STUDIES ON TRANSLATIONAL MECHANISMS OF RNA VIRUSES

WANG XIAOXING

NATIONAL UNIVERSITY OF SINGAPORE

2006

STUDIES ON TRANSLATIONAL MECHANISMS OF RNA VIRUSES

WANG XIAOXING

(*B. Sc.*, Fudan University)

A THESIS SUBMITTED

FOR THE DEGREE DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2006

Acknowledgements

I would like to thank my supervisors first – Professor Wong Sek Man and Associate Professor Liu Ding Xiang for their mentorship, guidance, encouragement and motivation, especially for providing me with this opportunity to collaborate between National University of Singapore (NUS) and Institute of Molecular and Cell Biology (IMCB). The collaboration makes it possible for me to experience different environments of doing research and to network with other scientists.

My heartfelt gratitude goes to my friends and colleagues of both the plant virology lab in NUS and the molecular virology and pathologenesis lab in IMCB for their assistance and encouragement. Special thanks to Haihe, Chunying, Srini and Jing Jing for their advice, help and warmth. My thanks also go to Dr. Fang Shouguo, Dr. Yamada, Dr. Nasir and Dr. Xu Linghui for their help and understanding. Special thanks to Law Yin Chern, Felicia, Benson, Siti, Rong Hua, Le Tra My, Xiao Han, Cheng Guang and Hui Hui for their friendships which brighten my days.

I would also like to thank NUS for providing me with a research scholarship and IMCB for giving me the chance to do my work there. Lastly, I want to express my appreciation to my parents for being the infinite source of love and support that I have so needed to stay grounded and focused.

Table of Contents

Abbreviations

BVDV	*Bovine viral diarrhea virus*
BYDV	*Barley yellow dwarf virus*
CaMV	*Cauliflower mosaic virus*
CrPV	*Cricket paralysis virus*
crTMV	*Crucifer-infecting tobamovirus*
CSFV	*Classical swine fever virus*
EMCV	*Encephalomyocarditis virus*
FMDV	*Foot-and-mouth disease virus*
HCV	*Hepatitis C virus*
HCRSV	*Hibiscus chlorotic ringspot virus*
HIV	*Human immunodeficiency virus*
HSV	*Herpes simplex virus*
IBV	*Infectious bronchitis virus*
PPV	*Plum pox virus*
PVM	*Potato virus M*
RHDV	*Rabbit hemorrhagic disease virus*
RTBV	*Rice tungro bacilliform virus*
SARS-CoV	*Severe acute respiratory syndrome coronavirus*
SBWV	*Soil-borne wheat mosaic virus*
SMYEV	*Strawberry mild yellow edge virus*
SV	*Sendai virus*
TCV	*Turnip crinkle virus*
TMEV	*Theiler's murine encephalomyelitis virus*
TMV	*Tobacco mosaic virus*
TYMV	*Turnip yellow mosaic virus*
3CLpro	3C-like protease
4E-BP	eIF4E binding protein
AdoMetDC	S-adenosylmethionine decarboxylase
A-site	aminoacyl-site
cdd	cytidine deaminase gene
C/EBP	CAAT/enhancer-binding protein
c-myc	cellular homologue of avian myelocytomatosis virus oncogene
CP	coat protein
dhfr	dihydrofolate reductase
E	envelope protein
E. coli	*Escherichia coli*
EGFP	enhanced green fluorescence protein
eIF	eukaryotic initiation factor
E-site	exit-site
FGF	fibroblast growth factor
GCN	general control non-derepressable

gRNA	genomic RNA
HE	hemaglutinin esterase
hsp	heat shock protein
IPTG	isopropyl-β-D-thiogalactopyranoside
IRES	internal ribosome entry site
IS element	insertional elements
ITAF	IRES transacting factor
kb	kilo-base
kDa	kilo-Dalton
M	membrane protein
M.O.I.	multiplicity of infection
N	nucleocapsid protein
nt	nucleotide
ODC	ornithine decarboxylase
OGP	osteogenic growth peptide
ORF	open reading frame
PABP	poly-A binding protein
PCR	polymerase chain reaction
PDGF-B	platelet-derived growth factor beta polypeptide
PLpro	papain-like protease
PRF	programmed ribosomal frame-shifting
P-site	peptidyl-site
PTB	polypyrimidine tract binding protein
RdRp	RNA-dependent RNA polymerase
RF	release factor
RT	reverse transcription
S	spike protein
SD	Shine-Dalgarno
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
sgRNA	subgenomic RNA
SL	stem-loop
Snrpn	small nuclear ribonucleoprotein polypeptide N
Snurf	Snrpn upstream reading frame protein
TAV	transactivator/viroplasmin protein
TEF	transcription enhancer factor
TK	thymidine kinase
TnT	coupled transcription and translation
TRS	transcription-regulatory sequence
uORF	upstream open reading frame
UTR	untranslated region
VEGF	vascular endothelial growth factor
v-FLIP protein	FLICE-inhibitory protein
VPg	viral protein, genome-linked

List of Figures

List of Tables

List of Publications

1. **Wang X.**, Wong S.M., Liu D.X. 2006. Identification of Hepta- and Octo-Uridine stretches as sole signals for programmed +1 and -1 ribosomal frameshifting during translation of SARS-CoV ORF 3a variants. *Nucleic Acids Res.* 34, 1250-60.
2. Koh, D.C.Y., **Wang X.**, Wong, S.M., Liu, D.X. 2006. Translation initiation at an upstream CUG codon regulates the expression of Hibiscus chlorotic ringspot virus coat protein. *Virus Res.* 122, 35-44.

Summary

Viruses have evolved a wide range of sophisticated mechanisms to optimize the ability to replicate or at least to survive the host defences. Regulation of gene expression is a key aspect of such processes and control of mRNA translation in particular represents an important focus for virus-host interactions. In this thesis, by studying two RNA viruses, *Severe acute respiratory syndrome coronavirus* (SARS-CoV) and *Hibiscus chlorotic ringspot virus* (HCRSV), the mechanisms of gene expression regulation are studied.

Programmed ribosomal frame-shifting is one of the translational recoding mechanisms that read the genetic code in alternative ways. This process is generally programmed by signals at defined locations in a specific mRNA. In the study of SARS-CoV, we report the identification of hepta- and octo-uridine stretches as sole signals for programmed +1/-2 and -1/+2 ribosomal frame-shifting during translation of SARS-CoV open reading frame (ORF) 3a variants. SARS-CoV ORF 3a encodes a minor structural protein of 274 amino acids. Over the course of cloning and expression of the gene, a mixed population of clones with six, seven, eight and nine uridine stretches located 14 nucleotides downstream of the initiation codon was found. *In vitro* and *in vivo* expression of clones with six, seven and eight Ts, respectively, showed the detection of the full-length 3a protein. Mutagenesis studies led to identification of the hepta- and octo-uridine stretches as slippery sequences for efficient frame-shifting. Interestingly, no stimulatory elements were found in the sequences upstream or downstream of the slippage site. When the hepta- and

octo-uridine stretches were used to replace the original slippery sequence of the SARS-CoV ORF 1a and 1b, efficient frame-shifting events were observed. The efficiencies of frame-shifting mediated by the hepta- and octo-uridine stretches were not affected by mutations introduced into a downstream stem-loop structure that totally abolished the frame-shifting event mediated by the original slippery sequence. Furthermore, the octo-uridine stretch was shown to direct frame-shifting of both -1/+2 and +1/-2. However, no -1/+2 frame-shifting was observed for the hepta-uridine stretch. Taken together, this study identifies the hepta- and octo-uridine stretches that function as sole elements for efficient +1/-2 and -1/+2 ribosomal frame-shifting events.

Most RNA viruses have evolved mechanisms to regulate the first step of translation – translation initiation. According to the conventional scanning model, only the 5'- proximal gene in the viral RNA is accessible to the ribosomes whereas other genes are silent. In this study, we use a model plant RNA virus, HCRSV, to investigate various translation mechanisms involved in the regulation of the expression of internal genes. The 3'-end 1.2 kb region of HCRSV genomic and subgenomic RNAs were shown to encode four polypeptides of 38 kDa, 27 kDa, 25 kDa and 22.5 kDa. Mutagenesis studies revealed that a CUG codon (^{2570}CUG) is the initiation codon for p27, the longest of the three co-C-terminal products (p27, p25 and p22.5), and translation of p25 and p22.5 was initiated at ^{2603}AUG, and ^{2666}AUG, respectively. Translation initiation of the p27 expression at the ^{2570}CUG codon is through a leaky scanning mechanism and regulates the expression of p38, the viral

coat protein. Mutational analysis of an upstream ORF demonstrated that initiation of the p27 expression at this CUG codon (instead of an AUG) may play a role in maintaining the ratio of p27 and p38. In addition, a previously identified internal ribosome entry site (IRES) was shown to control the expression of p27 and p38 in the subgenomic RNA 2. In summary, this study demonstrated that viral gene regulation is a very intricate event and a single gene can be regulated by multiple mechanisms.

CHAPTER 1. LITERATURE REVIEW

Viruses are intracellular parasites containing either DNA or RNA genomes which provide templates and relevant virally encoded proteins for replication and gene expression in an active biological system. In order for the system to function well, gene regulation is crucial as it ensures that proteins are produced at the right time, in the right place and with the right form. Failure of or improper gene regulation frequently leads to lethality or attenuated virulence (Petty *et al*., 1990; Slobodskaya *et al*., 1996). However, such alteration in gene expression may also lead to enhanced virulence which will result in more severe disease outbreaks (Brown *et al*, 2001; Cazzola and Skoda, 2000; Delépine *et al*., 2000; Han *et al*., 2001). Thus viral gene regulation becomes an attractive topic for study, which will provide more information on virus life cycles, virus pathogenesis and virus-host interactions. In addition, research on viral gene regulation helps to understand human diseases for effective treatments such as drug design and drug delivery.

Generally, viral genes are regulated at four different stages, including transcription/replication (regulation of RNA synthesis), post-transcription (RNA modification), translation (regulation of protein synthesis) and post-translation (protein modification) among which transcription and translation are most important. For RNA viruses, the translational regulation has been shown to be an essential contributor for gene regulation. Viruses do not harbor the translation machinery; therefore, they must rely on their host system for protein synthesis. To express their genes efficiently, viruses have evolved various mechanisms to compete with host cells for the translation machinery at different translation stages.

In this study, the translation initiation mechanism of *Hibiscus chlorotic ringspot virus* (HCRSV) and the recoding mechanism of the *Severe acute respiratory syndrome coronavirus* (SARS-CoV) were examined. In this chapter, a thorough review on viral translation initiation and programmed frame-shifting is presented.

1.1 TRANSLATION

Translation consists of three stages: initiation, elongation and termination. Among the three phases, initiation is the first event and it is the rate-limiting step. The generally accepted model of translation initiation in eukaryotes proposes that translation starts from the circularization of the mRNA in which the 5'-cap structure and 3'-poly (A) tail are brought to proximity through bridging proteins such as poly-A binding protein (PABP) (Blobel, 1973; Gallie, 1991; Sachs and Davis, 1989).

To initiate translation, the first step is to form a 43S complex which contains the small ribosomal subunit 40S, initiator methionine tRNA (Met-$tRNA_i^{Met}$), and various eukaryotic initiation factors (eIFs) including eIF2-GTP, eIF3, eIF1 and eIF1A (Fig 1.1). This 43S complex is subsequently recruited onto 5'-cap via eIF4A, eIF4E, eIF4G, and eIF4B, in which eIF4G acts as a scaffold interacting with the cap-binding protein eIF4E, helicase eIF4A, eIF3 as well as poly-A binding protein (PABP). In this way, the mRNA molecule forms a closed loop that is believed to confer stability to the mRNA and efficiency of ribosome recycles. Next, the 43S complex is believed to scan along the mRNA (Kozak, 1989a) from the 5' end until it reaches a proper initiation codon-AUG in most cases-which can base pair with an initiator tRNA

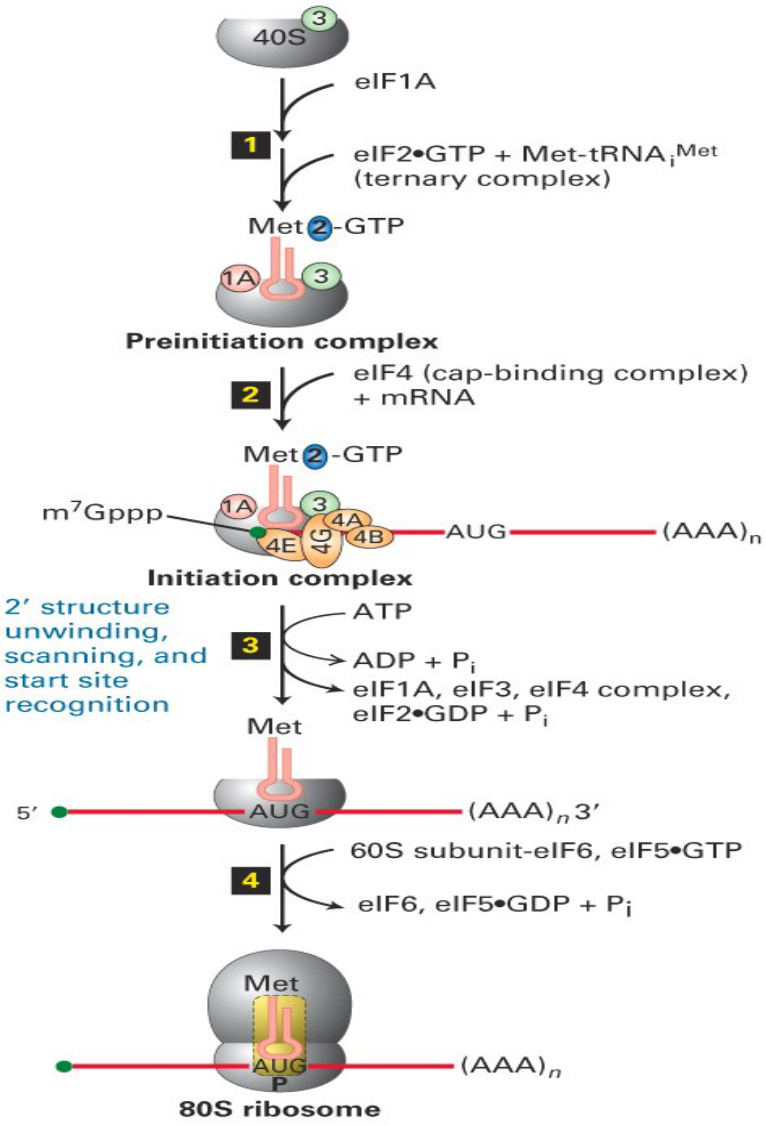

40S
3
eIF1A
1
eIF2•GTP + Met-tRNA$_i^{Met}$
(ternary complex)
Met
2
-GTP
1A
3
Preinitiation complex
2
eIF4 (cap-binding complex)
+ mRNA
Met
2
-GTP
m^{7}Gppp
1A
3
4A
4E
4G
4B
AUG
(AAA)$_n$
Initiation complex
2′ structure
unwinding,
scanning, and
start site
recognition
3
ATP
ADP + P$_i$
eIF1A, eIF3, eIF4 complex,
eIF2•GDP + P$_i$
Met
5′
AUG
(AAA)$_n$ 3′
4
60S subunit-eIF6, eIF5•GTP
eIF6, eIF5•GDP + P$_i$
Met
AUG
P
(AAA)$_n$
80S ribosome

Fig. 1.1

Fig 1.1 Diagram of eukaryotic translation initiation.

A 43S preinitiation complex is formed when a ternary complex of eIF2 bound to GTP and Met-tRNA$_i^{Met}$ associates with the small ribosomal 40S subunit, which is complexed with two other factors, eIF3 and eIF1A, that stabilize binding of the ternary complex. The 5' cap (m^7Gppp) of the mRNA to be translated is guided to the preinitiation complex by the multiprotein eIF4F complex (cap-binding complex), which unwinds any secondary structure at the 5' end of the mRNA and the initiation complex is thus formed. Subsequent scanning by the small ribosomal subunit positions the initiator tRNA at the AUG start codon, releasing eIF1A, eIF3, and eIF4F. With the initiator tRNA properly positioned at the start codon, another factor, eIF5, assists union of the 40S complex with the 60S subunit. Hydrolysis of GTP in eIF2-GTP provides the energy for this step. Factors eIF5 and eIF2-GDP are released, yielding the final 80S initiation complex, with the initiator tRNA at the P-site. The complex can now accept the second aminoacyl-tRNA. (Adapted from Molecular Cell Biology, 4th edition)

carrying a methionine. At the AUG site, a stable 48S complex is formed. Following the disassociation of the initiation factors, the 60S large ribosomal subunit joins the complex to form an 80S complex and translation starts. In most cases translation is initiated from the first AUG codon. Table 1.1 lists the eukaryotic initiation factors and their functions.

A ribosome contains three sites: a P-site (peptidyl-site), an A-site (aminoacyl-site) and an E-site (exit-site). As shown in Fig 1.2, during elongation, the polypeptide is positioned in the P-site and the charged tRNA molecules come into the A-site via a ternary complex with elongation factor (EF) 1A-GTP. If the anticodon of the tRNA at the A-site is base paired with the codon on mRNA, a peptide bond is formed between the P-site and the A-site, triggering GTP hydrolysis and the release of eEF1A bound to GDP from the ribosome. Hence, the peptide is transferred to the A-site, leaving an uncharged tRNA at the P-site. Subsequently, this tRNA will be moved to the E-site (exit site), while the ribosome will translocate the peptidyl-tRNA to the P-site with the help of EF2 and start the next synthesis cycle (Miller and Weissbach, 1977).

When ribosomes reach the stop codons, peptide synthesis is ceased. These codons do not encode for amino acids and cannot be recognized by tRNAs (there are exceptions, however, when stop codons are recognized by supressor tRNAs, which will be reviewed in section 1.2.3 of this chapter). Instead, release factors (RFs) will recognize the codons and induce the ribosome complex to dissociate from the mRNA, thus releasing the synthesized peptide (Fig 1.3).

Table 1.1 Summary of eukaryotic initiation factors

name	subunit	function
eIF1	1	Fidelity of AUG codon recognition, destabilizes aberrant initiation complexes
eIF1A	1	Catalytically promotes Met-tRNAi binding to 40S; required for strong binding of 40S subunit to mRNA
eIF2	3	GTPase, escorts Met-tRNAi onto 40S subunit
eIF2B	5	Guanine nucleotide exchange factor for eIF2
eIF3	11	Scaffold for the cap binding complex, binds 40S subunit; stabilizing Met-tRNAi and preventing association with 60S subunit
eIF4A	1	RNA dependent ATPase; essential for binding of ribosomes to mRNA
eIF4B	1	RNA binding protein; promotes eIF4A activity
eIF4E	1	Binds directly to the m^7G cap
eIF4F	3	Cap binding complex of eIF4A, 4E, and 4G
eIF4G	1	Binds mRNA, PABP, eIF4E, eIF4A, and eIF3
eIF4H	1	Similar to eIF4B
eIF5	1	AUG recognition and promote eIF2 GTPase activity
eIF5B	1	GTPase, mediates assembly of 80S from 40S and 60S
eIF6	1	Binds to 60S subunit and promotes dissociation

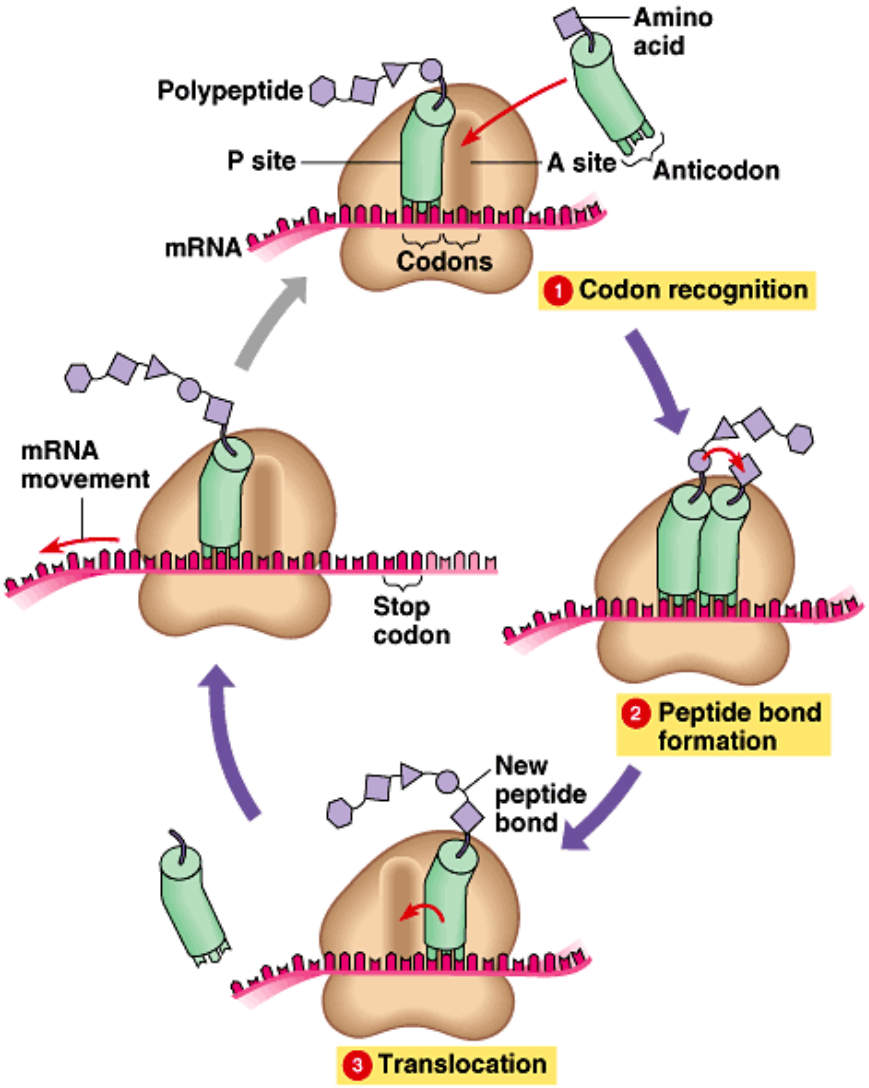

Amino acid
Polypeptide
P site
A site
Anticodon
mRNA
Codons
1 Codon recognition
mRNA movement
Stop codon
2 Peptide bond formation
New peptide bond
3 Translocation

Fig. 1.2

Fig 1.2 The elongation cycle in eukaryotic protein synthesis.

The ribosome contains three sites: a P-site, an A-site and an E-site (which is not shown in this picture). During elongation, the polypeptide is positioned in the P-site and the charged tRNA molecules come into the A-site. If the anticodon of the tRNA at the A-site is unable to base pair with the codon on mRNA, this tRNA will be rejected. However, if they can base pair with each other, a peptide bond is formed between P-site and A-site transferring the peptide to the A-site leaving an uncharged tRNA at the P-site. Subsequently, this tRNA will go to the E-site (exit site), while the ribosome will move to translocate the peptide to the P-site. (Adapted from Molecular Cell Biology, 4^{th} edition)

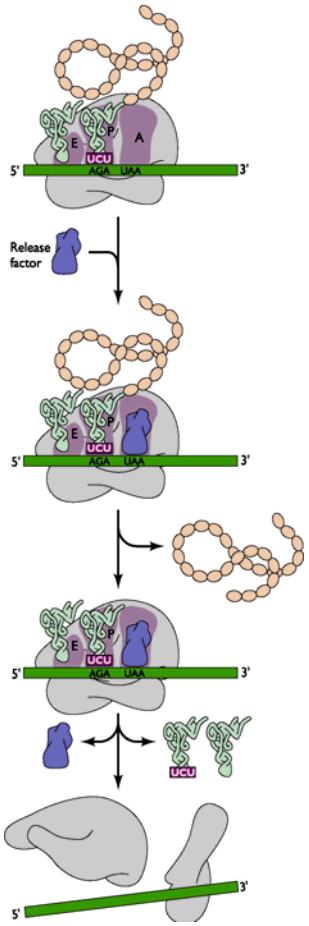


Fig. 1.3

Fig 1.3 Termination of translation.

When a ribosome bearing a nascent protein chain reaches a stop codon (UAA, UGA, UAG), release factors (RFs) enter the ribosomal complex, probably at or near the A-site. The peptide chain was cleaved from the tRNA and released along with the two ribosomal subunits. (Adapted from Molecular Cell Biology, 4^{th} edition)

1.2 OVERVIEW ON VIRAL REGULATION AT TRANSLATIONAL LEVEL

Viral genomes do not encode for any components of the translation machinery. Hence viral protein synthesis is wholly dependent on hosts. It is not so surprising that some viral RNAs have similar structures to the 5'-cap and 3'-poly (A) tail mimicking eukaryotic mRNA to compete for translation apparatus. However, most RNA viruses have been shown to evolve alternative translation initiation mechanisms. In addition to translation initiation, RNA viruses also express their overlapping open reading frames (ORFs) by recoding and read-through during the elongation and termination stages.

1.2.1 Translation initiation

Distinct from eukaryotic mRNAs, viral genomic RNAs (gRNAs) are usually poly-cistronic. Thus in eukaryotic hosts only the most 5'-proximal ORF is translatable while the rest are silent. In order to make full use of the RNA, viruses have various mechanisms to recruit host ribosomes for protein synthesis including leaky scanning, internal initiation, termination and re-initiation, non-AUG mediated initiation and ribosome shunting.

1.2.1.1 Leaky Scanning

In 1978, Marilyn Kozak proposed the scanning mechanism for translation initiation in eukaryotes (Kozak, 1978) stating that the 40S ribosomal subunit binds to the 5'-end of mRNA, migrates and stops at the first AUG codon in a favorable context to initiate translation. This "first-AUG rule" is true for most eukaryotic mRNAs.

Distinct from prokaryotes, eukaryotic ribosomes are restricted to initiating near the 5'-end which can be explained by the scanning model. Strong evidence for this includes the discovery of the 5' ^{m7}G cap (Shatkin, 1976) and that the 3'-ORFs in many viral mRNAs are not translatable *in vitro* and *in vivo* (Wong *et al*., 1987; Good *et al.*, 1988). This position effect can be seen in many cases where translation shifts from the normal initiation site to an upstream AUG newly introduced and where removal of the upstream start codon activates initiation from the next start codon.

However, initiation has been shown to be a context-dependent event. Extensive analysis of sequences flanking the initiation codon through alignment and comparison, a consensus sequence was identified as GCCGCCRCCAUGG in higher eukaryotes (Kozak, 1981, 1984a and b, 1987) (numbering begins with the A of AUG codon as position +1; nucleotides 5' to that site are assigned negative numbers and R is a purine). Site-directed mutagenesis data confirmed that G^{+4} as well as each of the consensus nucleotides from position -1 to -6 were very important (Kozak, 1986 and 1987). The importance of the purine in position -3 was demonstrated by targeting mutations to this position in α-globin from C**A**CCAUG to C**C**CCAUG (Morlé *et al.*, 1985). This mutation dramatically decreases the level of α-globin and resulted in a type of thalassemia. It has been established that a purine in position -3 is the most conserved nucleotide in eukaryotes such as plants (Heidecker and Messing, 1986) and fungi (Paluh *et al*., 1988) and a mutation on this purine has more deleterious effect on translation initiation than a point mutation anywhere else (Kozak, 1986). In the absence of the purine in position -3, efficient translation is dependent on G^{+4} (Kozak,

1986) and other nucleotides in the vicinity (Kozak, 1987a). In summary, an initiation codon can be considered "strong" or "weak" by only referring to position -3 and +4.

According to the scanning model, when the first AUG resides in a very weak context, some ribosomes start translation at that point but most continue scanning and initiate further downstream. This leaky scanning enables the production of two separate proteins from one mRNA. Leaky scanning is the most common phenomenon in eukaryotic mRNAs as well as viral mRNAs with overlapping ORFs allowing the translation from downstream initiation codons (Dinesh-Kumar and Miller, 1993; Fütterer *et al*., 1996 and 1997; Simon-Buela *et al*., 1997). It is dependent on the context of the first start codon. When the first initiation codon is weak or in a poor context or too close to the 5' end to be recognized efficiently, majority of the ribosomes initiated translation from a downstream start site. It is very striking that leaky scanning occurs even when the two initiation codons are far apart. Kozak (1998) reported that using synthetic transcripts no reduction in initiation from the downstream start codon was observed when the distance between the two AUGs was expanded from 11 to 251 nucleotides (nt) stepwise. In some viral mRNAs, the second functional start site is over 500 nt downstream from the first one (Herzog *et al*., 1995; Sivakumaran and Hacker, 1998). In addition, a recent study on *Turnip yellow mosaic virus* (TYMV) RNAs (Matsuda and Dreher, 2006) showed that close spacing between AUGs also contributed to the dicistronic character on a eukaryotic mRNA and increasing space resulted in a decrease in downstream initiation and increase in upstream initiation. In contrast to the above "maximally leaky" mRNAs, some

mRNAs are “minimally leaky” in which only a small fraction of ribosomes bypasses the first AUG in a strong but not perfect context and initiates translation downstream. Examples include the nucleocapsid protein and I-protein of bovine coronavirus (Senanayake and Brian, 1997), the rat histone H4 protein and the osteogenic growth peptide (OGP) (Bab *et al*., 1999). For many of the viruses, both proteins produced via leaky scanning are required for replication. In some viruses, the second protein is a virulence factor that compromises host defenses (Bridgen *et al*., 2001; Chen *et al*., 2001; Weber *et al*., 2002).

1.2.1.2 Internal initiation

1.2.1.2.1 Introduction

The majority of eukaryotic translation initiation is known to be cap-dependent. It is only recently accepted that initiation from internal region of mRNA is possible. The first work that conclusively showed internal initiation presence is on picornaviruses (Pelletier and Sonenberg, 1988b; Jang *et al*., 1988 and 1989). The well characterized 5’-UTR (untranslated region) of picornaviruses indicates that ribosomes can directly bind to an internal region of mRNA in a cap-independent manner (Jackson and Kaminski, 1995b). The site where ribosomes bind to is called internal ribosome binding site (IRES). Up to date, quite a few IRES elements have been discovered in both viral and cellular mRNAs such as the oncogene, *c-myc* and the hypoxia-induced factor, Vascular endothelial growth factor (VEGF) (Bernstein *et al*., 1997; Hellen and Sarnow, 2001; Huez *et al*., 1998; López-Lastra *et al*., 1997;

Martinez-Salas *et al*., 2001; Merrick, 2004; Nanbru *et al*., 1997; Vagner *et al*., 2001) and the best-characterized IRESs are from picornaviruses and *Hepatitis C virus* (HCV), *Classical swine fever virus* (CSFV), *Cricket paralysis virus* (CrPV) and *Bovine viral diarrhea virus* (BVDV) (Belsham and Sonenberg, 2000; Pestova *et al*., 1998b; Tsukiyama-Kohara *et al.*, 1992; Wilson *et al*., 2000). The ability of IRESs to promote internal initiation has facilitated the expression of two or more proteins from a polycistronic transcription unit.

1.2.1.2.2 Functions of IRESs

Viral IRESs have important functions in the viral life cycle, mostly to ensure efficient viral translation when components of the host translation machinery are limited due to virus-induced antiviral responses, some through modification of eukaryotic factors. For instance, a polycistronic transcript found in cells latently infected by *Kaposi's sarcoma-associated herpesvirus* is used to express the FLICE-inhibitory protein (v-FLIP protein) whose function is to counteract fatty acid synthase-induced apoptosis (Bieleski and Talbot, 2001; Grundhoff and Ganem, 2001). The well-studied poliovirus encodes a protease which can cleave eIF4G, thus decrease host cap-dependent translation and uses the IRES to initiate translation bypassing the requirement for eIF4E (Gradi *et al*., 1998a and b; Lamphear *et al*., 1993 and 1995), whereas the cardioviruses do not induce cleavage of eIF4G but is proposed to induce a change in ion concentration within the cell to bias the translational capacity of the cell toward the viral RNA (Alonso and Carrasco, 1981). Another advantage of cap-independent translation is that viruses do not have to devote a gene to encoding

capping enzymes which are usually found in the nucleus instead of cytoplasm where viruses replicate. In contrast to the high efficiency of IRES elements in most viruses, IRESs of cellular mRNAs usually have lower efficiency. It is now known that only under certain conditions, such as down-regulation of eIF4F activity through the sequestration of eIF4E by its binding protein (4E-BP) (Johannes *et al.*, 1999), these normally inefficient IRES become competitive when cap-dependent translation is reduced or eliminated.

Although IRES elements are of various sizes and shapes, a common feature they share is that they mediate translation in a cap-independent manner as IRES-mediated translation is insensitive to ^{m7}GTP inhibition and eIF-4E, a cap-binding factor, is usually dispensable for IRES activity (Ehrenfeld, 1996; Jackson, 1996; Jackson *et al.*, 1990, 1994, and 1995a; Jang *et al*., 1990; Kaminski *et al.*, 1994; Meerovitch *et al.*, 1989; Pelletier *et al.*, 1988a) as well as end-independent as *Encephalomyocarditis virus* (EMCV) IRES can direct efficient initiation within a circular RNA (Chen and Sarnow, 1995). IRES elements of animal viruses are usually longer and more structured than those in plant viruses. In addition, the requirement for various initiation factors during IRES-mediated initiation is also different from case to case. IRES elements in picornaviruses were shown to be capable of directly recruiting ribosomal 40S subunits with a reduced set of eIFs. Further detailed work demonstrated the interactions between IRES element and various eIFs and other protein factors such as the polypyrimidine tract binding protein (PTB) and IRES transacting factors (ITAFs) (Hellen *et al*., 2001).

1.2.1.2.3 Different groups of IRESs in picornaviruses

Within picornaviruses, several different IRES classes were identified based on secondary structure and cofactor requirements (Figure 1.4). The first group includes poliovirus and rhinovirus, and the second group includes EMCV, *Theiler's murine encephalomyelitis virus* (TMEV) and *Foot-and-mouth disease virus* (FMDV). These two groups of IRESs contain a pyrimidine-rich region in the 3' region and PTB is required for poliovirus IRES activity but is only stimulatory for EMCV IRES (Kaminski *et al*., 1995 and 1998), while $ITAF_{45}$ is strongly required by FMDV IRES (Pilipenko *et al*., 2000). In poliovirus, the initiation codon is ~160 nt downstream of the 3'-end of the IRES, and it is possible that the ribosome reaches it either by scanning or by shunting after initial attachment to the IRES (Hellen *et al*., 1994). The EMCV and TMEV initiation codons are located at the 3' border of the IRES, and ribosomes bind directly to them without scanning (Kaminski *et al*., 1990; Pilipenko *et al*., 1994). EMCV IRES was shown *in vitro* to require ATP hydrolysis, eIF 2, 3 and either eIF 4F or 4A and the central third of eIF 4G (Pestova *et al*., 1996a and b). A third group contains HCV, CSFV and BVDV, whose IRES elements are wholly distinct from both the EMCV-like and poliovirus-like groups of IRESs regarding length, sequence and structure. These HCV-like IRESs consist of four major structural domains (I-IV) and a complex pseudoknot between domains II, III, and IV. The boundaries of these IRESs extend beyond the 3'-end to the initiation codon, and IRES activity is affected by the coding sequence downstream of the initiation codon. *In vitro* reconstitution experiments showed that the minimum set of factors sufficient for

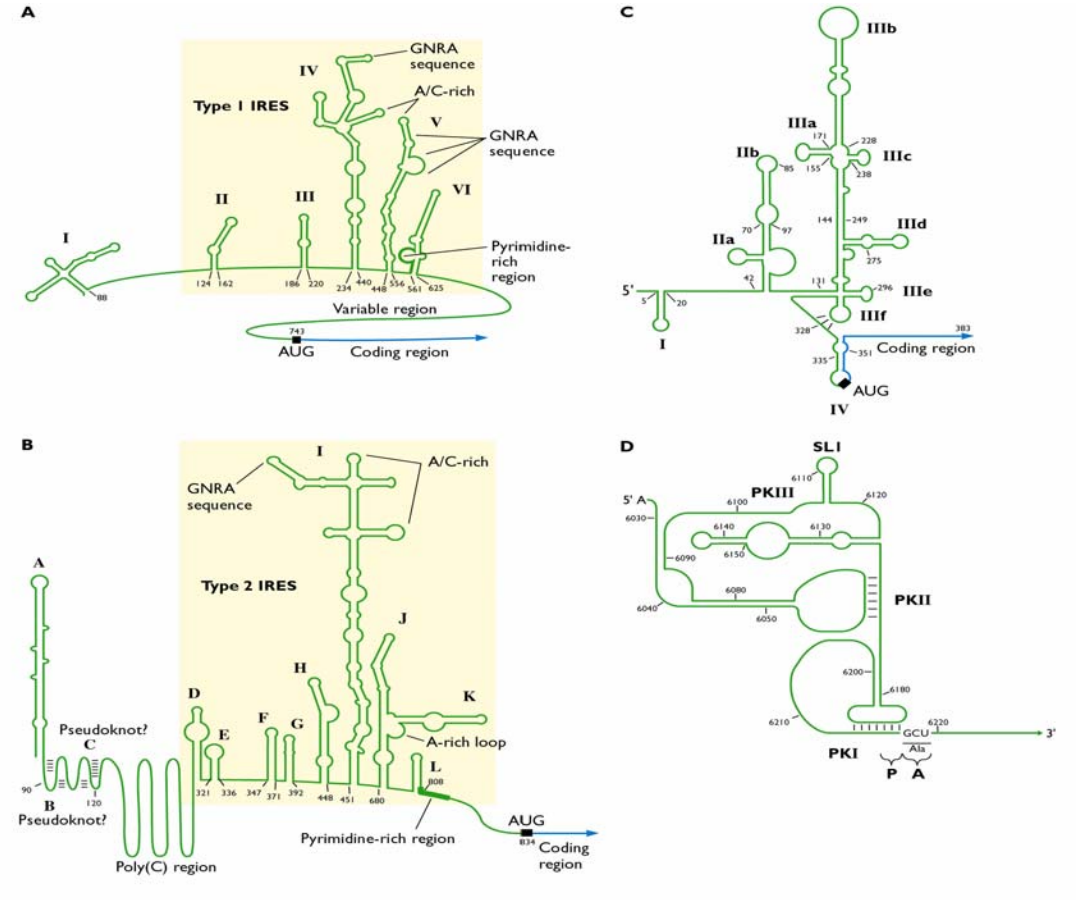

A
Type 1 IRES
GNRA sequence
A/C-rich
GNRA sequence
Pyrimidine-rich region
Variable region
AUG
Coding region
B
Type 2 IRES
GNRA sequence
A/C-rich
A-rich loop
Pseudoknot?
Pseudoknot?
Poly(C) region
Pyrimidine-rich region
AUG
Coding region
C
Coding region
AUG
D
SLI
PKIII
PKII
PKI
GCU
Ala
P
A
5'
3'

Fig. 1.4

Fig 1.4 Diagram of four types of IRES elements.

a. type 1 IRES typical for polioviruses which is located some distance upstream of the initiation codon.

b. type 2 IRES typical for EMCV and FMDV which is located close to the initiation codon.

c. type 3 IRES typical for HCV which extends beyond the initiation codon.

d. type 4 IRES typical for CrPV which involves a non-AUG initiation codon.

(Adapted from
http://www.rci.rutgers.edu/~bhillman/comparative_virology/Translation05)

the binding of 40S subunit onto the initiation codon is GTP-eIF2-Met-tRNA$_i$ (Pestova *et al*., 1998). However, eIF3 is likely to be required *in vivo* as it was reported to be associated with free 40S subunits in the cytoplasm (Goss and Rounds, 1988; Sizova *et al*., 1998). Notably, 48S complex formation on HCV-like IRESs has no requirement for eIF4A, 4B, 4E, or 4G, nor any requirement for ATP hydrolysis. Structural studies also confirmed the interactions between the 40S subunit and the HCV IRES (Spahn *et al*., 2001). However, a recent study suggested that IRES binding to the 40S subunit involves RNA-RNA base pairing with 18S rRNA (Chappell *et al*., 2004). A fourth group of IRESs contains CrPV, which remarkably requires neither initiator tRNA nor initiation factors (Wilson, *et al*., 2000). This IRES binds directly to 40S subunits but in a significantly different way with a pseudoknot in the P site inaccessible to the ternary complex and the non-AUG codon in the A site (Spahn *et al.*, 2004). Most recently, Herbreteau *et al*. reported the identification of an IRES in HIV-2 gRNA driving the production of the 5'-end gag and its two isoforms. Delineation of the RNA sequence revealed that the IRES is located entirely downstream of the first AUG codon and 5'UTR is not involved (Herbreteau *et al.*, 2005).

1.2.1.2.4 Animal virus IRESs and plant virus IRESs

Animal virus IRES elements are usually long (200-500 nt), structured and located in the 5' UTR. By contrast, those in plants are much smaller, less structured, and sometimes located in the 3' UTR (reviewed by Kneller, *et al*., 2006). So far the *Potyvirus*, *Tobamovirus*, *Polerovirus*, *Tombusvirus*, and *Luteovirus* have been reported to contain IRES elements (Gallie, 2001; Ivanov *et al.*, 1997; Jaag *et al.*, 2003;

Monkewich *et al.*, 2005). *Potyviruses* resemble picornaviruses in that they have a VPg (viral protein, genome-linked) at the 5' end and a poly (A) tail but their 5' UTRs are much shorter and less structured without multiple upstream AUGs as in picornaviruses. The VPg was proposed to play a role in direct recruiting translation factors, probably through the binding to eIF4E and eIFiso4E as supported by substantial evidence from protein interaction assays *in vitro* and *in vivo* (Grzela *et al*., 2006; Léonard *et al*., 2000, 2004; Wittmann *et al*., 1997). However, so far no direct evidence has been able to explain the role of these interactions in translation. In general, little sequence similarity was found among members in *Potyviridae* family (Maiss *et al*., 1989). Lacking extensive secondary structure and having low GC content (Simón-Buela *et al*., 1997), the simple IRESs may function depending on fewer host factors. Some other viruses such as *Tobacco mosaic virus U1* (TMV U1), *Crucifer-infecting tobamovirus* (crTMV), and the *Polerovirus* or *Enamovirus* genera in luteoviruses contain IRES elements within or between ORFs. Although *Polerovirus* and *Enamovirus* contain 5'-VPg, it was reported that a site within ORF 1 can initiate a small protein Rap1 in a different frame from ORF1 *in vitro* (Jaag *et al*., 2003). Previously, HCRSV (*Tombusviridae*) was reported to contain an IRES element upstream of the coat protein (CP) gene (Koh *et al*., 2003) and was shown to enhance translation in synergy with a hexa-nucleotide, GGGCAG (Koh *et al*., 2002). The authors also proposed that both elements function by direct base pairing to the ribosomal RNA.

1.2.1.3 Alternative initiation codon

It has been recognized that triplets other than AUG can function as translation initiation codon in bacteria (reviewed by Kozak, 1983). GUG, UUG, AUU and ACG have all been shown to be utilized as individual start codons in *Escherichia coli* and viruses. For example, protein isoforms can be generated from alternative initiation codon within a single mRNA from *Sendai virus* (Curran and Kolakofsky, 1988). The protein isoforms have distinct function in the viral replication cycle, indicating the importance of this process on gene expression. Subsequently, this feature is shown not to be restricted to bacterial and viral genes. In the past decades, quite a few eukaryotic cellular genes have been discovered to initiate translation at a non-AUG codon. Examples include a mutant version of the mouse gene dihydrofolate reductase (*dhfr*) with an ACG initiation codon (Peabody, 1987 and 1989), an isoform of the human *c-myc* gene initiated at a CUG triplet in exon 1 instead of the original initiator AUG in exon 2 due to methionine deprivation (Hann, 1988 and 1992) and many other proto-oncogenes (reviewed by Hann, 1994). Among plant viruses, *Soil-borne wheat mosaic virus* (SBWV) (Shirako, 1998), *Rice tungro bacilliform virus* (RTBV) (Fütterer *et al*., 1996) and *Strawberry mild yellow edge virus* (SMYEV) (Thompson and Jelkmann, 2004) have been reported to utilize this strategy. These protein isoforms may serve different functions or are localized in different subcellular compartments.

The efficiency of non-AUG initiation on natural transcripts *in vivo* varies considerably. Initiation from non-AUG codon in c-myc is about 10-15% as efficient as that from AUG codon in cell cultures (Hann *et al*., 1988), while the human

transcription enhancer factor-1 (TEF-1) protein appeared to initiate exclusively from a non-AUG codon (Xiao *et al*., 1991). Therefore, specific features of the natural transcripts seem to influence the efficiency of non-AUG initiation. Indeed, mutants of the authentic initiation codon of *Sendai virus* P/C mRNA is able to initiate translation from non-AUG codon while those of the human erythrocyte membrane protein 4.1 and platelet-derived growth factor beta polypeptide (PDGF-B chain) mRNA are not (Gupta *et al*., 1994), suggesting non-AUG initiation is mRNA species dependent.

In 1989, Peabody did experiments showing that almost all single nucleotide variants of AUG codon could direct translation initiation *in vitro*. At the same time, Dasso and Jackson (1989) and Kozak (1989b) also confirmed that other non-AUG codons could potentiate translation initiation in eukaryotic system. In addition to the intiator triplets, the flanking sequence also appears to affect efficiency of non-AUG initiation (Peabody, 1987; Kozak, 1989b), suggesting that interaction with nearby nucleotides might compensate for a weak codon-anticodon interaction. Initiation at a GUG codon was more efficient with A^{+5} and initiation at an AUG or CUG was shown to be favored by A^{+5} and U^{+6} (Boeck and Kolakofsky, 1994; Grunert and Jackson, 1994). However, Kozak (1997) refuted the above conclusion by showing that the augmentation was actually caused by G^{+4} but not A^{+5} or U^{+6}. These contradictory results imply that additional factors may contribute to recognition of non-AUG codon. It was further suggested that RNA structure or other *cis*-acting elements within the vicinity could determine non-AUG initiation sites and increase translation initiation at these codons (Kozak, 1990; Prats *et al*., 1992).

The detailed mechanism, however, still remains unclear. It is probably a very complicated process involving *cis*- as well as various *trans*-factors. According to the scanning model, when the 43S complex reaches the first AUG codon, GTP hydrolysis results in the dissociation of the eIF2-GDP complex stimulated by eIF5 (Asano *et al*., 2000; Das and Maitra, 2002; Hershay and Merrick, 2000) and signifies a codon-anticodon interaction between Met-tRNA$_i^{Met}$ and the start codon (Unbehaun *et al*., 2004). Recognition of a non-AUG initiator is possibly mediated by a weaker interaction between the start codon and initiator tRNA with a lower level of GTP hydrolysis. However, there are other evidences showing that eIF2-GTP-Met-tRNA$_i^{Met}$ is not the unique existing ternary complex. GUC and CAG were shown to be able to initiate translation in COS 1 cells, respectively, with valine or glutamine (Drabkin and RajBhandary, 1998). Therefore, it is possible that initiation can occur in the absence of Met-tRNA$_i^{Met}$.

Several eukaryotic initiation factors play critical roles on the selection of translation initiation codon. Biochemical evidence showed that eIF1 and eIF1A promote scanning and formation of 48S complex at the start codon (Asano *et al*., 2000; Hershey and Merrick, 2000; Pestova *et al*., 1998a). In addition, eIF1, eIF2, eIF3 and eIF5 have been implicated in AUG selection, as mutations in these factors increase initiation at non-AUG codons in yeast (Donahue *et al.*, 1988; Hashimoto *et al*., 2002; He *et al*., 2003; Valášek *et al*., 2004). Fekete *et al*. (2005) provided further genetic evidence that eIF1A C-terminus is important in initiation codon selection as the deletion mutations appeared to increase initiation at non-AUG triplets on *HIS4* mRNA

in vivo.

1.2.1.4 Termination reinitiation

Not all proteins in a cell are expressed in equivalent amounts regarding the various functions of individual proteins. Some of this phenomenon depends on rate of transcription of mRNA, which renders different levels of the mRNA. Alternatively, some mRNAs may be translated inefficiently. An example of this is the yeast *GCN4* (general control non-derepressable) mRNA that contains four small ORFs upstream of the *GCN4* start codon (Mueller and Hinnebusch, 1986; Hinnebusch, 1997). Expression of *GCN4* is via a 5' cap-dependent process termed re-initiation (see review by Hinnebusch, 1997). Further analysis indicated that the upstream ORFs (uORFs) were truly expressed (Hinnebusch *et al.*, 1988; Grant *et al*., 1994; Gaba *et al.*, 2001), although no evidence showed that the ribosomes which complete the uORF translation are exactly the same ones which initiate *GCN4* expression.

Current understanding of re-initiation may be summarized as follows: when 80S ribosome complex reaches the termination codon of uORF, the 60S subunit is thought to be released (although no direct evidence has been obtained), while the 40S subunit remains bound to the mRNA, resumes scanning, and may initiate another round of translation at a downstream AUG codon.

For the re-initiation to occur, the 40S subunit must reacquire Met-tRNA$_i$ which can be a point of control. Increasing the distance between the ORFs (Abastado *et al*., 1991a and b; Kozak, 1987b; Hinnebusch, 1997) can promote Met-tRNA$_i$ binding. Experiments in C/EBPβ showed that an AUG codon too close to the uORF

was skipped possibly because ribosomes have not yet reacquired Met-tRNA$_i$ (Calkhoven *et al*., 2000). Early experiments suggested that ribosome can move backwards to re-initiate at an AUG codon upstream of the termination site of uORF (Peabody *et al*., 1986a and b). However, results from more recent studies have refuted this view (Babik *et al*., 1999; Byrne *et al*., 1995; Ghilardi *et al*., 1998; Kos *et al*., 2002; Kozak, 2001; Lee *et al*., 1999) in that the overlapping stop codons of uORFs rendered the strongest inhibition on downstream ORF translation.

Efficiency of re-initiation is also dependent on the sizes of uORFs in eukaryotes. In most cases reinitiation occurs with a small uORF. Evidence supporting this view comes from plant *S-adenosylmethionine decarboxylase* (AdoMetDC), as shortening the uORF from 53 codons to 25 codons resulted in a five-fold increase in translation efficiency (Hanfrey *et al*., 2002). This result could be explained as an alternative shorter uORF with its AUG in a weak initiation context caused little inhibition, which is not contradictory. Another work on the mouse *Snurf-Snrpn* transcript also elegantly demonstrated the effect of uORF size on re-initiation (Gray *et al*., 1999; Tsai *et al*., 2002). With 71 codons of the upstream *Snurf* cistron, a very low level of re-initiation might account for translation of the downstream *Snrpn* cistron. When the start codon ATG of the *Snurf* was mutated to AGG, translation of *Snrpn* was elevated more than 15-fold. Currently the only apparent exception is *Cauliflower mosaic virus* (CaMV) (Park *et al*., 2001; Poogin *et al*., 2000). The reason why the size of uORF can restrict re-initiation is not known, but a possible explanation may be that certain initiation factors dissociate from the ribosome only gradually during the

course of elongation. If the elongation phase is brief, the factors required for re-initiation would still be present when the 40S subunit resumes scanning. However, it is hard to decide a cutoff size of uORF as the permissible size is likely to vary depending on features such as secondary structure and codon usage that affect the rate of elongation.

In some other cases, re-initiation can be regulated by *cis*- or *trans*-element. For example, re-initiation of ORF 2 of *Calicivirus rabbit hemorrhagic disease virus* (RHDV) is termination-dependent and is greatly influenced by an 84-nt RNA sequence upstream of the start codon of ORF2 (Meyers, 2003). Mammalian AdoMetDC mRNA produces a small peptide from uORF which is believed to interact with ribosomes and thus prevents re-initiation (Mize and Morris, 2001). In CaMV, the role of TAV (transactivator/viroplasmin protein) has been well characterized. Re-initiation on the polycistronic 35S RNA and its spliced versions is activated in the presence of TAV (Bonneville *et al*., 1989; Fütterer *et al*., 1990, 1991, and 1992; Kiss-László *et al*., 1995; Pooggin *et al*., 2000).

Many lines of evidence have shown the involvement of eukaryotic factors in re-initiation. Recent *in vitro* data showed that efficient re-initiation was largely driven by eIF4F or central eIF4G (Poyry *et al*., 2004). The *GCN4* case also clarified the role of eIF2 in recognition of the second start site for 40S subunit (Hinnebusch, 1997). In addition, eIF3 is also suggested to promote loading of the ternary complex onto the 40S subunit. In the case of CaMV, some host proteins such as subunit g of eIF3, ribosomal protein L24 and L18 were identified to interact with TAV allowing

translation of polycistronic mRNAs by re-initiation (Leh *et al*., 2000; Park *et al*., 2001).

1.2.1.5 Ribosome shunting

Ribosome shunting is a nonlinear scanning mechanism. It is cap-dependent but ribosomes can obviate a large intervening region which may contain substantial structural elements to directly "land" on to an initiation triplet. It is first described in CaMV (Fütterer *et al.*, 1990) and was shown to be dependent on a short upstream ORF. Shunting in CaMV can also be promoted by the multifunctional TAV. Other intensively studied cases of shunting include *Adenovirus*, *Sendai virus*, *Papillomaviruses*, and possibly some cellular genes such as hsp70 and *c-fos* (Yueh and Schneider, 2000). The shunting hypothesis was further supported by the fact that anti-sense oligonucleotides targeting 5' or 3' proximal region of leader inhibit translation but not those targeting the central part of the leader (Schmidt-Puchta *et al*., 1997), indicating a non-linear scanning model.

Extensive analysis of the effects of insertions of a strong stem interfering with scanning ribosomes, as well as start codon insertion, revealed shunt take-off and landing sites flanking the bypassed region of the leader. Mapping of the shunt landing site in CaMV or RTBV suggested that formation of the leader hairpin structure promotes shunting by bringing the shunt landing site upstream of the first main ORF into close proximity with a shunt take-off site downstream of the uORF (Dominguez *et al*., 1998).

Studies on CaMV as well as other pararetroviruses also showed that the

upstream ORF has critical effect on shunting based on the fact that replacement of the uORF start codon with UUG almost abolished the translation of the main ORF (Fütterer *et al*., 1993). By comparison of uORFs of different lengths revealed an optimal length of uORF for efficient shunting of 2 - 10 codons and optimal distance between the stop codon of uORF and the base of the leader hairpin of 5-10 nucleotides (Dominguez *et al*., 1998; Hemmings-Mieszczak *et al*., 2000; Pooggin *et al*., 1998).

However, it has not been reported from viruses other than *Caulimoviridae* that shunting is dependent on uORFs. With *Adenovirus*, a complementarity between the 5' UTR and 18S ribosomal RNA was believed to be essential for shunting (Yueh and Schneider, 2000), indicating a new clue that shunting might be induced by stalling, mediated by the complementarity region or the stem structure. Similar to CaMV, *Adenovirus* also encodes a protein L4-100K which is involved in ribosome shunting which is coupled with tyrosine phosphorylation (Xi *et al*., 2004 and 2005).

1.2.2 Programmed Ribosomal Frame-shifting (PRF)

1.2.2.1 Introduction

Translation elongation process is very complicated and potentially there are many ways this process can go wrong. Generally there are two kinds of errors during elongation: missense errors in which one amino acid is replaced by another, and processivity errors which lead to premature termination. Missense errors, 10^{-3} – 10^{-4} in frequency, occur during aminoacyl-tRNA selection and most are not harmful as they do not eliminate protein function. In contrast, processivity errors, 10^{-4} – 10^{-7} in

frequency, usually result in truncated and in most cases non-functional proteins (Donner and Kurland, 1972; Edelmann and Gallant, 1977a and b; Jelenc and Kurland, 1979; Jørgensen *et al*., 1993; Kurland and Gallant, 1996; Loftfield and Vanderjagt, 1972; Parker *et al*., 1983). There are two types of processivity errors: premature termination of translation and translational frame-shifting. Premature termination is far more common (Kurland *et al*., 1996) and often occurs by RNA editing (Menninger, 1977). As for frame-shifting, ribosomes change the reading frame either in the 5' direction (-1/-2 frame-shifting) or in the 3' direction (+1/+2 frame-shifting) on the mRNA. In this case, two proteins are produced with an identical N terminus up to the frame-shift point but different beyond that point since they have different read-out of the genes due to frame-shifting. In the shifted frame, ribosomes may encounter a stop codon early. The in-frame protein is always more abundant than the shifted protein because the frame shifting error occurs rarely.

Frame-shifting are of two types: "spontaneous" frame-shifting and programmed ribosomal frame-shifting (PRF). "Spontaneous" errors are 1,000-10,000 folds less frequent than PRF and efficiency of PRF can reach more than 90% (Farabaugh, 1996a and b). Unlike spontaneous errors, PRF occurs at specific mRNA sequence which is termed as the "slippery sequence". Efficient frame-shifting requires additional stimulatory elements, such as a secondary structure of RNA downstream, or a *cis*-sequence upstream.

PRF has been well characterized in retrotransposons (Belcourt *et al*., 1990; Farabaugh *et al*., 1993), bacteria (Blinkowa *et al*., 1990; Craigen *et al*., 1986; Gurvich

et al., 2003), animals (Beckenbach *et al*., 2005; Matsufuji *et al*., 1995) and animal viruses (Baranov *et al*., 2005; Brierley *et al*., 1989 and 1992; Horsburgh *et al*., 1996; Jacks *et al*., 1988a and 1988b), and is mechanistically diverse. Examples of +1 frame-shifting include GAG3 and POL3 (*GAG3-POL3*) genes of retrotransposon Ty3 of yeast (Farabaugh *et al*. 1993), mammalian ornithine decarboxylase (ODC) antizyme (Matsufuji *et al*., 1995) and thymidine kinase (TK) gene of *Herpes simplex virus* (HSV) (Horsburgh *et al*., 1996).

1.2.2.2 -1 frame-shifting

A -1 frame-shifting takes place in most cases, for example in many retroviruses such as the *gag-pol* gene of *Human immunodeficiency virus 1* (HIV-1; U UUU UUA) (Jacks *et al*., 1988b), the 1a/1b gene of coronavirus *Infectious bronchitis virus* (IBV; U UUA AAC) (Brierley *et al*., 1989) and SARS-CoV (U UUA AAC) (Baranov *et al*., 2005; Dos Ramos *et al*., 2004). In addition, *E. coli* chromosomal gene *dnaX*, bacterial insertional elements (IS elements), and some other viruses have been reported to use -1 frame-shifting (Flower and McHenry, 1990; Gramstat *et al*., 1994; Mejlhede *et al*., 2004; Prufer *et al*., 1992; Rettberg *et al*., 1999).

A well characterized example of -1 frame-shifting among viruses is the HIV-1, in which a conserved heptanucleotide motif was proposed as *X XXY YYZ* where *X*=A, G or U, *Y*=A or U and *Z*=A, C or U (Jacks *et al*., 1988b), acting as the signal for frame-shifting. This sequence is where tRNA disassociates from mRNA and re-pairs with the out-of-frame codon and it applies to most, if not all, cases of -1 frame-shifting. By mutagenesis study, Brierley *et al*. showed that monotonous runs of

uridines gave significant levels of frame-shifting (Brierley *et al*., 1992) suggesting that runs of uridines in mRNAs are particularly slippery.

In addition to the features of the frame-shifting signal itself, many other elements are found to affect frame-shifting. In some cases, a secondary structure or tertiary interaction, such as a 3' pseudoknot structure, could significantly increase frame-shifting efficiency. A mechanical explanation of the RNA structure in frame-shifting was recently proposed. From cryoelectron microscopic imaging of a ribosome-mRNA pseudoknot complex, the pseudoknot is able to interact with the ribosome in a way such that it blocks the mRNA entrance channel, compromising the translocation process and leading to a spring-like deformation of the P-site tRNA (Namy *et al*., 2006). However, Wilson *et al.* showed the HIV slippery site U UUA AAC could direct efficient -1 frame-shifting in mammalian and yeast systems independent of any secondary structure (Wilson *et al*., 1988). On the contrary, the structure in *Barley yellow dwarf virus* (BYDV) with a long bulged step-loop must base-pair to a sequence 4 kb downstream to promote frame-shifting (Barry and Miller, 2002). In other cases, sequence specific elements, such as the Shine-Dalgarno (SD) like sequence (Larsen *et al*., 1994 and 1995) complementary to the 3' end of 16S rRNA or sequences partially complementary to yeast 18S rRNA (Li *et al*., 2001), are required. Furthermore, space between the slippage site and the stimulatory element also plays a critical regulatory role in frame-shifting efficiency (Larsen *et al*., 1994). Frame-shifting can also be induced by starvation for particular amino acids or special tRNA structures. Mutagenesis on tRNA showed that certain features of the anticodon

loop were both necessary and sufficient to induce the shift (Bruce *et al*., 1986; Magliery *et al*., 2001).

In terms of mechanisms, a simultaneous slippage at both ribosomal A- and P-site accounts for -1 frame-shifting in *dnaX* (Tsuchihashi and Brown, 1992), IS911 (Chandler and Fayet, 1993) and the G-T ORF of bacteriophage λ (Levin *et al*., 1993). Previously mutagenesis studies in HIV *gag-pol* gene showed that U UUU UUA, the slippery sequence, induced -1 frame-shifting occurred at both P-site and (Jacks *et al*., 1988b). But later on an alternative mode of slippage, post-translocation simultaneous slippage in the E and P site in an *E. coli* system, was proposed by Horsfield *et al.* (1995). Most recently, a similar -1 frame-shifting was identified in the SARS-CoV which contains a slippery sequence U UUA AAC and a downstream pseudoknot (Dos Ramos *et al*., 2004). However, there are exceptions to this model. In *Potato virus M* (PVM) (Gramstat *et al.*, 1994), a single P-site slippage was proposed to be responsible for the recoding in which peptidyl-tRNA slips one nucleotide backward before the A-site is occupied by aminoacyl-tRNA. In *Bacillus subtilis* cytidine deaminase gene (*cdd*; CGA AAG), an A-site re-pairing without P-site slippage was suggested to cause the frame-shifting (Mejlhede *et al.*, 1999). Similarly, a recent report showed that IBV slippery sequence induced frame-shifting in *E. coli* system with an unusual mechanism where a single tRNA slipped in the A-site (Napthine *et al.*, 2003). How this can happen without P-site tRNA slippage and whether such unusual event is restricted to prokaryotic ribosomes is still not known.

1.2.2.3 +1 frame-shifting

Since -1 frame-shifting occurs in most cases, it is studied extensively. In contrast, +1 frame-shifting is less common and sequences of +1 frame-shifting share little similarities. The few known examples include the *prf B* gene in *E. coli* (Craigen *et al*., 1986), the retrotransposon element Ty1 and Ty3 in yeast (Belcourt and Farabaugh, 1990; Farabaugh *et al*., 1993), the mammalian ODC antizyme (Matsufuji *et al*., 1995), the HSV TK frame-shifting signal (Horsburgh *et al*., 1996).

Although the *cis*-acting signal for +1 frame-shifting is diverse, some stimulatory elements are required similar to those in -1 frame-shifting. In *E. coli prf B* gene encoding release factor 2 gene (*RF2*), a combination of a translation termination codon and an upstream SD-like sequence capable of interacting with 16S rRNA was shown to enhance +1 frame-shifting (Adamski *et al*., 1993; Craigen and Caskey, 1986; Donly *et al*., 1990a and b). A possibility for this interaction in promoting frame-shifting is that it may create a pause that expands the window of time for frame-shifting, or it may play a more direct realigning role in the mechanism. As for mammalian ODC antizyme, polyamine levels play a key regulatory role in its frame-shifting (Atkins *et al*., 1975; Matsufuji *et al*., 1995; Petros *et al*., 2005; Rom and Kahana, 1994). The polyamine sensing is through the 5' element. The stop codon included in the shift site was critical for frame-shifting. Similarly, in yeast Ty3 element (GCG AGU U) slow recognition of the rare serine codon AGU in 0 frame (F_0) induced a pause of the ribosomes which is important for frame-shifting. Interestingly, frame-shifting is not increased by a pause induced by non-sense codon (Li *et al*., 2001). In addition, this event was augmented by the 3' sequence of 12 nucleotides.

Recent study showed that this function depends on strict spacing from the shift site and complementarity between this region and ribosome Helix 18 implies that the stimulator may interfere with ribosomal error correction (Li *et al*., 2001). In HSV *TK* gene, no conventional stimulatory element adjacent to slippery site common in -1 frame-shifting was identified (Hwang *et al*., 1994; Horsburgh *et al.*, 1996). A recent report also showed that ribosomal pausing had little correlation with the frame-shifting efficiency (Kontos *et al*., 2001). The possibility that the presence of remote interactive element may help increase frame-shifting efficiency cannot be ruled out. Aside from the features on the mRNAs, changes in ribosomes (Widerak *et al*., 2005) as well as the balance of cognate and near-cognate tRNAs can also affect the recoding efficiency as demonstrated by O'Connor (2002).

Mechanistically, +1 frame-shifting occurs either dependent or independent on tRNA slippage. For example, in *prf B* gene of *E. coli*, a single slippage at the P-site was shown to take place before the A-site is occupied (Craigen and Caskey, 1986). In case of the mammalian antizyme (UCC UGA U), +1 shift occurs by occlusion of the fouth nucleotide, U of the UGA rather than re-pairing of the peptidyl-$tRNA^{Ser}$ with CCU (Matsufuji *et al.*, 1995). Conversely, frame-shifting in yeast Ty1 element happens by re-pairing of the peptidyl-$tRNA^{Leu}$ with the mRNA (Belcourt and Farabaugh, 1990). In yeast Ty3 element (Farabaugh *et al*., 1993), the tRNA decoding GCG (P-site) could not base pair with the +1 frame CGA codon. It was reasoned that no peptidyl-tRNA slippage was involved.

Among all the sequences able to induce frame-shift events, certain sequences

are highly prone to shifting. For instance, it is widely believed that monotonous runs of U are very shifty. Indeed, Brierley and co-workers (1992) made mutations on IBV slippery sequence U UUA AAC (-1 frame-shift event) and found that runs of U could give a significant level of +1 frame-shifting. It is interesting that the sequence identified in the HIV *gag-pol* region (U UUU UUA) inducing efficient -1 frame-shifting is almost identical to that of SARS-CoV 3a variants (UUU UUU UA) where a +1 frame-shifting occurred. What is more intriguing is that the same sequence can cause frame-shifting in either directions and it seems that no other stimulatory elements are necessary for -1 frame-shifting (Wilson *et al*., 1988). It is not known whether such stimulatory elements are required for +1 frame-shifting and little information is available on this shifty sequence. The mechanism of this bidirectional frame-shifting and the factors which decide the direction of frame-shifting are still in the dark.

1.2.3 Read-through

Among RNA viruses, an alternative strategy is employed to achieve similar end results to frame-shifting (reviewed in Dreher and Miller, 2006). Instead of reading in a different frame, ribosomes read through an in-frame stop codon and produce a C-terminal extended protein. This phenomenon is frequently encountered among plant viruses. A well characterized example is TMV. Expression of the C-terminal portion of the RNA-dependent RNA polymerase (RdRp) p183 requires read-through of the N-terminal p126 ORF stop codon (Beier *et al*., 1984; Pelham, 1978). Translation studies have revealed that the sequence UAGCARYYA (UAG is the stop codon of

p126, R is purine and Y is pyrimidine) was sufficient for 5% wild-type level of read-through (Namy *et al*., 2001; Skuzeski *et al*., 1991). A second well-known example is found in murine leukemia virus (MuLV) gag-pol protein, where the in-frame UAG codon of the gag gene is suppressed to produce a gag-pol fusion protein required for reverse transcription (Yoshinaka *et al*., 1985). There are other viruses using this strategy including most genera of *Tombusviridae* and many rod-shaped fungus-transmitted viruses (Brault *et al*., 1995 and 2005; Reavy *et al*., 1998; Tamada *et al*., 1996). In addition, many members in the alphavirus genus have a "leaky" UGA codon seperating the non-structural proteins nsP3 and nsP4 (Li and Rice, 1989).

This event is also known as suppression of termination which has been studied from two points of view, the *cis*-elements that favor read-through and the nature of the suppressor tRNAs. Different from the read-through signal in TMV, *Luteoviruses* employ a cytidine-rich repeat $(CCNNNN)_{8-16}$ as *cis*-acting signal for ribosomes to read through the stop codon. In the monopartite RNA genome of carom-, necro-, and tombusviruses, a conserved GDD motif in RdRp was characterized but its implication is still in the dark. So far many suppressor tRNAs have been isolated from tobacco, wheat and lupin (Beier *et al*., 1984; Urban and Beier, 1995; Zerfass and Beier, 1992a and b). Interestingly, up to date only UAG and UGA suppressible codons have been described although experiments in TMV suggest that there may be a tRNA capable of suppressing the UAA codon (Ishikawa *et al*., 1986). However, compared to other translational regulatory events, read-through is poorly understood and it is still not

known how ribosomes can read through upon the diverse read-through signals.

1.3 SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS (SARS-COV)

It was until 2003 a world wide outbreak of severe acute respiratory syndrome (SARS) did coronaviruses draw great attention from the world. Before that coronaviruses are thought to cause serious disease in farm animals and pets but nothing more than the common cold in humans. Finally, the culprit behind SARS was identified as a new member of the coronavirus genus, SARS-CoV (Enserink, 2003; Enserink and Vogel, 2003; Fleck, 2003; Krokhin *et al*., 2003; Navas-Martin and Weiss, 2003; WHO report, 2003).

Coronaviruses are typically associated with respiratory, enteric, hepatic and central nervous system diseases. In humans, they cause mainly upper-respiratory-tract infections, and are responsible for a large proportion of all common colds. Apparently, SARS-CoV spreads in droplets but its efficiency of infection seems to be low. It was indicated that the virus could have transmitted from palm civets and raccoon dogs based on the discovery of a virus closely related to SARS-CoV from these animals. The reverse could also happen.

1.3.1 Classification of SARS-CoV

Coronaviruses are enveloped plus-strand RNA viruses with about 30,000 nucleotides of genome, the largest RNA genomes, and characterized by a corona-like appearance in the electron microscope (Fig 1.5). They have been grouped into three

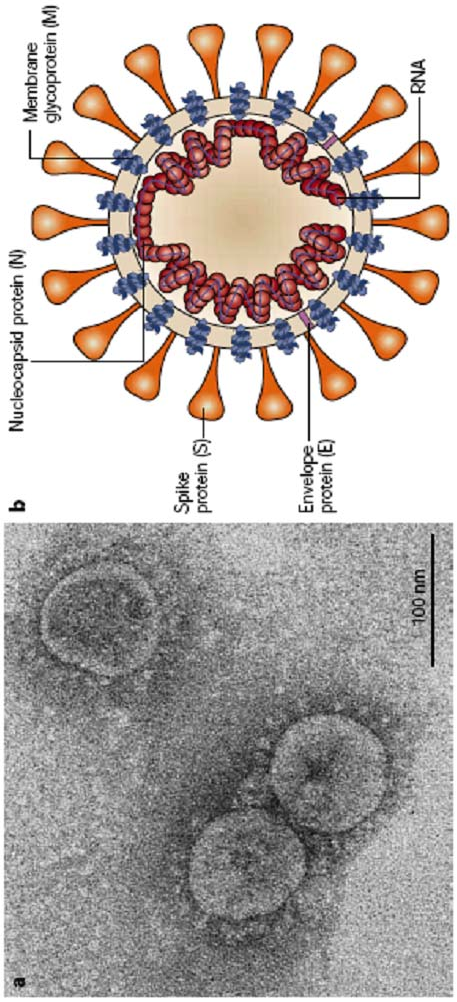

a
100 nm
b
Nucleocapsid protein (N)
Membrane glycoprotein (M)
RNA
Spike protein (S)
Envelope protein (E)

Fig. 1.5

Fig 1.5 Morphology of the SARS coronavirus.

a. Electron micrograph of the virus that was cultivated in Vero cells.

b. Schematic representation of the virus. A lipid bilayer comprising the spike protein, the membrane glycoprotein and the envelope protein encloses the helical nucleocapsid, which consists of the nucleocapsid protein associated with the viral RNA. (Adapted from Stadler *et al.*, 2003)

categories based on cross-reactivity of antibodies backed up by genetic data. Phylogenetic analysis of the predicted viral proteins indicates that the SARS-CoV does not closely resemble any of the three groups of coronaviruses (Marra *et al*., 2003; Rota *et al*., 2003) (Fig 1.6a). However, using rooted phylogenetic trees, analysis of ORF1b which is the most conserved region in the SARS-CoV genome indicates that the SARS-CoV represents an early split-off from group 2 (Snijder *et al*., 2003) (Fig 1.6b). Rappuoli group also reached this conclusion using a different approach studying less conserved proteins such as the PLpro, S, M and N proteins (Eickmann *et al*., 2003; Stadler *et al*., 2003). The close relationship of SARS-CoV with group 2 coronaviruses is reinforced by the observation that 19 out of 20 cysteine residues in the S1 domain of SARS-CoV S protein are spatially conserved compared with the group 2 consensus sequence, but much less conservation with group 1 or 3. Now there is growing number of evidence showing that SARS-CoV is closely related to group 2 coronavirus.

1.3.2 Genome organization of SARS-CoV

SARS-CoV genomic RNA is around 29.7 kb in length. At the 5'-end lies a 192-nt UTR with a 72-nt leader sequence (Fig 1.7). Two thirds of SARS-CoV genome following the 5'UTR is two overlapping ORFs 1a and 1b coding for various components involved in viral replication. The remaining 3' part of the genome encodes four structural proteins: spike protein (S); envelope protein (E); membrane glycoprotein (M); and nucleocapsid protein (N). Furthermore, eight additional non-structural proteins known as "accessory genes" are encoded in this region as well.

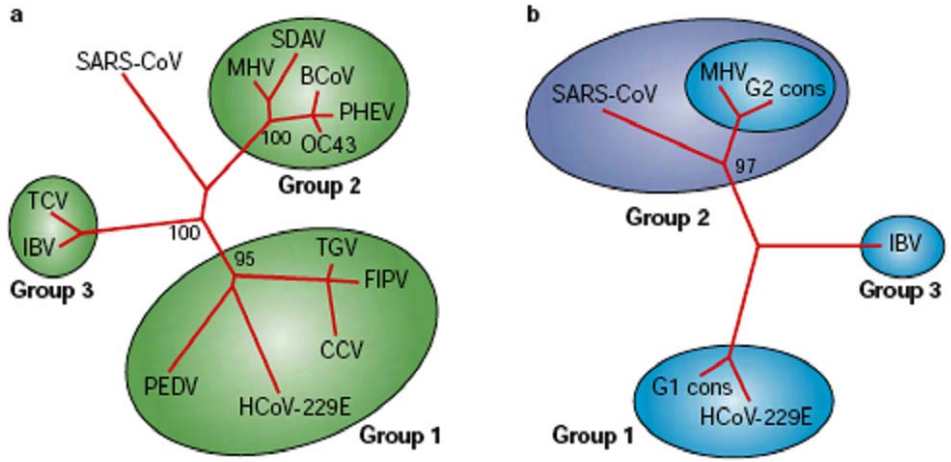

a
SARS-CoV
SDAV
MHV
BCoV
PHEV
100
OC43
Group 2
TCV
IBV
Group 3
100
95
TGV
FIPV
CCV
PEDV
HCoV-229E
Group 1
b
SARS-CoV
MHV
G2 cons
97
Group 2
IBV
Group 3
G1 cons
HCoV-229E
Group 1

Fig. 1.6

Fig 1.6 Relationship between SARS-CoV and other coronaviruses using different phylogenetic strategies.

a. Unrooted tree obtained by comparing the well-conserved polymerase protein sequence. According to this approach, SARS-CoV belongs to a new group. The tree has been constructed using the protein sequences of the RNA-dependent RNA polymerase of the following coronaviruses: *porcine epidemic diarrhea virus* (PEDV), *human coronavirus 229E* (HCoV-229E), *canine coronavirus* (CCV), *feline infectious peritonitis virus* (FIPV), *transmissible gastroenteritis virus* (TGEV), *mouse hepatitis virus* (MHV), *bovine coronavirus* (BCoV), *syaloacryoadenitis virus of rats* (SDAV), *human coronavirus OC43* (OC43), *haemagglutinating encephalomyelitis virus of swine* (PHEV), *turkey coronavirus* (TCV), *avian infectious bronchitis virus* (IBV) and SARS-CoV.

b. Tree obtained using the sequenced of the S1 domain of the spike protein. The multiple sequence alignment was constructed using consensus sequences generated from group 1 and group 2 coronaviruses, the sequence of IBV (group 3) and of SARS-CoV. The neighbour-joining algorithm was used to build the tree. Numbers represent the result of a bootstrap analysis performed with 100 replicates.

(Adapted from Stadler *et al.*, 2003)

SARS-CoV genome organization

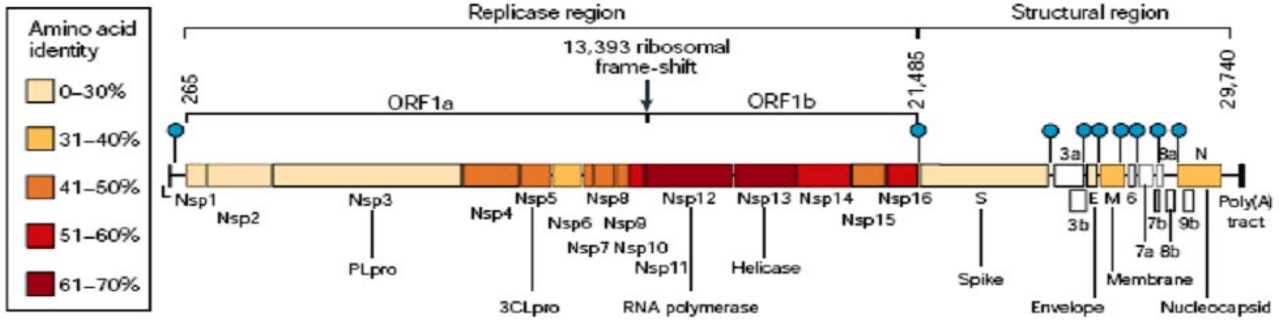


Fig. 1.7

Fig 1.7 Genome structure of SARS-CoV.

Replicase and structural regions are shown together with the predicted cleavage products in ORF1a and ORF1b. The position of the leader sequence (L), the 3' poly (A) tract and the ribosomal frame-shift site between ORF1a and ORF1b are also indicated. Each box represents a protein product (Nsp, non-structural protein). Colours indicate the level of amino-acid identity with the best-matching protein of other coronaviruses (HCoV-229E, TEGV, PEDV, MHV, BCoV and IBV). The SARS-CoV accessory genes are white. Filled circles indicate the positions of the nine transcription-regulatory sequences (TRSs) that are specific for SARS-CoV (5'ACGAAC3'). (Adapted from Stadler *et al.*, 2003)

Some of these molecules are group specific genes such as ORF 3a, 3b, 6, 7a, 7b, 8a, 8b and 9b with no homology to other coronaviruses (Marra *et al*., 2003; Rota *et al*., 2003; Ziebuhr, 2004) and they seem to be dispensable for virus viability both *in vitro* and *in vivo* since mutant viruses with these genes deleted are able to replicate (de Haan *et al*., 2002; Sarma *et al*., 2002). Distinct from other members in group 2 coronaviruses, SARS-CoV lacks the hemaglutinin esterase (HE) gene.

1.3.3 Viral replication and subgenomic RNA (sgRNA) synthesis

Figure 1.8 shows that the life cycle of a coronavirus starts when the S protein interacts with a receptor through its S1 domain. The entry, probably mediated by the S2 domain, occurs by membrane fusion. The RNA genome is then released into the cytoplasm where the replication takes place. The host translation machinery translates the overlapping ORF1a and 1b to produce a single polyprotein. After assembly, the virus particle is released by exocytosis.

Coronaviruses can synthesize a set of nested sgRNAs with common 5' and 3' sequences through a unique strategy (Lai, 1990; Lai *et al*., 1982; Pasternak *et al*., 2001; Sawicki and Sawicki, 1998; Siddel, 1995). Although most of them are polycistronic, only the 5'-most ORF of each sgRNA is translatable (Lai and Cavanagh, 1997). SARS-CoV contains nine species of mRNAs 1 – 9 (Marra *et al*., 2003; Rota *et al*., 2003; Thiel *et al*., 2003) of which mRNA 1 is the longest with an equivalent size to the genomic RNA (Fig 1.9). Each of the mRNAs contains at the 5' end a short leader sequence derived from the 5' end of the gRNA (Thiel *et al*., 2003). The leader sequence is acquired dependent on a transcription-regulatory sequence (TRS) found at

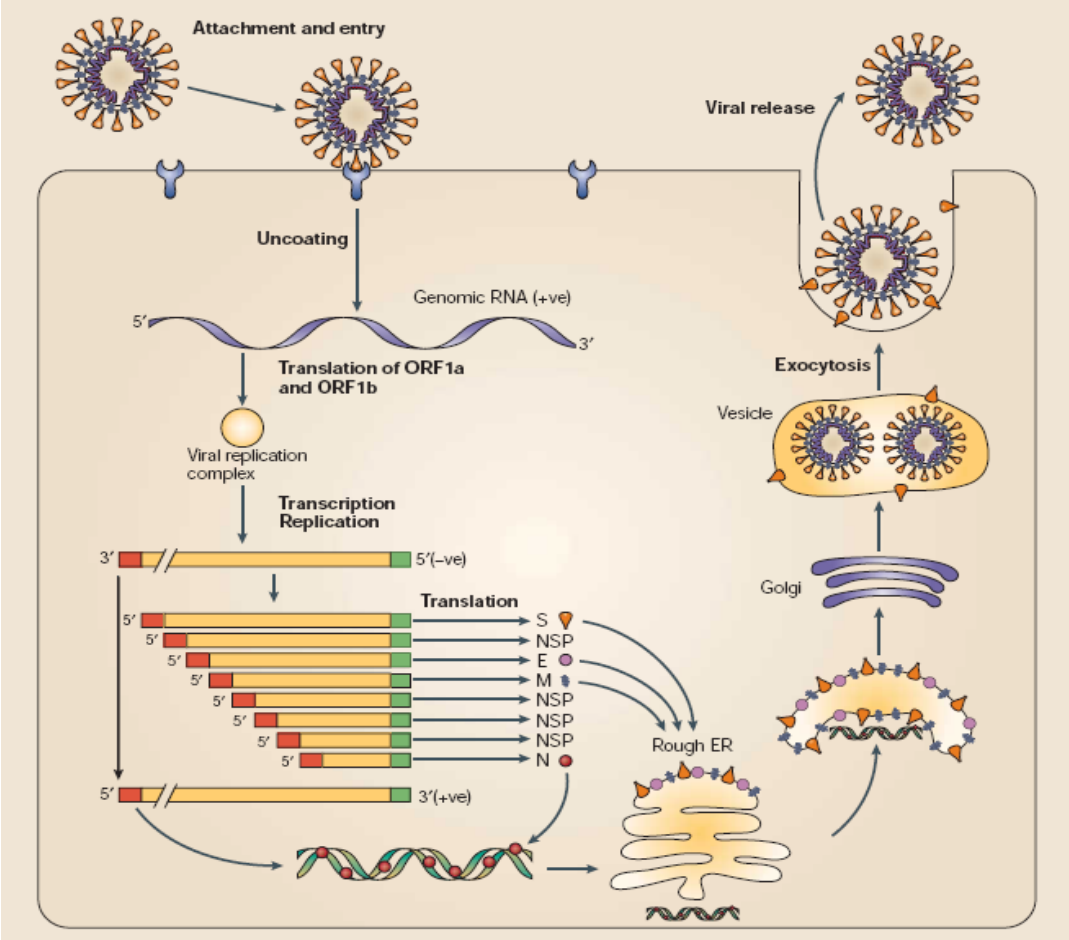

Attachment and entry
Viral release
Uncoating
Genomic RNA (+ve)
5′
3′
Translation of ORF1a
and ORF1b
Exocytosis
Vesicle
Viral replication
complex
Transcription
Replication
3′
5′(–ve)
Translation
S
NSP
E
M
NSP
NSP
NSP
N
Golgi
Rough ER
5′
3′(+ve)

Fig. 1.8

Fig 1.8 Life cycle of coronaviruses.

The virus enters the cell by membrane fusion mediated by the spike protein probably S2 domain. The RNA genome is then released into the cytoplasm where replication takes place. ORF1a/1b is first translated by host translation machinery through ribosomal frame-shifting mechanism to produce a single polyprotein, which is cleaved by virally encoded proteases. A discontinuous transcription strategy during positive-strand synthesis produces a nested set of subgenomic mRNAs, which then act as templates for the synthesis of viral proteins. Nucleocapsid (N) protein and genomic RNA assemble in the cytoplasm to form the helical nucleocapsid. This core structure acquires its envelope by budding through intracellular membranes between the endoplasmic reticulum (ER) and the Golgi apparatus. The M, E and S proteins are transported through the ER to the budding compartment. During the transport of the virus through the Golgi apparatus, sugar moieties are modified and in some coronaviruses, the S protein is cleaved into S1 and S2 domains. Finally, the virus is released from the host cell by fusion of virion-containing vesicles with the plasma membrane. The common leader sequence on the 5' end of each mRNA is shown in red (Adapted from Stadler *et al.*, 2003).

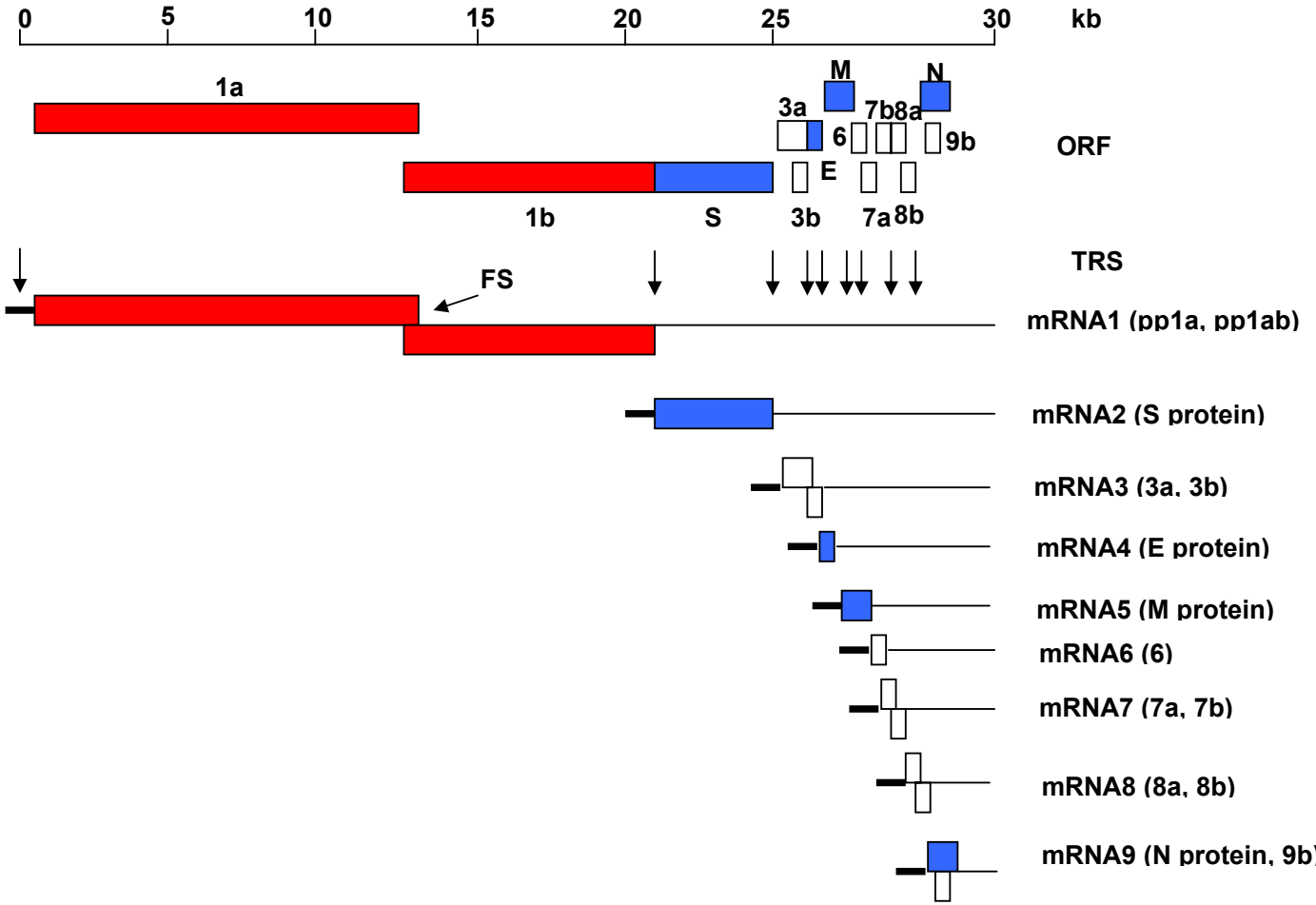

0
5
10
15
20
25
30
kb
1a
M
N
3a
7b
8a
6
9b
ORF
E
1b
S
3b
7a
8b
TRS
FS
mRNA1 (pp1a, pp1ab)
mRNA2 (S protein)
mRNA3 (3a, 3b)
mRNA4 (E protein)
mRNA5 (M protein)
mRNA6 (6)
mRNA7 (7a, 7b)
mRNA8 (8a, 8b)
mRNA9 (N protein, 9b)

Fig. 1.9

Fig 1.9 SARS-CoV genome organization and expression.

The SARS-CoV ORFs, frameshift (FS) and TRS elements, and genomic and subgenomic mRNAs are shown. Black boxes represent the 72 nt leader RNA sequence located at the 5' end of each viral mRNA. Also indicated are the viral proteins predicted to be expressed from a given mRNA's "unique" region. The non-structural genes are in red, the structural genes are in blue and the accessory genes are in white. (Adapted from Thiel *et al.*, 2003)

the 3' end of the leader RNA and preceding each translated ORF (Lai and Holmes, 2001). Currently a discontinuous model for SARS-CoV sgRNA synthesis is generally accepted that is mediated by the 5' leader sequence, the 5' UTR and 3' UTR. In SARS-CoV, a minimal conserved sequence (5' ACGAAC 3') in front of nine predicted ORFs in genomic and eight subgenomic RNAs was identified (Thiel *et al*., 2003) as TRS that is sufficient to direct sgRNA synthesis. By complementary base-pairing, this core sequence assists in the transfer of the nascent minus strand to the leader TRS involving protein-protein interactions that keep the 5' end of the genome in close proximity to the site of ongoing minus strand synthesis (Zúñiga *et al*., 2004). The number of identical nucleotides in leader TRS and body TRS regions varies, but there is no clear correlation between the extent of sequence complementarity and abundance of a given mRNA, suggestive of involvement of additional factors.

1.3.4 Expression of SARS-CoV proteins and characterization of 3a

Translation of the polyprotein 1a1b is through a -1 frame-shifting and is coupled with proteolytic processing by virus-encoded proteinases, namely the two papain-like cysteine proteases (PL1pro and PL2pro) and the 3C-like cystein protease (3CLpro) (Thiel *et al*., 2003). The resulting cleaved proteins form replicase complex and synthesize negative strand RNA and sgRNAs (Ziebuhr *et al*., 2000).

Most SARS-CoV proteins are translated by cap-dependent mechanism with the exceptions of some accessory proteins such as 3b or 7b. For example, protein 3b is predicted to be generated by IRES whereas protein7b and 9b are predicted to be generated through leaky scanning (Snijder *et al*., 2003). However, so far only protein

3a, 7a, 8a and 9b have been detected in patients' sera (Guo *et al*., 2004; Qiu *et al*., 2005; Tan *et al.,* 2004a and b).

SARS-CoV ORF 3a is on sgRNA 3 encoding a protein of 274 amino acids. Recent studies demonstrated that 3a protein is a minor structural protein and is associated with the spike protein in virions (Ito *et al*., 2005; Shen *et al*., 2005; Yuan *et al*., 2005; Zeng *et al*., 2004). Protein 3a is located in the perinuclear region and Golgi apparatus and it can interact with structural proteins S, M and E (Qiu *et al*., 2005; Yuan *et al*., 2005). In addition, it is transported to the surface of host cells and undergoes endocytosis (Tan *et al*., 2004b). Shen *et al*. (2005) showed that apoptosis in Vero E6 cells was mediated by 3a and it is also able to induce the expression of fibrinogen (Tan YJ *et al*. 2005).

During the course of cloning and expression of ORF 3a, a mixed population of clones with six, seven, eight and nine U stretches located 14 nucleotides downstream of the initiation codon was found. The existence of these ORF 3a variants was recently confirmed by isolation of a heterogeneous population of sgRNA 3 transcripts with six, seven, eight and nine Us from the sera of SARS patients (Tan TH *et al*., 2005). The significance, however, is not known yet.

1.4 HIBISCUS CHLOROTIC RINGSPOT VIRUS (HCRSV)

HCRSV is a monopartite positive-stranded RNA virus. It was first described

in a *Hibiscus* cultivar imported to US from El Salvador. The virus infects kenaf (*Hibiscus cannabinus L*.), a plant of interest to the wood pulp industry in the United States (Johnson, 2001), and is found world-wide where *Hibiscus* is cultivated (Jones and Behncken, 1980; Waterworth, 1976 and 1980) including Singapore (Wong and Chng, 1992). HCRSV can cause various symptoms on plants ranging from generalized mottle, chlorotic ringspots and vein-banding patters, to severe stunting and flower distortion (Waterworth *et al*., 1976). Based on virion morphology, genome organization, physico-chemical properties and amino-acid sequence, the virus is classified in the genus *Carmovirus* in the family *Tombusviridae*.

1.4.1 Genome organization of HCRSV

HCRSV virus particle is isometric, measuring 28 nm in diameter which contains either the genomic RNA of 3,911 nucleotides or one of the two putative 3'-coterminated subgenomic RNAs (sgRNAs, 1.5 kbp and 1.7 kbp; Huang *et al*., 2000) whose cDNA clones are denoted as pHCRSV129 and pHCRSV80. Recently, the promoter regions of each sgRNA are mapped in detail (Li and Wong, 2006). The RNA species from the virions are 5'-capped but not polyadenylated at the 3'-terminus. However, both capped and uncapped transcripts of the full-length genomic cDNA clone were infectious. Potentially, five major ORFs plus two novel ORFs p23 and p27 are produced from gRNA and two sgRNAs (Huang *et al*., 2000; Koh *et al*., 2006) (Fig 1.10).

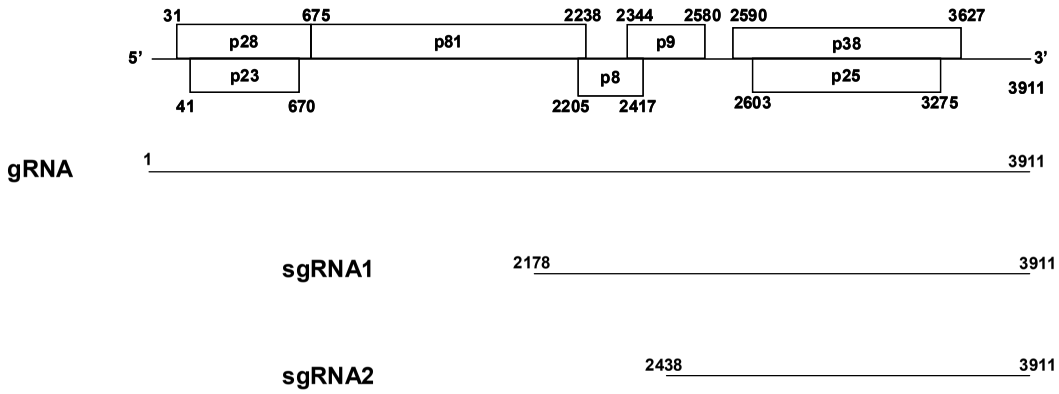

31
675
2238
2344
2580
2590
3627
p28
p81
p9
p38
5'
3'
p23
p8
p25
3911
41
670
2205
2417
2603
3275
gRNA
1
3911
sgRNA1
2178
3911
sgRNA2
2438
3911

Fig. 1.10

Fig 1.10 Genome organization of HCRSV.

The ORFs are represented by open boxes with the name of each ORF. The first and last nucleotides of each ORF are indicated. Below are the genomic and two subgenomic RNAs. The first nucleotides of each mRNA species are indicated.

Among the five major ORFs that are highly conserved in the *carmovirus* group, the most 5'-proximal two ORFs (p28 and p81) downstream the 5'UTR encode essential proteins involved in virus replication (Koonin *et al*., 1991). ORF p23 overlaps with p28 (Huang *et al*., 2000) and the expression of p23 was shown essential for viral replication in kenaf protoplasts that can be abolished by mutations in p23 (Liang *et al*., 2002a and b).The central region contains two small ORFs p8 and p9 which are predicted to function in systemic and cell-to-cell movement (Hacker et al., 1992; Huang *et al*., 2000). At the 3'-end of the genome lies the ORF p38 coding for the nucleocapsid protein followed by the 3'UTR region which acts as *cis*-element in both sgRNA and protein synthesis.

1.4.2 Expression of HCRSV proteins

Termination suppression of p28 with the Gly-Asp-Asp (GDD) motif near the amber stop codon of p28 results in the production of a C-terminal extended protein p81. This extended protein contains motifs common to RdRp (Kamer and Argos, 1984). Mutagenesis studies of *Turnip crinkle virus* (TCV) RdRp in which the expression of the 5'-proximal p28 but not p88 was abolished greatly impairs virus accumulation, suggestive of a putative role of p28 involved in TCV replication (Hacker *et al*., 1992). Complementation experiments revealed two non-viable mutant viruses which expressed only p28 or p88 can be rescued by each other (White *et al*., 1995). This further confirmed that p28 plays a very important role in TCV

replication. However, it is not confirmed yet in HCRSV whether both proteins are necessary for virus replication. Previous studies on TCV suggested that the two small ORFs p8 and p9 were conserved among all carmoviruses. They are proposed to play important roles in cell-to-cell movement (Hacker *et al*., 1992; Wobbe *et al*., 1998) and can complement in *trans* (Li *et al*., 1998). However, the translations of these two proteins were only confirmed in an *in vitro* system (Huang *et al*., 2000).

Translation of the 3'-end ORF p38 was proposed by various mechanisms. They can be generated from both gRNA and sgRNA2. In contrast to the 5'-end ORFs p28, p23 and p81 which are translated depending on 5'-UTR, the 3'-end ORFs including p38, p27 and its isoform indicated a 5'-end independence, suggesting a different strategy used for the translation of these ORFs. Previous study identified an IRES element of around 100 nucleotides residing in p9 which could regulate the translation of p38 in synergism with the 3'UTR (Koh *et al*., 2003).

Deletion studies showed that the first half of the IRES rendered the strongest initiation efficiency while the second half did not. Detailed mutagenesis studies within each part of the IRES revealed that specific sequence rather than secondary structure or a GNRA motif was crucial for the IRES function. However, no other factors involved in this mechanism have been identified and as for those putative proteins especially those overlapping with the nucleocapsid protein, the expression mechanisms are poorly studied.

1.5 OBJECTIVES AND SIGNIFICANCE

This work focuses on mechanisms of viral gene regulation at translational level using two RNA viruses – a plant virus, HCRSV, and a human coronavirus, SARS-CoV. With HCRSV, the regulation of p27 and p38 expression was studied. For SARS-CoV, frame-shift events occuring during the expression of ORF 3a variants were investigated. By studying these two examples, we aimed to advance our understanding of viral gene regulation and its role in the virus life cycles. Specifically, the objectives are:

1. To characterize the frame-shifting mechanism controlling the expression of SARS-CoV 3a variants;
2. To propose a biological function of 3a frame-shifting in both virus life cycle and viral pathogenesis;
3. To characterize the translation initiation mechanism of HCRSV p27 and p38;
4. To understand the relationship between p27 and p38 in translation initiation and the expression of these two proteins.

The mechanisms studied here could provide useful information for anti-viral drug design or anti-viral therapy. For example, studies on HCRSV may provide more details in multiple mechanisms to regulate the expression of viral genes. Work of SARS-CoV 3a frame-shifting may increase our understanding on frame-shifting mechanism and further examination of factors involved might provide additional

information on frame-shifting at molecular level.

For HCRSV, an *in vitro* translation system and a plant protoplast system were chosen instead of a plant because the latter is more complicated and more interference on viral gene expression may be present. For SARS-CoV, due to limitation to the use of full-length cDNA clone and live virus, only ORF 3a cDNA clones were used. While viral gene regulation can be studied at both transcriptional and translational levels, in this thesis I will only address the translational regulation aspects of those genes studied.

In Chapter 1, the background on translation regulation as well as the latest findings and developments in the field of translation was covered, while Chapter 2 describes the materials and methods used. Chapter 3 describes the study of SARS-CoV 3a frame-shifting mechanism. In Chapter 4, the translational control of HCRSV p27 and p38 by multiple mechanisms is covered. Finally in Chapter 5, the findings of this project are summarized, general discussion is made and future work is proposed.

CHAPTER 2. MATERIALS AND METHODS

2.1 LIST OF CHEMICALS, ANTIBODIES AND REAGENTS

Trypsin solution	Invitrogen (USA)
Penicillin-Streptomycin stock solution	Invitrogen (USA)
Fetal bovine serum (FBS)	Hyclone (USA)
bovine serum albumin (BSA)	Sigma-Aldrich (USA)
Luria-Bertani (LB) medium	Sigma-Aldrich (USA)
Antibiotics ampicillin and kanamycin	Sigma-Aldrich (USA)
5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-Gal)	Bio-rad (USA)
isopropyl-β-D-thiogalactopyranoside (IPTG)	Bio-rad (USA)
Dithiothreitol (DTT) and lysozyme	Sigma-Aldrich (USA)
Restriction enzymes	New England Biolabs (USA)
1 kb DNA ladder and 100 bp DNA ladder	New England Biolabs (USA)
Pfu DNA polymerase and *Taq* DNA polymerase	Fermentas
Tri-reagent	Molecular Research Center
Goat-anti-mouse, goat-anti-rabbit, rabbit-anti-goat antibody	DakoCytomation (Denmark)
6×His monoclonal antibody/HRP conjugates	BD Biosciences (USA)

anti-EGFP monoclonal antibody	BD Biosciences (USA)
anti-HA tag polyclonal antibody	BD Biosciences (USA)
anti-luciferase monoclonal antibody	Sigma-Aldrich (USA)
anti-FLAG M2 agarose affinity gel	Sigma-Aldrich (USA)
protein A-agarose beads	Kirkegaard&Perry Laboratories
L-[^{35}S]-methionine	Amersham Biosciences (USA)
rabbit reticulocyte lysates	Promega (USA)
RNasin®	Promega (USA)
Enhanced chemiluminescence (ECL) reagents	Amersham Biosciences (USA)

The phosphate-buffered saline (PBS) used in all experiments is composed of 137 mM NaCl, 2.7 mM KCl, 4.3 mM KH_2PO_4 and 1.4 mM $K_2HPO_4 \cdot 7H_2O$, pH 7.3.

2.2 CELL CULTURE

2.2.1 Cell lines

COS7 cells (African green monkey kidney cells) and HeLa cells (human cervical adenocarcinoma cells) were used for all the expression assays.

2.2.2 Cell culture

COS7 and HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM) (Sigma) with 4500 mg/L glucose, supplemented with 10% fetal bovine serum (FBS) (Hyclone), 100 units/ml penicillin and 100 units/ml streptomycin (Invitrogen). The cells are incubated at 37°C and 90% humidity level with 5% CO_2.

For subculturing, cells were washed with ice-cold PBS once. Subsequently, a buffered salt solution (Invitrogen) containing 0.5% (w/v) trypsin and 0.2% (w/v) EDTA was used to dissociate adherent cells from tissue culture flasks. Cells were further detached by DMEM medium and collected by centrifugation at 500×g for 5 min. After re-suspension in fresh DMEM medium with 10% FBS and 1% PS, cells were seeded in new culture flasks.

2.2.3 Cell stock preparation

To prepare cell stocks, cells were trypsinized and collected by centrifugation at 500×g for 5 min. After re-suspension in cell stock medium (DMEM medium with 10% FBS, 1% PS, and 20% glycerol), cells were transfer to 2 ml cryo-tube (NUNC) by aliquots and stored in liquid nitrogen.

2.3 MOLECULAR CLONING

2.3.1 Preparation of *E. coli* competent cells

Escherichia coli strain DH5α or BL21DE3 was streaked onto a LB plate (1% bacto-tryptone, 0.5% bacto-yeast extract, 1% NaCl and 1.5% agar, pH 7.0) and

incubated at 37°C overnight. Single colony was picked and grown in 2 ml LB broth at 37°C overnight with shaking at 220 rpm. An inoculum of the culture was diluted 1:100 into 400 ml LB medium and incubated at 37°C with vigorous shaking until the absorbance at 660 nm (A_{660}) reached 0.6-0.8. The culture was chilled on ice for 30 min and bacterial cells were collected with centrifugation at 4,000 g for 15 min at 4°C. The pellet was re-suspended in 40 ml pre-cooled 0.1 M $CaCl_2$ solution and incubated on ice for 1 hr. After centrifugation at 4,000 g for 15 min at 4°C, the pellet was re-suspended in 4 ml solution of 0.1 M $CaCl_2$ with 20% glycerol. The competent cells were stored in aliquots at -80°C.

2.3.2 Transformation of competent cells

An aliquot of *E. coli* competent cells were pre-thawed on ice before transformation. DNA plasmid (10 ng) or ligation products were added to competent cells and incubated on ice for 15-20min. The mixture was heat shocked at 42°C for 1 min and immediately chilled on ice for 2 min. One ml 2×LB broth was then added into the mixture and incubated at 37°C with vigorous shaking for 45-60 min. The suspension was subsequently spread onto LB plates containing 100 μg/μl ampicillin or 50 μg/μl kanamycin. The LB plates were incubated at 37°C overnight and single colonies were picked for mini-scale DNA preparation and further characterization.

2.3.3 Restriction enzyme digestion of DNA

The restriction endonuclease digestion mixture was prepared as follows: 1-3 μg of DNA, appropriate 1×restriction digestion buffer and 10 units of each restriction enzyme. The mixture was topped up with distill H_2O to a total volume of 20-50 μl and incubated at 37°C for 2 h. Experimental details are subjected to change according to the manufacturer's instruction for different restriction enzymes.

To prevent singly-digested vector from self-ligation, calf intestinal alkaline phosphatase (CIP) (New England Biolabs) was used to remove the phosphate group from the 5' end of the DNA strand. Typically, 1 unit of CIP was added directly to the restriction reaction mixture in the end of the 2 h digestion and incubated for another 30 min at 37°C.

2.3.4 End-filling of DNA fragment

End-filling of digested DNA fragment with 5' or 3' overhangs were carried out in a total volume of 50 μl reaction containing 2 μg of linearized DNA fragments, 1×buffer, 33 μM of each dNTP and 5 units of Klenow Fragment and incubate at 25°C for 15 min.

2.3.5 Polymerase chain reaction (PCR)

Primers used in PCR were synthesized by Proligos (Sigma). Briefly, PCR reaction mixture included 10-20 ng of DNA template, 50 pM of the forward and the reverse primers, 2.5 units of DNA polymerase, 1×PCR buffer supplemented with

1.75 mM $MgCl_2$, together with 250 μM each of deoxy-adenine (dATP), deoxy-cytosine (dCTP), deoxy-guanine (dGTP) and deoxy-thymidine (dTTP) nucleotides (Promega). The reaction was carried out by 1 cycle of denaturation at 94°C for 2 min, followed by 30-35 cycles of denaturation at 94°C for 30 sec, primer annealing for 30-60 sec and DNA extension at 72°C. The final cycle of DNA extension was carried out at 72°C for 7 min. The reagent mixture and thermal cycle set-up were slightly adjusted to optimize the product yield and specificity.

2.3.6 Site-directed mutagenesis PCR

Site-directed mutagenesis PCR was achieved using the QuikChange site-directed mutagenesis kit (Strategene). Typically, PCR reaction containing 2.5 units of *Pfu* Turbo™ DNA polymerase and 1×buffer, 30 ng of plasmid DNA template, 125 ng of each mutagenic primers and 250 μM of each deoxynucleotides (dNTPs) was carried out by 1 cycle of denaturation at 95°C for 30 sec followed by 18 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 1 min and extension at 68°C for 5-7 min. The PCR reaction was then digested by 20 units of DpnI (New England Biolabs) at 37°C for 1 hr and was ready for transformation into *E. coli* cells.

2.3.7 Gel purification of DNA

Gel purification of DNA was performed by using the Qiaquick™ Gel

Extraction Kit from QIAGEN. Restriction reaction products were separated in 1% agarose gel. Visualized under long-wavelength UV-transilluminator, the DNA bands of interest were excised with scalpels and the gel slices were dissolved in appropriate volume of dissolving buffer QG according to the manufacturer's instruction. Complete melting of gel slice was achieved by incubation at 55°C for 10 min with occasional vortex. The dissolved mixture was applied onto the Qiaquick^{TM} column for DNA purification. After washing, the DNA fragments of interest were eluted with nuclease free water to a desired volume.

2.3.8 PCR purification

PCR products were purified by Qiaquick PCR purification Kit (Qiagen). Briefly, 250 µl of PB buffer were added to a 50 µl PCR reaction. The mixture was applied onto Qiaquick column. After washing, the DNA fragments of interest were eluted with nuclease free water to a desired volume.

2.3.9 Agarose gel electrophoresis

DNA samples were separated by electrophoresis in the 1.0-2.0% (w/v) agarose gels prepared with TAE buffer (40 mM Tris-acetate and 2 mM EDTA), supplemented with 1 µg/ml ethidium bromide. 6×DNA loading buffer [0.25% (w/v) of bromophenol blue and 0.25% (w/v) xylene cyanol FF and 40% (w/v) sucrose in H_2O] was added to each DNA sample before loading to the agarose gel. DNA bands

were visualized under UV illumination.

2.3.10 DNA Ligation

The purified vector and insert DNA fragment were mixed at a molar ratio of 1:3, together with 1 unit of T4 DNA ligase (New England Biolabs) and 1×T4 ligation buffer in a total volume of 10 µl. Ligation mixture was incubated at 16°C waterbath overnight.

2.3.11 DNA preparation

The QIAprep® Spin Miniprep Kit was used for small-scale DNA preparation from *E. coli* cells. Single colony was inoculated into 5 ml LB medium with proper antibiotics at 37°C with shaking overnight. Cells were pelleted by centrifugation at 14,000 rpm for 1 min and re-suspended in P1 solution containing 100 mg/ml RNase A. Cells were subsequently lysed in P2 solution and neutralized by P3 solution prior to centrifugation at 14,000 rpm for 10 min. The supernatant was applied onto the QIAprep® spin column. After washing, the DNA was eluted with 50 µl nuclease free water.

The QIAGEN® Plasmid Midi Kit was used for large-scale DNA preparation. Single colony was inoculated into 5 ml LB medium with proper antibiotics at 37°C with shaking for 8 h and the culture was inoculated into large volume of LB medium with antibiotics with a 100 times dilution shaking at 37°C overnight. Cells were

pelleted by centrifugation at 4,000 g for 15 min at 4°C and re-suspended in P1 solution containing 100 mg/ml RNase A. Cells were subsequently lysed in P2 solution and neutralized by P3 solution prior to loading onto to the QIAfilter™ Midi Cartridge. The filtrate was subsequently applied onto the QIAGEN-tips. After washing with QC buffer, the DNA was eluted with 5 ml QF solution. DNA was then precipitated with isopropanol and re-suspended in 100-200 µl nuclease free water.

2.3.12 Automated DNA sequencing

Plasmid DNA (300 ng) was added to a mixture of total volume of 10 µl containing 3.2 pmol of primer, 1×Big Dye termination (ver 3.1). Sequencing PCR reaction was carried out by 25 cycles of denaturation at 96°C for 10 sec, annealing at 50°C for 5 sec and extension at 65°C for 4 min. The reaction mixture was purified by adding 2 volumes of 100% ethanol, 0.1 volume of 3M sodium acetate pH 5.2 and incubated on ice for 10 min before centrifugation at 14,000 rpm for 20 min. The pellet was washed twice with 70% ethanol and air-dried. Sequence was determined on an Applied Biosystems Model 3730XL automatic DNA sequencer and analyzed by Chromas program.

2.4 IN VITRO TRANSCRIPTION

Five µg plasmid DNAs were linearized and purified by phenol/chloroform and ethanol precipitated before transcription. The reaction was performed using

mMESSAGE mMACHINE® RNA transcription kit (Ambion). The mixture contains 1 μg linearized DNA, 1×transcription buffer, 2.5 mM ribonucleotides (NTPs), 2 μl Enzyme Mix in a total volume of 20 μl. Transcription was done by incubating the mixture at 37°C for 1 hr prior to DNase I digestion at 37°C for another 15 min. For digoxigenin (DIG)-labeled RNA production, the NTPs were replaced by DIG-labeled NTPs (Roche).

2.5 RNA MANIPULATION

2.5.1 Isolation of total RNA from mammalian cells

Monolayer cells were washed twice with ice-cold PBS prior to the addition of Tri-Reagent. After incubation at room temperature for 10 min with shaking, cells were scraped and transferred to microcentrifuge tubes. Chloroform (0.2 volume of Tri-Reagent) was added to the lysates and vortex vigorously before centrifugation at 14,000 rpm for 20 min at 4°C. The aqueous phase was transferred to a new tube and was mixed with equal volume of chloroform. After centrifugation at 14,000 rpm for 10 min at room temperature, the aqueous phase was transferred to a new tube and RNA was precipitated with 100% ethanol at -20°C for at least 30 min. The mixture was centrifuged at 14,000 rpm for 20 min at 4°C and the pellet was washed with RNase free 70% ethanol. RNA was re-suspended with proper volume of RNase-free water. To avoid DNA contamination in subsequent analysis, DNase treatment was

carried out and further purified using RNeasy® Mini Kit (Qiagen) according to manufacturer's instructions.

2.5.2 Reverse transcription

RNA was reverse transcribed using a first-strand cDNA synthesis kit (Expand™, Roche). One μg of total RNA was mixed with 20 pmol primer and denatured at 65° C for 10 min. The mixture was immediately chilled on ice before the addition of 4 μl of 5×Expand reverse transcriptase buffer, 2 μl of 100 mM DTT, 20 units of RNase inhibitor, 2 μl of 10 mM dNTPs and 50 units of Expand reverse transcriptase. The reaction was performed by incubating the mixture at 43° C for 1 h. An aliquot of the reaction was used for PCR.

2.5.3 RNA secondary structure prediction

Computer program MFOLD (made available from the Bioinformatics Centre, National University of Singapore) was used to predict RNA secondary structure.

2.6 EXPRESSION AND ANALYSIS OF PROTEINS

2.6.1 Transient expression of plasmid DNA in mammalian cells

COS7 cells of 80-90% confluence in 6-well plates were transfected using a recombinant vaccinia virus system. Cells were infected with recombinant vaccinia/T7 virus at M.O.I 1 for 1 h. The supernatant was removed and cells were

washed once with PBS before transfection with 0.4 µl plasmid DNA using the Effectene Transfection Reagent (Qiagen) according to the manufacturer's instruction. After incubation at 37^{o} C, 5% CO_2 for 18 h, cells were harvested for further analysis.

2.6.2 Coupled *in vitro* transcription and translation

Plasmid DNA was translated in rabbit reticulocyte lysates (RRL) using a transcription coupled-translation (TnT) system (Promega). Briefly, 1 µl circularized plasmid DNA was incubated with 25 µl RRL together with RNase inhibitor, minus-methionine amino acids mix in the presence of 50 µCi of [^{35}S]-methionine (Amersham Biosciences) at 30°C for 60-90 min. The mixture was stored at -80°C.

2.6.3 Induction of protein in *E. coli* BL21DE3 cells

Single colony of transformed BL21DE3 cells was inoculated into 5 ml LB medium with antibiotics and incubated at 37°C with shaking overnight. An inoculum of the overnight culture was diluted 1:100 into large volume of LB medium with antibiotics shaking at 37°C until A_{600} reached 0.6-0.8. Protein was induced by 1 mM IPTG at 37°C shaking for 3-4 h. Cells were pelleted by centrifugation at 4,000 g for 15 min at 4°C and lysed in 0.05 volume of PBS with 1 mg/ml lysozyme at room temperature before sonication on ice. After centrifugation at 14,000 rpm for 20 min at 4°C, supernatant and pellet were both analysed by SDS-PAGE.

2.6.4 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were prepared by adding 2×Laemmli's sample buffer [100 mM Tris-HCl pH 6.8, 4% SDS, 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol and 10% (v/v) ß-mercaptoethanol] and were heated at 95°C for 5 min before loading onto SDS-PAGE. The resolving gels of various concentrations (10%, 12% or 15%) and 5% stacking gels were cast between two glass plates. The gels were run in the Bio-Rad Mini-PROTEAN II system in a reservoir of running buffer [25 mM Tris-HCl pH 8.3, 192 mM glycine and 0.1% SDS] at 20 mA per gel.

After electrophoresis, gels containing [^{35}S]-methionine were fixed in 50% methanol and 10% acetic acid for 45 min before the signal was enhanced with Amplify^{TM} (Amersham Biosciences) for 15 min. Gels were dried under vacuum at 80^{o}C for 1 h and exposed to X-ray film (Amersham Biosciences) for autoradiography at -80^{o}C overnight.

Gels for western blotting were electroblotted to polyvinylidene difluoride (PVDF) membrane (Bio-Rad) at 4^{o}C using the Bio-Rad Trans-Blot^{TM} system. The transfer buffer contains 24 mM Tris-base, 192 mM glycine and 20% methanol.

2.6.5 Western Blotting

After the membrane was blocked in blocking buffer [PBS with 10% skim milk powder and 0.1% Tween 20] at room temperature for 1 h, it was incubated for 1 h at room temperature in a dilution of a specific antiserum in blocking buffer. After washing thrice for 10 min each with PBST [PBS with 0.1% Tween 20], the

membrane was incubated with proper IgG conjugated with hoseradish peroxidase (HRP) diluted in blocking buffer for 45 min at room temperature. After three washes with PBST, the membrane was subjected to chemiluminescence detection with the ECL kit (Amersham Biosciences). To re-probe the same membrane with another antibody, the membrane was incubated in stripping buffer [62.5 mM Tris-HCl pH 6.7, 100 mM ß-mercaptoethanol and 2% SDS] at 55^{o}C for 30-45 min, followed by blocking again.

2.6.6 Coomassie Blue staining and silver staining

After gel electrophoresis, the SDS-PAGE was incubated in Coomassie Blue staining buffer [0.05% (w/v) Coomassie blue R-250 (Sigma), 50% (v/v) methanol and 10% (v/v) acetic acid] for 1 h and de-stained in 50% (v/v) methanol and 10% (v/v) acetic acid before drying.

For silver staining, SDS-PAGE was first incubated in 50% (v/v) methanol and 10% (v/v) acetic acid for at least 3 h followed by incubation in 50% (v/v) methanol for 15 min shaking. After washing with Milli-Q water five times, the gel was incubated in 0.02% (w/v) sodium thiosulfate for 1-2 min. After another three washes with Milli-Q water, the gel was incubated with 0.2% (w/v) silver nitrate at room temperature for 45 min. After three washes with Milli-Q water, the gel was developed in developing solution [3% (w/v) sodium carbonate, 0.05% (v/v) formaldehyde (37%)]. The reaction was stopped by incubation with 1.4% (w/v)

EDTA for 10 min.

2.6.7 Densitometry

The intensities of the protein bands were measured by a GS-700 imaging densitometer (Bio-Rad) and analyzed using Quantity One^{TM} program (Bio-Rad) after normalized to a consistent background band.

2.6.8 Immunoprecipitation

Transfected mammalian cells were washed once with ice-cold PBS, and lysed with lysis buffer [140 mM NaCl, 10 mM Tris-HCl pH 8.0, 0.5% (v/v) NP-40]. Cell debris was removed by centrifugation at 14,000 rpm for 20 min at 4^{o}C. The supernatant was incubated with 1 μg of appropriate antiserum for 2 h at room temperature with shaking before incubation with 50 μl protein A-agarose beads for 2 h at room temperature. After washing with lysis buffer three times, the beads were incubated with 2×Laemmli's sample buffer and boiled for 5 min.

2.6.9 Luciferase assay

HeLa cells were seeded in 6-well dishes and incubated in at 37^{o}C, 5% CO_2 for 16-18 h. Four μg of each plasmid DNA containing luciferase genes were introduced into cells in serum-free DMEM by Lipofectamine 2000 reagent (Invitrogen) according to manufacturer's instruction. After 4 h post-transfection, the medium was replaced with fresh DMEM supplemented with 10% FBS, 1%

ampicillin and 1% streptomycin, and continue incubation for 16-18 h. Cells were washed three times with PBS, lysed with 100-300 µl of passive lysis buffer (Promega). A luciferase reporter assay was performed according to the manufacturer's instructions using the TD-20/20 luminometer (Turner Biosystems). Each transfection and luciferase assay was repeated at least three times.

CHAPTER 3. CHARACTERIZATION OF FRAME-SHIFTING MECHANISMS IN SARS-COV 3A VARIANTS

3.1 Introduction

It has long been believed that coronaviruses can cause no more than cold in human beings until a world-wide outbreak of severe acute respiratory syndrome in 2002. The 'culprit' was discovered to be a new member of the coronavirus family-SARS-CoV.

SARS-CoV contains a mono-partite positive sensed RNA of 29.7 kb in length. The genomic RNA encapcidated by the nucleocapsid protein is enclosed in a protein shell comprised of the membrane protein (M), the envelope protein (E), and the spike protein (S). These structural proteins are coded by the 3' one third of the genome RNA. The 5' two-third of the genomic RNA encodes for a poly protein 1a/1b which is produced by a -1 frame-shifting. This polyprotein is subjected to proteolytic process by virally encoded proteases. The resulting individual non-structural proteins (nsp) are believed to be involved in viral replication and subgenomic RNAs synthesis. In addition to the above structural and non-structural proteins, SARS-CoV also produces some small non-essential proteins which are called accessory proteins.

ORF 3a, located downstream of S and upstream of E, is one of the accessory genes. The 3a protein is produced from sgRNA3 species and contains 274 amino acids. It was shown that 3a was associated with the spike protein in virions indicating its minor structural character (Ito *et al*., 2005; Shen *et al*., 2005; Yuan *et*

al., 2005; Zeng *et al*., 2004). Other evidence also showed that it can interact with structural proteins S, M and E. Subcellular localization studies reveals that the 3a protein is mainly in the perinuclear region and Golgi apparatus (Qiu *et al*., 2005; Yuan *et al*., 2005) and is transported to the surface of host cells and undergoes endocytosis (Tan *et al*., 2004b). The expression of 3a *in vivo* was supported by the fact that antibodies against the 3a protein were detected in patients' sera.

3.2 Identification of initiation site of SARS-CoV 3a ORF

The predicted translation initiation site of ORF 3a on sgRNA3 is at nt 25252. According to the Kozak consensus sequence, the ^{25252}AUG is in a favorable initiation context with a G at +4 position and a C at -3 position. In addition, 18 nt downstream of ^{25252}AUG is a stretch of six uridines. However, 12 nt downstream of this ^{25252}AUG, another ^{25264}AUG is present in a less favorable context with an A at +4 position and a T at -3 position. To identify the initiation site of ORF 3a, full-length sgRNA3 containing wild type ORF 3a and overlapping ORF 3b was reverse transcribed, amplified by PCR and cloned into a vector under the control of T7 promoter, yielding construct pKTO-3a3b (Figure 3.1a). ORF 3a is 825 nt long coding for a 31 kDa protein of 274 amino acids (initiated from ^{25252}AUG) and ORF 3b is 465 nt long encoding a 17 kDa protein of 154 amino acids. The *in vitro* translation of pKTO-3a3b showed a strong band corresponding to a molecular mass

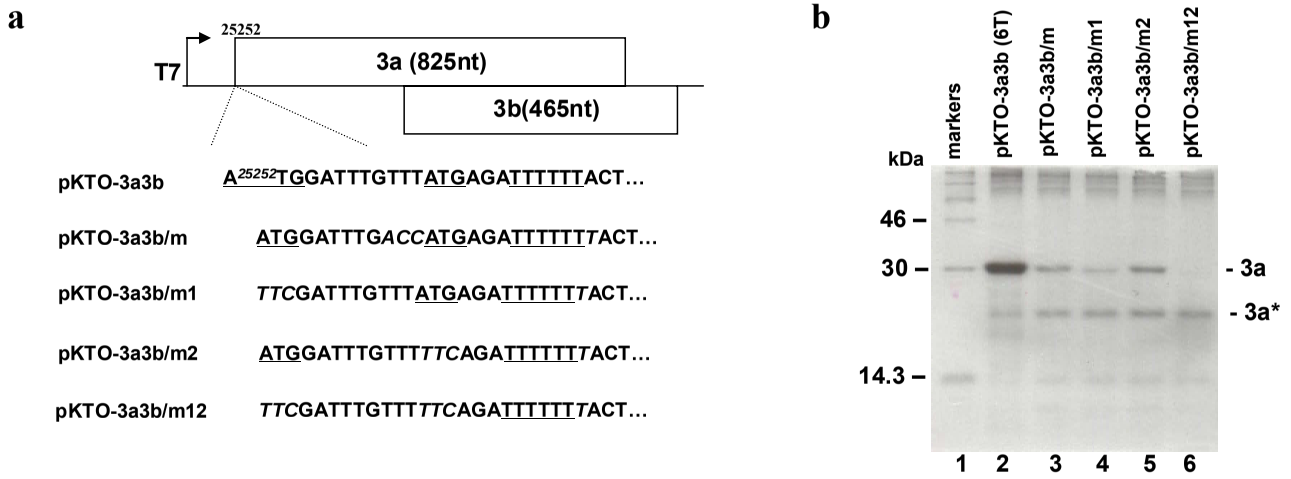

a
25252
T7
3a (825nt)
3b(465nt)
pKTO-3a3b A25252TGGATTTGTTTATGAGATTTTTTACT…
pKTO-3a3b/m ATGGATTTGACCATGAGATTTTTTTACT…
pKTO-3a3b/m1 TTCGATTTGTTTATGAGATTTTTTTACT…
pKTO-3a3b/m2 ATGGATTTGTTTTTCAGATTTTTTTACT…
pKTO-3a3b/m12 TTCGATTTGTTTTTCAGATTTTTTTACT…
b
markers
pKTO-3a3b (6T)
pKTO-3a3b/m
pKTO-3a3b/m1
pKTO-3a3b/m2
pKTO-3a3b/m12
kDa
46 –
30 –
14.3 –
- 3a
- 3a*
1 2 3 4 5 6

Fig. 3.1

Fig 3.1 Expression of SARS-3a3b mutants.

a. Diagram of SARS-CoV 3a3b constructs under the control of T7 promoter. The first nucleotide of ORF3a is indicated. Also indicated are the lengths of ORF3a and ORF3b. The six-uridine stretches and the ATG codons are underlined. All the mutations are shown in italic.

b. Expression of pKTO-3a3b and its mutants in rabbit reticulocyte lysates *in vitro*. Equal amount of plasmid DNA was used for each reaction. Polypeptides were labeled with [^{35}S]-methionine, seperated on 12% SDS-PAGE and detected by autoradiography. Bands corresponding to the full-length 3a and 3a* are indicated on the right. Numbers on the left indicate molecular masses in kilodaltons.

of 31 kDa. However, the band corresponding to 17 kDa was missing (Figure 3.1b, lane 1). Sequence analysis revealed that a point mutation at nt 25748 from C to T introduced a stop codon within ORF 3b resulting in a truncated version of the protein.

Based on pKTO-3a3b, several mutants was made to abolish either one or both of the initiation sites resulting in constructs pKTO-3a3b/m1, m2 and m12 (Figure 3.1a). In addition, the context of the second ^{25264}AUG was changed to a more favorable one at the positions of -3 to -1 from UUU to ACC, generating constructs pKTO-3a3b/m (Figure 3.1a). Interestingly, nucleotide sequencing of these constructs revealed an additional pyrimidine insertion in the six T stretch in all the mutant constructs (Figure 3.1a). Expressions of all these constructs *in vitro* showed a weak band of approximately 31 kDa with similar gel mobility to the band detected in pKTO-3a3b (Figure 3.1b, lanes 2 and 4), except in construct m12, in which both initiation sites were abolished. However, the efficiency of the translation of the 31-kDa species was much lower than that of pKTO-3a3b. In mutant pKTO-3a3b/m1 in which the ^{25252}AUG was mutated to ^{25252}TTC (Figure 3.1a), a protein band migrating a little more rapidly than the 31-kDa band, and with lesser intensity was detected (Figure 3.1b, lane 3 compared to lanes 1, 2 and 4). This suggests that the second AUG could also serve as an initiation site even with a weak context. Thus, leaky scanning is likely to occur at this position. However, in mutant pKTO-3a3b/m,

this species was not observed even when the second AUG codon is in a better context, indicating that ^{25264}AUG is not favored for translation initiation with the presence of ^{25252}AUG codon. In summary, the initiation site for ORF 3a is ^{25252}AUG instead of ^{25264}AUG.

3.3 Expression of ORF 3a variants

Analysis on SARS-CoV sgRNA3 isolated from cultured Vero E6 cells discovered a heterogeneous population of sgRNA3 with different lengths of uridine stretch (Tan TH *et al*., 2005). The additional nucleotide insertion in ORF 3a was supposed to change the ribosome reading frame and produce a much smaller gene product instead of full-length protein. The translation of the full-length 3a *in vitro* from these mutant 3a forms led us to consider the possibility of frame-shifting. It is likely that during translation of the mutant 3a transcripts, the ribosomes restore the original reading frame by frame-shifting.

To confirm the expression of the full-length 3a protein, ORF 3a with six, seven and eight Ts was cloned under the control of T7 promoter, generating constructs pSARS-3a/6T, 7T and 8T (Figure 3.2). *In vitro* expression showed the full-length 3a expression in pSARS-3a/6T, about 31 kDa in molecular mass (Figure 3.3a, lane 1). Interestingly, this protein expression was also observed in pSARS-3a/7T and 8T (Figure 3.3a, lanes 2 and 3). The efficiency of full-length 3a

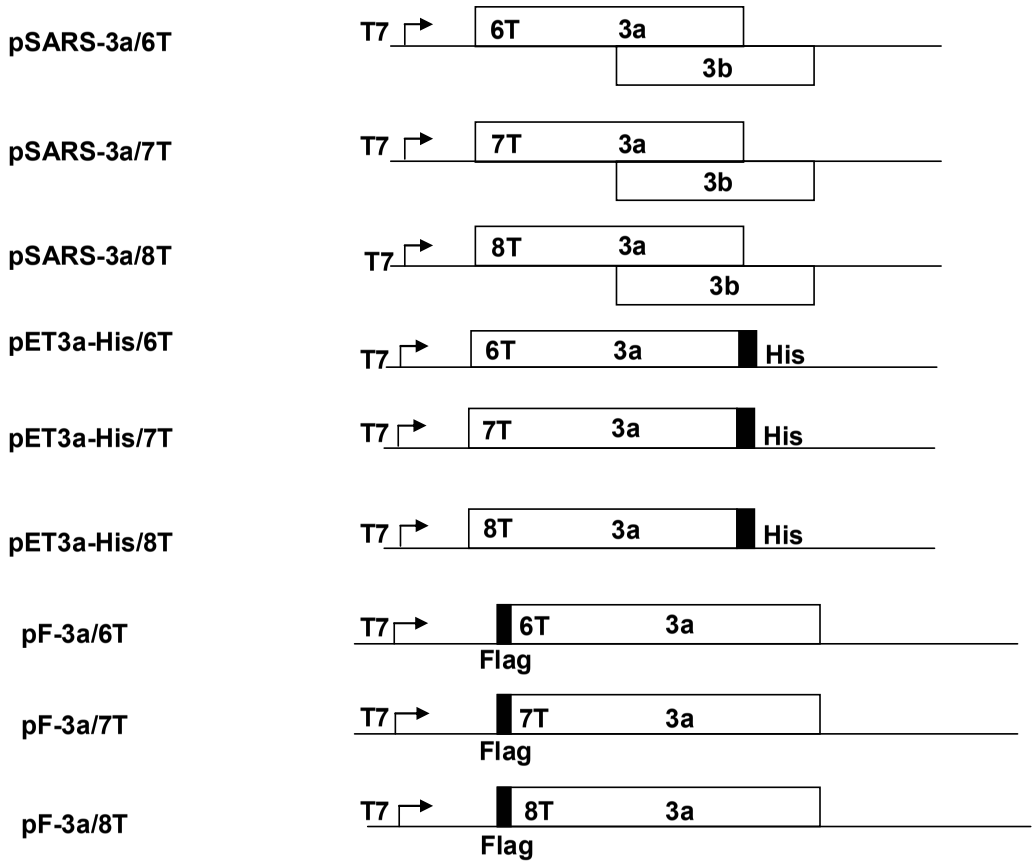

pSARS-3a/6T
T7
6T
3a
3b
pSARS-3a/7T
T7
7T
3a
3b
pSARS-3a/8T
T7
8T
3a
3b
pET3a-His/6T
T7
6T
3a
His
pET3a-His/7T
T7
7T
3a
His
pET3a-His/8T
T7
8T
3a
His
pF-3a/6T
T7
6T
3a
Flag
pF-3a/7T
T7
7T
3a
Flag
pF-3a/8T
T7
8T
3a
Flag

Fig. 3.2

Fig 3.2 Schematic diagram of SARS-CoV ORF 3a variants with six-, seven- and eight-T stretches under the control of T7 promoter.

Black boxes represent the His-tag in constructs pET3a-His/6T, pET3a-His/7T and pET3a-His/8T or Flag-tag in constructs pF-3a/6T, pF-3a/7T and pF-3a/8T. The six-, seven-, and eight-T stretches are indicated.

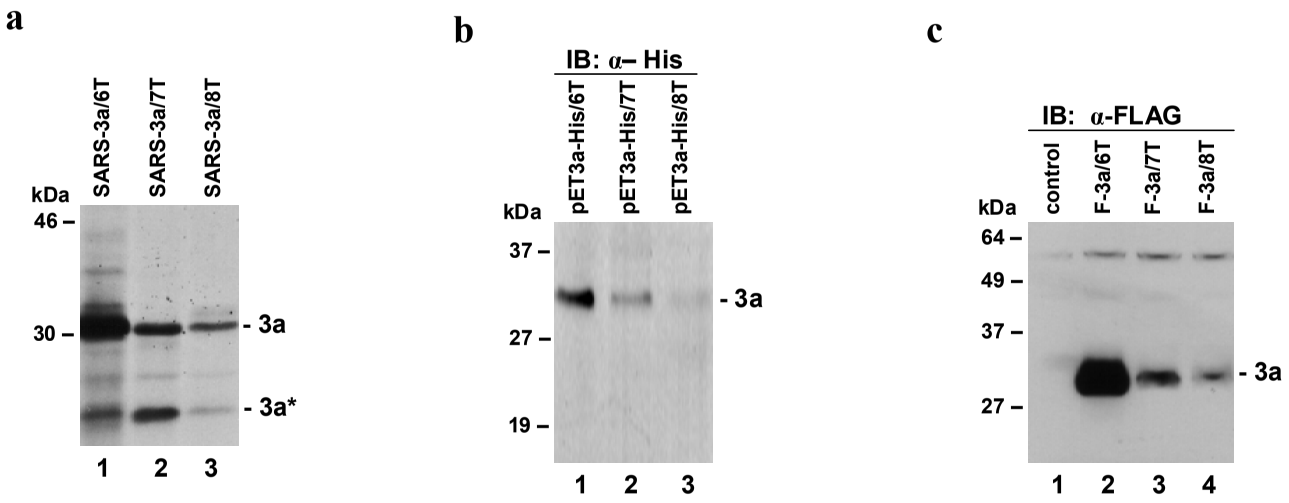

a
SARS-3a/6T
SARS-3a/7T
SARS-3a/8T
kDa
46 –
30 –
- 3a
- 3a*
1 2 3
b
IB: α– His
pET3a-His/6T
pET3a-His/7T
pET3a-His/8T
kDa
37 –
27 –
19 –
- 3a
1 2 3
c
IB: α-FLAG
control
F-3a/6T
F-3a/7T
F-3a/8T
kDa
64 –
49 –
37 –
27 –
- 3a
1 2 3 4

Fig. 3.3

Fig 3.3 Expression of SARS-CoV ORF 3a variants with six-, seven- and eight-T stretches in *in vitro* system (a), in bacteria cells (b) and in mammalian cells (c).

a. Expression of pSARS-3a/6T (lane 1), pSARS-3a/7T (lane 2) and pSARS-3a/8T (lane 3) in rabbit reticulocyte lysates *in vitro*. Equal amount of each plasmid DNA was used. Polypeptides were labeled with [^{35}S]-methionine, separated on 12% SDS-PAGE and detected by autoradiography.
b. Expression of pET3a-His/6T (lane 1), pET3a-His/7T (lane 2) and pET3a-His/8T (lane 3) in bacterial cells. Plasmid DNA was transformed into *E. coli* strain BL21, and the protein expression was induced by 1 mM IPTG. After induction for 2 h, total cell lysates were prepared. Proteins were resolved on 12% SDS-PAGE, and analyzed by western blot with anti-His antibody.
c. Expression of pF-3a/6T (lane 2), pF-3a/7T (lane 3) and pF-3a/8T (lane 4) in Cos-7 cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1) and the three Flag-tagged 3a constructs, respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-FLAG antibody.

Bands corresponding to the full-length 3a and a minor species 3a* representing an internal initiation product are all indicated on the right of each panel. Numbers on the left indicate molecular masses in kilodaltons.

expression in these mutant 3a constructs were 37% and 10%, respectively, compared to that in pSARS-3a/6T. In addition, a minor band of approximately 27 kDa, indicated as 3a*, was present in all three constructs. This band was also observed in the previous 3a mutants (Figure 3.1b). According to previous study, this minor band could have resulted from internal initiation at a downstream AUG site within the 3a ORF (Tan *et al.*, 2004b).

To confirm the expression in bacterial cells, a 6×His-tag was fused to the 3' end of ORF 3a with six, seven and eight Ts, respectively (Figure 3.2) and cloned into bacteria expression vector pET24a (Qiagen) vector. *E. coli* strain BL21 was transformed with these plasmids and proteins were induced with 1 mM IPTG. Total proteins were resolved on 12%SDS-PAGE and probed with 6×His monoclonal antibody (Santa Cruz) (Figure 3.3b). The results showed the full-length expression of 3a in all three constructs and similar level of efficiency of pET-3a/7T and 8T to the *in vitro* data (35% and 18%, respectively, of that from pET-3a/6T). Subsequently, expression from COS 7 cells transfected with the FLAG-tagged constructs also disclosed a consistent profile by Western blot with an anti-FLAG antibody (Figure 3.3c). The efficiencies of full-length 3a expression in pFLAG-3a/7T and 8T were 30% and 12%, respectively, of that in pFLAG-3a/6T. All these results confirmed that full-length 3a are translated from mutant 3a with extra uridine insertions.

The expression of full-length 3a protein in pFLAG-3a/7T was further

confirmed by immunoprecipitation experiment. Total proteins from COS 7 cells transfected with either empty vector pFLAG or pFLAG-3a/7T plasmid were incubated with FLAG M2 agarose affinity gel (Sigma). After extensive wash, the FLAG beads were incubated with 2×SDS loading buffer and were boiled. Protein expression of 3a present in total cell lysate, in the solution after binding to the beads and in the final eluate were examined on 12% SDA-PAGE and probed by a FLAG monoclonal antibody. The results in Figure 3.4 showed the detection of full-length 3a protein in fractions of pFLAG-3a/7T (lanes 2, 4, and 6) but not in mock (lanes 1, 3, and 5). All the above data suggest that ribosome frame-shifting may account for the expression of full-length 3a protein.

However, we cannot rule out the possibility of transcriptional slippage. As a matter of fact, SARS-CoV generates a mixture of sgRNA3 with different number of uridines in ORF 3a during replication. It is possible that during transcription of the plasmid DNA in cells, a heterogeneous population of 3a was also produced. Some preliminary data also suggested this possibility. However, care must be taken before a conclusion is drawn.

3.4 Identification of slippery sequences

Since the above mutations were in the uridine stretch, we proposed that this stretch was the slippery sequence. Mutagenesis study in the seven uridines was first

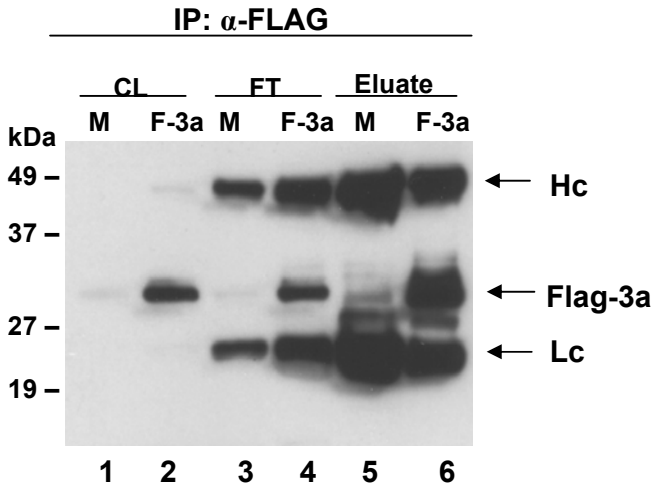

IP: α-FLAG
CL
FT
Eluate
M
F-3a
M
F-3a
M
F-3a
kDa
49 –
37 –
27 –
19 –
Hc
Flag-3a
Lc
1
2
3
4
5
6

Fig. 3.4

Fig 3.4 Immunoprecipitation of 3a proteins from COS7 cells expressing pF-3a/7T (3a/7T) or empty vector (M) by anti-FLAG M2 agarose gel.

COS7 cells were lysed in lysis buffer 18 h post transfection. After removal of the cell debris, the lysates were incubated with FLAG beads for 1 h. The unbound proteins were separated by centrifugation. After extensive washing, the 3a proteins were eluted with elution buffer. Aliquots from cell lysates, unbound portions and the final eluates were analysed on 12% SDS-PAGE and detected with anti-FLAG antibody. The bands corresponding to 3a protein are indicated on the right. Also indicated are the bands representing heavy chains (Hc) and light chains (Lc) of the antibody. Numbers on the left indicate the molecular masses in kilodaltons.

carried out based on the construct pFLAG-3a/7T, in which ORF 3a with seven were fused to a Flag-tag at the 5' terminus and cloned under the control of T7 promoter. Point mutations within the T stretch were made from T to C, resulting in constructs pF-3a/M1 to M7. To investigate whether the nucleotide A immediately downstream of the U stretch could affect full-length 3a expression, point mutation on this A to C was also made as M8 (Figure 3.5a). Figure 3.5b shows the *in vitro* translation of these plasmids in TnT system. As control, F-3a/7T showed a band equivalent to full-length 3a of ~31 kDa. A similar band was also observed in M8. On the contrary, full-length 3a expression was reduced to an undetectable level in mutant M1 to M8. In addition, the minor product 3a* was present in all the constructs (Figure 3.5b, lanes 1 to 9). To confirm the band observed is truly the 3a protein, COS 7 cells were transfected with these plasmids using a recombinant Vaccinia virus system and total proteins were separated on 12% SDS-PAGE and probed with either 3a antiserum or anti-FLAG monoclonal antibody. Cells transfected with a vector DNA pFLAG was used as negative control. As expected, full-length 3a expression was only seen in F-3a/7T and M8 (Figure 3.5c, lanes 2 and 10) but not in the mutant M1 to M7 or in the control (Figure 3.5c, lane 1 and lanes 3 to 9). These results indicate that the full-length 3a expression is dependent on the uridine stretch and any mutations within the stretch can reduce the expression significantly. These experiments provided evidence that the stretch of seven uridines is probably the slippery

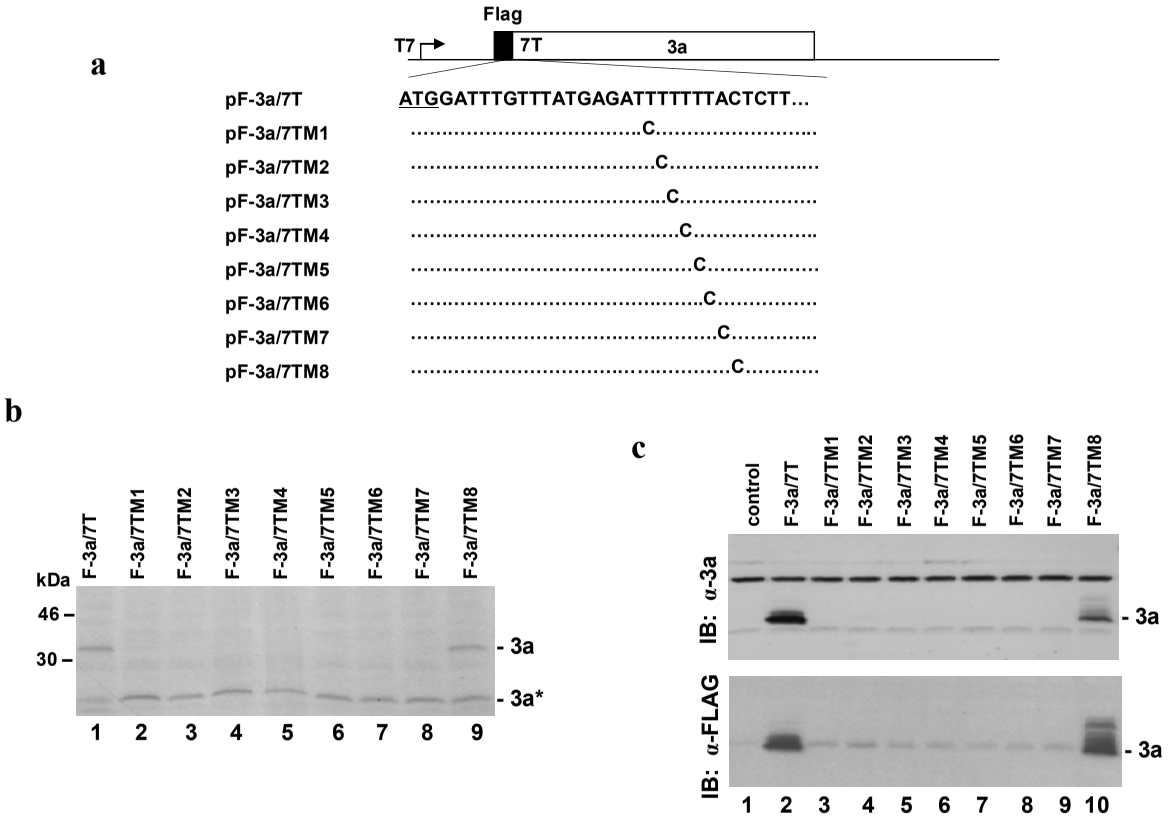

a
Flag
T7
7T
3a
pF-3a/7T
ATGGATTTGTTTATGAGATTTTTTTACTCTT…
pF-3a/7TM1
pF-3a/7TM2
pF-3a/7TM3
pF-3a/7TM4
pF-3a/7TM5
pF-3a/7TM6
pF-3a/7TM7
pF-3a/7TM8
b
F-3a/7T
F-3a/7TM1
F-3a/7TM2
F-3a/7TM3
F-3a/7TM4
F-3a/7TM5
F-3a/7TM6
F-3a/7TM7
F-3a/7TM8
kDa
46
30
- 3a
- 3a*
1 2 3 4 5 6 7 8 9
c
control
F-3a/7T
F-3a/7TM1
F-3a/7TM2
F-3a/7TM3
F-3a/7TM4
F-3a/7TM5
F-3a/7TM6
F-3a/7TM7
F-3a/7TM8
IB: α-3a
- 3a
IB: α-FLAG
- 3a
1 2 3 4 5 6 7 8 9 10

Fig. 3.5

Fig 3.5 Mutational analysis of the slippery sequence in pF-3a/7T.

a. Diagram of pF-3a/7T and the eight mutants (pF-3a/M1-M8) under the control of T7 promoter. The black box represents the Flag-tag. The initiator AUG codon for 3a is underlined. Also shown are the point mutations introduced into the seven T stretch in each mutant constructs.

b. Expression of pF-3a/7T (lane 1) and mutant constructs (lane 2-9) in rabbit reticulocyte lysates *in vitro*. Polypeptides were labeled with [^{35}S]-methionine, separated on 12% SDS-PAGE and detected by autoradiography.

c. Expression of pF-3a/7T and mutant constructs in COS7 cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with an empty control plasmid (lane 1), pF-3a/7T (lane 2) and the eight mutant constructs (lanes 3-10), respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with either anti-3a antibody (upper panel) or anti-FLAG antibody (lower panel). Bands corresponding to the full-length 3a and a minor species 3a* representing an internal initiation product are indicated. Numbers on the left indicate molecular masses in kilodaltons.

sequence inducing a +1/-2 frame-shifting.

To rule out the possibility that reversion occurs during transcription, the sequence of transcripts was analyzed by RT-PCR. Figure 3.6 showed the region of uridine stretch of each construct. No reversion or transcriptional slippage was observed, reinforcing the above conclusion.

Similarly, point mutations were also made on the uridines of pF-3a/8T as well as the nucleotide A immediately downstream of the uridines, yielding pF-3a/8TM1 to M8 and MA (Figure 3.7a). COS 7 cells were then transfected with these mutants and total proteins were analyzed with immunoblots. Similar level of full-length 3a expression in pF-3a/8T, M8 and MA was observed (Figure 3.7b, lanes 2, 10 and 11), while no 3a expression was detected in mutants M1 to M7. ß-tubulin was probed for loading control. This suggests that mutation of the first seven uridines could significantly affect frame-shifting and only the first seven uridines but not the last seem to be involved in the recoding. It is interesting that the sequence in pF-3a/8TM8 was exactly the same as that in pF-3a/7TM8 (UUU UUU UC), except that the latter lacked one nucleotide compared to pF-3a/8TM8.

3.5 Effect of 5'-extension on the frame-shifting mediated by UUU UUU U

Although the full-length 3a was detected in mutant 3a constructs, the terminated product of the zero frame (F_0) was undetectable due to its small size and


GATTTTTTTACT
3a/7T
GACTTTTTTACT
3a/7TM1
GATCTTTTTACT
3a/7TM2
GATTCTTTTACT
3a/7TM3
GATTTCTTTACT
3a/7TM4
GATTTTCTTACT
3a/7TM5
GATTTTTCTACT
3a/7TM6
GATTTTTTCACT
3a/7TM7
GATTTTTTTCCT
3a/7TM8

Fig. 3.6

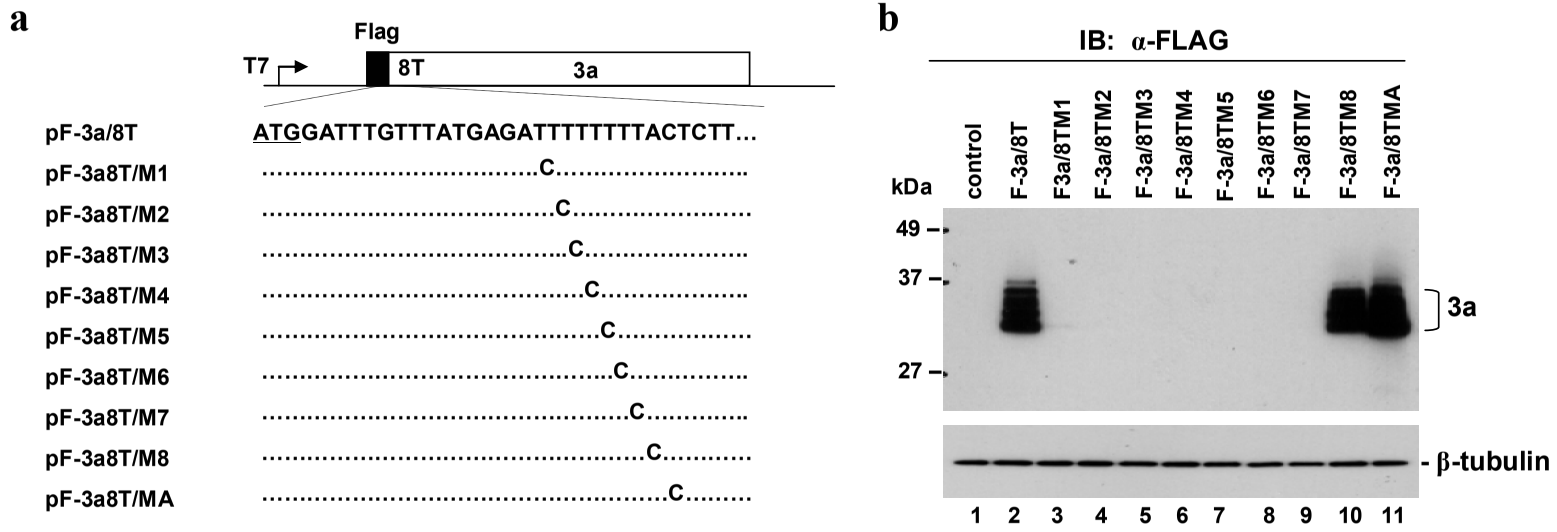

a
Flag
T7
8T
3a
pF-3a/8T
ATGGATTTGTTTATGAGATTTTTTTTACTCTT…
pF-3a8T/M1
pF-3a8T/M2
pF-3a8T/M3
pF-3a8T/M4
pF-3a8T/M5
pF-3a8T/M6
pF-3a8T/M7
pF-3a8T/M8
pF-3a8T/MA
b
IB: α-FLAG
kDa
control
F-3a/8T
F3a/8TM1
F-3a/8TM2
F-3a/8TM3
F-3a/8TM4
F-3a/8TM5
F-3a/8TM6
F-3a/8TM7
F-3a/8TM8
F-3a/8TMA
49
37
27
3a
β-tubulin
1 2 3 4 5 6 7 8 9 10 11

Fig. 3.7

Fig 3.6 RT-PCR results of F-3a/7T and its mutants.

COS7 cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with the plasmid DNAs. At 18 h post transfection, cells were lysed by Trizol reagent. Total RNA was extracted and 1μg RNA was used as template for RT-PCR. The poly-uridine region of each RT-PCR products is shown.

Fig 3.7 Mutational analysis of the slippery sequence in pF-3a/8T.

a. Diagram of pF-3a/8T and the nine mutants (pF-3a/8TM1-MA) under the control of T7 promoter. The black box represents the Flag-tag. The initiation codon AUG for 3a is underlined. Also shown are the point mutations introduced into the eight-T stretch in each mutant constructs.
b. Expression of pF-3a/8T and mutant constructs in COS7 cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1), pF-3a/8T (lane 2) and the nine mutant constructs (lanes 3-11), respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-3a polyclonal (upper panel) and β-tubulin antibody (lower panel). Bands corresponding to the full-length 3a and β-tubulin are indicated. Numbers on the left indicate molecular masses in kilodaltons.

limitation of the SDS-PAGE. In order to detect both products from F_0 and +1 frame (F_{+1}), the 5' end of 3a ORF was extended by fusing it to the 3' end of EGFP (*enhanced green fluorescence protein*) ORF and cloned into pKTO vector under the control of a T7 promoter (Figure 3.8), yielding pEGFP-3a/7T. The position of the nucleotide A of the 3a initiation codon AUG was designated +1 and the nucleotide positions upstream and downstream of the AUG codon are indicated by minus and plus numbers, respectively. The T stretch is underlined. The stop codon of F_0 was italicized and the stop codon of F_{+1} was indicated. The expected translation product of F_0 is approximately 28 kDa and the frame-shifting product is about 55 kDa. Indeed, *in vitro* translation of this construct showed both the termination and frame-shifting products of expected molecular mass. To study if the 5'-end extension affects the frame-shifting, constructs with different lengths of 5'-portions to 3a ORF were made by deleting 20, 40, and 60 amino acids at the C-terminal of the EGFP (Figure 3.8), resulting in pEGFPΔ1 to Δ3-3a. Translation *in vitro* showed both termination and frame-shifting products in Figure 3.9a. Expression of these deletion constructs in COS 7 cells showed the frame-shifting product with different levels of expression, when probed with 3a antiserum (Figure 3.9b). The proteins were not probed with EGFP antibody since the commercial antibody purchased from Sigma is against the C-terminus of EGFP and it is unlikely to detect the C-terminal deletion constructs.

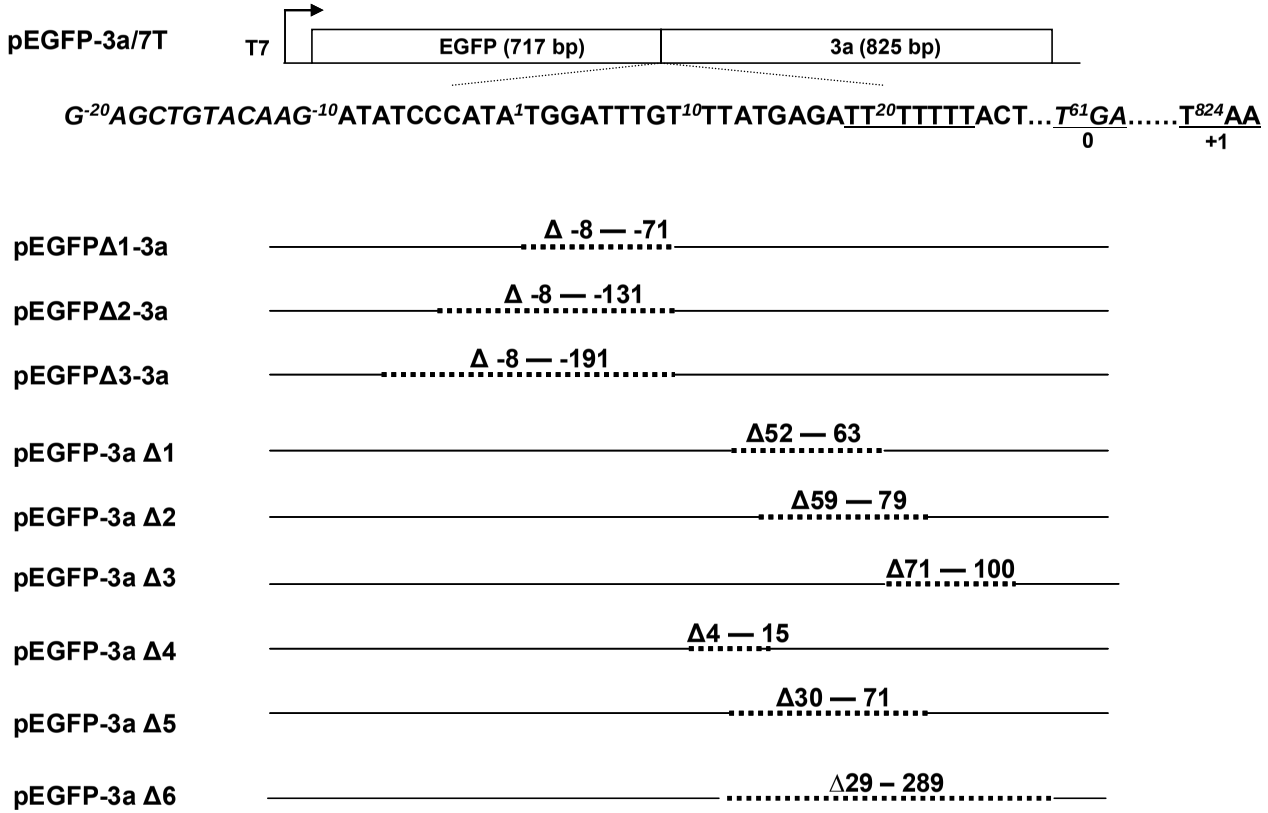

pEGFP-3a/7T
T7
EGFP (717 bp)
3a (825 bp)
G-20AGCTGTACAAG-10ATATCCCATA1TGGATTTGT10TTATGAGATT20TTTTTACT…T61GA……T824AA
0
+1
pEGFPΔ1-3a
Δ -8 — -71
pEGFPΔ2-3a
Δ -8 — -131
pEGFPΔ3-3a
Δ -8 — -191
pEGFP-3a Δ1
Δ52 — 63
pEGFP-3a Δ2
Δ59 — 79
pEGFP-3a Δ3
Δ71 — 100
pEGFP-3a Δ4
Δ4 — 15
pEGFP-3a Δ5
Δ30 — 71
pEGFP-3a Δ6
△29 – 289

Fig. 3.8

Fig 3.8 Diagram showing the structures of pEGFP-3a/7T and eight derivative constructs with deletion at different regions.

Sequence covering the fusion region between EGFP and 3a is shown. The nucleotides of the 3a ORF are indicated in bold, and nucleotides of EGFP gene are italic. The position of nucleotide A in AUG of 3a ORF is designated +1 and the nucleotide positions upstream and downstream of the AUG codon are indicated by minus and plus numbers, respectively. The seven-T stretch is underlined, the termination codon UGA for F_0 is italic and underlined, and the termination codon UAA for F_{+1} is underlined and in bold. Also shown are the nucleotides deleted in each construct.

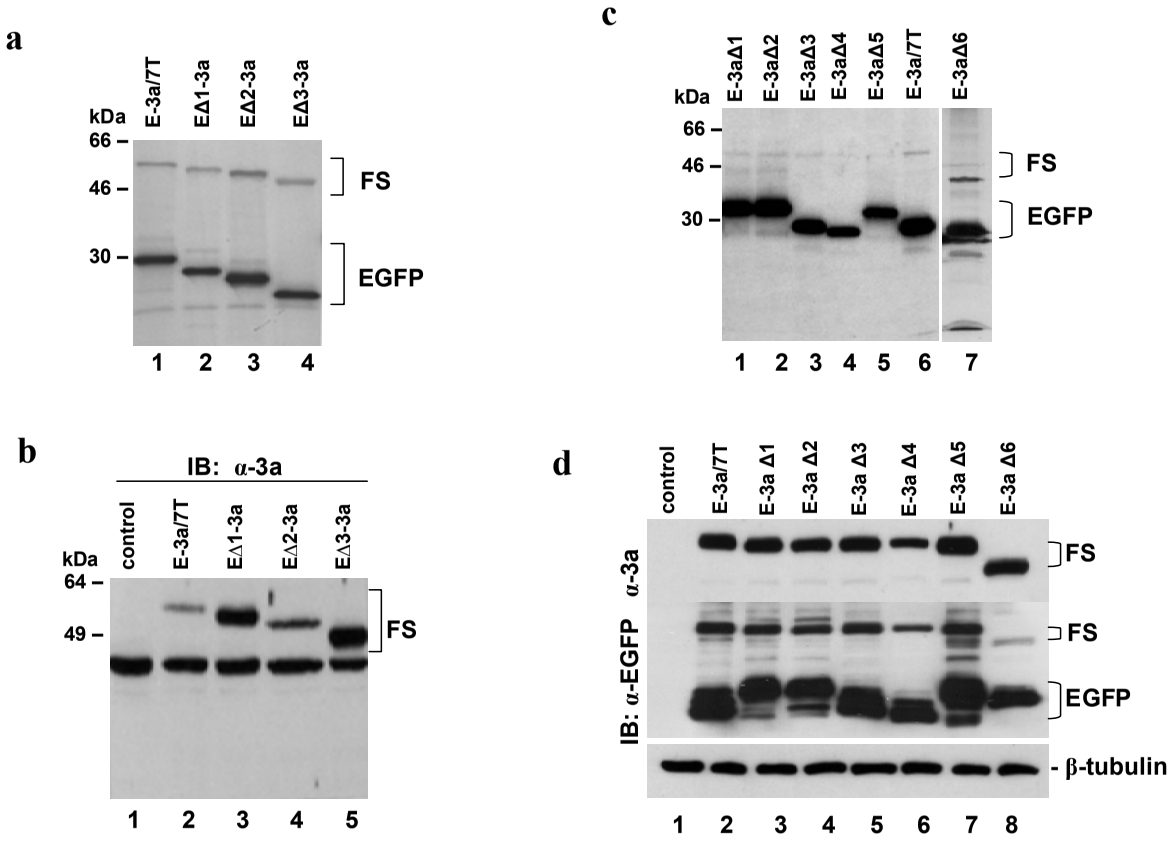

a
kDa
E-3a/7T
EΔ1-3a
EΔ2-3a
EΔ3-3a
66
46
30
FS
EGFP
1 2 3 4
b
IB: α-3a
kDa
control
E-3a/7T
EΔ1-3a
EΔ2-3a
EΔ3-3a
64
49
FS
1 2 3 4 5
c
kDa
E-3aΔ1
E-3aΔ2
E-3aΔ3
E-3aΔ4
E-3aΔ5
E-3a/7T
E-3aΔ6
66
46
30
FS
EGFP
1 2 3 4 5 6 7
d
control
E-3a/7T
E-3a Δ1
E-3a Δ2
E-3a Δ3
E-3a Δ4
E-3a Δ5
E-3a Δ6
α-3a
IB: α-EGFP
FS
FS
EGFP
β-tubulin
1 2 3 4 5 6 7 8

Fig. .3.9

Fig 3.9 Expression of deletion constructs of EGFP-3a/7T.

a. Expression of pEGFP-3a/7T (lane 1), pEGFPΔ1-3a (lane 2), pEGFPΔ2-3a (lane 3) and pEGFPΔ3-3a (lane 4) in rabbit reticulocyte lysates *in vitro*. The polypeptides were labeled with [^{35}S]-methionine, separated on 12% SDS-PAGE and detected by autoradiography.

b. Expression of pEGFP-3a/7T (lane 2), pEGFPΔ1-3a (lane 3), pEGFPΔ2-3a (lane 4) and pEGFPΔ3-3a (lanes 5) in COS7 cells. COS7 cells were transfected with plasmid DNAs with deletions (lanes 2 to 5) or a control plasmid (lane 1). At 18 h post transfection, cell lysates were prepared. Proteins were separated on 12% SDS-PAGE and analyzed by western blot with anti-3a antibody. Bands corresponding to about 38 kDa in the right panel are background bands.

c. Expression of pEGFP-3a/7T (lane 6), pEGFP-3aΔ1 (lane 1), pEGFP-3aΔ2 (lane 2) pEGFP-3aΔ3 (lane 3), pEGFP-3aΔ4 (lane 4), pEGFP-3aΔ5 (lane 5) and pEGFP-3aΔ6 (lane 7) in *in vitro* system. Polypeptides were labeled with [^{35}S]-Methionine and separated on 12% SDS-PAGE and visualized by autoradiography.

d. Expression of 3a deletion constructs in COS7 cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1), pEGFP-3a/7T (lane 2) and 3a deletion constructs pEGFP-3a/7TΔ1 to Δ6 (lanes 3 to 8). Total proteins were harvested at 18 h post transfection and analyzed by western blot with anti-3a antibody (upper panel), anti-EGFP antibody (mid panel), and anti-β-tubulin antibody (lower panel). Bands corresponding to EGFP, the frame-shifting products (FS) and β-tubulin are indicated. Numbers on the left indicate molecular masses in kilodaltons.

Combining the *in vitro* and *in vivo* data, it suggests that 5'-end extension did not render much effect on frame-shifting.

3.6 Characterization of sequences upstream and downstream of the slippery sequence UUU UUU U

There are many stimulatory elements residing downstream of the slippery sequence. For example, a typical pseudoknot found in HIV-1 *gag-pol* is able to enhance -1 frame-shifting (Bidou *et al*., 1997; Dinman *et al*., 2002; Le *et al.*, 1991; Namy *et al*., 2006). To investigate the +1 frame-shift event in 3a/7T, deletion studies were carried out in search of certain stimulating element(s) based on the fusion construct EGFP-3a/7T. Five constructs with different length of deletions downstream of the seven T stretch were made as pEGFP-3aΔ1 to Δ3, Δ5 and Δ6 (Figure 3.8). Deletion 1 has 12 nucleotides deleted (nt 52 to 63), deletion 2 with 21 nucleotides (nt 59 to 79), deletion 3 with 32 nucleotides (nt 71 to 100), deletion 5 with 42 nucleotides (nt 30 to 71) and deletion 6 with 261 nucleotides (nt 29 to 289). Translation of these constructs *in vitro* was shown in Figure 3.9c (lanes 1 to 3 and 5 to 7, respectively). The frame-shifting and termination product was indicated on the right. As seen in the figure, no apparent change in frame-shifting efficiency was observed. Expressions in COS 7 cells showed similar profile as in Figure 3.9d.

In *E. coli*, frame-shifting efficiency was elevated by base-pairing between the

ribosome 16S subunit and a specific sequence upstream of the slippery site. Thus, to examine the possible *cis*-element in the 5'-end of the U stretch, another deletion construct Δ4 was made with 12 nucleotides deleted (nt 4 to 15) which is between the initiation codon AUG of 3a and the U stretch (Figure 3.8). However, neither the *in vitro* nor the *in vivo* expression showed significant change in frame-shifting efficiency (Figure 3.9c, lane 4; Figure 3.9d, lane 6).

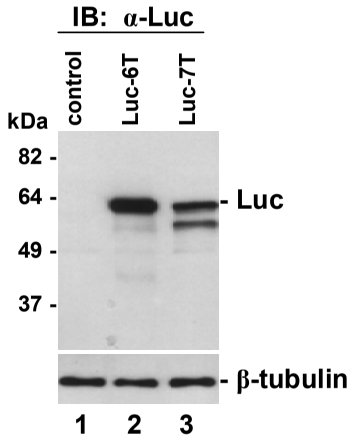
In summary, results presented in this section suggest that no stimulatory element was found in the vicinity of the slippery sequence and it is very likely that frame-shifting induced by uridine stretches is independent of additional elements. However, we cannot rule out the possibility that this kind of element may exist far from the slippery sequence on the transcripts. Thus, to test this possibility, the six and seven T stretches were transplanted into a Firefly luciferase gene 18 nucleotides downstream of the initiation codon ATG, resulting in a pLuc-6T and a pLuc-7T (Figure 3.10a) in the vector pcDNA 3.1. The full-length luciferase gene is 1653 nt long coding for a protein of ~60 kDa. The stop codons of F_0 are underlined and italicized and the stop codons of F_{+1} are underlined. COS 7 cells were transfected with either empty vector plasmid pcDNA 3.1 (negative control) or the above two plasmids and total proteins were analyzed on 10% SDS-PAGE and probed with luciferase monoclonal antibody (Sigma). While no signal was detected in the negative control (Figure 3.10b, lane 1), a strong band corresponding to the size of

a

pLuc-6T A[1]TG GAA GAC GCC AAA AAC TTT TTT ATA AAG AAA GGC ……GTG[1650] *TAA*
0

pLuc-7T A[1]TG GAA GAC GCC AAA AAC TTT TTT TAT AAA …GCA[80] *TAA*……GTG[1650]TAA
0 +1

b



c

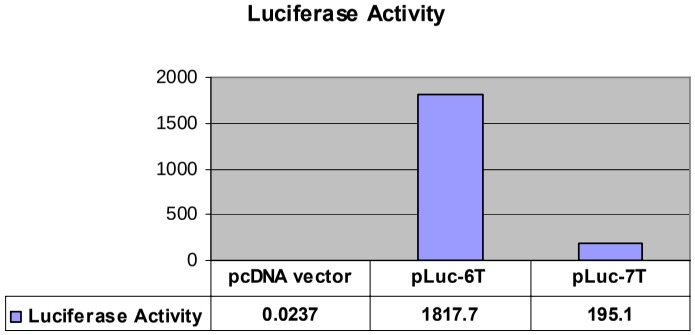


Fig. 3.10

Fig 3.10 Analysis of frame-shifting efficiencies mediated by hepta-uridine stretch in a heterogeneous ORF.

a. Diagram showing the sequences of luciferase gene with insertions of six- and seven-uridine stretches (pLuc-6T and pLuc-7T). The vector used is pcDNA3.1. The slippery sequences (the uridine stretch) are underlined, the UAA termination codon for F_0 is italic and underlined, and the UAA termination codon for F_{+1} is underlined.
b. Expression of pLuc-6T and pLuc-7T *in vivo*. COS7 cells were transfected with a control plasmid (lane 1), pLuc-6T (lane 2) and pLuc-7T (lane 3) using Effectene reagent (Qiagen). At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-luciferase antibody. Bands corresponding to the full-length luciferase and β-tubulin are indicated. Numbers on the left indicate molecular masses in kilodaltons.
c. Summary of luciferase activity assay. HeLa cells were tranfected with the empty vector, pLuc-6T and pLuc-7T. At 18 h post transfection, cells were lysed in passive lysis buffer (Promega) and luciferase activity was analyzed by luminometer. Numbers below the bar graph indicate the readings of luciferase activity in each construct.

luciferase was present in Luc-6T (Figure 3.10b, lane 2). As expected, a similar signal was also detected in Luc-7T with about 2.5 fold less intensity than Luc-6T (lane 3). Interestingly, a lower minor band of ~56 kDa was observed in Luc-7T as well, which appeared much stronger than that in Luc-6T. One possibility is that this could be a product of an internal initiation codon as sequence analaysis revealed three AUG codons downstream, which are at nt 88-90, nt 175-177, and nt 199-201. Luciferase activity assay was then carried out. HeLa cells were transfected with the empty vector pcDNA 3.1 as well as the two constructs pLuc-6T and 7T. The total cell lysates were diluted with passive lysis buffer. As in Figure 3.10c, the activity of Luc-6T was 10-fold higher than that of Luc-7T. The data suggest it is not likely that long distance interaction exists in ORF 3a to enhance frame-shifting and the uridine stretch can induce frame-shifting regardless of any other stimulating elements.

3.7 Involvement of the codon immediately downstream of the hepta-uridine stretch

Many frame-shifting signals include a stop codon or a hungry codon downstream of the slippery site. During the slow recognition of these rare codons, pause of ribosomes at these sites favors the frame-shifting (Atkinson *et al*., 1997; Barak *et al*., 1996). To study the mechanism of the frame-shifting induced by UUU UUU U, a stop codon was introduced into the F_0 of pF-3a/7T by mutation of ^{25278}C

to ^{25278}G (Figure 3.11a), giving rise to pF-3a/7TM. The stop codon TAG is expected to induce a pause of the moving ribosome. If this codon is involved in frame-shifting, presumably positioned at ribosome A-site, the frame-shifting efficiency will be affected. Expression of the construct in COS 7 cells, however, showed comparable level of 3a protein expression in F-3a/7TM to that in the wild type 3a/7T (Figure 3.11b, lanes 2 and 3). Considering the previous result that point mutation of ^{25277}A only marginally affected 3a expression (Figure 3.5), these data suggest that this codon is unlikely to be involved in recoding, suggesting that ribosome pausing on this codon has little effect on frame-shifting.

3.8 Effects of pseudoknot structure on the frame-shifting mediated by uridine stretches

Although no stimulatory elements were identified close to the slippery site, it would be interesting to study the effect of these elements on frame-shifting induced by uridine stretch. In SARS-CoV 1a1b, a typical signal for -1 frame-shifting has been identified as U UUA AAC, followed by an atypical pseudoknot served as stimulatory element. A region containing this slippery sequence from nt 12711 to nt 14100 was fused with a FLAG-tag at the 5' end and cloned under the control of T7 promoter (Figure 3.12a). This construct pF-S1ab was for positive control with the original slippery sequence U UUA AAC. The seven or eight pyrimidines were

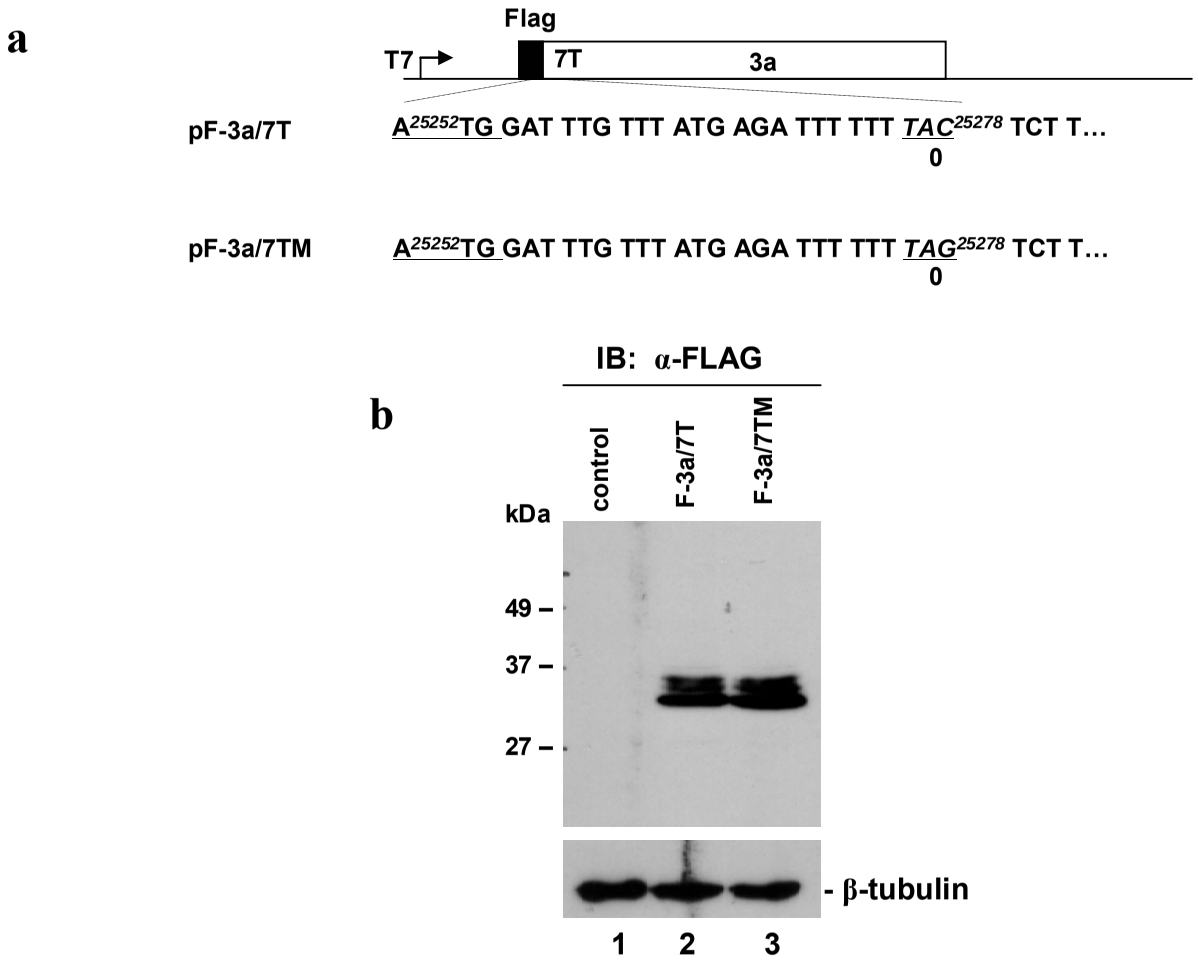

a
Flag
T7
7T
3a
pF-3a/7T
A25252TG GAT TTG TTT ATG AGA TTT TTT TAC25278 TCT T…
0
pF-3a/7TM
A25252TG GAT TTG TTT ATG AGA TTT TTT TAG25278 TCT T…
0
b
IB: α-FLAG
control
F-3a/7T
F-3a/7TM
kDa
49 –
37 –
27 –
- β-tubulin
1
2
3

Fig. 3.11

Fig 3.11 Effect of ribosome pausing on frame-shifting in ORF 3a/7T.

a. Diagram of ORF3a with seven uridines in pFLAG vector under the control of a T7 promoter. The black box indicates the Flag-tag. The sequences of the first 31 nt of ORF 3a in pF-3a/7T and pF-3a/7TM are showed below. The initiation codons of ORF 3a are underlined. The codons where the mutation was introduced are italic and underlined. Numbers indicate nucleotide position.

b. Expression of pF-3a/7T and its mutants *in vivo*. COS7 cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1), pF-3a/7T (lane 2) and pF-3a/7TM (lane 3). Total cell lysates were prepared at 18 h post transfection. Proteins were analyzed on 12% SDS-PAGE and detected by western blot using anti-FLAG antibody (upper panel) and anti-β-tubulin antibody (lower panel). The bands corresponding to 3a proteins and β-tubulin are indicated. Numbers on the left indicate the molecular masses in kilodaltons.

a

pF-S1ab FLAG-G^{12711}GT……ACG TTT^{13376} TTA AAC GGG TTT GCG GTG *TAA*……GCT^{14110} TAA
(0 under GGG; 0 under *TAA*; -1 under TAA)

pF-S1ab/7T FLAG-G^{12711}GT……ACG TTT^{13376} TTT TGG GTT TGC GGT GTA A……GCT^{14110} *TAA*
(+1 under TGG GTT; +1 under GTA A; 0 under *TAA*)

pF-S1ab/8T FLAG-G^{12711}GT……ACG TTT^{13376} TTT TTC GGG TTT GCG GTG *TAA*……GCT^{14110} TAA
(-1 under TTC GGG; 0 under *TAA*; -1 under TAA)

b

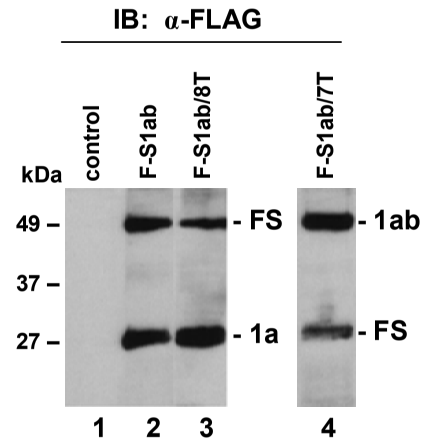


Fig. 3.12

Fig 3.12 Analysis of frame-shifting efficiencies mediated by hepta- and octo-uridine stretches in SARS-CoV ORF1a/1b.

a. Diagram showing the structures of constructs pF-S1ab, pF-S1ab/7T and pF-S1ab/8T. The S1a/1b region cloned is from nt 12722 to nt 14110. The slippery sequences are underlined, the UAA termination codon for F_0 is italic and underlined, and the UAA termination codon for F_{+1} or F_{-1} is underlined.

b. Expression of pF-S1ab, pF-S1ab/7T and pF-S1ab/8T in COS7 cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1), pF-S1ab (lane 2), pF-S1ab/7T (lane 3) and pF-S1ab/8T (lane 4). At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-FLAG antibody. Bands corresponding to 1a (or 1ab) and the frame-shifting products (FS) are indicated. Numbers on the left indicate molecular masses in kilodaltons.

cloned into the same site of S1ab to replace the original sequence (Figure 3.12a) resulting in pF-S1ab/7T and pF-S1ab/8T and the slippery sequences were underlined, stop codons of F_0 underlined and italicized, stop codons of +1 (F_{+1}) or -1 frame (F_{-1}) underlined and indicated. In pF-S1ab/7T, two nucleotides AC^{25232} were deleted so that the gene products of individual frames are of different sizes and can be differentiated on SDS-PAGE. COS 7 cells were transfected with these three plasmids as well as empty vector pFLAG. Total proteins were separated on 12% SDS-PAGE and immunoblotted with anti FLAG antibody. Figure 3.12b shows the results in which the products from F_0 (1a or 1ab) and F_{+1} or F_{-1} (FS) are indicated on the right. Compared to the wild type slippery sequence, the uridine stretches induced frame-shifting at an equivalent level (Figure 3.12b).

Previous mutagenesis on the pseudoknot confirmed its stimulatory role in -1 frame-shifting of SARS-CoV 1a1b (Baranov *et al.*, 2005; Su *et al.*, 2005). This pseudoknot structure contains 3 stem-loops (SL) and disruption of SL2 stem by destabilizing the base-pairing (SARSPK5) reduced the frame-shifting efficiency almost to a basal level (Baranov *et al.*, 2005). To examine the role of the pseudoknot in frame-shifting mediated by uridine stretches, we made similar mutations to destabilize the stem 2 region yielding constructs with suffix M (Figure 3.13a). COS7 cells were transfected with these plasmid DNA or empty vector and the total proteins were immunoblotted with FLAG antibody. As expected, the mutation in stem 2

a

pF-S1ab	FLAG-GGT……ACGTTTTTAAACGGGTTTGCGGTGTAAGTGCAGCCCGTCTT……
pF-S1abM	FLAG-GGT……ACGTTTTTAAACGGGTTTGCGGTGTAAGTGC*GGGCT*GTCTT…….
pF-S1ab/7T	FLAG-GGT……ACGTTTTTTTGGGTTTGCGGTGTAAGTGCAGCCCGTCTT……
pF-S1ab/7TM	FLAG-GGT……ACGTTTTTTTGGGTTTGCGGTGTAAGTGC*GGGCT*GTCTT……
pF-S1ab/8T	FLAG-GGT……ACGTTTTTTTTCGGGTTTGCGGTGTAAGTGCAGCCCGTCTT……
pF-S1ab/8TM	FLAG-GGT……ACGTTTTTTTTCGGGTTTGCGGTGTAAGTGC*GGGCT*GTCTT……

Fig. 3.13

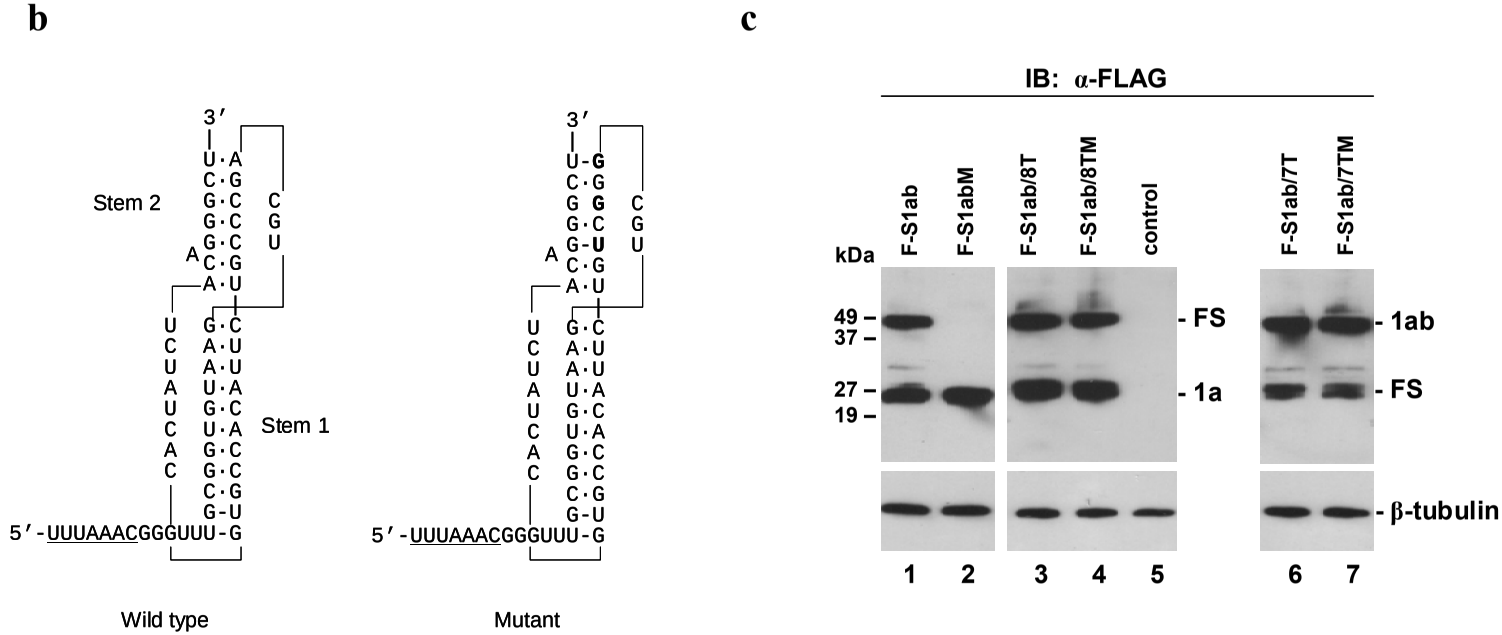

b
Stem 2
Stem 1
5′-UUUAAACGGGUUU-G
Wild type
5′-UUUAAACGGGUUU-G
Mutant
c
IB: α-FLAG
kDa
F-S1ab
F-S1abM
F-S1ab/8T
F-S1ab/8TM
control
F-S1ab/7T
F-S1ab/7TM
49 –
37 –
27 –
19 –
- FS
- 1a
- 1ab
- FS
- β-tubulin
1
2
3
4
5
6
7

Fig. 3.13

Fig 3.13 Mutational analysis of a downstream stimulator on frame-shifting efficiencies mediated by the hepta- and octo-uridine stretches.

a. Diagram showing the S1ab sequences covering the slippery sequence and stem 2 regions. The slippery sequence is underlined and the mutated nucleotides are italic and in red.
b. Diagram showing the pseudoknot structure of wild type (right) and mutant (left) S1a/1b with the slippery sequence underlined. The stem regions are indicated. Also indicated are the mutations introduced into stem 2 (in bold). The canonical base pairing is represented by dot and the weak base pair between U and G is represented by dash.
c. Expression of pF-S1ab (lane 1), pF-S1abM (lane 2), pF-S1ab/7T (lane 6), pF-S1ab/7TM (lane 7), pF-S1ab/8T (lane 3) and pF-S1ab/8T (lane 4) *in vivo*. COS7 cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 5) and the six constructs, respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-FLAG antibody. Bands corresponding to 1a (or 1ab), the frame-shifting products (FS) and β-tubulin are indicated. Numbers on the left indicate molecular masses in kilodaltons.

resulted in a substantial reduction in frame-shifting efficiency in the S1ab construct (Figure 3.13c, lanes 1 and 2). Surprisingly, this same mutation rendered only marginal effect on frame-shifting mediated by uridine stretches, indicating a much lower sensitivity of the frame-shifting to the base pairing disruption (Figure 3.13c, lands 3 and 4, 6 and 7). In addition, double bands were observed in the 27 kDa regions of S1ab/7T, 8T and their mutants but not in S1ab. The above results demonstrate that typical stimulatory elements like pseudoknot have little effect on frame-shifting caused by poly-U tracts.

3.9 Downstream pseudoknot has differential effect on frame-shifting by uridine stretches with point mutations at different positions

Since mutation of any of the hepta-uridines in an ORF3a variant severely reduced the frame-shifting efficiency mediated by the seven U slippery sequence (Figure 3.5), we set up to test the effect of a downstream stimulatory element on the frame-shift event mediated by the uridine stretches with point mutations at different positions. As shown in Figure 3.14a, mutants containing mutation of the first T to an A (pF-Slab/7TM1A) or to a C (pF-Slab/7TM1C), mutation of the seventh T to an A (pF-Slab/7TM7A) or to a C (pF-S1ab/7TM7C), mutation of the first and fourth Ts to Cs (pF-Slab/7TM14C), the second and sixth Ts to Cs (pF-S1ab/7TM26C) as well as the third (pF-S1ab/7TM3C) and the sixth (pF-S1ab/7TM6C) T to C were made based

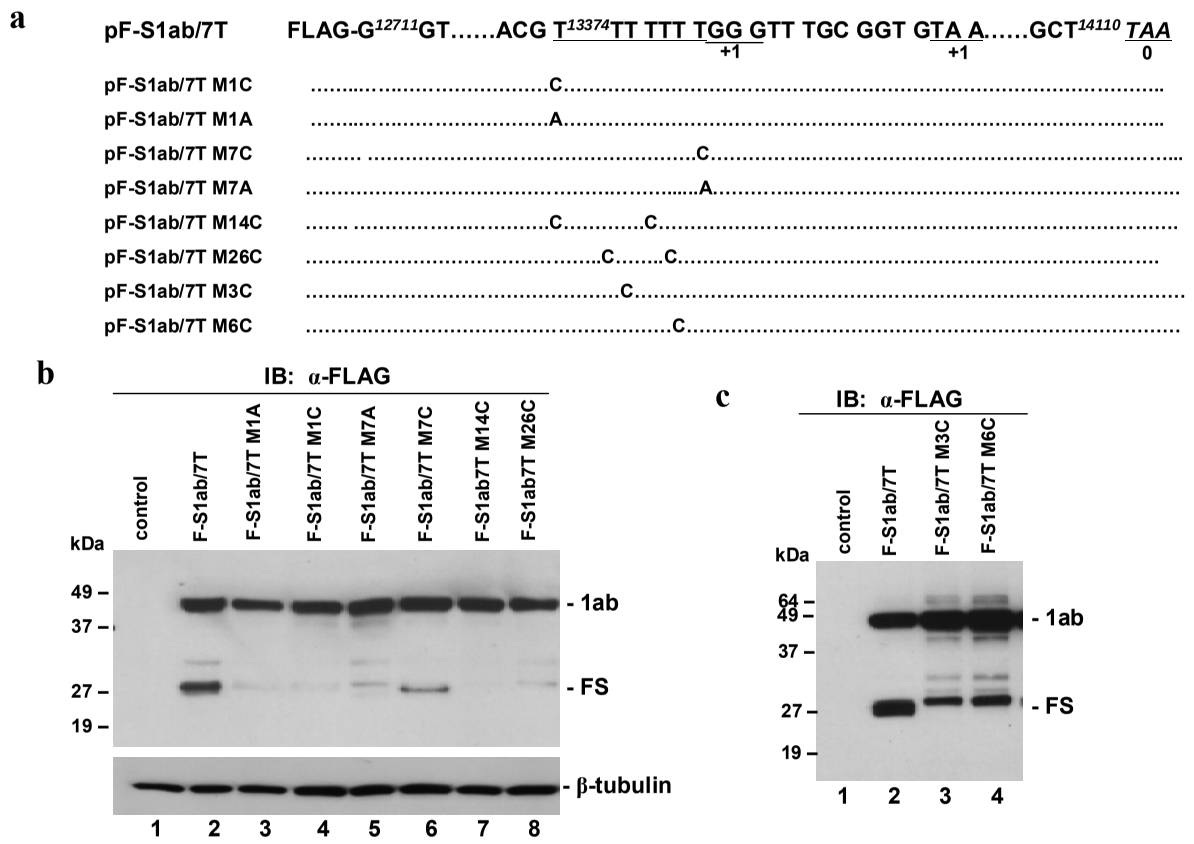

a
pF-S1ab/7T FLAG-G12711GT……ACG T13374TT TTT TGG GTT TGC GGT GTA A……GCT14110 TAA
+1
+1
0
pF-S1ab/7T M1C ……C……
pF-S1ab/7T M1A ……A……
pF-S1ab/7T M7C ……C……
pF-S1ab/7T M7A ……A……
pF-S1ab/7T M14C ……C……C……
pF-S1ab/7T M26C ……C……C……
pF-S1ab/7T M3C ……C……
pF-S1ab/7T M6C ……C……
b
IB: α-FLAG
control
F-S1ab/7T
F-S1ab/7T M1A
F-S1ab/7T M1C
F-S1ab/7T M7A
F-S1ab/7T M7C
F-S1ab7T M14C
F-S1ab7T M26C
kDa
49
37
27
19
- 1ab
- FS
- β-tubulin
1 2 3 4 5 6 7 8
c
IB: α-FLAG
control
F-S1ab/7T
F-S1ab/7T M3C
F-S1ab/7T M6C
kDa
64
49
37
27
19
- 1ab
- FS
1 2 3 4

Fig. 3.14

Fig 3.14 Differential effects of a downstream stimulator on frame-shifting efficiencies mediated by wild type and mutant hepta-uridine stretch.

a. Diagram showing the structure of constructs pF-S1ab/7T, pF-S1ab/7TM1C, pF-S1ab/7TM1A, pF-S1ab/7TM7C, pF-S1ab/7TM7A, pF-S1ab/7TM14C, pF-S1ab/7TM26C, pF-S1ab/7TM3C, and pF-S1ab/7TM6C. Numbers indicate nucleotide positions. The uridine stretch and the UAA termination codon for F_{+1} are underlined. The UAA termination codon for F_0 is italic and underlined. The mutations introduced into the uridine region are also shown.

b. Expression of pF-S1ab/7T (lane 2), pF-S1ab/7TM1A (lane 3), pF-S1ab/7TM1C (lane 4), pF-S1ab/7TM7A (lane 5), pF-S1ab/7TM7C (lane 6), pF-S1ab/7TM14C (lane 7), and pF-S1ab/7TM26C (lane 8) in COS7 cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1) and the seven constructs, respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-FLAG antibody (upper panel) or anti-β-tubulin antibody (lower panel). Bands corresponding to 1ab, the frame-shifting products (FS) and β-tubulin are indicated. Numbers on the left indicate molecular masses in kilodaltons.

c. Expression of pF-S1ab/7T (lane 2), pF-S1ab/7TM3C (lane 3), and pF-S1ab/7TM6C (lane 4) in COS7 cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1) and the three constructs, respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-FLAG antibody. Bands corresponding to 1ab and the frame-shifting products (FS) are indicated. Numbers on the left indicate molecular masses in kilodaltons.

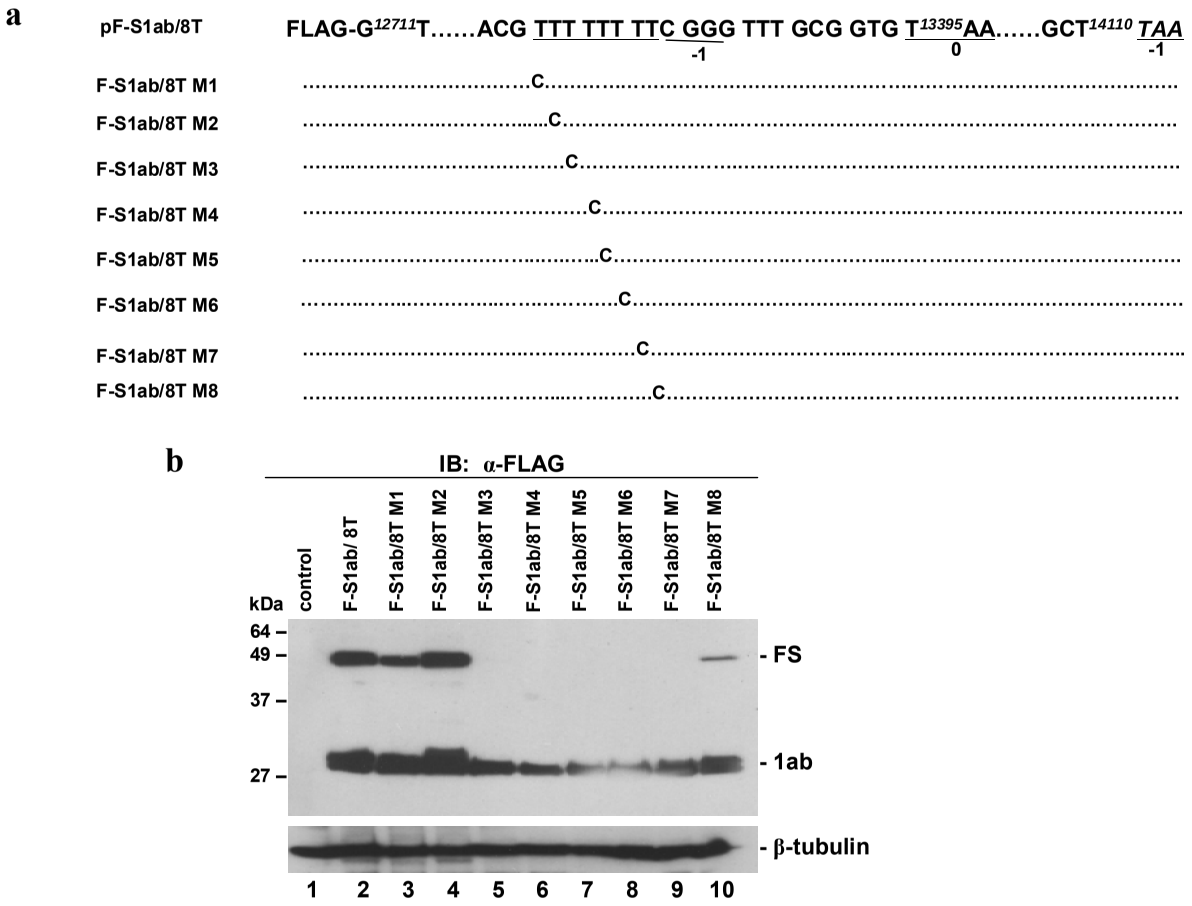

a
pF-S1ab/8T
FLAG-G12711T……ACG TTT TTT TTC GGG TTT GCG GTG T13395AA……GCT14110 TAA
-1
0
-1
F-S1ab/8T M1
F-S1ab/8T M2
F-S1ab/8T M3
F-S1ab/8T M4
F-S1ab/8T M5
F-S1ab/8T M6
F-S1ab/8T M7
F-S1ab/8T M8
b
IB: α-FLAG
control
F-S1ab/ 8T
F-S1ab/8T M1
F-S1ab/8T M2
F-S1ab/8T M3
F-S1ab/8T M4
F-S1ab/8T M5
F-S1ab/8T M6
F-S1ab/8T M7
F-S1ab/8T M8
kDa
64
49
37
27
- FS
- 1ab
- β-tubulin
1 2 3 4 5 6 7 8 9 10

Fig. 3.15

Fig 3.15 Effects of a downstream stimulator on frame-shifting efficiencies mediated by wild type and mutant octo-uridine stretch.

a. Diagram showing the structure of constructs pF-S1ab/8T, pF-S1ab/8TM1toM8. Numbers indicate nucleotide positions. The uridine stretch and the UAA termination codon for F_0 are underlined. The UAA termination codon for F_{-1} is italic and underlined. The mutations introduced into the uridine region are also shown.

b. Expression of pF-S1ab/8T (lane 2), pF-S1ab/8TM1 to M8 (lanes 3-10) in COS7 cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1) and the nine constructs, respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-FLAG antibody (upper panel) or anti-β-tubulin antibody (lower panel). Bands corresponding to 1ab, the frame-shifting products (FS) and β-tubulin are indicated. Numbers on the left indicate molecular masses in kilodaltons.

on pF-slab/7T. Expression of pF-Slab/7T showed the frame-shifting efficiency of approximately 35% (Figure 3.14b, lane 2). Mutation of the first T to either an A or a C significantly reduced the frame-shifting efficiency to approximately 1% of wild type (Figure 3.14b, lanes 3 and 4). Mutations of the first and fourth Ts to Cs or both the second and sixth Ts to Cs totally abolished the frame-shifting (Figure 3.14b, lane 7). However, mutation of the seventh T to C resulted in the detection of about 15% of frame-shifting product (Figure 3.14b, lanes 6). Mutation of this T to A introduced an amber stop codon TAG in F_{-1}. It seemed to render some different effect on the frame-shifting from M7C. In addition, strong signals were detected in M3 and M6 but with a slightly larger molecular weight as compared to S1ab-7T (Figure 3.14c, lanes 2 to 4). As bands of this size were also present in other mutations such as M7A or M26C, it is probably not the recoding product but a background noise due to inefficient washing during the immuno-detection.

Similar point mutations were made based on construct pF-S1ab/8T (Figure 3.15a). The thymidine at each position was mutated to a C resulted in M1 to M8. Expression of mutants M3 to M7 in COS 7 cells was consistent with those corresponding mutants in F-3a/8T. Mutant M8 also showed the frame-shifting product though with a decreased level of frame-shifting (Figure 3.15b, lane 10). Different from the previous F-3a/8T mutants, M1 and M2 showed comparable levels of frame-shifting to F-S1ab8T (Figure 3.15b, lanes 2 to 4). In addition, double bands

corresponding to the termination products were observed in all constructs, except for the negative control and M3 to M6 which only showed a single band. These results confirmed that the downstream stimulator has differential effects on the frame-shifting efficiencies mediated by the uridine stretches with mutations at different positions.

3.10 Detection of products from all frames in the octa-uridine mediated frame-shifting but not in the hepta-uridine mediated frame-shifting

As in Figure 3.15b, double bands corresponding to the terminated product were observed. Sequence analysis revealed that the TAG stop codon for F_{+1} frame is at nt 13432-13434, which is close to the TAA stop codon for the F_0 frame at nt 13395-13397. The F_{+1} and F_0 produects are both about 30 kDa. Therefore, the double bands may represent these two gene products. It appeared that UUU UUU UU could mediate both -1 and +1 frame-shifting. To examine the directionality of the frame-shifting mediated by UUU UUU UU, mutations were made based on the construct pEGFP-3a/8T. In pEGFP-3a/8T, products from F_0 and F_{+1} were of similar size of ~28 kDa and protein of F_{-1} was expected to be ~60 kDa. To differentiate the product of F_0 and F_{+1}, the stop codon of F_{+1} was mutated from TGA to CGA, extending the F_{+1} product to next stop codon TAA (Figure 3.16a), resulting in plasmid pEGFP-3a/8TM. Thus, in pEGFP-3a/8TM, product of F_{+1} is expected to be

33 kDa. When these plasmids were expressed in COS 7 cells, both the -1 frame-shifting product (P_{-1}) and termination product (P_0) were observed (Figure 3.16b left, lanes 1 and 3). In addition, a strong band corresponding to F_{+1} product (P_{+1}) was detected in EGFP-3a/8TM (Figure 3.16b left, lane 3). To further clarify that P_{+1} was truly produced, an HA-tag was inserted into F_{+1} in EGFP-3a/8T, yielding EGFP-3a/8THA_{+1} (Figure 3.16a). When the proteins were probed with anti-EGFP antibody, both P_{-1} and P_0/P_{+1} were observed (Figure 3.16b left, lane 2). When the proteins were probed with anti-HA antibody, signals were only detected in EGFP-3a/8THA_{+1} (Figure 3.16b middle, lane 2) but not in EGFP-3a/8T or 8TM (Figure 3.16b middle, lanes 1 and 3). These data clearly indicated that both -1 and +1 frame-shifting could have occurred during translation mediated by UUU UUU UU.

Previously in constructs F-3a/7TM8 and F-3a/8TM8, the slippery sequences of these two mutants are the same (UUU UUU UC) (Figures 3.5a and 3.7a). The full-length 3a were detected in both constructs (Figures 3.5 and 3.7b). As it is proposed that 3a is produced from +1 frame-shifting in F-3a/7TM8 and -1 frame-shifting in F-3a/8TM8, this result implied that the hepta-uridine may also direct different frame-shift events.

To examine whether the hepta-uridines can induce frame-shifting at different direction, we tested the different frame-shifting products based on the construct pF-S1ab/7T. Two extra nucleotides AC were inserted immediately downstream the

a

pEGFP-3a/8T *EGFP*…….A[1]TG……TTT TTT TTA…. T[53]AA…. T[60]GA…….T[823]AA
0 +1 -1

pEGFP-3a/8TM *EGFP*…….A[1]TG……TTT TTT TTA…. T[53]AA…. C[60]GA…. T[185]AA…. T[823]AA
0 +1 +1 -1

pEGFP-3a/8THA$_{+1}$ *EGFP*…….A[1]TG……TTT TTT TTA…. T[53]AA….HAtagT[60]GA……T[823]AA
0 +1 -1

b

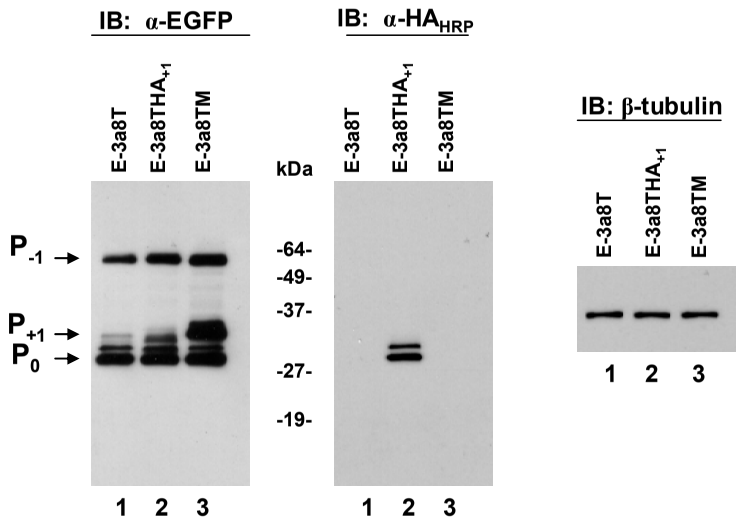


Fig. 3.16

Fig 3.16 Detection of products from each frame in pEGFP-3a/8T.

a. Diagram showing the construct of pEGFP-3a/8T and its derivative constructs. The uridine stretch is underlined. Numbers indicate nucleotide positions. The termination codons in each frame are underlined. The HA-tag in pEGFP-3a/8THA$_{+1}$ is indicated.

b. Expression of pEGFP-3a/8T (lane 2), pEGFP-3a/8THA$_{+1}$ (lane 2), and pEGFP-3a/8TM (lane 3) in COS7 cells. Cells are infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected the three constructs, respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-EGFP antibody (left panel), anti-HA antibody (mid panel), and anti-β-tubulin antibody (right panel). Proteins from all three frames are indicated. Numbers in the middle indicate molecular masses in kilodaltons.

uridines stretch, giving rise to pF-S1ab/7TA (Figure 3.17a). The stop codons for each frame was underlined and indicated. In pF-S1ab/7T, products from the F_{-1} and F_{+1} frames were expected to be ~28 kDa and the product of F_0 was ~48 kDa. In pF-S1ab/7TA, product of F_{-1} was expected to be ~48 kDa and the products of F_{+1} and F_0 were ~28 kDa. COS 7 cells were transfected with these two plasmids as well as the empty vector pFLAG. Total proteins were resolved on 12% SDS-PAGE and probed with anti-FLAG antibody. As expected, both the 48 kDa and 28 kDa bands were observed in F-S1ab/7T but absent in control (Figure 3.17b, lanes 1 and 2). However, only the band of 28 kDa was present in F-S1ab/7TA, indicating that no -1 frame-shifting occurred (Figure 3.17b, lane 3) in frame-shifting mediated by hepta-uridine stretch.

3.11 Discussion

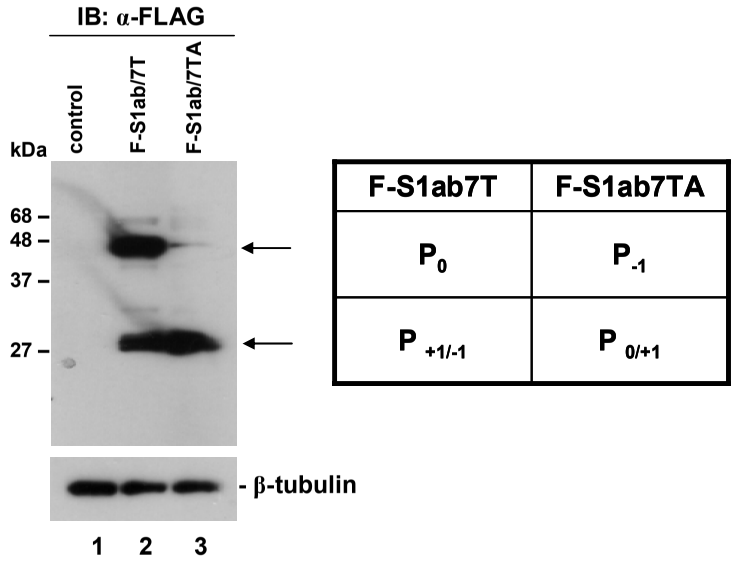
Here we presented the work which identified two ribosomal frame-shift events with hepta- (UUU UUU U) and octa- (UUU UUU UU) uridines as sole signals for efficient +1/-2 and -1/+2 frame-shifting, respectively, during translation of SARS-CoV ORF 3a variants. Expression of clones with six, seven and eight Ts in an *in vitro* expression system, in bacteria and intact cells showed the detection of the full-length 3a protein. *In vitro* and *in vivo* experiments revealed an average of 30% and 10% expression efficiency for 3a/7T and 3a/8T, respectively. This efficiency was

a

pF-S1ab/7T FLAG-G^{12711}….TTT TTT TGG G…. $\underset{+1}{T^{13395}AA}$……. $\underset{-1}{T^{13432}AG}$…….. $\underset{0}{T^{14111}AA}$

pF-S1ab/7TA FLAG-G^{12711}….TTT TTT TAC GGG…. $\underset{0}{T^{13395}AA}$……. $\underset{+1}{T^{13432}AG}$…… $\underset{-1}{T^{14111}AA}$

b



F-S1ab7T	F-S1ab7TA
P_0	P_{-1}
$P_{+1/-1}$	$P_{0/+1}$

Fig. 3.17

Fig 3.17 Detection of products from each frame in pF-S1ab/7T.

a. Diagram showing the construct of pF-S1ab/7T and its mutant. The sequences covering the slippery site are shown. The termination codons in each frame are underlined. Numbers indicate nucleotide positions. The two nucleotides inserted in pF-S1ab/7TA are italic and underlined.

b. Expression of pF-S1ab/7T (lane 2) and pF-S1ab/7TA (lane 3) in COS7 cells. Cells are infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with a control plasmid (lane 1) and the two constructs, respectively. At 18 h post transfection, cells were harvested and lysates prepared. Polypeptides were separated on 12% SDS-PAGE and analyzed by western blot with anti-FLAG antibody (upper panel) or anti-β-tubulin antibody (lower panel). Bands corresponding to β-tubulin are indicated. Numbers on the left indicate molecular masses in kilodaltons. The Table on the right shows the gene products (initial P) from different frames in each construct.

calculated as the amount of 3a expressed in 3a/7T or 3a/8T over that in 3a/6T. These data strongly suggeste that ribosomal frame-shifting accounts for the expression of full-length 3a. However, we couldn't rule out the possibility that transcriptional slippage also existed and contributed to the expression of full-length 3a. Some preliminary data on sequences of transcripts by RT-PCR in cells transfected with 3a/7T and 3a/8T constructs revealed a mixed population of transcripts. However, to ensure that these experiments are free from contamination, more careful measures should be taken.

To study the mechanisms of frame-shifting, site-directed mutagenesis was carried out within the slippery sequence based on the construct with seven Ts (TTT TTT T) or eight Ts (TTT TTT TT). The mRNA sequences showed no reversion at transcriptional level since no minor peaks were observed. However, we could not rule out the possibility that some small portions contained different number of uridines as a result from transcriptional slippage. The results showed that mutation of any U in the seven-uridine stretch or the first seven uridines in the eight-uridine stretch significantly reduced the expression of the full-length 3a protein. However, mutation of the nucleotide A immediately downstream of the uridine stretch did not affect the expression of 3a, suggesting that UUU UUU U and UUU UUU UU is the slippery sequence and the full-length 3a protein is expressed by a +1/-2 and -1/+2 frame-shifting mechanism, respectively.

In addition to the slippery sequence, the majority of the known frame-shifting signals also contain a stimulatory element. In some cases, a secondary structure or tertiary interaction, such as a pseudoknot structure downstream of the slippage site, could significantly increase the frame-shifting efficiency. In other cases, sequence specific elements, such as the SD-like sequence (Larsen *et al.*, 1994 and 1995) or sequences partially complementary with yeast 18S rRNA (Li *et al.*, 2001), are required. Furthermore, space between the slippage site and the stimulatory element also plays a critical regulatory role in the frame-shifting efficiency (Larsen *et al.*, 1994). Attempts were subsequently made to identify upstream and downstream regulatory elements by deletion analysis. Neither downstream stem-loop structures and pseudoknots, nor upstream sequences with stimulatory function were identified. Although it is possible that certain stem-loop structures and potential pseudoknot interaction downstream of the slippage site in ORF 3a variants may be formed, deletion of sequence within 260 nucleotides downstream of the uridine stretch showed that these potential structural elements did not affect the frame-shifting mediated by seven-U stretch, indicating that either long range RNA interaction may play a role or stimulatory elements may be dispensable for these frame-shifting events. However, the fact that frame-shifting was still efficient in other ORF context, such as the luciferase gene, argues against the first possibility. This conclusion was reinforced by the observation that mutation of stem 2 in the SARS-CoV 1a/1b region

rendered no effect on the frame-shifting event mediated either by the seven or eight uridines. The same mutation, however, totally abolished the frame-shifting efficiency mediated by the original slippery sequence of ORF 1a and 1b, consistent with previous study (Baranov *et al*., 2005). Previously, a G-rich sequence was reported to induce net +1 frame-shifting in HSV *TK* gene (Horsburgh *et al.*, 1996). In this case, the frame-shifting efficiency was not augmented by downstream structures and ribosomal pausing. A recent report also showed that ribosomal pausing had little correlation with the frame-shifting efficiency (Kontos *et al.*, 2001). Our results were in agreement with these observations. The stop codon introduced did not have much effect on frame-shifting, suggesting that the codon immediately downstream of the uridines in F_0 is not involved in recoding and that ribosome pausing is not a crucial factor for the frame-shifting mediated by uridine stretch. Taken together, this study demonstrated that the hepta- and octo-uridine stretches could function as sole elements for efficient ribosomal frame-shifting.

Expression of the full-length 3a protein from the construct with seven Ts (with a single nucleotide insertion) requires a +1/-2 frame-shifting at the slippage site (UUU UUU U). At present, we do not know exactly whether +1 or -2 frame-shifting occurs. Attempts are being made to purify the frame-shifting products from various expression systems and to determine the precise sequence across the slippage site. Generally, +1 frame-shifting is considered the major event taking place

at the slippage site since the phenylalanine $tRNA_{GAA}$, the only cognate tRNA for codon UUU, could form base pairing in F_{+1} with UUU at P-site, but only one base pairing could form in the -2 frame with GAU. Another interesting question is whether the putative +1 frame-shifting at the seven-uridine stretch requires ribosomal slippage at both P and A sites. The dramatic reduction by the mutations introduced into the slippery sequence in pF-3a/7TM1, M2, M3 and M4 from U to C provided evidence for a P-site slippage as imperfect base-pairing formed at the P site in F_{+1} abolished the expression of F_{+1}. If no A-site slippage occurs, meaning that A-site is not occupied before the P-site slippage, other tRNAs rather than $tRNA_{GAA}$ will base pair with the new codon in the A-site and full-length 3a is still expressed without compromise. However, this was not the case. In fact, mutations in the second triplet of F_{+1} also greatly affect the expression of 3a. Thus we proposed that a double slippage mechanism is responsible for frame-shifting mediated by UUU UUU U. At present, however, we cannot rule out the possibility that a slippage-independent mechanism occurs. For example, in GAG3-POL3 genes of yeast Ty3, frame-shifting is dependent on slow decoding of certain codons in F_0 and does not involve peptidyl-tRNA slippage (Farabaugh *et al.*, 1993; Vimaladithan and Farabaugh, 1994). The frame-shifting in mammalian ODC antizyme does not involve slippage, as shown by mutagenesis studies (Matsufuji *et al.*, 1995).

Similarly, detection of the full-length 3a protein from the construct with eight

Ts (with two uridines insertion) would require a -1/+2 frame-shifting event with double slippage at the slippery sequence (UUU UUU UU). However, since the mutation in the last uridine (in construct F-3a/8T) did not appear to affect the expression of 3a greatly, regardless of the coding context, we propose that a +2 frame-shifting with double slippage is responsible for the full-length expression. Interestingly, considerably less amounts of the full-length product in all three expression systems were detected from this construct than that from the construct with seven Ts. However, when the original slippery sequence of SARS-CoV ORF 1a/1b was replaced by eight uridines, similar or even higher frame-shifting efficiency was observed. In addition, mutations that destroyed the downstream stimulatory structure did not alter/reverse the frame-shifting efficiency of S1ab/7T and 8T, suggesting that either additional sequences located in the region or long range RNA interaction may favor -1 frame-shifting. More systematic deletion and mutagenesis studies would be required to address this possibility further. Another possibility is that the unexpected different efficiency could have resulted from a different mechanism. The evidence supporting this is that in point mutant of S1ab/8T, efficient frame-shifting was observed in M1 (CUU UUU UUC) and M2 (UCU UUU UUC). In these mutants, a -1 frame-shift event could have occurred with the slippery sequence becoming U UUU UUC instead of UUU UUU UU and with a strong stimulatory pseudoknot downstream, mimicking the -1 frame-shifting in the well

characterized HIV *gag-pol*. However, in the original 3a ORF, no stimulatory element similar to this pseudoknot in SARS-CoV ORF 1a/1b region could elevate this particular -1 frame-shifting. One experiment that can be done to prove this is to analyze the S1ab/8TM1 and M2 with the mutation in PKS2.

Another interesting issue here is the directionality of the frame-shifting mediated by uridine stretches. For the eight-uridine stretch, it is conceivable that the migrating ribosomes can treat this sequence as either eight-uridines or seven-uridines. Combined with transcriptional slippage which generates mixed transcripts with seven or eight uridines, it is possible to detect both F_{-1} and F_{+1} shifting products. Indeed, these were confirmed by Western blot using antibody against tags in different reading frames. However, some conflicting results were generated with the seven-uridine stretch. Two different mutants, pF-3a/7TM8 and pF-3a/8TM8, share the same sequence in the slippery region UUU UUU UC, but pF-3a/8TM8 contains one extra nucleotide. Expression of full-length 3a detected from these two constructs required different frame-shifting: a +1/-2 frame-shifting in pF-3a/7TM8 and a -1/+2 frame-shifting in pF-3a/8TM8. It appears that seven-uridine stretch could induce frame-shifting in both directions. However, when we tried to differentiate products encoded from different reading frames by insertion of two additional nucleotides based on the construct pS1ab/7T, we were not able to identify both products from F_{+1} and F_{-1}. As these experiments were only carried out once, further confirmation is

needed before a definite conclusion can be made.

CHAPTER 4. TRANSLATIONAL CONTROL OF HCRSV P38, P27 AND ITS ISOFORMS

4.1 Introduction

The full-length genomic RNA sequence of HCRSV was determined and two co-3'-terminal sgRNAs were identified (Huang *et al.*, 2000). The sgRNA2, starting from nt 2438 (Li and Wong, 2006), encode four polypeptides including the viral coat protein p38 and three co-C-terminal proteins, p25 (^{2603}ATG), p24 (^{2630}ATG), and p22.5 (^{2666}ATG) (Huang *et al.*, 2000). *In vitro* translation of the cDNA clone of sgRNA2 – pHCRSV80 – in wheat germ showed a distinct band corresponding to p38 and three bands with smaller molecular mass which were predicted to be p25, p24 and p22.5 (the ORFs were named according to the molecular mass of their gene products) (Figure 4.1). However, subsequent mutagenesis studies revealed a mis-assignment of each ORF. Mutations were made based on the cDNA clone of HCRSV sgRNA2 p80 to abolish the translation initiation codons for each ORF (Figure 4.2a). Initiation codon ^{2603}ATG of p25 was changed to ^{2603}GTT, yielding pHCRSV80M1; ^{2630}ATG of p24 changed to ^{2630}GTT, yielding M2; of p22.5 changed to ^{2666}GTT, yielding M3. As expected, the smallest-sized band corresponded to p22.5 and the ^{2666}ATG codon is the initiation site, since mutation of this codon abolished the expression of p22.5 (Figure 4.2b, lane 4). Surprisingly, *in vitro* translation of M1 showed a missing p24 and no missing band in M2 (Figure 4.2, lanes 1 to 3). This suggests that the protein band missing in M1 initiates from ^{2603}ATG and should be re-assigned as p25. Detailed examinations on the sequence revealed a ^{2570}CUG

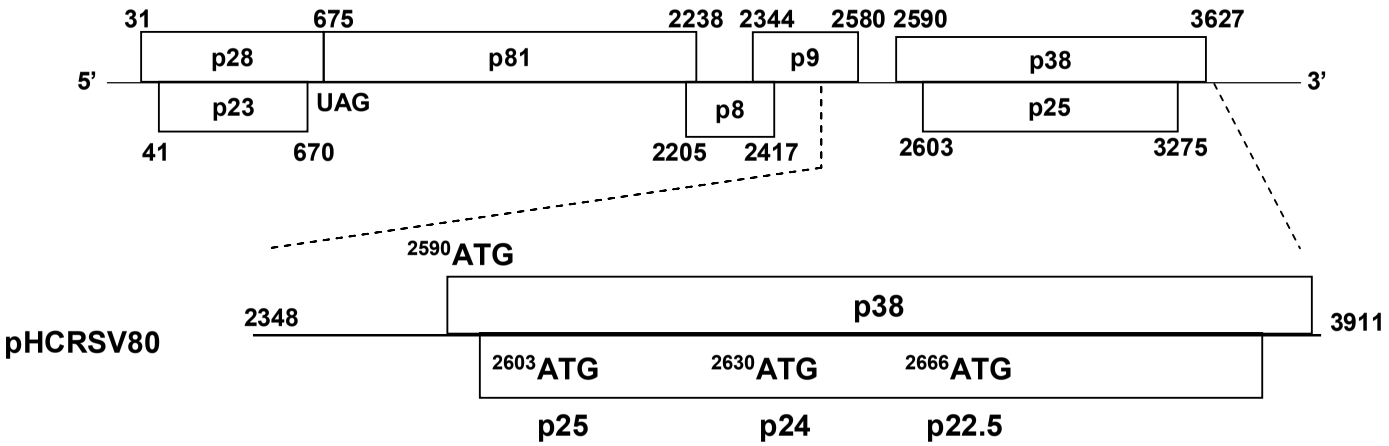
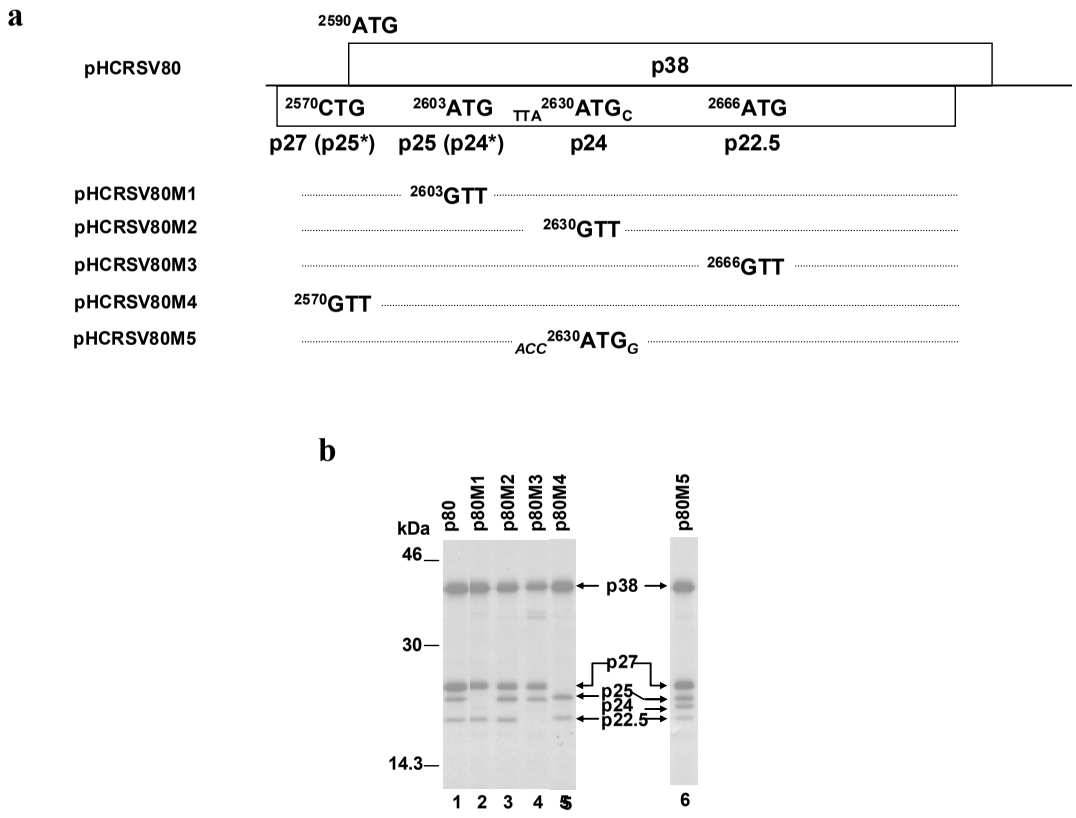

31
675
2238
2344
2580
2590
3627
p28
p81
p9
p38
5'
3'
p23
UAG
p8
p25
41
670
2205
2417
2603
3275
2590ATG
pHCRSV80
2348
p38
3911
2603ATG
2630ATG
2666ATG
p25
p24
p22.5

Fig. 4.1



Koh et al., 2006. Virus Research.

Fig. 4.2

Fig 4.1 Schematic representation of HCRSV genome organization and construct of pHCRSV80.

ORFs encoding p38, p25, p24 and p22.5 are highlighted. Numbers indicate nucleotide positions. The AUG codons for p38, p25, p24 and p22.5 are indicated.

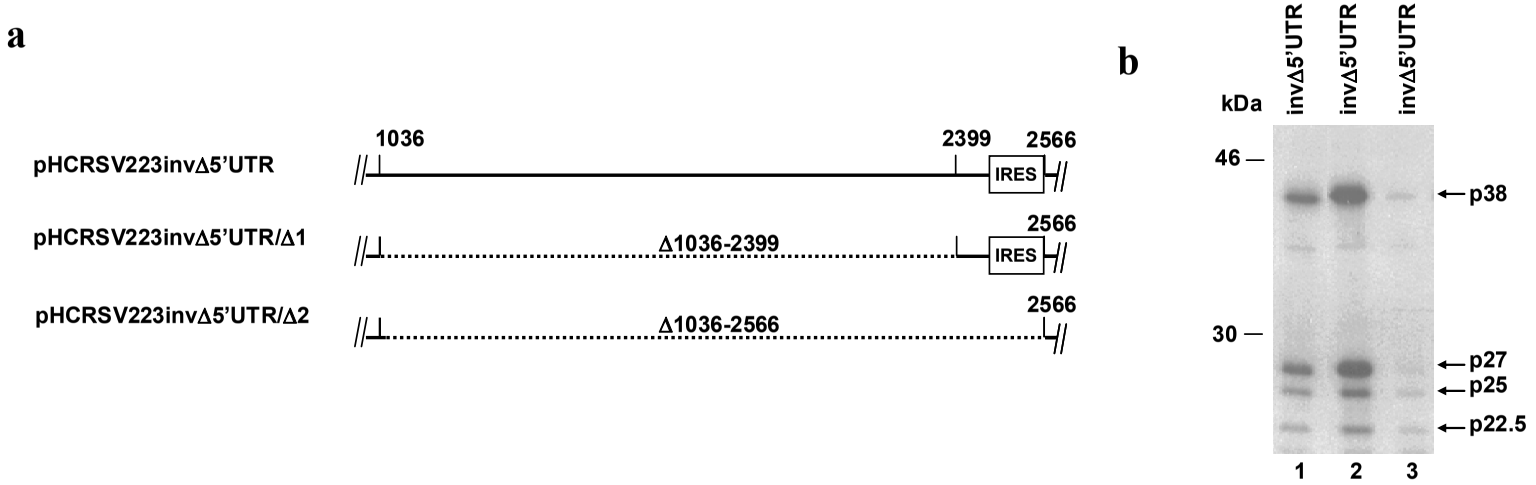
Fig 4.2 Re-assignment of ORFs encoding p38, p27, p25, p24, and p22.5.

a. Schematic representation of the cDNA clones of pHCRSV80. The initiation codons for p38, p27, p25, p22.5 are indicated. Previous assignment of p27 to p25 and p25 to p24 is in brackets and indicated by (*). Numbers indicate nucleotide positions. Below are the mutants pHCRSV80M1, pHCRSV80M2, pHCRSV80M3, pHCRSV80M4 and pHCRSV80M5. The mutations introduced into each construct are indicated.

b. Analysis of *in vitro* translation products from pHCRSV80 (lane 1), pHCRSV80M1 (lane 2), pHCRSV80M2 (lane 3), pHCRSV80M3 (lane 4), pHCRSV80M4 (lane 5) and pHCRSV80M5 (lane 6) in wheat germ extract with the TnT coupled translation system. Equal amounts of non-linearized templates were used. [^{35}S]methionine-labeled translation products were separated on 17.5% SDS-PAGE and visualized by autoradiography. Bands corresponding to p38, p27, p25, p24 and p22.5 are indicated. Numbers on the left indicate molecular masses in kilodaltons.

codon upstream of the p25 ORF (previously p24, indicated as p24*). Mutation of this CUG to GTT (M4; Figure 4.2a) abolished the translation of previous ORF p25 (indicated as p25*) (Figure 4.2b, lane 5). This ORF was re-assigned as p27, whose translation is initiated from a non-AUG codon. Further analysis of p24 initiation codon showed that it resides in a weak context with a C^{+4} and a T^{-3}. When this context was changed to a favorable one with a G^{+4} and an A^{-3} (M5; Figure 4.2a), additional translation product was observed with a size between p25 and p22.5 (Figure 4.2b, lane 6). This confirmed the correct re-assignment and suggests that p24 is usually silent due to its weak initiation context.

In a previous study, p38, p27, p25 and p22.5 were reported to be expressed from construct pHCRSV223invΔ5'UTR, which contained a deletion of the 5' UTR and inversion of the first 1000 nucleotides (Koh *et al.*, 2002). Subsequently deletion constructs pHCRSV223invΔ5'UTR/Δ1 and pHCRSV223invΔ5'UTR/Δ2 were made to investigate if expression of these proteins is truly independent of 5'-UTR and is under the control of the previously identified IRES element (Figure. 4.3a) (Koh *et al.*, 2006). Expression of pHCRSV223invΔ5'UTR showed that translation of p38, p27 and its isoforms are independent of the 5'UTR region. Expression of pHCRSV223invΔ5'UTR/Δ1, which contains the IRES element, resulted in enhanced expression of the four proteins (Fig. 4.3b, lanes 1 and 2). However, deletion of the IRES element in pHCRSV223invΔ5'UTR/Δ2 resulted in a significant reduction of



Koh et al., 2006. Virus Research.

Fig. 4.3

Fig 4.3 Mapping of the IRES element.

a. Schematic representation of constructs of pHCRSV223invΔ5'UTR, pHCRSV223invΔ5'UTR/Δ1, pHCRSV223invΔ5'UTR/Δ2. Numbers indicate nucleotide positions. The deleted regions are in dotted line. The IRES element is in open box.

b. Analysis of *in vitro* translation products from pHCRSV223invΔ5'UTR (lane 1), pHCRSV223invΔ5'UTR/Δ1 (lane 2), and pHCRSV223invΔ5'UTR/Δ2 (lane 3) in wheat germ extract with the TnT coupled translation system. Equal amounts of non-linearized templates were used. [^{35}S]methionine-labeled translation products were separated on 15% SDS-PAGE and visualized by autoradiography. Bands corresponding to p38, p27, p25 and p22.5 are indicated on the right. Numbers on the left indicate the molecular masses in kilodaltons.

the p38 and p27 expression (Fig. 4.3b, lanes 1 and 3). Interestingly, reduction of p25 and p22.5 expression was less dramatic (Fig. 4.3b, lanes 1 and 3), suggesting that expression of these four polypeptides by HCRSV genomic and subgenomic RNAs may be regulated by different mechanisms.

4.2 Translation of p38 is regulated by p27 through a leaky scanning mechanism

Previously *in vitro* analysis revealed that p27 translation is initiated from a non-AUG codon, ^{2570}CUG (Koh *et al.*, 2006) which is only 10 nt upstream of p38 initiation codon ^{2580}AUG (numbers in superscript are the locations of the first nucleotides on gRNA). To test if this non-AUG codon has any effect on translation of p38, the start codon of p27 CUG was mutated to AUG in sgRNA2 cDNA clone (Figure 4.4), designated as pHCRSV80M (or p80M). Expression of plasmid p80 in wheat germ extract through a TnT system (Promega) showed that p38 is the most abundant protein translated (Figure 4.5a, lane 1). However, an enhanced level of p27 (almost 3-fold increase) was observed in p80M compared to the wild type pHCRSV80 (or p80) (Figure 4.5a, lanes 1 and 3). On the contrary, a drastic reduction was observed in the translation of p38 and a moderate reduction of p22.5, but the translation of p25 did not appear to be affected significantly. To further confirm that these bands truly represented p38, p27 and its isoforms, a 6×His-tag was inserted into the 3'-terminus of p27 in both wild type p80 and p80M. The

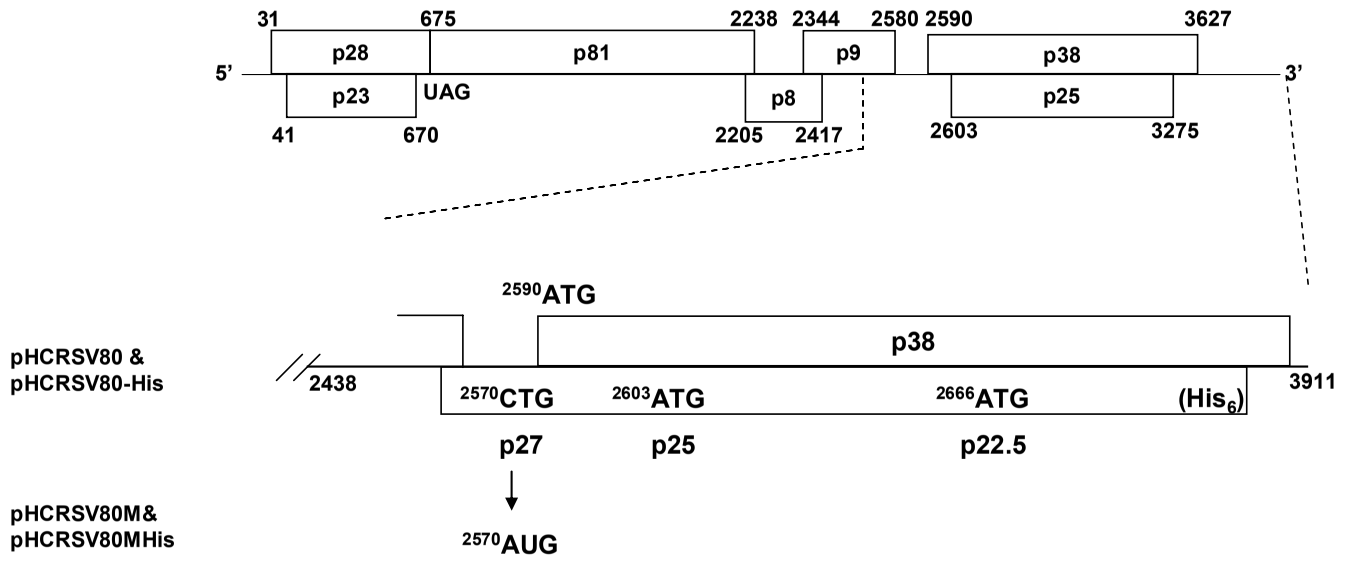

31
675
2238
2344
2580
2590
3627
p28
p81
p9
p38
5'
3'
p23
UAG
p8
p25
41
670
2205
2417
2603
3275
2590ATG
pHCRSV80 &
pHCRSV80-His
p38
2438
2570CTG
2603ATG
2666ATG
(His6)
3911
p27
p25
p22.5
pHCRSV80M&
pHCRSV80MHis
2570AUG

Fig. 4.4

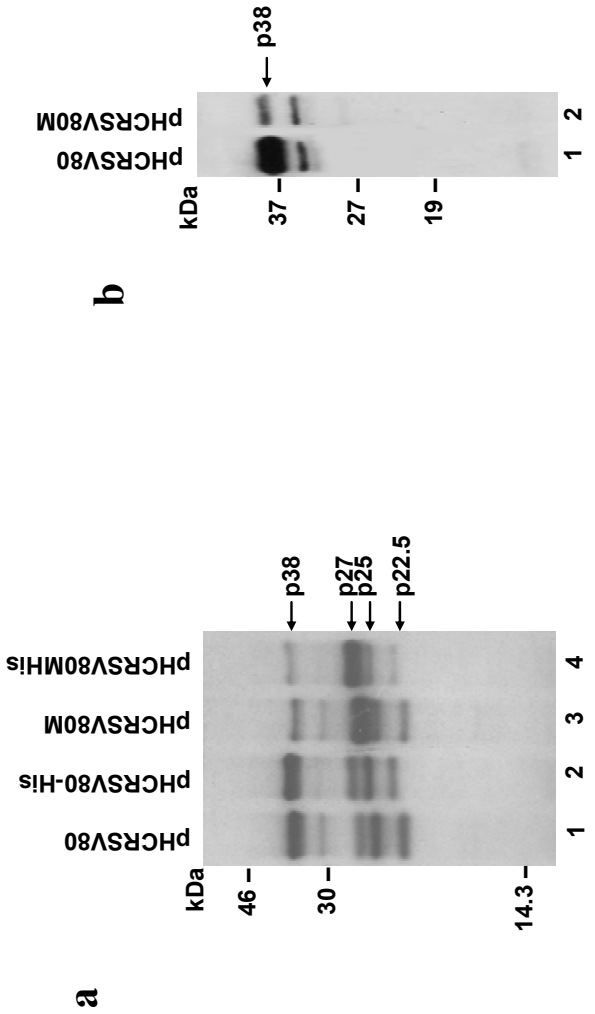

a
kDa
46
30
14.3
p38
p27
p25
p22.5
pHCRSV80
pHCRSV80-His
pHCRSV80M
pHCRSV80MHis
1
2
3
4
b
kDa
37
27
19
p38
pHCRSV80
pHCRSV80M
1
2

Fig. 4.5

Fig 4.4 Schematic representation of HCRSV genome organization and constructs of pHCRSV80, pHCRSV80-His and the mutants.

ORFs encoding p38, p27, p25 and p22.5 are highlighted. Numbers indicate nucleotide positions. The 6×His-tag and the mutation of p27 CUG are indicated.

Fig 4.5 Effect of p27 CUG on the expression of p38 in pHCRSV80.

a. Analysis of *in vitro* translation products from pHCRSV80 (lane 1), pHCRSV80His (lane 2), pHCRSV80M (lane 3), and pHCRSV80MHis (lane 4) in wheat germ extract with the TnT coupled translation system. Equal amounts of non-linearized templates were used. [^{35}S]methionine-labeled translation products were separated on 17.5% SDS-PAGE and visualized by autoradiography. Bands corresponding to p38, p27, p25, and p22.5 are indicated on the right. Numbers on the left indicated the molecular masses in kilodaltons.
b. Expression of pHCRSV80 (lane 1) and pHCRSV80M (lane 2) in HeLa cells. Cells were infected with the recombinant vaccinia/T7 virus at M.O.I. 1, and transfected with these two plasmid DNAs. At 18 h post transfection, total cell lysates were resolved on 12% SDS-PAGE and analyzed by western blot using anti-p38 antibody. Bands corresponding to p38 are indicated on the right. Numbers on the left indicate the molecular masses in kilodaltons.

resulting plasmids were carrying a suffix "His" (Figure 4.4). The resulting gene products migrated slower on SDS-PAGE and the bands were slightly higher than those of constructs without the His-tag (Figure 4.5a, lanes 2 and 4 compared to lanes 1 and 3). *In vitro* translation of these constructs further supported the view that translation of p38 may be regulated by p27 through a leaky scanning mechanism. To confirm if this effect could be observed *in vivo*, plasmids p80 and p80M were transfected into HeLa cells and the p38 was probed with its rabbit antiserum. Similar to the *in vitro* system, translation of p38 was greatly impaired when p27 CUG was mutated to AUG. All the above data suggested that: firstly, CUG codon of p27 is a relatively weak initiation codon; and secondly, this weak initiation codon may regulate downstream ORF translation through a leaky scanning mechanism.

4.3 An IRES element plays a role in p38 translation

A previous study reported that p38, p27, p25 and p22.5 could be expressed independent of the 5'UTR based on the observation that these proteins were all translated from a construct in which the 5'UTR was deleted and the first 1000 nucleotides were inverted (Koh *et al*., 2003). Subsequently, deletion study identified a putative IRES element within ORF p9 of about 100-nucleotide long (Koh *et al*., 2003). To study the effect of this IRES element on the translation of downstream ORFs in HCRSV sgRNA2, a *green fluorescence protein* (GFP) gene was inserted

upstream of the IRES element in p80, giving rise to pHCRSV80-GFP (Figure 4.6a). Expression of this construct *in vitro* showed the production of GFP as well as p38, p27, p25 and p22.5 (Figure 4.6b, lane 2). Compared to the expression of p80, the GFP protein was the most abundant in pHCRSV80-GFP since it is at the most 5' end and p38, p27, p25 and p22.5, were all reduced to some extent. As a control, in another construct pHCRSV80Δ-GFP, the whole region of IRES was deleted (Figure 4.6a bottom). While the expression of GFP was comparable to that in pHCRSV80-GFP, a drastic reduction was observed in p38 and p22.5 expression (Figure 4.6b, lanes 2 and 3), indicating that IRES may regulate the translation of p38 and p22.5 in sgRNA2. Interestingly, not much changes in the translation of p27 and p25 were observed (Figure 4.6b, lane 3). This suggests that IRES has less effect on p27 and p25, implying that these two protein expression are of different mechanisms from that of p38 and p22.5. One explanation could be that p27 translation is from a termination and re-initiation mechanism since ORF GFP and ORF p27 are relatively close to each other. Translation of p25 could have resulted from leaky scanning from p27 for their start codons which are only 33 nucleotides away with a weak initiation codon for p27. In summary, these data indicate that this IRES element is functional in sgRNA2 and plays a role in translation of p38 and p22.5.

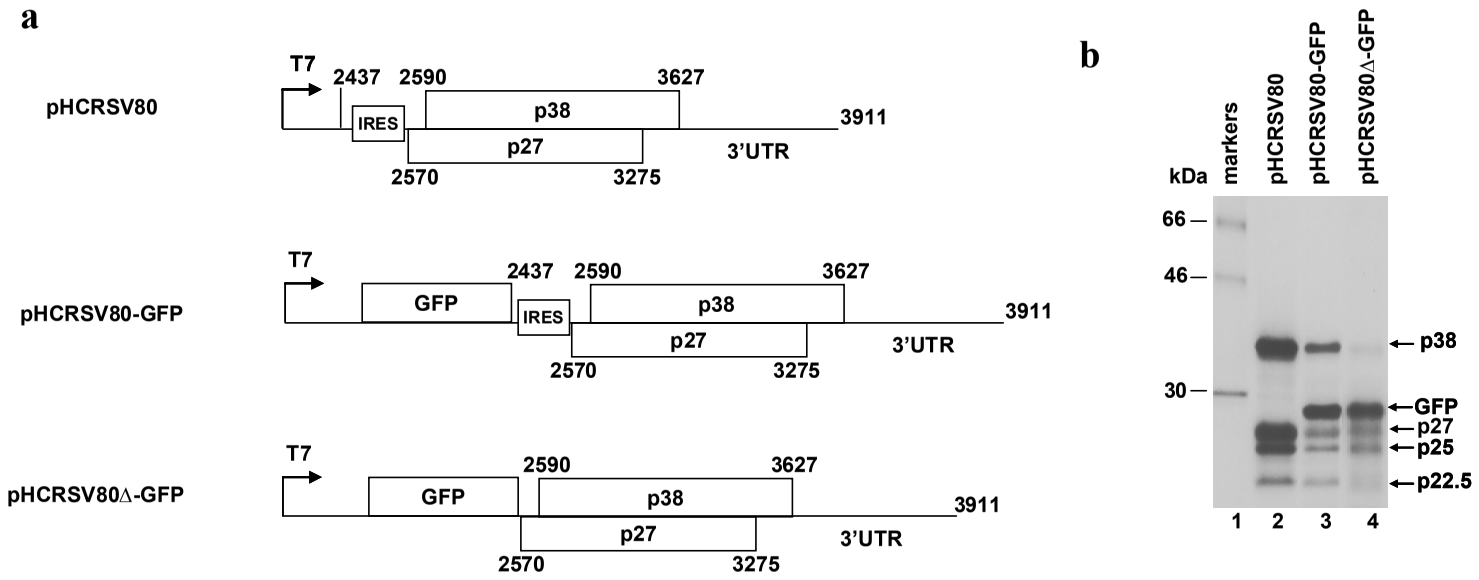

a
pHCRSV80
T7
2437
2590
3627
IRES
p38
3911
p27
3'UTR
2570
3275
pHCRSV80-GFP
T7
2437
2590
3627
GFP
IRES
p38
3911
p27
3'UTR
2570
3275
pHCRSV80Δ-GFP
T7
2590
3627
GFP
p38
3911
p27
3'UTR
2570
3275
b
kDa
markers
pHCRSV80
pHCRSV80-GFP
pHCRSV80Δ-GFP
66
46
30
p38
GFP
p27
p25
p22.5
1
2
3
4

Fig. 4.6

Fig 4.6 Analysis of the IRES element.

a. Schematic representation of constructs of pHCRSV80, pHCRSV80-GFP and pHCRSV80Δ-GFP. GFP gene, p38, and p27 as well as the IRES element are in open boxes. Numbers indicate nucleotide positions.

b. Analysis of *in vitro* translation products from pHCRSV80 (lane 1), pHCRSV80-GFP (lane 2) and pHCRSV80Δ-GFP (lane 3) in wheat germ extract. Equal amount of non-linearized DNA templates were used. [^{35}S]methionine-labeled translation products were separated on 17.5% SDS-PAGE and visualized by autoradiography. Bands corresponding to p38, GFP, p27, p25, and p22.5 are indicated on the right. Numbers on the left indicate the molecular masses in kilodaltons.

4.4 Effect of upstream small ORF p9 on the translation of downstream ORFs

ORF p9 encodes for a small protein of only about 9 kDa. The 3' terminus of p9 is overlapped with the 5' terminus of p27, with the CUG codon of p27 10 nt upstream of the TAA codon of p9 (Figure 4.7a). To investigate whether p9 affects the translation of downstream p27, p38, p25 and p22.5, we made various mutations in ORF p9 as well as p27 based on the construct pHCRSV223invΔ5'UTR. To reduce the effect of 5'UTR on the translation, this construct is the full-length cDNA clone of HCRSV with a deletion of the 5'UTR and an inversion of the first 1,000 nucleotides.

Firstly, p9 start codon AUG was mutated to GTG in both constructs with p27 CUG and p27 AUG (pHCRSV223invΔ5'UTR/M1, M2 and M3) to analyze whether translation of p9 is important for the translation of downstream ORFs (Figure 4.7a). Figure 4.7b is the *in vitro* translation of these plasmids. The relevant proteins are indicated by arrows on the left. Translation of p38 was only moderately affected in both constructs with p27 CUG and p27 AUG (comparing lanes 2 and 1, lanes 4 and 3). In constructs with the wild type p27 initiation codon CUG (pHCRSV223invΔ5'UTR and M3), p9 translation did not appear to influence the downstream ORFs and the ratio of p27/p38 kept relatively constant. Consistent with previous studies, a strong p27 initiation codon AUG (M1 and M2) induced significant reduction of p38 and elevation of p27 (Figure 4.7b, lanes 1 and 3), with

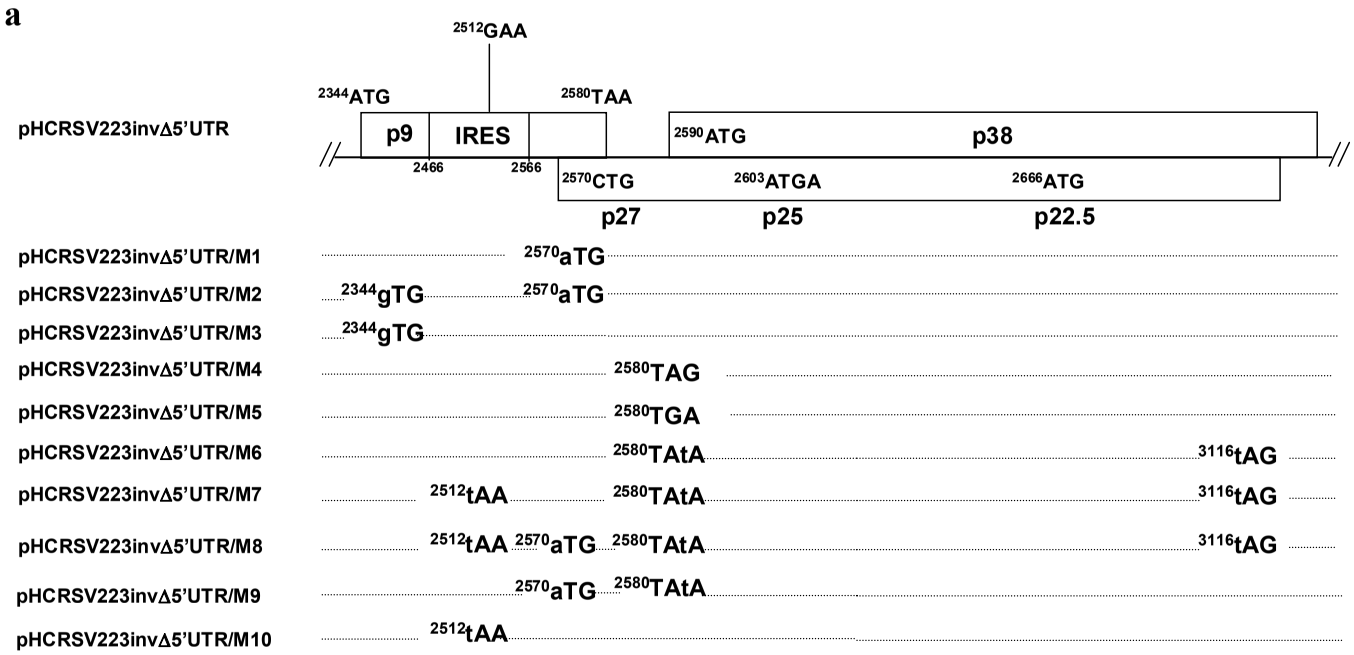

a
2512GAA
2344ATG
2580TAA
pHCRSV223invΔ5'UTR
p9
IRES
2590ATG
p38
2466
2566
2570CTG
2603ATGA
2666ATG
p27
p25
p22.5
pHCRSV223invΔ5'UTR/M1
2570aTG
pHCRSV223invΔ5'UTR/M2
2344gTG
2570aTG
pHCRSV223invΔ5'UTR/M3
2344gTG
pHCRSV223invΔ5'UTR/M4
2580TAG
pHCRSV223invΔ5'UTR/M5
2580TGA
pHCRSV223invΔ5'UTR/M6
2580TAtA
3116tAG
pHCRSV223invΔ5'UTR/M7
2512tAA
2580TAtA
3116tAG
pHCRSV223invΔ5'UTR/M8
2512tAA
2570aTG
2580TAtA
3116tAG
pHCRSV223invΔ5'UTR/M9
2570aTG
2580TAtA
pHCRSV223invΔ5'UTR/M10
2512tAA

Fig. 4.7

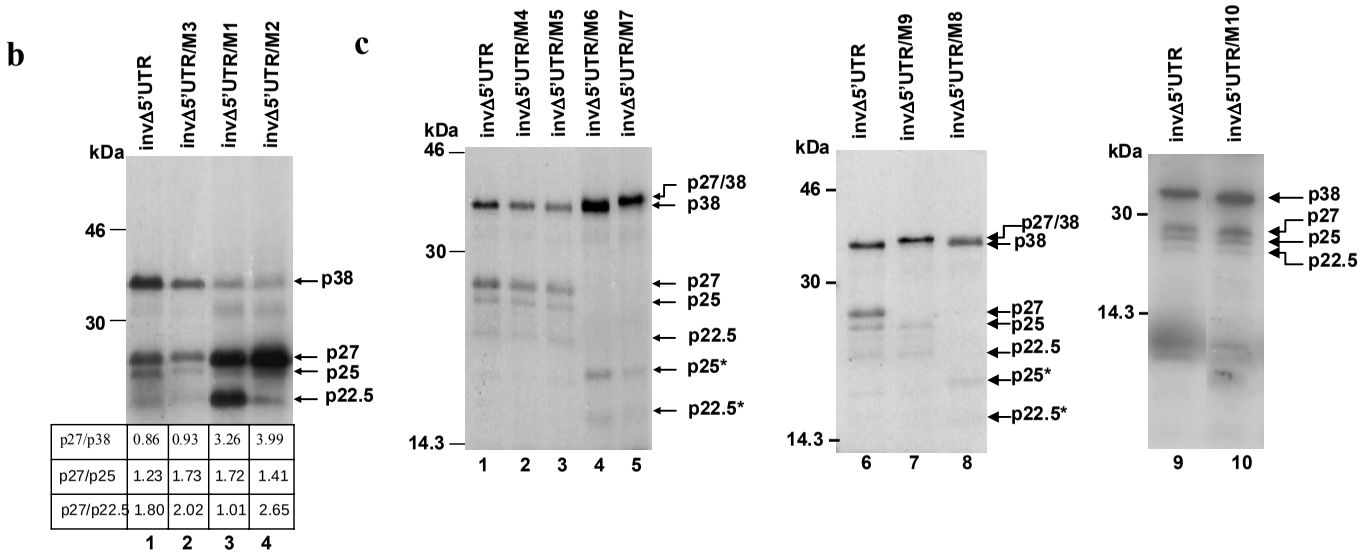

b
invΔ5'UTR
invΔ5'UTR/M3
invΔ5'UTR/M1
invΔ5'UTR/M2
kDa
46
30
p38
p27
p25
p22.5
p27/p38 0.86 0.93 3.26 3.99
p27/p25 1.23 1.73 1.72 1.41
p27/p22.5 1.80 2.02 1.01 2.65
1 2 3 4
c
invΔ5'UTR
invΔ5'UTR/M4
invΔ5'UTR/M5
invΔ5'UTR/M6
invΔ5'UTR/M7
kDa
46
30
14.3
p27/38
p38
p27
p25
p22.5
p25*
p22.5*
1 2 3 4 5
invΔ5'UTR
invΔ5'UTR/M9
invΔ5'UTR/M8
kDa
46
30
14.3
p27/38
p38
p27
p25
p22.5
p25*
p22.5*
6 7 8
invΔ5'UTR
invΔ5'UTR/M10
kDa
30
14.3
p38
p27
p25
p22.5
9 10

Fig. 4.7

Fig 4.7 Effect of small upstream ORF p9 on the expression of downstream ORFs.

a. Schematic representation of constructs pHCRSV223invΔ5'UTR, pHCRSV223invΔ5'UTR/M1 to M10. ORFs p9, p38, p27, p25, and p22.5 are highlighted in open boxes. IRES element is also indicated. Numbers indicate nucleotide positions. Mutations in each construct are indicated and the mutated nucleotides are indicated as small caps.

b. Analysis of *in vitro* translation products from pHCRSV223invΔ5'UTR (lane 1), pHCRSV223invΔ5'UTR/M1 (lane 3), M2 (lane 4) and M3 (lane 2) in wheat germ extract. Equal amount of non-linearized templates were used. [^{35}S]methionine-labeled translation products were separated on 17.5% SDS-PAGE and visualized by autoradiography. Bands corresponding to p38, p27, p25, and p22.5 are indicated on the right. Numbers on the left indicate the molecular masses in kilodaltons. Bands intensities are read by densitometer and the table below shows the ratios of p27/p38, p27/p25, and p27/p22.5.

c. Analysis of *in vitro* translation products from pHCRSV223invΔ5'UTR (lanes 1, 6, and 9), pHCRSV223invΔ5'UTR/M4 (lane 2), M5 (lane 3), M6 (lane 4), M7 (lane 5), M8 (lane 8), M9 (lane 7), and M10 (lane 10) in wheat germ extract. Equal amount of non-linearized DNA templates were used. [^{35}S]methionine-labeled translation products were separated on 17.5% SDS-PAGE and visualized by autoradiography. Bands corresponding to p38, p27, p25, p22.5 are indicated on the right of each panel. The truncated version of p25 and p22.5 are indicated with (*). The fusion product of p27 and p38 is indicated as p27/38. Numbers on the left of each panel indicate the molecular masses in kilodaltons.

the ratio from 0.86 to 3.26. However, translation of p22.5 in M1 was also increased which was not observed in p80M, suggesting that other component(s) absent in sgRNA 2 might regulate the translation of p22.5 (Figure 4.7b, lanes 1 and 3). When p9 expression was abolished (M2, Figure 4.7b, lane 4), p27 translation was dramatically increased. In contrast, expression levels of p38, p25 and p22.5 decreased. Nevertheless, the ratio of p27/p38 did not change greatly compared to that of M1 (Figure 4.7b, lane 3). These data suggest that expression of p9 did not have much impact on the translation of downstream p27, p38 and p25. However, expression of p22.5 can be greatly affected by p9, if p27 carries a strong initiation codon.

Next, the effect on p9 stop codon on translation of downstream ORFs was studied. When the stop codon of p9 ^{2580}TAA was changed to amber or ophre codon (^{2580}TAG or ^{2580}TGA) (M4 or M5, respectively, Figure 4.7a), the expression levels of p27, p38, p25 and p22.5 in rabbit reticulocyte did not alter greatly (Figure 4.7c, lanes 1-3). The stop codon of p9 was then abolished by inserting an additional nucleotide T from ^{2580}TAA to ^{2580}TATA (M6, Figure 4.7a), making p27 and p38 in the same reading frame, resulting in a fusion protein p27/38 as indicated in Figure 4.7c. The fusion protein is 7 amino acids longer than p38 and indeed a slightly higher band was observed. However, during the course of cloning, an additional point mutation was introduced at nucleotide 3116 from G to T, leading to a truncated version of p25

as well as p22.5. These mutations did not affect p27/38 and were denoted as p25* and p22.5*. *In vitro* translation of this construct showed a doublet around the location equivalent to 38 kDa, the lower one of which is believed to result from leaky scanning. Considering p38 is translated by leaky scanning from p27, which means a portion of the ribosomes landing on the transcripts initiate from p38 instead of p27 or p27/38. The total number of ribosomes therefore should be the sum of those translating p27 or p27/38 and those of p38. There was a comparable level of p27/38 and p38 than that of p38 and p27 in control plasmid (pHCRSV223invΔ5'UTR) (Figure 4.7c, lanes 1 and 4), indicating that this mutation obviously did not affect the expression efficiency. Similar levels were also observed for p25 and p22.5 expression. In plasmid M7, a point mutation at nt 2512 from G to T introduced a premature termination codon in ORF p9 (Figure 4.7a), which is located 68 nucleotides upstream of ^{2580}TAA. The expression of this construct showed similar expression level to construct M6 (Figure 4.7c, lane 5). To confirm these results, a construct carrying only the ^{2512}TAA mutation was made, yielding M10 (Figure 4.7a). The *in vitro* translation of M10 showed that the expression levels of p38, p27, p25 and p22.5 were comparable to wild type, suggesting no effect of p9 stop codon could be observed on the translation of downstream ORFs (Figure 4.7c, lanes 9 and 10). Based on M6 and M7, the start codon of p27 was mutated from ^{2570}CUG to ^{2570}AUG (Figure 4.7a), yielding M9 and M8, respectively. However, M9

did not carry the point mutation at nt 3116, while M8 did. The *in vitro* translation of M9 revealed that the lower band of the doublet, corresponding to p38 probably from leaky scanning, was greatly reduced (Figure 4.7c, lane 7). Consistent with M6, change of p9 stop codon did not affect the efficiency of downstream ORFs translation even with a strong p27 initiation codon. Translation of M8 showed similar results to that of M7 (Figure 4.7c, lanes 5 and 8). All the above data suggest that although there is a small ORF p9 upstream of p27, it does not appear to regulate the expressions of downstream ORFs p27, p38, p25 and p22.5.

4.5 Discussion

RNA viruses employ various mechanisms to regulate their gene expression at the translation level. In this study, we identified a CUG codon (^{2570}CUG) that is responsible for the initiation of p27. Consistent with the leaky scanning model, initiation of p38 expression was demonstrated to be regulated by leaky scanning of ribosomes at the ^{2570}CUG, the initiation codon for p27. Recently, presence of p27 *in vivo* was confirmed in the protoplast system (Zhou *et al*., 2006). We also showed that additional regulatory mechanisms are involved in the control of the expression of p38, p27, p25 and p22.5.

Besides SBWV (Shirako, 1998) and RTBV (Futterer *et al.*, 1996), HCRSV is the first carmovirus shown to utilise a non-AUG initiation codon. Utilization of

alternative initiation codons could increase gene diversity and control the relative expression of proteins with different locations and functions encoded by the same mRNA. For example, multiple isoforms of eIF4GI with different N-termini can be generated by use of alternate translation initiation codons. These isoforms possess alternative binding sites for different ancillary translation factors that allow the regulation of translation (Byrd *et al.*, 2002). The choice of initiation codons can also determine the subcellular fate and function of proteins, such as the localization of CUG-initiated isoforms of human fibroblast growth factor-2 (FGF-2) in nucleus, and AUG-initiated isoforms in the cytosol (Arnaud *et al*., 1999). So far, we do not know if the three isoforms, p27, p25 and p22.5, play distinct roles in HCRSV replication cycles and pathogenesis. Analysis of the functions of p27 and p25 in kenaf plants revealed that they might be involved in symptom severity and virus movement, respectively (Zhou *et al*., 2006).

The efficiency of translation initiation at non-AUG codons, such as CUG, GUG, ACG and AUU, is generally lower than that of the AUG codon (Peabody, 1989). Non-AUG initiation codons, in spite of being surrounded by strong contexts, are generally considered weak initiation sites and could facilitate leaky scanning of ribosomes (Portis *et al*., 1994). The significance of the presence of non-AUG codons as initiator in an alternative translation initiation can be distinguished by the abolishment of the expression of a downstream-initiated isoform resulting from

replacement of CUG with an AUG codon. Over-expression of a 34-kDa CUG-initiated isoform of FGF-2 resulted in the maintenance of cell proliferation that could be harmful to the cells (Arnaud *et al.*, 1999). Therefore, the utilization of weak initiation codon CUG to initiate the synthesis of the 34-kDa isoform of FGF-2 is necessary for maintaining the correct ratio of isoforms. In this study, replacement of the ^{2570}CUG with ^{2570}AUG led to the over-expression of p27 and drastic reduction in the expression of viral coat protein. Plants inoculated with such mutant transcripts revealed that systemic spread of the virus was impeded (Zhou *et al*., 2006), possibly due to the reduction of viral coat protein synthesis that is essential for viral packaging and movement.

Leaky scanning is a common mechanism employed by plant viruses to direct multiple gene expression from one mRNA. Examples include RTBV (Futterer *et al*., 1996 and 1997), BYDV (Dinesh-Kumar and Miller, 1993), TYMV (Weiland and Dreher, 1989) and *Plum pox virus* (PPV) (Simon-Buela *et al.*, 1997). In this study, we showed that expression of an important viral structural component, viral coat protein (p38) is regulated by leaky scanning at the ^{2570}CUG initiation codon of p27. This was clearly demonstrated when replacement of the weak CUG initiation codon with an AUG codon resulted in a drastically reduced expression of downstream p38.

In a recent study, close proximity of AUGs was shown to account for the efficient expression of polypeptides from extensively overlapping ORFs encoded in

the genomic RNA of TYMV (Matsuda and Dreher, 2006). In addition to a translation initiation event occurring at an optimal upstream AUG in TYMV, translation was also initiated at an AUG codon located 7-nt downstream, indicating the selection of initiation codons is not strictly sequential (i.e. 5'-3' direction) but competitive. Such observation was suggested to be inconsistent with the leaky scanning model in which only the first 5'-proximal initiation codon in an optimal sequence context would be selected for translation initiation and not other optimal initiation codons downstream. It was proposed in TYMV, that ribosome movement involves small amplitude of forward and backward oscillations allowing competition for translation initiation between optimal AUGs, averaging about 15 nucleotides. In HCRSV, when expression of p27 was initiated from a strong AUG codon, increased expression of downstream ORF (p22.5) was also observed. This could also result from a similar ribosome oscillation. However, the long distance (96 nucleotides) between p27 ^{2570}CUG and p22.5 ^{2666}AUG argues against this possibility. Instead, it could be suggestive of additional regulatory element(s) which is absent in sgRNA2 since the same mutation in construct p80 did not result in the same increment of p22.5.

Upstream ORFs (uORFs) have been shown to play a role in regulating translation reinitiation, especially in the translation of cytokines, transcription factors and other potent proteins that are required only in small amounts (Kozak, 2002). Depending on the availability of nutrients, the expression of yeast *GCN4* is regulated

by four uORFs (Gaba *et al*., 2001). In many cases, mutations targeted to the upstream AUG codons confirmed their role in restricting downstream ORF translation. In this study, abolishment of p9 AUG codon led to dramatic enhancement of the initiation efficiency of p27 which possessed the ^{2570}AUG codon. This suggests that both AUGs of p9 and p27 may be competing for the same pool of ribosomes. However, only marginal enhancement was observed when the initiation codon of p27 remained as ^{2570}CUG. It is likely that these intricate regulatory mechanisms would ensure the production of p27 and p38 in their correct ratio during HCRSV replication cycles. The relatively constant ratio between p27 and p38 would possibly render evolutional advantages to the virus.

A uORF can also affect the expression of downstream ORFs by termination-reinitiation. Depending on the length of the uORF, reinitiation can be categorized into two groups. Protein expression of yeast *GCN4* and mammalian AdoMetDC gene are regulated by short uORFs that are less than 30 codons (Hinnesbusch, 1997; Morris and Geballe, 2000) with the efficiency of translation reinitiation decreasing with increasing length of the uORF. A long intercistronic region between the uORF and the main ORF is a characteristic feature of such regulation. On the contrary, there are also examples of functional proteins regulated by long uORFs of more than 30 codons such as *Influenza B virus* RNA segment 7 (Peabody *et al.*, 1986), CaMV 35S RNA (Bonneville *et al*., 1989; Godwa *et al*., 1989;

Horvath *et al.*, 1990), ORF2 of *Human respiratory syncytial virus* (Ahmadian *et al.*, 2000) and calicivirus RHDV (Meyers, 2003). These RNAs not only possess long uORFs but also have short intercistronic region and some even contain overlaps of the upstream and downstream ORFs. It has been suggested that such RNAs utilize a set of translation initiation mechanisms different from RNAs with short uORFs and long intercistronic regions. In addition, the position of the uORF termination codon is usually crucial for reinitiation as certain features might prevent the resumption of scanning (Grant and Hinnebusch, 1994; Vilela *et al.*, 1998). Mutagenesis studies performed in this investigation showed that the mutation of the initiation and termination codons for p9 affect the expression downstream proteins especially p22.5 in the presence of the ^{2570}AUG codon but not ^{2570}CUG. Therefore, it has yet to be established if a termination-reinitiation event happens in the expression of naturally-occurring HCRSV genomic RNA.

CHAPTER 5. CONCLUDING REMARKS AND FUTURE WORK

The main focus of the work presented in this thesis is the investigation of translational regulation mechanisms of two RNA viruses (SARS-CoV and HCRSV). In this chapter, results of the research work are summarized, and discussed and suggestions are made for future directions.

5.1 Frame-shift events in the expression of SARS-CoV 3a ORF variants

SARS-CoV 3a protein is one of the accessory proteins encoded from the genome. The expression of 3a protein was previously detected during both *in vitro* and *in vivo* infection (Tan YJ *et al*., 2005). When analyzing the viral transcripts in the sera of infected patients, researchers found a mixed population of sgRNA3 transcripts (Tan TH *et al*., 2005). Although the full-length 3a protein is not essential for virus replication, it is possible that 3a would either contribute to viral pathogenesis as it was shown to be associated with the spike protein or confer a fitness gain under selective pressure.

Programmed ribosomal frame-shifting allows viruses to produce a second protein from a single mRNA. Frame-shifting occurs at a much higher frequency than natural errors, and is responsive to specific mRNA sequences. The resulting gene products, in some cases, are essential for virus replication or assembly. The ratios between the two proteins are kept relatively constant which is a means of viral gene regulation. The efficiency of most -1 frame-shifting can be enhanced by ancillary

elements, such as an RNA pseudoknot or a *cis*-element that can base-pair with ribosome RNAs. The aim of this work was to characterize the frame-shifting mechanisms that account for the full-length 3a expression in ORF 3a variants. By looking into this, we aim to further understand the mechanism for programmed frame-shifting as well as the decisive factors.

In Chapter 3, hepta- and octo-uridine stretches (UUU UUU U and UUU UUU UU) were identified to be capable of inducing +1/-2 and -1/+2 frame-shifting, respectively. By deletion studies, no other *cis*- or *trans*-element was found to be required for efficient frame-shifting. These findings are consistent with the mutagenesis study by Brierley *et al.* on IBV frame-shifting sequence, supporting the general view that mono-runs of uridines are very slippery. Based on mutagenesis studies of the hepta-uridine stretch, all the nucleotides in this slippery sequence appeared to be involved in frame-shifting. Mutagenesis studies of the octo-uridine stretch showed that the first seven uridines were essential for frame-shifting. These studies suggest a double slippage at both ribosomal A and P sites for the frame-shifting to occur.

The exact mechanisms, however, seemed to vary in different contexts. Mutagenesis studies on octo-uridine stretch in the S1ab context suggest a different frame-shifting mechanism from that in ORF 3a. Since neither the phenylalanine $tRNA_{GAA}$ is a hungry codon in mammals nor that a stop codon is present in the

proximity to the slippage site, it is unlikely that frame-shifting is induced by pausing at the slippage site. By examining the effect of RNA secondary structure on frame-shifting, we found that it rendered no effect on frame-shifting mediated by hepta- and octo-uridine stretches. In contrast, this RNA structure can enhance conventional -1 frame-shifting of SARS-CoV 1a/1b, indicating that frame-shifting mechanism caused by the uridine stretches is distinct from that mediated by other -1 frame-shifting signals. Therefore, determination of the amino acid sequence spanning the frame-shifting site is necessary to tell which event actually takes place to produce the full-length 3a protein.

Another interesting finding is that octo-uridine stretch can mediate the production of proteins from multiple reading frames. This could have resulted from both ribosomal frame-shifting and transcriptional slippage. Future work can focus on the nature of programmed frame-shifting, such as various factors including the nature of tRNA and the elongation factors that affect frame-shifting efficiency. Other protein partners that may interact with the poly-pyrimidine tract are of interest as well. In addition, it would be more accurate if the effect of ribosomal frame-shifting could be evaluated separately from transcriptional slippage.

5.2 Translational control of HCRSV p38, p27 and its isoforms

Translation of mRNAs requires numerous enzymatic components which are

not encoded within viral genomes, although the complexity of viruses varies enormously. Consequently, viruses compete with the host cell at the translational level. To optimize viral mRNA translation, many viruses have evolved remarkably sophisticated mechanisms, especially at the first step of translation for the recruitment of the host translation initiation factors. The purpose of this study was to characterize the mechanisms that regulate the translation initiation of the coat protein and other polypeptides, by using HCRSV as a model.

In Chapter 4, we showed that a polypeptide p27 overlapping with the coat protein (p38) was translated through a non-conventional initiation codon CUG. This initiation codon was able to control p38 expression probably through a leaky scanning mechanism since mutation from CUG to AUG led to significant decrease of p38 production. In addition, a previously identified IRES element upstream of p38 was shown to regulate the translation of p38 as well. These data suggest that production of p38 could be the result of two or possibly more regulatory events. Since the coat protein is essential in virus particle assembly and replication, especially the copy number of this protein would directly affect virus morphology, steady production of the protein by sophisticated regulation is very critical. Another important finding is that the ratio of p38 and p27 was relatively stable and resistant to mutations. It is likely that this ratio is also essential in viral replication and assembly, as the constant ratio of the two proteins may be advantageous for the virus

to maintain a balance in the genes being expressed.

As shown in Chapter 4, mechanisms that regulate the expression of the polypeptides p25 and p22.5 were also studied. By mutagenesis analysis, we found that expression of p22.5 *in vitro* was greatly affected when translation of p9 was abolished and with a CUG initiation codon for p27. Interestingly, p25 expression was not influenced by all these mutations. It suggests that translation regulation of p22.5 and p25 is of distinct mechanisms. However, further detailed analysis is required to fully explore the mechanisms for expression of p22.5 and p25.

Future work would be carried out in the following areas: 1) studies of the host factors involved in these regulations. For example, identification of host factors interacting with the IRES element will help us to understand the role of host proteins in translational control of viral proteins. 2) *in vivo* studies are indispensable. As the work in this thesis was mostly done in an *in vitro* system, *in vivo* results may differ. In addition, the effects in *in vitro* studies may be under-estimated due to the more complicated situation in whole plants. 3) mechanisms that regulate the expression of p22.5 and p25 can be studied further. Based on the observation that introduction of p27 AUG did not affect the expression of p25 and p22.5, ribosome shunting could be one of the mechanisms that regulate the translation of these two proteins. This possibility can be tested by insertion of a stable stem-loop structure upstream the AUG for p25, since this structure will affect linear scanning but not ribosome

shunting.

5.3 Main conclusions

- The investigation of frame-shift events adds to a better understanding of frame-shifting and viral gene regulation. Frame-shifting is relatively common in viral gene expression and is one of the strategies used by viruses to expand the coding capacity in most cases. However, in the ORF 3a variants, frame-shifting is to maintain certain level of full-length 3a expression. In this sense, it can be a means of anti-mutation.
- Understanding of translation regulation by multiple mechanisms. The study on p38 expression *in vitro* suggests that multiple regulation events could happen in HCRSV protein production. Similar regulatory mechanisms may also be applicable to other RNA viruses and even mammalian genes.
- Further understanding of translational regulation at different stages. Viral gene regulation is an intricate event which may involve mechanisms at multiple controling points during translation. By examination of the initiation regulation and frame-shift events, it enables us to understand further how viruses take advantage of host translational apparatus in favor of their own gene expression. The studies could also provide useful target for drug design or vaccine development for the control of viral pathogens.

REFERENCES:

Abastado, J.P., Miller, P.F., Hinnebusch, A.G. 1991a. A quantitative model for translational control of the GCN4 gene of Saccharomyces cerevisiae. New Biol. 3, 511-24

Abastado, J.P., Miller P.F., Jackson, B.M., Hinnebusch, A.G. 1991b. Suppression of ribosomal reinitiation at upstream open reading frames in amino acid-starved cells forms the basis for GCN4 translational control. Mol. Cell. Biol. 11, 486-96

Adamski, F.M., Donly, B.C., Tate, W.P. 1993. Competition between frameshifting, termination and suppression at the frameshift site in the Escherichia coli release factor-2 mRNA. Nucleic Acids Res. 21, 5074-8

Alonso, M.A., and Carrasco, L. 1981. Reversion by hypotonic medium of the shutoff of protein synthesis induced by encephalomyocarditis virus. J. Virol. 37, 535-40

Arnaud, E., Touriol, C., Boutonnet, C., Gensac, M.C., Vagner, S., Prats, H., Prats, A.C. 1999. A new 34-kilodalton isoform of human fibroblast growth factor 2 is cap dependently synthesized by using a non-AUG start codon and behaves as a survival factor. Mol. Cell Biol. 19, 505-14

Asano, K., Clayton, J., Shalev, A., Hinnebusch, A.G. 2000. A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. Genes Dev. 14, 2534-46

Atkins, J.F., Lewis, J.B., Anderson, C.W., Gesteland, R.F. 1975. Enhanced differential synthesis of proteins in a mammalian cell-free system by addition of polyamines. J. Biol. Chem. 250, 5688-95

Atkinson, J., Dodge, M., Gallant, J. 1997. Secondary structures and starvation-induced frameshifting. Mol. Microbiol. 26, 747-53

Bab, I., Smith, E., Gavish, H., Attar-Namdar, M., Chorev, M., Chen, Y.C., Muhlrad, A., Birnbaum, M.J., Stein, G., Frenkel, B. 1999. Biosynthesis of osteogenic growth peptide via alternative translational initiation at AUG85 of histone H4 mRNA. J. Biol. Chem. 274, 14474-81

Babik, J.M., Adams, E., Tone, Y., Fairchild, P.J., Tone, M., Waldmann, H. 1999. Expression of murine IL-12 is regulated by translational control of the p35 subunit. J.

Immunol. 162, 4069-78

Barak, Z., Lindsley, D., Gallant, J. 1996. On the mechanism of leftward frameshifting at several hungry codons. J. Mol. Biol. 256, 676-84

Baranov, P.V., Henderson, C.M., Anderson, C.B., Gesteland, R.F., Atkins, J.F., Howard, M.T. 2005. Programmed ribosomal frameshifting in decoding the SARS-CoV genome. Virology 332, 498-510

Barry, J.K., and Miller, W.A. 2002. A -1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. Proc. Natl. Acad. Sci. 99, 11133-8

Beckenbach, A.T., Robson, S.K., Crozier, R.H. 2005. Single nucleotide +1 frameshifts in an apparently functional mitochondrial cytochrome b gene in ants of the genus Polyrhachis. J. Mol. Evol. 60, 141-52

Beier, H., Barciszewska, M., Krupp, G., Mitnacht, R., Gross, H.J. 1984. UAG readthrough during TMV RNA translation: isolation and sequence of two tRNAs with suppressor activity from tobacco plants. EMBO J. 3, 351-356

Belcourt, M.F., Farabaugh, P.J. 1990. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. Cell 62, 339-52

Belsham, G.J., and Sonenberg, N. 2000. Picornavirus RNA translation: roles for cellular proteins. Trends Microbiol. 8, 330-5

Bernstein, J., Sella, O., Le, S.Y., Elroy-Stein, O. 1997. PDGF2/c-sis mRNA leader contains a differentiation-linked internal ribosomal entry site (D-IRES). J. Biol. Chem. 272, 9356-62

Bidou, L., Stahl, G., Grima, B., Liu, H., Cassan, M., Rousset, J.P. 1997. In vivo HIV-1 frameshifting efficiency is directly related to the stability of the stem-loop stimulatory signal. RNA 3, 1153-8

Bieleski, L., and Talbot, S.J. 2001. Kaposi's sarcoma-associated herpesvirus vCyclin open reading frame contains an internal ribosome entry site. J. Virol. 75, 1864-9

Blinkowa, A.L., Walker, J.R. 1990. Programmed ribosomal frameshifting generates

the Escherichia coli DNA polymerase III gamma subunit from within the tau subunit reading frame. Nucleic Acids Res. 18, 1725-9

Blobel, G. 1973. A protein of molecular weight 78,000 bound to the polyadenylate region of eukaryotic messenger RNAs. Proc. Natl. Acad. Sci. 70, 924

Boeck, R., and Kolakofsky, D. 1994. Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis. EMBO J. 13, 3608-17

Bonneville, J.M., Sanfacon, H., Futterer, J., Hohn, T. 1989. Posttranscriptional trans-activation in cauliflower mosaic virus. Cell 59, 1135-43

Brault, V., Perigon, S., Reinbold, C., Erdinger, M., Scheidecker, D., Herrbach, E., Richards, K., Ziegler-Graff, V. 2005. The polerovirus minor capsid protein determines vector specificity and intestinal tropism in the aphid. J. Virol. 79, 9685-93

Brault, V., van den Heuvel, J.F., Verbeek, M., Ziegler-Graff, V., Reutenauer, A., Herrbach, E., Garaud, J.C., Guilley, H., Richards, K., Jonard, G. 1995. Aphid transmission of beet western yellows luteovirus requires the minor capsid read-through protein P74. EMBO J. 14, 650-9

Bridgen, A., Weber, F., Fazakerley, J.K., Elliott, R.M. 2001. Bunyamwera bunyavirus nonstructural protein NSs is a nonessential gene product that contributes to viral pathogenesis. Proc. Natl. Acad. Sci. 98, 664-9

Brierley, I., Digard, P., Inglis, S.C. 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell 57, 537-47

Brierley, I., Jenner, A.J., Inglis, S.C. 1992. Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal. J. Mol. Biol. 227, 463-79

Brown, E.G., Liu, H., Kit, L.C., Baird, S., Nesrallah, M. 2001. Pattern of mutation in the genome of influenza A virus on adaptation to increased virulence in the mouse lung: identification of functional themes. Proc. Natl. Acad. Sci. 98, 6883-8

Bruce, A.G., Atkins, J.F., Gesteland, R.F. 1986. tRNA anticodon replacement

experiments show that ribosomal frameshifting can be caused by doublet decoding. Proc. Natl. Acad. Sci. 83, 5062-6

Byrd, M.P., Zamora, M., Lloyd, R.E. 2002. Generation of multiple isoforms of eukaryotic translation initiation factor 4GI by use of alternate translation initiation codons. Mol. Cell Biol. 22, 4499-511

Byrne, P.C., Sanders, P.G., Snell, K. 1995. Translational control of mammalian serine hydroxymethyltransferase expression. Biochem. Biophys. Res. Commun. 214, 496-502

Calkhoven, C.F., Muller, C., Leutz, A. 2000. Translational control of C/EBPalpha and C/EBPbeta isoform expression. Genes Dev. 14, 1920-32

Cazzola, M., Skoda, R.C. 2000. Translational pathophysiology: a novel molecular mechanism of human disease. Blood 95, 3280-8

Centers for Disease Control and Prevention. 2003. Severe acute respiratory syndrome (SARS) and coronavirus testing--United States. Morb. Mortal. Wkly. Rep. 52, 297-302. Erratum in: Morb. Mortal. Wkly. Rep. 52, 345

Chandler, M., Fayet, O. 1993. Translational frameshifting in the control of transposition in bacteria. Mol. Microbiol. 7, 497-503

Chappel, S.A., Edelman, G.M., Mauro, V.P. 2004. Biochemical and functional analysis of a 9-nt RNA sequence that affects translation efficiency in eukaryotic cells. Proc. Natl. Acad. Sci. 101, 9590-4

Chen, C.Y., and Sarnow, P. 1995. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. Science 268, 415-7

Chen, W., Calvo, P. A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., Basta, S., O'Neill, R., Schickli, J., Palese, P., Henklein, P., Bennink, J.R., Yewdell, J.W. 2001. A novel influenza A virus mitochondrial protein that induces cell death. Nat. Med. 7, 1306-12

Craigen, W.J., and Caskey, C.T. 1986. Expression of peptide chain release factor 2 requires high-efficiency frameshift. Nature 322, 273-5

Curran, J., and Kolakofsky, D. 1988. Ribosomal initiation from an ACG codon in the

Sendai virus P/C mRNA. EMBO J. 7, 245-51

Das, S., and Maitra, U. 2002. Functional significance and mechanism of eIF5-promoted GTP hydrolysis in eukaryotic translation initiation. Prog. Nucl. Acids Res. Mol. Biol. 70, 207-31

Dasso, M.C., and Jackson, R.J. 1989. Efficient initiation of mammalian mRNA translation at a CUG codon. Nucleic Acids Res. 17, 6485-97

de Haan, C.A., Masters, P.S., Shen, X., Weiss, S., Rottier, P.J. 2002. The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. Virology 296, 177-89

Delépine, P., Guillaume, C., Floch, V., Loisel, S., Yaouanc, J., Clement, J., Des Abbayes, H., Ferec, C. 2000. Cationic phosphonolipids as nonviral vectors: in vitro and in vivo applications. J. Pharm. Sci. 89, 629-38

Dinesh-Kumar, S.P., Miller, W.A. 1993. Control of start codon choice on a plant viral RNA encoding overlapping genes. Plant Cell 5, 679-92

Dinman, J.D., Richter, S., Plant, E.P., Taylor, R.C., Hammell, A.B., Rana, T.M. 2002. The frameshift signal of HIV-1 involves a potential intramolecular triplex RNA structure. Proc. Natl. Acad. Sci. 99, 5331-6

Dominguez, D.I., Ryabova, L.A., Pooggin, M.M., Schmidt-Puchta, W., Futterer, J., Hohn, T. 1998. Ribosome shunting in cauliflower mosaic virus. Identification of an essential and sufficient structural element. J. Biol. Chem. 273, 3669-78

Donahue, T.F., Cigan, A.M., Pabich, E.K., Valavicius, B.C. 1988. Mutations at a Zn(II) finger motif in the yeast eIF-2 beta gene alter ribosomal start-site selection during the scanning process. Cell 54, 621-32

Donly, B.C., Edgar, C.D., Adamski, F.M., Tate, W.P. 1990. Frameshift autoregulation in the gene for Escherichia coli release factor 2: partly functional mutants result in frameshift enhancement. Nucleic Acids Res. 18, 6517-22

Donly, C., Williams, J., Richardson, C., Tate, W. 1990. Frameshifting at the internal stop codon within the mRNA for bacterial release factor-2 on eukaryotic ribosomes. Biochim. Biophys. Acta. 1050, 283-7

Donner, D., and Kurland, C.G. 1972. Changes in the primary structure of a mutationally altered ribosomal protein S4 of Escherichia coli. Mol. Gen. Genet. 115, 49-53

Dos Ramos, F., Carrasco, M., Doyle, T., Brierley, I. 2004. Programmed -1 ribosomal frameshifting in the SARS coronavirus. Biochem. Soc. Trans. 32, 1081-3

Drabkin, H.J., and RajBhandary, U.L. 1998. Initiation of protein synthesis in mammalian cells with codons other than AUG and amino acids other than methionine. Mol. Cell Biol. 18, 5140-7

Dreher, T.W., and Miller, W.A. 2006. Translational control in positive strand RNA plant viruses. Virology 344, 185-97

Edelmann, P., Gallant, J. 1977a. Mistranslation in E. coli. Cell 10, 131-7

Edelmann, P., Gallant, J. 1977b. On the translational error theory of aging. Proc. Natl. Acad. Sci. 74, 3396-8

Ehrenfeld, E. 1996. Initiation of translation by picornavirus RNAs. *In* J.W.B. Hershey, M.B. Mathews, and N. Sonenberg (ed.), Translational control. Cold Spring harbor Laboratory Press, Cold Spring Harbor, N.Y. pp549-573

Eickmann, M., Becker, S., Klenk, H.D., Doerr, H.W., Stadler, K., Censini, S., Guidotti, S., Masignani, V., Scarselli, M., Mora, M., Donati, C., Han, J.H., Song, H.C., Abrignani, S., Covacci, A., Rappuoli, R. 2003. Phylogeny of the SARS coronavirus. Science 302, 1504-5

Enserink, M. 2003. Infectious diseases. Calling all coronavirologists. Science 300, 413-4

Enserink, M., and Vogel, G. Infectious diseases. Deferring competition, global net closes in on SARS. Science 300, 224-5

Farabaugh, P.J. 1996a. Programmed translational frameshifting. Annu. Rev. Genet. 30, 507-28

Farabaugh, P.J. 1996b. Programmed translational frameshifting. Microbiol. Rev. 60, 103-34

Farabaugh, P.J., Zhao, H., Vimaladithan, A. 1993. A novel programed frameshift expresses the POL3 gene of retrotransposon Ty3 of yeast: frameshifting without tRNA slippage. Cell 74, 93-103. Erratum in: Cell 75, 826

Fekete, C.A., Applefield, D.J., Blakely, S.A., Shirokikh, N., Pestova, T., Lorsch, J.R., Hinnebusch, A.G. 2005. The eIF1A C-terminal domain promotes initiation complex assembly, scanning and AUG selection in vivo. EMBO J. 24, 3588-601

Fleck, F. 2003. How SARS changed the world in less than six months. Bull World Health Organ. 81, 625-6

Flower, A.M., and McHenry, C.S. 1990. The gamma subunit of DNA polymerase III holoenzyme of Escherichia coli is produced by ribosomal frameshifting. Proc. Natl. Acad. Sci. 87, 3713-7

Futterer, J., Gordon, K., Sanfacon, H., Bonneville, J.M., Hohn, T. 1990. Positive and negative control of translation by the leader sequence of cauliflower mosaic virus pregenomic 35S RNA. EMBO J. 9, 1697-707

Futterer, J., Hohn, T. 1991. Translation of a polycistronic mRNA in the presence of the cauliflower mosaic virus transactivator protein. EMBO J. 10, 3887-96

Futterer, J., Hohn, T. 1992. Role of an upstream open reading frame in the translation of polycistronic mRNAs in plant cells. Nucleic Acids Res. 20, 3851-7

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

Futterer, J., Hohn, T. 1996. Translation in plants – rules and exceptions. Plant Mol. Biol. 32, 159-89.

Futterer, J., Potrykus, I., Bao, Y., Li, L., Burns, T.M., Hull, R., Hohn, T. 1996. Position-dependent ATT initiation during plant pararetrovirus rice tungro bacilliform virus translation. J. Viol. 70, 2999-3010

Futterer, J., Rothnie, H.M., Hohn, T., Potrykus, I. 1997. Rice tungro bacilliform virus open reading frames II and III are translated from polycistronic pregenomic RNA by leaky scanning. J. Virol. 71, 7984-9

Gaba, A., Wang, Z., Krishnamoorthy, T., Hinnebusch, A.G., Sachs, M.S. 2001.

Physical evidence for distinct mechanisms of translational control by upstream open reading frames. EMBO J. 20, 6453-63

Gallie, D.R. 1991. The cap and poly(A) tail function synergistically to regulation mRNA translational efficiency. Genes Dev. 5, 2108

Gallie, D.R. 2001. Cap-independent translation conferred by the 5' leader of tobacco etch virus is eukaryotic initiation factor 4G dependent. J. Virol. 75, 12141-52

Ghilardi, N., Wiestner, A., Skoda, R.C. 1998. Thrombopoietin production is inhibited by a translational mechanism. Blood 92, 4023-30

Godwa, S., Wu, F.C., Scholthof, H.B., Shepherd, R.J. 1989. Gene VI of figwort mosaic virus (caulimovirus group) functions in posttranscriptional expression of genes on the full-length RNA transcript. Proc. Natl. Acad. Sci. 86, 9203-7

Good, P.J., Welch, R.C., Barkan, A., Somasekhar, M.B., Mertz, J.E. 1988. Both VP2 and VP3 are synthesized from each of the alternative spliced late 19S RNA species of simian virus 40. J. Virol. 62, 944-53

Goss, D.J., Rounds, D.J. 1988. A kinetic light-scattering study of the binding of wheat germ protein synthesis initiation factor 3 to 40S ribosomal subunits and 80S ribosomes. Biochemistry 27, 3610-3

Gradi, A., Imataka, H., Svitkin, Y.V., Rom, E., Raught, B., Morino, S., Sonenberg, N. 1998a. A novel functional human eukaryotic translation initiation factor 4G. Mol. Cell Biol. 18, 334-42

Gradi, A., Svitkin, Y.V., Imataka, H., Sonenberg, N. 1998b. Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. Proc. Natl. Acad. Sci. 95, 11089-94

Gramstat, A., Prufer, D., Rohde, W. 1994. The nucleic acid-binding zinc finger protein of potato virus M is translated by internal initiation as well as by ribosomal frameshifting involving a shifty stop codon and a novel mechanism of P-site slippage. Nucleic Acids Res. 22, 3911-7

Grant, C.M., Hinnebusch, A.G. 1994. Effect of sequence context at stop codons on efficiency of reinitiation in GCN4 translational control. Mol. Cell Biol. 14, 606-18

Grant, C.M., Miller, P.F., Hinnebusch, A.G. 1994. Requirements for intercistronic distance and level of eukaryotic initiation factor 2 activity in reinitiation on GCN4 mRNA vary with the downstream cistron. Mol. Cell Biol. 14, 2616-28

Gray, T.A., Saitoh, S., Nicholls, R.D. 1999. An imprinted, mammalian bicistronic transcript encodes two independent proteins. Proc. Natl. Acad. Sci. 96, 5616-21

Grundhoff, A., and Ganem, D. 2001. Mechanisms governing expression of the v-FLIP gene of Kaposi's sarcoma-associated herpesvirus. J. Virol. 75, 1857-63

Grunert, S., and Jackson, R.J. 1994. The immediate downstream codon strongly influences the efficiency of utilization of eukaryotic translation initiation codons. EMBO J. 13, 3618-30

Grzela, R., Strokovska, L., Andrieu, J.P., Dublet, B., Zagorski, W., Chroboczek, J. 2006. Potyvirus terminal protein VPg, effector of host eukaryotic initiation factor eIF4E. Biochimie. 88, 887-96

Guo, J.P., Petric, M., Campbell, W., McGeer, P.L. 2004. SARS corona virus peptides recognized by antibodies in the sera of convalescent cases. Virology 324, 251-6

Gupta, K.C., Ono, E., Ariztia, E.V., Inaba, M. 1994. Translation initiation from non-AUG codons in COS1 cells is mRNA species dependent. Biochem. Biophys. Res. Commun. 201, 567-73

Gurvich, O.L., Baranov, P.V., Zhou, J., Hammer, A.W., Gesteland, R.F., Atkins, J.F. 2003. Sequences that direct significant levels of frameshifting are frequent in coding regions of Escherichia coli. EMBO J. 22, 5941-50

Hacker, D.L., Petty, I.T., Wei, N., Morris, T.J. 1992. Turnip crinkle virus genes required for RNA replication and virus movement. Virology 186, 1-8

Han, M.G., Kim, S.J. 2001. Analysis of Korean strains of infectious laryngotracheitis virus by nucleotide sequences and restriction fragment length polymorphism. Vet. Microbiol. 83, 321-31

Hanfrey, C., Franceschetti, M., Mayer, M.J., Illingworth, C., Michael, A.J. 2002. Abrogation of upstream open reading frame-mediated translational control of a plant S-adenosylmethionine decarboxylase results in polyamine disruption and growth

perturbations. J. Biol. Chem. 277, 44131-9

Hann, S.R., King, M.W., Bentley, D.L., Anderson, C.W., Eisenman, R.N. A non-AUG translational initiation in c-myc exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. Cell 52, 185-95

Hann, S.R., Sloan-Brown, K., Spotts, G.D. 1992. Translational activation of the non-AUG-initiated c-myc 1 protein at high cell densities due to methionine deprivation. Genes Dev. 6, 1229-40

Hann, S.R. 1994. Regulation and function of non-AUG-initiated proto-oncogenes. Biochimie. 76, 880-6

Hashimoto, N.N., Carnevalli, L.S., Castilho, B.A. 2002. Translation initiation at non-AUG codons mediated by weakened association of eukaryotic initiation factor (eIF) 2 subunits. Biochem. J. 367, 359-68

He, H., von der Haar, T., Singh, C.R., Ii, M., Li, B., Hinnebusch, A.G., McCarthy, J.E., Asano, K. 2003. The yeast eukaryotic initiation factor 4G (eIF4G) HEAT domain interacts with eIF1 and eIF5 and is involved in stringent AUG selection. Mol. Cell Biol. 23, 5431-45

Heidecker, G., and Messing, J. 1986. Structural analysis of plant genes. Annu. Rev. Plant Physiol. 37, 439-66

Hellen, C.U., Pestova, T.V., Wimmer, E. 1994. Effect of mutations downstream of the internal ribosome entry site on initiation of poliovirus protein synthesis. J. Virol. 68, 6312-22

Hellen, C.U., Sarnow, P. 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. 15, 1593-612

Hemmings-Mieszczak, M., Hohn, T., Preiss, T. 2000. Termination and peptide release at the upstream open reading frame are required for downstream translation on synthetic shunt-competent mRNA leaders. Mol. Cell Biol. 20, 6212-23

Herbreteau, C.H., Weill, L., Decimo, D., Prevot, D., Darlix, J.L., Sargueil, B., Ohlmann, T. 2005. HIV-2 genomic RNA contains a novel type of IRES located downstream of its initiation codon. Nat. Struct. Mol. Biol. 12, 1001-7

Hershey, J.W.B., and Merrick, W.C. 2000. Pathway and mechanism of initiation of protein synthesis. In: Sonenberg, N., Hershey, J.W.B., Mathews, M. (Eds.), Translation control of gene expression. Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp33-88

Herzog, E., Guilley, H., Fritsch, C. 1995. Translation of the second gene of peanut clump virus RNA2 occurs by leaky scanning in vitro. Virology 208, 215-25

Hinnebusch, A.G. 1997. Translational regulation of yeast GCN4. A window on factors that control initiator-trna binding to the ribosome. J. Biol. Chem. 272, 21661-4.

Hinnebusch, A.G., Jackson, B.M., Mueller, P.P. 1988. Evidence for regulation of reinitiation in translational control of GCN4 mRNA. Proc. Natl. Acad. Sci. 85, 7279-83

Horsburgh, B.C., Kollmus, H., Hauser, H., Coen, D.M. 1996. Translational recoding induced by G-rich mRNA sequences that form unusual structures. Cell 86, 949-59

Horsfield, J.A., Wilson, D.N., Mannering, S.A., Adamski, F.M., Tate, W.P. Prokaryotic ribosomes recode the HIV-1 gag-pol-1 frameshift sequence by an E/P site post-translocation simultaneous slippage mechanism. Nucleic Acids Res. 23, 1487-94

Horvath, C.M., Williams, M.A., Lamb, R.A. 1990. Eukaryotic coupled translation of tandem cistrons: identification of the influenza B virus BM2 polypeptide. EMBO J. 9, 2639-47

Huang, M., Koh, D.C.Y., Weng, L.J., Chang, M.L., Zhang, L., and Wong, S.M. 2000. Complete nucleotide sequence and genome organization of hibiscus chlorotic ringspot virus, a new member of the genus Carmovirus: evidence for the presence and expression of two novel open reading frames. J. Virol. 74, 3149-55

Huez, I., Creancier, L., Audigier, S., Gensac, M.C., Prats, A.C., Prats, H. Two independent internal ribosome entry sites are involved in translation initiation of vascular endothelial growth factor mRNA. Mol. Cell Biol. 18, 6178-90

Hwang, C.B., Horsburgh, B., Pelosi, E., Roberts, S., Digard, P., Coen, D.M. 1994. A net +1 frameshift permits synthesis of thymidine kinase from a drug-resistant herpes simplex virus mutant. Proc. Natl. Acad. Sci. 91, 5461-5

Ishikawa, M., Meshi, T., Motoyoshi, F., Takamatsu, N., Okada, Y. 1986. In vitro mutagenesis of the putative replicase genes of tobacco mosaic virus. Nucleic Acids Res. 14, 8291-305

Ito, N., Mossel, E.C., Narayanan, K., Popov, V.L., Huang, C., Inoue, T., Peters, C.J., Makino, S. 2005. Severe acute respiratory syndrome coronavirus 3a protein is a viral structural protein. J. Virol. 79, 3182-6

Ivanov, P.A., Karpova, O.V., Skulachev, M.V., Tomashevskaya, O.L., Rodionova, N.P., Dorokhov YuL, Atabekov, J.G. 1997. A tobamovirus genome that contains an internal ribosome entry site functional in vitro. Virology 232, 32-43

Jaag, H.M., Kawchuk, L., Rohde, W., Fischer, R., Emans, N., Prufer, D. 2003. An unusual internal ribosomal entry site of inverted symmetry directs expression of a potato leafroll polerovirus replication-associated protein. Proc. Natl. Acad. Sci. 100, 8939-44

Jacks, T., Madhani, H.D., Masiarz, F.R., Varmus, H.E. 1988a. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. Cell 55, 447-58

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., Varmus, H.E. 1988b. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331, 280-3

Jackson, R.J., Howell, M.T., Kaminski, A. 1990. The novel mechanism of picornavirus translation. Trends Biochem. Sci. 15, 477-83

Jackson, R.J., Hunt, S.L., Gibbs, C.L., Kaminski, A. 1994. Internal initiation of translation of picornavirus RNAs. Mol. Biol. Rep. 19, 147-59

Jackson, R.J., Hunt, S.L., Reynolds, J.E., Kaminski, A. 1995a. Cap-dependent and cap-independent translation: operational distinctions and mechanistic interpretations. Curr. Top. Microbiol. Immunol. 203, 1-30

Jackson, R.J., and Kaminski, A. 1995b. Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyong. RNA 1, 985-1000

Jackson, R.J. 1996. A comparative view of initiation site selection mechanisms, p.71-112. *In* J.W.B. Hershey, M.B. Mathews, and N. Sonenberg (ed.), Translational

control. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., Wimmer, E. 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J. Virol. 62, 2636-43

Jang, S.K., Davies, M.V., Kaufman, R.J., Wimmer, E. 1989. Initiation of protein synthesis by internal entry of ribosomes into the 5' nontranslated region of encephalomyocarditis virus RNA in vivo. J. Virol. 63, 1651-60

Jang, S.K., Pestova, T.V., Hellen, C.U., Witherell, G.W., Wimmer, E. 1990. Cap-independent translation of picornavirus RNAs-structure and function of the internal ribosome entry site. Enzyme 44, 292-309

Jelenc, P.C., and Kurland, C.G. 1979. Nucleoside triphosphate regeneration decreases the frequency of translation errors. Proc. Natl. Acad. Sci. 76, 3174-8

Johannes, G., Carter, M.S., Eisen, M.B., Brown, P.O., Sarnow, P. 1999. Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. Proc. Natl. Acad. Sci. 96, 13118-23

Johnson, J. 2001. Alternative Agriculture: What Is Kenaf? Rural Enterprise and Alternative Development Initiative Report 1. Carbondale, IL, USA: Southern Illinois University Carbondale

Jones, D.R., and Behncken, G.M. 1980. Hibiscus chlorotic ringspot, a widespread virus disease in the ornamental Hibiscus rosa-sinensis. Australian Journal of Plant Pathology 9, 4

Jorgensen, F., Adamski, F.M., Tate, W.P., Kurland, C.G. 1993. Release factor-dependent false stops are infrequent in Escherichia coli. J. Mol. Biol. 230, 41-50

Kamer, G., and Argos, P. 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. Nucleic Acids Res. 12, 7269-82

Kaminski, A., Howell, M.T., Jackson, R.J. 1990. Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning

mechanism. EMBO J. 9, 3753-9

Kaminski, A., Belsham, G.J., Jackson, R.J. 1994. Translation of encephalomyocarditis virus RNA: parameters influencing the selection of the internal initiation site. EMBO J. 13, 1673-81

Kaminski, A., Hunt, S.L., Patton, J.G., Jackson, R.J. 1995. Direct evidence that polypyrimidine tract binding protein (PTB) is essential for internal initiation of translation of encephalomyocarditis virus RNA. RNA 1, 924-38

Kaminski, A., Jackson, R.J. 1998. The polypyrimidine tract binding protein (PTB) requirement for internal initiation of translation of cardiovirus RNAs is conditional rather than absolute. RNA 4, 626-38

Kiss-Laszlo, Z., Blanc, S., Hohn, T. 1995. Splicing of cauliflower mosaic virus 35S RNA is essential for viral infectivity. EMBO J. 14, 3552-62

Kneller, E.L., Rakotondrafara, A.M., Miller, W.A. 2006. Cap-independent translation of plant viral RNAs. Virus Res. 119, 63-75

Koh, D.C., Liu, D.X., Wong, S.M. 2002. A six-nucleotide segment within the 3' untranslated region of hibiscus chlorotic ringspot virus plays an essential role in translational enhancement. J Virol. 76, 1144-53.

Koh, D.C., Wang, X., Wong, S.M., Liu, D.X. 2006. Translation initiation at an upstream CUG codon regulates the expression of Hibiscus chlorotic ringspot virus coat protein. Virus Res. July 17; Epub ahead of print

Koh, D.C., Wong, S.M., Liu, D.X. 2003. Synergism of the 3'-untranslated region and an internal ribosome entry site differentially enhances the translation of a plant virus coat protein. J. Biol. Chem. 278, 20565-73

Kontos, H., Napthine, S., Brierley, I. 2001. Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. Mol. Cell Biol. 21, 8657-70

Koonin, E.V. 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J. Gen. Virol. 72, 2197-206

Kos, M., Denger, S., Reid, G., Gannon, F. 2002. Upstream open reading frames

regulate the translation of the multiple mRNA variants of the estrogen receptor alpha. J. Biol. Chem. 277, 37131-8

Kozak, M. 1978. How do eukaryotic ribosomes select initiation regions in messenger RNA? Cell 15, 1109-23

Kozak, M. 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. 9, 5233-52

Kozak, M. 1983. Comparison of initiation of protein synthesis in prokaryotes, eukaryotes, and organelles. Microbiol. Rev. 47, 1-45

Kozak, M. 1984a. Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo. Nature 308, 241-6

Kozak, M. 1984b. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. 12, 857-72

Kozak, M. 1986. Point mutations define a sequence flanking the AUG intiator codon that modulates translation by eukaryotic ribosomes. Cell 44, 283-92

Kozak, M. 1987a. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. J. Mol. Biol. 196, 947-50

Kozak, M. 1987b. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. Mol.Cell Biol. 7, 3438-45

Kozak, M. 1989a. The scanning model for translation: an update. J. Cell Biol. 108, 229-41

Kozak, M. 1989b. Context effects and inefficient initiation at non-AUG codons in eukaryotic cell-free translation system. Mol. Cell Biol. 9, 5073-80

Kozak, M. 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. Proc. Natl. Acad. Sci. 87, 8301-5

Kozak, M. 1997. Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. EMBO J. 16, 2482-92.

Kozak, M. 1998. Primer extension analysis of eukaryotic ribosome-mRNA complexes. Nucleic Acids Res. 26, 4853-9

Kozak, M. 2001. Constraints on reinitiation of translation in mammals. Nucleic Acids Res. 29, 5226-32

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

Krokhin, O., Li, Y., Andonov, A., Feldmann, H., Flick, R., Jones, S., Stroeher, U., Bastien, N., Dasuri, K.V., Cheng, K., Simonsen, J.N., Perreault, H., Wilkins, J., Ens, W., Plummer, F., Standing, K.G. Mass spectrometric characterization of proteins from the SARS virus: a preliminary report. Mol. Cell Proteomics 2, 346-56

Kurland, C., and Gallant, J. 1996. Errors of heterologous protein expression. Curr. Opin. Biotechnol. 7, 489-93

Lai, M.M. 1990. Coronavirus: organization, replication and expression of genome. Annu Rev Microbiol. 44, 303-33

Lai, M.M., and Cavanagh, D. 1997. The molecular biology of coronaviruses. Adv. Virus Res. 48, 1-100

Lai, M.M., and Holmes, K.V. 2001. Coronaviridae: the viruses and their replication. In Fields Virology, 4th edn, pp. 1163–1185. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.

Lai, M.M., Patton, C.D., Stohlman, S.A. 1982. Replication of mouse hepatitis virus: negative-stranded RNA and replicative form RNA are of genome length. J. Virol. 44, 487-92

Lamphear, B.J., Yan, R., Yang, F., Waters, D., Leibig, H.D., Klump, H., Kuechler, E., Skern, T., Rhoads, R.E. 1993. Mapping the cleavage site in protein synthesis initiation factor eIF-4 gamma of the 2A proteases from human Coxsackievirus and rhinovirus. J. Biol. Chem. 268, 19200-3

Lamphear, B.J., Kirchweger, R., Skern, T., Rhoads, R.E. 1995. Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. J. Biol. Chem. 270, 21975-83

Larsen, B., Peden, J., Matsufuji, S., Matsufuji, T., Brady, K., Maldonado, R., Wills, N.M., Fayet, O., Atkins, J.F., Gesteland, R.F. 1995. Upstream stimulators for recoding. Biochem. Cell Biol. 73, 1123-9

Larsen, B., Wills, N.M., Gesteland, R.F., Atkins, J.F. 1994. rRNA-mRNA base pairing stimulates a programmed -1 ribosomal frameshift. J. Bacteriol. 176, 6842-51

Le, S.Y., Shapiro, B.A., Chen, J.H., Nussinov, R., Maizel, J.V. 1991. RNA pseudoknots downstream of the frameshift sites of retroviruses. Genet. Anal. Tech. Appl. 8, 191-205

Lee, Y.C., Chang, C.W., Su, C.W., Lin, T.N., Sun, S.H., Lai, H.L., Chern, Y. 1999. The 5' untranslated regions of the rat A2A adenosine receptor gene function as negative translational regulators. J. Neurochem. 73, 1790-8

Leh, V., Yot, P., Keller, M. 2000. The cauliflower mosaic virus translational transactivator interacts with the 60S ribosomal subunit protein L18 of Arabidopsis thaliana. Virology 266, 1-7

Leonard, S., Plante, D., Wittmann, S., Daigneault, N., Fortin, M.G., Laliberte, J.F. 2000. Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity. J. Virol. 74, 7730-7

Leonard, S., Viel, C., Beauchemin, C., Daigneault, N., Fortin, M.G., Laliberte, J.F. 2004. Interaction of VPg-Pro of turnip mosaic virus with the translation initiation factor 4E and the poly(A)-binding protein in planta. J. Gen. Virol. 85, 1055-63

Levin, M.E., Hendrix, R.W., Casjens, S.R. 1993. A programmed translational frameshift is required for the synthesis of a bacteriophage lambda tail assembly protein. J. Mol. Biol. 234, 124-39

Li, G., and Rice, C.M. 1989. Mutagenesis of the in-frame opal termination codon preceding nsP4 of sindbis virus: studies of translation readthrough and its effect on virus replication. J. Virol. 63, 1326-37

Li, W., and Wong, S.M. 2006. Analyses of subgenomic promoters of Hibiscus chlorotic ringspot virus and demonstration of 5' untranslated region and 3'-terminal sequences functioning as subgenomic promoters. J. Virol. 80, 3395-405

Li, W.Z., Qu, F., Morris, T.J. 1998. Cell-to-cell movement of turnip crinkle virus is controlled by two small open reading frames that function in trans. Virology 244, 405-16

Li, Z., Stahl, G., Farabaugh, P.J. 2001. Programmed +1 frameshifting stimulated by complementarity between a downstream mRNA sequence and an error-correcting region of rRNA. RNA 7, 275-84

Liang, X.Z., Lee, B.T., Wong, S.M. 2002. Covariation in the capsid protein of hibiscus chlorotic ringspot virus induced by serial passaging in a host that restricts movement leads to avirulence in its systemic host. J. Virol. 76, 12320-4

Liang, X.Z., Lucy, A.P., Ding, S.W., Wong, S.M. 2002. The p23 protein of hibiscus chlorotic ringspot virus is indispensable for host-specific replication. J. Virol. 76, 12312-9

Loftfield, R.B., Vanderjagt, D. 1972. The frequency of errors in protein biosynthesis. Biochem. J. 128, 1353-6

Lopez-Lastra, M., Gabus, C., Darlix, J.L. 1997. Characterization of an internal ribosomal entry segment within the 5' leader of avian reticuloendotheliosis virus type A RNA and development of novel MLV-REV-based retroviral vectors. Hum. Gene Ther. 8, 1855-65

Lordish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J. 2000. Molecular Cell Biology. New York: W. H. Freeman and Company.

Magliery, T.J., Anderson, J.C., Schultz, P.G. 2001. Expanding the genetic code: selection of efficient suppressors of four-base codons and identification of "shifty" four-base codons with a library approach in Escherichia coli. J. Mol. Biol. 307, 755-69

Maiss, E., Timpe, U., Brisske, A., Jelkmann, W., Casper, R., Himmler, G., Mattanovich, D., Katinger, H.W. 1989. The complete nucleotide sequence of plum pox virus RNA. J. Gen. Virol. 70, 513-24

Marra, M.A., *et al.* 2003. The Genome sequence of the SARS-associated coronavirus. Science 300, 1399-404

Martinez-Salas, E., Ramos, R., Lafuente, E., de Quinto, S.L. 2001. Functional

interactions in internal translation initiation directed by viral and cellular IRES elements. J. Gen. Virol. 82, 973-84

Marvil, D. K., Nowak, L. and Szer, W. 1980. A single-stranded nucleic acid-binding protein from Artemia salina. I. Purification and characterization. J. Biol. Chem. 255, 6466 – 6472

Matsuda, D., Dreher, T.W. 2006. Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. RNA 12, 1338-49

Matsufuji, S., Matsufuji, T., Miyazaki, Y., Murakami, Y., Atkins, J.F., Gesteland, R.F., Hayashi, S. 1995. Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. Cell 80, 51-60

Meerovitch, K.R., Nicholson, R., Sonenberg, N. 1991. In vitro mutational analysis of *cis*-acting RNA translational elements within the poliovirus type 2 5' untranslated region. J. Virol. 65, 5859-901

Mejlhede, N., Atkins, J.F., Neuhard, J. 1999. Ribosomal -1 frameshifting during decoding of Bacillus subtilis cdd occurs at the sequence CGA AAG. J. Bacteriol. 181, 2930-7

Mejlhede, N., Licznar, P., Prere, M.F., Wills, N.M., Gesteland, R.F., Atkins, J.F., Fayet, O. 2004. -1 frameshifting at a CGA AAG hexanucleotide site is required for transposition of insertion sequence IS1222. J. Bacteriol. 186, 3274-7

Menninger, J.R. 1977. Ribosome editing and the error catastrophe hypothesis of cellular aging. Mech. Ageing Dev. 6, 131-142

Merrick, W.C. 2004. Cap-dependent and cap-independent translation in eukaryotic systems. Gene 332, 1-11

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, 34051-60

Miller, D.L., and Weissbach, H. 1977. Factors involved in the transfer of aminoacyl-tRNA to the ribosome. In Molecular Mechanisms of Protein Biosynthesis. (Weissbach, H. and Petska, S., eds) pp. 323-373, Academic Press

Mize, G.J., and Morris, D.R. 2001. A mammalian sequence-dependent upstream

open reading frame mediates polyamine-regulated translation in yeast. RNA 7, 374-81

Monkewich, S., Lin, H.X., Fabian, M.R., Xu, W., Na, H., Ray, D., Chernysheva, O.A., Nagy, P.D., White, K.A. 2005. The p92 polymerase coding region contains an internal RNA element required at an early step in Tombusvirus genome replication. J. Virol. 79, 4848-58

Morlé, F., Lopez, B., Henni, T., Godet, J. 1985. Alpha-Thalassaemia associated with the deletion of two nucleotides at position -2 and -3 preceding the AUG codon. EMBO J. 4, 1245-50

Morris, D.R., Geballe, A.P. 2000. Upstream open reading frames as regulators of mRNA translation. Mol. Cell Biol. 20, 8635-8642

Mueller, P.P., and Hinnebusch, A.G. 1986. Multiple upstream AUG codons mediate translational control of GCN4. Cell 45, 201-7

Namy, O., Hatin, I., Rousset, J.P. 2001. Impact of the six nucleotides downstream of the stop codon on translation termination. EMBO Rep. 2, 787-93

Namy, O., Moran, S.J., Stuart, D.I., Gilbert, R.J., Brierley, I. 2006. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. Nature 441, 244-7

Nanbru, C., Lafon, I., Audigier, S., Gensac, M.C., Vagner, S., Huez, G., Prats, A.C. 1997. Alternative translation of the proto-oncogene c-myc by an internal ribosome entry site. J. Biol. Chem. 272, 32061-6

Napthine, S., Vidakovic, M., Girnary, R., Namy, O., Brierley, I. 2003. Prokaryotic-style frameshifting in a plant translation system: conservation of an unusual single-tRNA slippage event. EMBO J. 22, 3941-50

Navas-Martin, S., and Weiss, S.R. 2003. SARS: lessons learned from other coronaviruses. Viral Immunol. 16, 461-74

O'Connor, M. 2002. Imbalance of tRNA(Pro) isoacceptors induces +1 frameshifting at near-cognate codons. Nucleic Acids Res. 30, 759-65

Paluh, J.L., Orbach, M.J., Legerton, T.L., Yanofsky, C. 1988. The cross-pathway

control gene of Neurospora crassa, cpc-1, encodes a protein similar to GCN4 of yeast and the DNA-binding domain of the oncogene v-jun-encoded protein. Proc. Natl. Acad. Sci. 85, 3728-32

Park, H.S., Himmelbach, A., Browning, K.S., Hohn, T., Ryabova, L.A. 2001. A plant viral "reinitiation" factor interacts with the host translational machinery. Cell 106, 723-33

Parker, J., Johnston, T.C., Borgia, P.T., Holtz, G., Remaut, E., Fiers, W. 1983. Codon usage and mistranslation. In vivo basal level misreading of the MS2 coat protein message. J. Biol. Chem. 258, 10007-12

Pasternak, A.O., van den Born, E., Spaan, W.J., Snijder, E.J. 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J. 20, 7220-8

Peabody, D.S., and Berg, P. 1986a. Termination-reinitiation occurs in the translation of mammalian cell mRNAs. Mol. Cell Biol. 6, 2695-703

Peabody, D.S., Subramani, S., Berg, P. 1986b. Effect of upstream reading frames on translation efficiency in simian virus 40 recombinants. Mol. Cell Biol. 6, 2704-11

Peabody, D.S. 1987. Translation initiation at an ACG triplet in mammalian cells. J. Biol. Chem. 262, 11847-51

Peabody, D.S. 1989. Translation initiation at non-AUG triplets in mammalian cells. J. Biol. Chem. 264, 5031-5

Pelham, H.R. 1978. Leaky UAG termination codon in tobacco mosaic virus RNA. Nature 272, 469–471

Pelletier, J., Kaplan, G., Racaniello, V.R., Sonenberg, N. 1988a. Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5’ noncoding region. Mol. Cell Biol. 8, 1103-12

Pelletier, J., and Sonenber, N. 1988b. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature 334, 320-5

Pestova, T.V., Hellen, C.U., Shatsky, I.N. 1996a. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. Mol. Cell Biol.

16, 6859-69

Pestova, T.V., Shatsky, I.N., Hellen, C.U. 1996b. Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. Mol. Cell Biol. 16, 6870-8

Pestova, T.V., Borukhov, S.I., Hellen, C.U. 1998a. Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. Nature 394, 854-9

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

Petros, L.M., Howard, M.T., Gesteland, R.F., Atkins, J.F. 2005. Polyamine sensing during antizyme mRNA programmed frameshifting. Biochem. Biophys. Res. Commun. 338, 1478-89

Petty, I.T., Edwards, M.C., and Jackson, A.O. 1990. Systemic movement of an RNA plant virus determined by a point substitution in a 5' leader sequence. Proc. Natl. Acad. Sci. 87, 8894-7

Pilipenko, E.W., Gmyl, A.P., Maslova, S.V., Belov, G.A., Sinyakov, A.N., Huang, M., Brown, T.D., Agol, V.I. 1994. Starting window, a distinct element in the cap-independent internal initiation of translation on picornaviral RNA. J. Mol. Biol. 241, 398-414

Pilipenko, E.V., Pestova, T.V., Kolupaeva, V.G., Khitrina, E.V., Poperechnaya, A.N., Agol, V.I., Hellen C.U. 2000. A cell cycle-dependent protein serves as a template-specific translation initiation factor. Genes Dev. 14, 2028-45

Pooggin, M.M., Hohn, T., Futterer, J. 1998. Forced evolution reveals the importance of short open reading frame A and secondary structure in the cauliflower mosaic virus 35S RNA leader. J Virol. 72, 4157-69

Pooggin, M.M., Hohn, T., Futterer, J. 2000. Role of a short open reading frame in ribosome shunt on the cauliflower mosaic virus RNA leader. J. Biol. Chem. 275, 17288-96

Portis, J.L., Spangrude, G.J., McActee, F.J. 1994. Identification of a sequence in the unique 5' open reading frame of the gene encoding glycosylated gag which influences the incubation period of neurodegenerative disorder disease induced by murine retrovirus. J. Virol. 68, 3879-87

Poyry, T.A., Kaminski, A., Jackson, R.J. 2004. What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? Genes Dev. 18, 62-75

Prats, A.C., Vagner, S., Prats, H., Amalric, F. 1992. Cis-acting elements involved in the alternative translation initiation process of human basic fibroblast growth factor mRNA. Mol. Cell Biol. 12, 4796-805

Prufer, D., Tacke, E., Schmitz, J., Kull, B., Kaufmann, A., Rohde, W. 1992. Ribosomal frameshifting in plants: a novel signal directs the -1 frameshift in the synthesis of the putative viral replicase of potato leafroll luteovirus. EMBO J. 11, 1111-7

Qiu, M., Shi, Y., Guo, Z., Chen, Z., He, R., Chen, R., Zhou, D., Dai, E., Wang, X., Si, B., Song, Y., Li, J., Yang, L., Wang, J., Wang, H., Pang, X., Zhai, J., Du, Z., Liu, Y., Zhang, Y., Li, L., Wang, J., Sun, B., Yang, R. 2005. Antibody responses to individual proteins of SARS coronavirus and their neutralization activities. Microbes Infect. 7, 882-9

Reavy, B., Arif, M., Cowan, G.H., Torrance, L. 1998. Association of sequences in the coat protein/readthrough domain of potato mop-top virus with transmission by Spongospora subterranea. J. Gen. Virol. 79, 2343-7

Rettberg, C.C., Prere, M.F., Gesteland, R.F., Atkins, J.F., Fayet, O. 1999. A three-way junction and constituent stem-loops as the stimulator for programmed -1 frameshifting in bacterial insertion sequence IS911. J. Mol. Biol. 286, 1365-78

Rom, E., and Kahana, C. 1994. Polyamines regulate the expression of ornithine decarboxylase antizyme in vitro by inducing ribosomal frame-shifting. Proc. Natl. Acad. Sci. 91, 3959-63. Erratum in: Proc. Natl. Acad. Sci. 91, 9195

Rota, P.A., *et al*. 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300, 1394-9

Sachs, A.B., and Davis, R.W. 1989. The poly(A) binding protein is required for

poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. Cell 58, 857

Sarma, J.D., Scheen, E., Seo, S.H., Koval, M., Weiss, S.R. 2002. Enhanced green fluorescent protein expression may be used to monitor murine coronavirus spread in vitro and in the mouse central nervous system. J. Neurovirol. 8, 381-91

Sawicki, S.G., and Sawicki, D.L. 1998. A new model for coronavirus transcription. Adv. Exp. Med. Biol. 440, 215-9

Schmidt-Puchta, W., Dominguez, D., Lewetaq, D., Hohn, T. 1997. Plant ribosome shunting in vitro. Nucleic Acids Res. 25, 2854-60

Senanayake, S.D., Brian, D.A. 1997. Bovine coronavirus I protein synthesis follows ribosomal scanning on the bicistronic N mRNA. Virus Res. 48, 101-5

Shatkin, A.J. 1976. Capping of eukaryotic mRNAs. Cell 9, 645-53

Shen, S., Lin, P.S., Chao, Y.C., Zhang, A., Yang, X., Lim, S.G., Hong, W., Tan, Y.J. 2005. The severe acute respiratory syndrome coronavirus 3a is a novel structural protein. Biochem. Biophys. Res. Commun. 330, 286-92

Shirako, Y. 1998. Non-AUG translation initiation in a plant RNA virus: a forty-amino-acid extension is added to the N terminus of the soil-borne wheat mosaic virus capsid protein. J. Virol. 72, 1677-82

Siddell, S.G. 1995. The Coronaviridae. New York: Plenum Press

Simon-Buela, L., Guo, H.S., Garcia, J.A. 1997. Cap-independent leaky scanning as the mechanism of translation initiation of a plant viral genomic RNA. J. Gen. Virol. 78, 2691-9

Sivakumaran, K., Hacker, D.L. 1998. Identification of viral genes required for cell-to-cell movement of southern bean mosaic virus. Virology 252, 376-86.

Sizova, D.V., Kolupaeva, V.G., Pestova, T.V., Shatsky, I.N., Hellen, C.U. 1998. Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. J. Virol. 72, 4775-82

Skuzeski, J.M., Nichols, L.M., Gesteland, R.F., Atkins, J.F. 1991. The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. J. Mol. Biol. 218, 365-73

Slobodskaya, O.R., Gmyl, A.P., Maslova, S.V., Tolskaya, E.A., Viktorova, E.G., Agol, V.I. 1996. Poliovirus neurovirulence correlates with the presence of a cryptic AUG upstream of the initiator codon. Virology 221, 141-50

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L., Guan, Y., Rozaov, M., Spaan, W.J., Gorbalenya, A.E. 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. 331, 991-1004

Spahn, C.M., Kieft, J.S., Grassucci, R.A., Penczek, P.A., Zhou, K., Doudna, J.A., Frank, J. 2001. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. Science 291, 1959-62

Spahn, C.M., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P., Frank, J. 2004. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. Cell 118, 465-75

Stadler, K., Masignani, V., Eichmann, M., Becker, S., Abrignani, S., Klenk, H.D., Rappuoli, R. 2003. SARS-beginning to understand a new virus. Nat. Rev. Microbiol. 1, 209-18

Su, M.C., Chang, C.T., Chu, C.H., Tsai, C.H., Chang, K.Y. 2005. An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. Nucleic Acids Res. 33, 4265-75

Tamada, T., Schmitt, C., Saito, M., Guilley, H., Richards, K., Jonard, G. 1996. High resolution analysis of the readthrough domain of beet necrotic yellow vein virus readthrough protein: a KTER motif is important for efficient transmission of the virus by Polymyxa betae. J. Gen.Virol. 77, 1359-67

Tan, T.H., Barkham, T., Fielding, B.C., Chou, C.F., Shen, S., Lim, S.G., Hong, W., Tan, Y.J. 2005. Genetic lesions within the 3a gene of SARS-CoV. Virol. J. 2, 51

Tan, Y.J., Fielding, B.C., Goh, P.Y., Shen, S., Tan, T.H., Lim, S.G., Hong, W. 2004a. Overexpression of 7a, a protein specifically encoded by the severe acute respiratory syndrome coronavirus, induces apoptosis via a caspase-dependent pathway. J. Virol.

78, 14043-7

Tan, Y.J., Teng, E., Shen, S., Tan, T.H., Goh, P.Y., Fielding, B.C., Ooi, E.E., Tan, H.C., Lim, S.G., Hong, W. 2004b. A novel severe acute respiratory syndrome coronavirus protein, U274, is transported to the cell surface and undergoes endocytosis. J. Virol. 78, 6723-34

Tan, Y.J., Lim, S.G., Hong, W. 2005. Characterization of viral proteins encoded by the SARS-coronavirus genome. Antiviral Res. 65, 69-78

Thiel, V., Ivanov, K.A., Putics, A., Hertzig, T., Schelle, B., Bayer, S., Weissbrich, B., Snijder, E.J., Rabenau, H., Boerr, H.W., Gorbalenya, A.E., Ziebuhr, J. 2003. Mechanisms and enzymes involved in SARS coronavirus genome expression. J. Gen. Virol. 84, 2305-15

Thompson, J.R., and Jelkmann, W. 2004. Strain diversity and conserved genome elements in Strawberry mild yellow edge virus. Arch. Virol. 149, 1897-909

Tsai, T.F., Chen, K.S., Weber, J.S., Justice, M.J., Beaudet, A.L. 2002. Evidence for translational regulation of the imprinted Snurf-Snrpn locus in mice. Hum. Mol. Genet. 11, 1659-68

Tsuchihashi, Z., Brown, P.O. 1992. Sequence requirements for efficient translational frameshifting in the Escherichia coli dnaX gene and the role of an unstable interaction between tRNA(Lys) and an AAG lysine codon. Genes Dev. 6, 511-9

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M., Nomoto, A. 1992. Internal ribosome entry site within hepatitis C virus RNA. J. Virol. 66, 1476-83

Unbehaun, A., Borukhov, S.I., Hellen, C.U., Pestova, T.V. 2004. Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP. Genes Dev. 18, 3078-93

Urban, C., and Beier, H. 1995. Cysteine tRNAs of plant origin as novel UGA suppressors. Nucleic Acids Res. 23, 4591-7

Vagner, S., Galy, B., Pyronnet, S. 2001. Irresistible IRES. Attracting the translational machinery to internal ribosome entry sites. EMBO Rep. 21, 893-8

Valasek, L., Nielsen, K.H., Zhang, F., Fekete, C.A., Hinnebusch, A.G. 2004. Interactions of eukaryotic translation initiation factor 3 (eIF3) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. Mol. Cell Biol. 24, 9437-55

Vilela, C., Linz, B., Rodrigues-Pousada, C., McCarthy, J.E.G. 1998. The yeast transcription factor gene YAP1 and YAP2 are subject to differential control at the levels of both translation and mRNA stability. Nucleic Acids Res. 26, 1150-9

Vimaladithan, A., Farabaugh, P.J. 1994. Special peptidyl-tRNA molecules promote translational frameshifting without slippage. Mol. Cell Biol. 14, 8107-16

Waterworth, H. 1980. Hibiscus Chlorotic ringspot virus. CMI/AAB Descriptions of Plant Viruses No. 227

Waterworth, H.E., Lawson, R.H., Monroe, R.L. 1976. Phytopathology 64, 570-575

Weber, F., Bridgen, A., Fazakerley, J.K., Streitenfeld, H., Kessler, N., Randall, R.E., Elliott, R.M. 2002. Bunyamwera bunyavirus nonstructural protein NSs counteracts the induction of alpha/beta interferon. J. Virol. 76, 7949-55

Weiland, J.J., Dreher, T.W. 1989. Infectious TYMV RNA from cloned cDNA: effects in vitro and in vivo of point substitutions in the initiation codons of two extensively overlapping ORFs. Nucleic Acid Res. 17, 4675-87

White, K.A., Skuzeski, J.M., Li, W., Wei, N., Morris, T.J. 1995. Immunodetection, expression strategy and complementation of turnip crinkle virus p28 and p88 replication components. Virology 211, 525-34

Widerak, M., Kern, R., Malki, A., Richarme, G. 2005. U2552 methylation at the ribosomal A-site is a negative modulator of translational accuracy. Gene 347, 109-14

Wilson, J.E., Powell, M.J., Hoover, S.E., Sarnow, P. 2000. Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. Mol. Cell Biol. 20, 4990-9

Wilson, W., Braddock, M., Adams, S.E., Rathjen, P.D., Kingsman, S.M., Kingsman, A.J. 1988. HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. Cell 55, 1159-69

Wittmann, S., Chatel, H., Fortin, M.G., Laliberte, J.F. 1997. Interaction of the viral protein genome linked of turnip mosaic potyvirus with the translational eukaryotic initiation factor (iso) 4E of Arabidopsis thaliana using the yeast two-hybrid system. Virology 234, 84-92

Wobbe, K.K., Akgoz, M., Dempsey, D.A., Klessiq, D.F. 1998. A single amino acid change in turnip crinkle virus movement protein p8 affects RNA binding and virulence on Arabidopsis thaliana. J. Virol. 72, 6247-50

Wong, S.M., and Chng, C.G. 1992. Occurrence of hibiscus chlorotic ringspot virus in Singapore. Phytopathology 82, 722

Wong, T.C., Wipf, G., Hirano, A. 1987. The measles virus matrix gene and gene product defined by in vitro and in vivo expression. Virology 157, 497-508

Xi, Q., Cuesta, R., Schneider, R.J. 2004. Tethering of eIF4G to adenoviral mRNAs by viral 100k protein drives ribosome shunting. Genes Dev. 18, 1997-2009

Xi, Q., Cuesta, R., Schneider, R.J. 2005. Regulation of translation by ribosome shunting through phosphotyrosine-dependent coupling of adenovirus protein 100k to viral mRNAs. J. Virol. 79, 5676-83

Xiao, J.H., Davidson, I., Matthes, H., Garnier, J.M., Chambon, P. 1991. Cloning, expression, and transcriptional properties of the human enhancer factor TEF-1. Cell 65, 551-68

Yoshinaka, Y., Katoh, I., Copeland, T.D., Oroszlan, S. 1985. Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. Proc. Natl. Acad. Sci. USA 82, 1618-22

Yuan, X., Li, J., Shan, Y., Yang, Z., Zhao, Z., Chen, B., Yao, Z., Dong, B., Wang, S., Chen J., Cong, Y. 2005. Subcellular localization and membrane association of SARS-CoV 3a protein. Virus Res. 109, 191-202

Yueh, A., and Schneider, R.J. 2000. Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to 18S rRNA. Genes Dev. 14, 414-21

Zeng, R., *et al*. 2004. Characterization of the 3a protein of SARS-associated coronavirus in infected vero E6 cells and SARS patients. J. Mol. Biol. 341, 271-9

Zerfass, K., and Beier, H. 1992. The leaky UGA termination codon of tobacco rattle virus RNA is suppressed by tobacco chloroplast and cytoplasmic tRNAs(Trp) with CmCA anticodon. EMBO J. 11, 4167-73

Zerfass, K., and Beier, H. 1992. Pseudouridine in the anticodon G psi A of plant cytoplasmic tRNA(Tyr) is required for UAG and UAA suppression in the TMV-specific context. Nucleic Acids Res. 20, 5911-8

Zhou, T., Fan, Z.F., Li, H.F. and Wong, S.M. 2006. Hibiscus chlorotic ringspot virus p27 and its isoforms affect symptom expression and potentiate virus movement in kenaf (Hibiscus cannabinus L.). Mol. Plant-Microbe Interact. 19, In press.

Ziebuhr, J. 2004. Molecular biology of severe acute respiratory syndrome coronavirus. Curr. Opin. Microbiol. 7, 412-9

Ziebuhr, J., Snijder, E.J., Gorbalenya, A.E. 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. J. Gen. Virol. 81, 853-79

Zuniga, S., Sola, I., Alonso, S., Enjuanes, L. 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J. Virol. 78, 980-94