lambden PR, & Clarke IN (2008) Translation termination reinitiation between open reading frame 1 (ORF1) and ORF2 enables capsid expression in a bovine norovirus without the need for production of viral subgenomic RNA. J Virol 82(17):8917-8921.

98. Sanford BA, de Feijter AW, Wade MH, & Thomas VL (1996) A dual fluorescence technique for visualization of Staphylococcus epidermidis biofilm using scanning confocal laser microscopy. J Ind Microbiol 16(1):48-56.

99. Yamamoto N, et al. (2004) Cellular dynamics visualized in live cells in vitro and in vivo by differential dual-color nuclear-cytoplasmic fluorescent-protein expression. Cancer Res 64(12):4251-4256.

100. Stewart MD, Jang CW, Hong NW, Austin AP, & Behringer RR (2009) Dual fluorescent protein reporters for studying cell behaviors in vivo. Genesis 47(10):708-717.

101. Martin K, et al. (2009) Transient expression in Nicotiana benthamiana fluorescent marker lines provides enhanced definition of protein localization, movement and interactions in planta. Plant J 59(1):150-162.

102. Barelle CJ, et al. (2004) GFP as a quantitative reporter of gene regulation in Candida albicans. Yeast 21(4):333-340.

103. Soboleski MR, Oaks J, & Halford WP (2005) Green fluorescent protein is a quantitative reporter of gene expression in individual eukaryotic cells. Faseb J 19(1):440-442.

104. Cardno TS, Poole ES, Mathew SF, Graves R, & Tate WP (2009) A homogeneous cell-based bicistronic fluorescence assay for high-throughput identification of drugs that perturb viral gene recoding and read-through of nonsense stop codons. Rna 15(8):1614-1621.

105. Rakauskaite R, Liao PY, Rhodin MHJ, Lee K, & Dinman JD (2011) A rapid, inexpensive yeast-based dual-fluorescence assay of programmed-1 ribosomal frameshifting for high-throughput screening. Nucleic Acids Res 39(14).

106. Maor R, Puyesky M, Horwitz BA, & Sharon A (1998) Use of green fluorescent protein (GFP) for studying development and fungal-plant interaction in Cochliobolus heterostrophus. Mycol Res 102:491-496.

107. Mikkelsen L, Sarrocco S, Lubeck M, & Jensen DF (2003) Expression of the red fluorescent protein DsRed-Express in filamentous ascomycete fungi. Fems Microbiol Lett 223(1):135-139.

108. Li H, Havens WM, Nibert ML, & Ghabrial SA (2011) RNA Sequence Determinants of a Coupled Termination-Reinitiation Strategy for Downstream Open Reading Frame Translation in Helminthosporium victoriae Virus 190S and Other Victoriviruses (Family Totiviridae). J Virol 85(14):7343-7352.

109. Pooggin MM, Ryabova LA, He XY, Futterer J, & Hohn T (2006) Mechanism of ribosome shunting in Rice tungro bacilliform pararetrovirus. Rna 12(5):841-850.

110. Kronstad LM, Brulois KF, Jung JU, & Glaunsinger BA (2013) Dual Short Upstream Open Reading Frames Control Translation of a Herpesviral Polycistronic mRNA. Plos Pathog 9(1).

111. Szamecz B, et al. (2008) eIF3a cooperates with sequences 5 ' of uORF1 to promote resumption of scanning by post-termination ribosomes for reinitiation on GCN4 mRNA. Gene Dev 22(17):2414-2425.

112. Munzarova V, et al. (2011) Translation Reinitiation Relies on the Interaction between eIF3a/TIF32 and Progressively Folded cis-Acting mRNA Elements Preceding Short uORFs. Plos Genet 7(7).

VITA

HUA LI

EDUCATION:

2002-2005: M.S. Department of Plant Pathology, College of Agriculture, Huazhong Agriculture University, China.
M.S. dissertation title: Down-regulation of *Sclerotinia sclerotiorum* gene expression in response to infection with *Sclerotinia sclerotiorum* debilitation-associated RNA virus.

1998-2002: B.S. Department of Horticulture, College of Agriculture, Huazhong Agriculture University, China.

PUBLICATION:

1, **Li H**, Havens WM, Nibert ML, Ghabrial SA. An H-type pseudoknot structure is essential for coupled termination-reinitiation translation of the downstream ORF of HvV190S. (**In process**).
2, Ghabrial SA, Dunn SE, **Li H**, Xie J, Baker TS. Viruses of *Helminthosporium* (*Cochlioblus*) *victoriae*. **Adv Virus Res.2013**, **86**: 289-325.
3, Dunn SE, **Li H**, Cardone G, Nibert ML, Ghabrial SA, Baker TS. Three-dimensional structure of *victorivirus* HvV190S suggests coat proteins in most totiviruses share a conserved core. **PLoS Pathog 2013**, **9**(3): e1003225.
4, **Li H**, Havens WM, Nibert ML, Ghabrial SA: RNA Sequence Determinants of a Coupled Termination-Reinitiation Strategy for Downstream Open Reading Frame Translation in *Helminthosporium victoriae* Virus 190S and Other Victoriviruses (Family *Totiviridae*). **J Virol 2011**, **85**: 7343-7352.
5, de Sa PB, **Li H**, Havens WM, Farman ML, Ghabrial SA: Overexpression of the victoriocin gene in *Helminthosporium* (*Cochliobolus*) *victoriae* enhances the antifungal activity of culture filtrates. **Phytopathology 2010**, **100**: 890-896.
6, Tang J, Ochoa WF, **Li H**, Havens WM, Nibert ML, Ghabrial SA, Baker TS: Structure of *Fusarium poae* virus 1 shows conserved and variable elements of *partitivirus* capsids and evolutionary relationships to *picobirnavirus*. **J Struct Biol 2010**, **172**: 363-371.

7, **Li H**, Fu Y, Jiang D, Li G, Ghabrial SA, Yi X. Down-regulation of *Sclerotinia sclerotiorum* gene expression in response to infection with *Sclerotinia sclerotiorum* debilitation-associated RNA virus. **Virus Res 2008**, **135**(1):95-106.

AWARDS/FELLOWSHIPS:

2009: Awarded travel grant by University of Kentucky Graduate School to APS meeting held in Hawaii, USA.
2002-2005: Fellowship during M.S period.
1998-2002: Fellowship during B.S. period.