CHARACTERIZATION OF TWO TOMATO RINGSPOT VIRUS REPLICATION PROTEINS: THE NTB PROTEIN AND THE RNA-DEPENDENT RNA POLYMERASE

by

TING WEI

B.Sc., Northwest Agriculture & Forestry University, Shaanxi, China, 2006
M.Sc., Northwest Agriculture & Forestry University, Shaanxi, China, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Botany)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2015

Abstract

Tomato ringspot virus (ToRSV) replicates in large protein complexes that are associated with modified endoplasmic reticulum membranes. The ToRSV RNA-dependent RNA polymerase (Pol) and integral membrane protein NTB-VPg are essential components of these complexes. Membrane-associated modifications of NTB-VPg (N-glycosylation and a putative signal peptidase cleavage) were previously observed *in vitro* but were not well characterized. Two forms of the polymerase were detected in infected plants: the full-length Pol, which accumulates at low concentration and the VPg-Pro-Pol' polyprotein, which includes a 15kDa C-terminal truncation of Pol and accumulates to higher levels. My specific objectives were to characterize the signal peptidase cleavage in NTB-VPg and investigate the stability and function of various forms of the polymerase.

Using *in vitro* translation assays and plant transient expression assays, I detected signal peptidase processing in the NTB-VPg of three ToRSV isolates. Using site-directed mutagenesis, I mapped a suboptimal GAAGG cleavage site (Rasp2 isolate) and identified key amino acids that regulate the efficiency of cleavage. Compared to typical signal peptides, the NTB-VPg sequence has an unusually long distance between the end of the hydrophobic region and the cleavage site, indicating that it adopts a unique topology in the membrane. This is the first detailed characterization of signal peptidase cleavage of a plant virus replication protein. This cleavage may alter the conformation of NTB-VPg in the membrane and influence the architecture of the replication complexes.

Using agroinfiltration assays, I show that the full-length Pol and VPg-Pro-Pol are unstable when ectopically expressed in *N. benthamiana*. Truncation of the C-terminal 15 kDa from Pol or VPg-Pro-Pol increased their stability in plants, which is consistent with the accumulation of VPg-Pro-Pol' in infected plants. In spite of repeated attempts, I was unable to establish an *in vitro* assays to compare the activity of Pol and VPg-Pro-Pol', possibly because an essential plant factor is missing. The instability of VPg-Pro-Pol and Pol may regulate the rate of virus replication, allowing the virus to keep its genome integrity and reducing the chance of being recognized by host defense responses.

Preface

The work presented in this thesis was done in Dr. Hélène Sanfaçon's lab from September 2009 to February 2015 by the candidate. The published manuscript and the ones that are in preparation are listed below. My contribution for each chapter is mentioned.

Chapter 1: Literature review

The candidate wrote the chapter and Dr. Hélène Sanfaçon contributed to the editing of the chapter .

Chapter 2: Membrane-associated glycosylation and signal peptidase processing of the cNTB-VPg protein of ToRSV

A manuscript describing the sequencing of three ToRSV isolates has been published.

Walker M, Chisholm J, Wei T, Ghoshal B, Saeed H, Rott M, Sanfacon H. 2015. Complete genome sequence of three tomato ringspot virus isolates: evidence for reassortment and recombination. Arch Virol 160:543-547.

In this manuscript, I contributed the sequence of the X2-NTB-VPg-Pro-Pol coding region for the Rasp1 isolate and the sequence of the X2-NTB-VPg coding region for the GYV isolate. Other co-authors contributed other sections of the ToRSV genome sequence.

A manuscript describing the second part of Chapter 2 (characterization of membrane-associated modification of NTB-VPg) is under preparation.

Dr. Hélène Sanfaçon produced the PYB and GYV pCITE-cNV constructs,

performed the initial *in vitro* translation experiment, and the inhibitor treatment. The candidate produced all the mutants, repeated the *in vitro* work and did all the *in vivo* work.

The candidate designed the experiments, performed most of the experiments (as described above) and wrote the chapter. Dr. Hélène Sanfaçon provided supervisory support and contributed to the editing of the chapter.

Chapter 3: Investigation into the stability of ToRSV RNA-dependent RNA polymerase

A manuscript will be submitted for publication.

The candidate designed, performed all the experiments and wrote the chapter. Dr. Hélène Sanfaçon provided supervisory support and contributed to the editing of the chapter.

Chapter 4: Investigation into the activity of different forms of the ToRSV polymerase

The candidate designed, performed all the experiments and wrote the chapter. Dr. Hélène Sanfaçon provided supervisory support and contributed to the editing of the chapter.

Chapter 5: General discussion

The candidate wrote the chapter and Dr. Hélène Sanfaçon contributed to the editing of the chapter.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

3’	Three prime
5’	Five prime
3CD	Pro-Pol
AGO	ARGONAUTE
BMV	Brome mosaic virus
cNV	C-terminal region of NTB-VPg
CP	Coat protein
CPMV	*Cowpea mosaic virus*
c-region	Carboxyl-terminal region
C-terminal	Carboxyl-terminal
Dpa	Days post-agroinfiltration
E. coli	*Escherichia coli*
ER	Endoplasmic reticulum
FCV	*Feline calicivirus*
Fig	Figure
GDD	Gly-Asp-Asp motif
GFP	Green fluorescence protein
GFLV	*Grapevine fanleaf nepovirus*
GW motif	Glycine / Tryptophane motif
GYV	Grape yellow vein
HCV	Hepatitis C virus
h-region	Hydrophobic core region
IRES	Internal ribosome entry site
MM	Microsomal membrane
MP	Movement protein
mRNA	Messenger RNA
N. benthamiana	*Nicotiana benthamiana*
n-region	Amino-terminal region
nt	Nucleotide
NTB	Nucleoside triphosphate-binding proteins
N-terminal	Amino-terminal
NV	*Norovirus*
OST	Oligosaccharyltransferase
P1	Picornavirus 1ABCD polyprotein
P2	Picornavirus 2ABC polyprotein
P3	Picornavirus 3ABCD polyprotein
p19	19 kDa protein of Tomato bushy stunt virus
PD	Plasmodesmata
Poly(A)	polyadenylate or polyadenylic acid
Pol	RNA-dependent RNA polymerase
Pro	Proteinase
PV	*Poliovirus*
PVX	*Potato virus X*
PYB	Peach yellow bud

RC	Replication complexes
RISC	RNA-induced silencing complexes
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGT1	Suppressor of G2 allele of SKP1
siRNA	Small interfering RNA
$^{[35S]}$-methionine	Radioactive ^{35}S labeled methionine
SPC	Signal peptidase complex
SPP	Signal peptide peptidase
SRP	Signal recognition particle
TBRV	*Tomato black ring virus*
TBSV	*Tomato bushy stunt virus*
ToRSV	*Tomato ringspot virus*
TRV	*Tobacco rattle virus*
TRSV	*Tobacco ringspot virus*
TYMV	*Turnip yellow mosaic virus*
UTR	Untranslated region
VIGS	Virus-induced gene silencing
VPg	Viral genome-linked protein
VPg-Pro-Pol	ToRSV polyprotein including VPg, Pro and Pol
VPg-Pro-Pol'	VPg-Pro-Pol that lacks the C-terminal 15 kDa of Pol
VSR	Viral suppressor of RNA silencing

Acknowledgments

First of all, I would like to express my sincere thanks to my supervisor Dr. Helene Sanfaçon for giving me the opportunity to do my Ph.D. in her lab. Her excellent guidance and critical thinking helped me to grow quickly and professionally over the last few years. I also want to thank her for all the support and encouragement during the hard days. Her enthusiasm for science also inspired me to explore more in the scientific world.

I wish to thank my supervisory committee members Dr. D'Ann Rochon, Dr. George Haughn and Dr. François Jean for their valuable suggestions and discussions on my research work. I also want to thank Dr. Xin Li and Dr. Eric Jan for being the external examiners for my comprehensive exam and Dr. Keith Adams for being my graduate advisor.

I would like to give a special thanks to Melanie Walker and Joan Chisholm. They have provided excellent technical assistance and given good advice on my project. I also appreciate their enduring support and encouragement during the challenges of graduate school and life. I also want to thank my lab members Basudev Ghoshal, Rajita Karran, Dinesh Babu Paudel and Sushma Jossey for their suggestions and discussions on the lab meetings. Many thanks to the staff members of the Department of Botany especially Mrs. Veronica Oxtoby for her assistance over the years.

I want to thank Dr. Guus Bakkeren, Jane Theilmann, Les Willis, Ron Reade, Rob Linning, Mike Bouthillier and Robyn DeYoung for letting me use their lab facilities. Many thanks to Yunfeng Wu, Yu Xiang, Susan Wahlgren, Jim and Linda Lancaster, Paul and Eunice Randle,

Don Walker, Changwen Lu for their encouragement and support. I would also like to thank Yingchao Nie, Nadia Sokal, Kankana Ghoshal, Shawkat Ali, Kamni Naidu, Julie Boule, Zhenpu Liang, Xiao Song, Ming Chi, Xiaotang Yang, Li Su, Hala Khalil, Olga Shaposhnikova, Siddartha Biswas, Syed Benazir Alam, Ana Montenegro and Sarah Wong. Their friendship makes my life in Summerland more colorful and memorable. I am grateful to my friends Lianzhen Xu, Yuan Ruan, Fei Shen and Jinguang Hu, they have provided lots of help and support when I was in Vancouver.

I would like to thank the Chinese Science Council for giving me the doctoral scholarship. Work in the Sanfacon lab is supported by funding from an NSERC discovery grant. I also want to thank the Pacific Agriculture and Agri-Food Canada (PARC), Summerland for the experimental facilities. I am thankful to all the staff at PARC Summerland for their help in the last few years.

I am very grateful to my husband Anyong Wang; his unconditional love, encouragement and support helped me go through all the difficult days. I am truly thankful for having him in my life. I also want to show my sincere gratitude to my parents and my in-laws, who have always encouraged me to work hard for the things that I aspire to achieve.

.

Chapter 1

Literature Review

1.1 Introduction

In this thesis, I have studied proteolytic cleavage events that regulate the function and stability of two replication proteins from tomato ringspot virus (ToRSV, a plant picorna-like virus). Viruses are small, obligate infectious parasites of living cells. Most plant viruses, including ToRSV, contain positive strand RNA genome, which act as template for both translation and replication. Viruses have evolved various strategies to maximize the coding capacity of their genomes, including the polyprotein strategy, which is used by ToRSV and other picorna-like viruses and allows regulated release of mature viral proteins and intermediate polyprotein precursors (1). Polyprotein processing is directed by viral proteases and in some cases also by host proteases. Virus infection induces rearrangement of intracellular membranes to form membrane-bound replication compartments which are specialized for viral genome replication (2-4). Viral-encoded membrane proteins directly interact with intracellular membranes through their hydrophobic domain and act as scaffolds for the formation of viral replication complexes (RC) (2, 5). The viral RNA-dependent RNA polymerase (Pol), which catalyzes the synthesis of RNA molecules, localizes to the RC along with the double-stranded RNA replication intermediate, other viral replication proteins and host factors (6-8). Soluble viral replication proteins and host factors are likely brought to the RC via direct or indirect interaction with membrane anchor proteins.

Our lab is using ToRSV as a model for understanding the mechanisms of virus replication. ToRSV is a serious pathogen of small fruit and fruit trees and belongs to the genus *Nepovirus*

(family *Secoviridae*, order *Picornavirales*) (9, 10). Previous studies have determined that the ToRSV nucleoside triphosphate-binding proteins (NTB) is an integral membrane protein, which is associated with endoplasmic reticulum (ER) membranes and probably acts as an anchor for the RC (11). A putative signal peptidase cleavage site was identified in the C-terminal region of the NTB protein (12). The first part of this thesis was aimed at characterizing this cleavage and other membrane-associated modifications of the NTB protein and their impact on the stability of the protein. The ToRSV polymerase (Pol) was previously detected in various forms in ToRSV-infected plants, including the mature Pol and an polyprotein that contains a C-terminal truncation of the polymerase (13). The second part of my thesis was aimed at investigating the function and stability of various forms of Pol.

In this chapter, I will introduce intracellular membranes, in particular how proteins associate with ER membranes and how ER membrane-associated enzymes modify these proteins. I will briefly discuss the virus replication cycle and then focus on how picorna- and picorna-like viruses express their genome, and how viral-encoded proteins associate with intracellular membrane organelles for replication. The biology and molecular characteristics of ToRSV will also be discussed.

1.2 Cellular membranes and protein-membrane interactions

1.2.1 Biological membranes and intracellular membrane organelles

Membranes are important components of the cell. Biological membranes are composed of the lipid bilayer and associated membrane proteins (14). Membrane lipid molecules have bipolar properties. The formation of a lipid bilayer in an aqueous environment allows the

hydrophilic heads of the lipids to be oriented toward the polar aqueous environment and the hydrophobic fatty acyl tails to form the hydrophobic interior. Membrane proteins associate with the lipid bilayer and perform various functions on the membranes. Association with different sets of membrane proteins enables the membrane to exhibit different properties. The plasma membrane acts as the boundary between the cytosol and the extracellular space. Intracellular membranes are compartmentalized to form different intracellular organelles, which include the ER, Golgi apparatus, nucleus, mitochondria, peroxisome, endosomes, lysosomes (only in mammalian cells), vacuoles and chloroplasts (only in plant cells). Each organelle has its own specific function. In eukaryotic cells, the ER, Golgi apparatus, endosome, lysosome and plasma membrane are interconnected via membrane vesicles to function as a network and form the secretory and endocytosis pathways (15, 16). Since ToRSV replicates on ER-derived membrane (11), I will focus this chapter on the ER and protein-membrane interactions of the ER.

1.2.2 Endoplasmic reticulum (ER)

The ER membrane accounts for almost half of the total membranes in eukaryotic cells. It is composed of convoluted, flattened sacs or tube-like membrane structures (cisternae). These cisternae are continuous with the membrane of the nuclear envelope and are held together by the cytoskeleton. The membrane enclosed space is called the ER lumen (17). Two types of ER are identified in both animal cells and plant cells: the smooth ER and the rough ER. The smooth ER is mainly responsible for lipid metabolism and produces lipids for all other intracellular organelles. The rough ER is studded with ribosomes and acts as a manufacturing site for protein synthesis. Additionally, the rough ER acts as an important gateway for

protein transport. The cellular secretory proteins first enter into the ER. From there, they are destined to other intracellular compartments, or to the plasma membrane (18).

1.2.3 Membrane proteins

1.2.3.1 Classification of membrane proteins

Membrane proteins have different structural properties and they associate with biological membranes in different ways (19). Based on this, membrane proteins are generally classified into two major groups: the integral membrane proteins (intrinsic proteins) and the peripheral membrane proteins (extrinsic proteins). In this section, I will focus the discussion on the ER-associated proteins and orientation of their domains in relation to the cytosolic or luminal faces of the ER membrane. Integral membrane proteins contain one or more hydrophobic domains that span the lipid bilayer by directly interacting with the interior fatty acyl group. The hydrophilic regions of these proteins are located either on the cytosolic side, or on the luminal side of the lipid bilayer. However, not all integral membrane proteins carry a transmembrane domain. Some membrane proteins anchor themselves to the membrane by forming a covalent bond with the fatty acid. In this case, the polypeptide chain does not enter into the phospholipid bilayer but is located outside. These proteins are named lipid-anchor proteins. Peripheral membrane proteins do not penetrate the interior of the lipid bilayer, but associate with the membranes through interactions with integral membrane proteins or with the lipid head group. Salt treatment can successfully disassociate peripheral membrane proteins from the membrane, whereas integral membrane proteins can only be removed from the membrane via detergent treatment.

1.2.3.2 Lipid-protein interaction domains

Integral membrane proteins contain hydrophobic regions that connect hydrophilic regions. Association with membranes is driven by a hydrophobic interaction between their hydrophobic regions and the lipid bilayer. The hydrophobic domains of the membrane protein traverse the lipid bilayer and leave the hydrophilic loops in the aqueous environment. Two membrane interacting motifs have been well characterized in membrane proteins: the amphipathic α-helix and the transmembrane α-helix.

1.2.3.2.1 Amphipathic α-helices

The amphipathic helix is a unique lipid interacting motif first described by Segrest *et al.* in 1974 (20). It mediates weak, reversible association between proteins and membrane bilayers (21). The amphipathic helix is characterized by forming an α-helix structure with hydrophobic amino acid residues located on one side and hydrophilic amino acid residues segregated on the other side of the helix (20). Due to this property, the amphipathic helix is able to orient itself horizontally on the lipid monolayer with its hydrophobic side embedded in the hydrophobic core of the phospholipid bilayer and its hydrophilic moieties facing the cytosolic environment or the membrane lumen (22). Amphipathic helices also serve as a protein-protein interaction motif. It has been observed that amphipathic helix containing proteins tend to form oligomers. When associated with membranes, this process is usually accompanied by topological changes to the amphipathic helices and formation of aqueous pore that spans the lipid bilayer (23). In this case, the hydrophilic residues of the amphipathic helix are facing the aqueous pore and the hydrophobic amino acid residues are pointing towards the lipid fatty acid chain (23). Association of the amphipathic helix with local

membrane and formation of the amphipathic helix oligomers on the lipid monolayer also result in displacement of the lipid head group and reorganization of the fatty acid chains, which lead to curvature of the membrane (24, 25)

1.2.3.2.2 Transmembrane α-helices

Transmembrane domains of integral membrane proteins usually contain approximately 20 hydrophobic amino acid residues and adopt the α-helix conformation when traversing the membrane (26). Formation of the α-helix helps to expose the hydrophobic side chain to the interior membrane fatty acid group and establish a stable hydrophobic interaction between the transmembrane domain and the lipid bilayer. Some proteins contain one hydrophobic α-helix that spans the membrane with either the N-terminus or the C-terminus in the cytosol, while other proteins contain several hydrophobic α-helixes and traverse the membrane in and out multiple times. The hydrophobic α-helix also acts as a protein-protein interaction domain, both intramolecular and intermolecular helix-helix interaction are observed. The intramolecular interaction has been found to stabilize the protein and allow the protein to form higher structures (27). The intermolecular helix-helix interaction leads to the formation of high molecular oligomeric complexes (28).

1.2.4 Targeting of proteins to the ER

In eukaryotic cells, most secretory proteins and membrane proteins associate with ER co-translationally. For secretory proteins and for some membrane-associated proteins, targeting to the ER is directed by signal peptides that are located at the N-terminal region of the proteins. These signal peptides contain a stretch of hydrophobic amino acid residues. Once produced from the ribosome, the signal peptides are readily recognized by signal recognition

particles (SRP) and form an SRP-ribosome–nascent chain complex, which results in transient inhibition of mRNA translation (29-31). This whole complex is subsequently brought to the ER by specific interactions between the SRP and the rough ER located SRP receptors. This receptor binding process promotes the release of the SRP from the ribosome and polypeptide chain and allows them to interact with the ER-bound sec61 translocation complex (32). The signal peptide then inserts into the translocon channel and translation resumes. The nascent polypeptide is directed into the ER lumen, where the signal peptide is cleaved off from the growing polypeptide by a luminal enzyme called signal peptidase (see section 1.2.6.2) (33). The mature proteins are released into the ER lumen or on the membrane if they contain more than one hydrophobic region. Some membrane proteins possess internal signal sequences that are recognized by SRP but no signal peptidase cleavage event occurs. These internal signal sequences usually form α-helix structures to traverse the membrane and they are referred as membrane anchor sequences or transmembrane domains (34).

There are also proteins that target to the ER after the translation is finished. These proteins are usually synthesized on free ribosomes in the cytosol and their association with ER membrane is not SRP mediated. Instead, their signal peptides are directly recognized by different membrane-bound receptor proteins (Sec62 and Sec63) that form complexes with sec61 on the ER membrane (35). Molecular chaperones in the cytosol help to keep the polypeptide chain in an unfolded state so that they can easily get into the translocon channel. The Bip protein, which is located in the lumen, binds and pulls the polypeptide towards the lumen (36).

1.2.5 Topology of integral membrane proteins

Integral membrane proteins contain various numbers of transmembrane domains with the first transmembrane domain usually called signal peptide or signal sequence (34) (see 1.2.6.2). These domains traverse in and out of the lipid bilayer and form a special topology in the membrane. The protein topology is used to interpret the number of transmembrane domains and their orientation relative to the lipid bilayer. A number of factors determine the membrane protein topology. First, positively charged amino acid residues that flank the hydrophobic core region of the transmembrane domain usually locate on the cytosolic side of the membrane, which is referred as the positive inside rule (37). The second influential factor is the folding state of the upstream N-terminal region. It is found that only unfolded polypeptides transit through the translocon channel (38). Thirdly, the general hydrophobicity of the transmembrane domain also plays a role in determining its orientation. When the transmembrane domain is highly hydrophobic, the positive inside rule can be broken; resulting in a luminal localization of the N-terminal positively charged flanking sequence. In contrast, a decrease in the hydrophobicity results in the opposite orientation of the transmembrane domain (39).

Secretory proteins usually contain a single signal peptide in their N-terminal region that mediates membrane association and directs the downstream polypeptide into the ER lumen, this signal peptide is removed by the signal peptidase and the mature protein is released into the ER lumen (34).

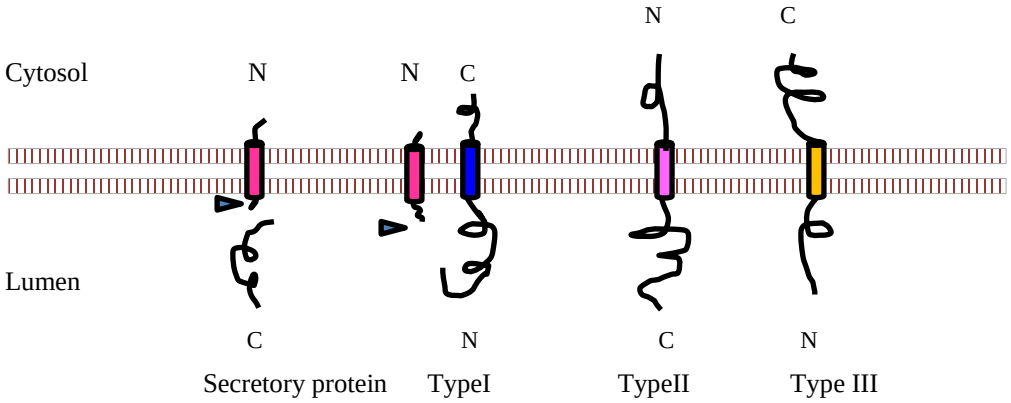
Single spanning-membrane proteins are divided into three categories based on their

topologies (Fig. 1.1). Type I membrane proteins possess a cleavable N-terminal signal peptide that directs translocation of C-terminal region into the ER lumen (40). Removal of the signal peptide by signal peptidase produces a new N-terminus in the lumen. The downstream transmembrane domain is responsible for anchoring the protein on the ER membrane and directs the C-terminus to the cytosol. Type II membrane proteins contain an uncleavable N-terminal or internal hydrophobic domain that directs the C-terminus or the protein into the ER lumen and the N-terminus in the cytosol (34). Type III integral membrane proteins have a reverse signal anchor that translocates their N-terminus in the lumen and C-terminus in the cytosol (34).

Some membrane proteins traverse the ER membrane multiple times. Their first transmembrane domain can be a cleavable signal peptide, an uncleavable signal anchor or sometimes a reverse signal anchor (34) (Fig. 1.1). For these proteins, the topology is mainly determined by this signal peptide. However, the downstream transmembrane domains also contribute to the overall topology. Von Heijne suggested that the downstream hydrophobic fragments also follow the positive inside rule, but not stringently (41). They insert into the membrane one after the other with alternating orientation. In some cases, the downstream transmembrane domains overrule the topology of the preceding transmembrane domain and forces it to fold back and reorient itself in the translocation channel (42). There is also evidence which suggests that N-glycosylation (see below, section 1.2.6.1) can influence the membrane protein topology (43). Some transmembrane domains only show weak hydrophobicity and depend on other transmembrane domains to tightly associate with the membrane (44). Both reverse signal anchors and cleavable signal peptides are observed in

some multi-transmembrane proteins. There are also proteins that contain uncleavable signal anchors which direct the C-terminal region into the ER lumen (Fig. 1.1).

Secretory protein and single-spanning membrane proteins



Multi-spanning membrane proteins

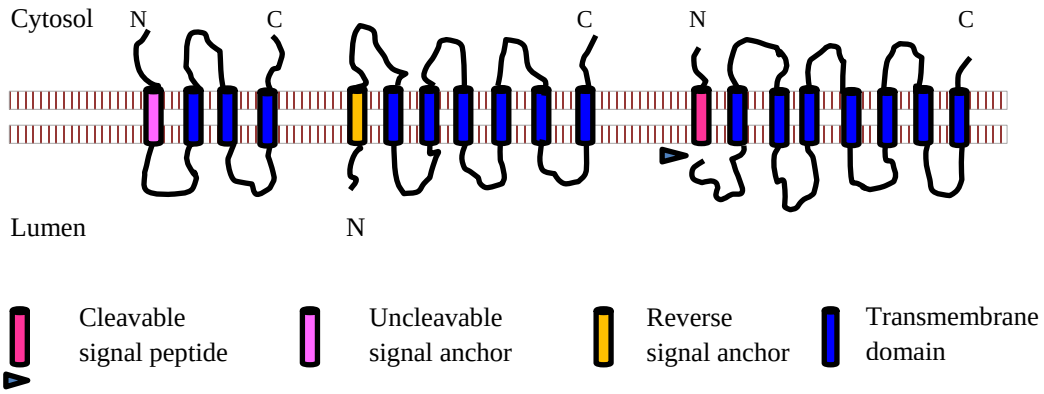


Figure 1.1: Topology of secretory proteins and membrane proteins. Three types of signal sequences are observed in membrane proteins: the cleavable signal peptide, the uncleavable signal anchor and the reverse signal anchor, these signal sequence are specified at the bottom of the figure. Please refer to text for further detail on the different types of membrane proteins. Adapted with permission from (Higy M, Junne T, Spiess M. 2004. Topogenesis of membrane proteins at the endoplasmic reticulum. Biochemistry 43:12716-12722.). Copyright (2004) American Chemical Society.

1.2.6 ER-associated protein modification and processing

The ER is not only an important site for protein synthesis. Association with different enzymes also makes it an important place for protein modification and processing. These membrane-associated enzymes including the oligosaccharyltransferase (OST) that catalyzes N-linked glycosylation, the signal peptidase that specifically cleaves the signal peptide and the signal peptide peptidase (SPP) that processes the signal peptide.

1.2.6.1 Asparagine (N)-linked glycosylation

N-linked glycosylation is a co-translational modification that commonly occurs on membrane proteins in both prokaryotic and eukaryotic cells (45, 46). It plays an important role in assisting proper protein folding and regulating protein function (47). N-linked glycosylation is performed by a large membrane protein complex called the oligosaccharyltransferase (OST) (48, 49). The yeast OST contains eight polypeptides, the active site of the enzyme is located on a subunit called Stt3p and is directed into the ER lumen by its transmembrane domain (50). Other subunits are involved in regulation of the catalytic activity or selection of a client protein. OST is found attached to both ribosome and translocon (51, 52). As long as the nascent polypeptide is translocated into the lumen and reaches the catalytic center of the OST, cleavage of the oligosaccharide from the lipid carrier will occur and covalently attach it to the polypeptide chain. The attachment of the glycan occurs through a selected asparagine (N) residue in the context of the conserved motif Asn-X-Ser/Thr (X can be any amino acids except Pro) (53). The efficiency of glycosylation is strongly affected by the flanking sequence, especially the amino acid at the +2 position. Both Ser and Thr are accepted at +2 positions. However, the glycosylation of Asn-X-Thr is much

more efficient than Asn-X-Ser (53, 54). N-linked glycosylation modification mostly occurs co-translationally when the polypeptide is still growing from the ribosome. However, post–translational glycosylation is also observed and a different OST isoform catalyzes this process (55). In both cases, the substrate polypeptide needs to be in the unfolded state to ensure that the glycosylation sequon is exposed to the OST. Because N-glycosylation only occurs on the luminal side of the ER, it is a good indicator of the membrane protein topology. A number of factors influence whether or not a particular N-glycosylation sequon is glycosylated, including the binding affinity between the flanking sequence and OST, local folding of the protein and concentration of the oligosaccharide donors. In addition, N-glycosylation only occurs when there is a distance of at least 12-14 amino acids between the glycosylation site and the transmembrane domain (56).

1.2.6.2 Signal peptidase processing

Signal peptidases belong to the group of serine proteases and are associated with ER, mitochondria and chloroplast-derived membranes (33). Eukaryotic ER signal peptidases consist of several polypeptides, and are referred to as the signal peptidase complex (SPC) (33). These polypeptides contain transmembrane domains that mediate their association with the membrane. The SPC from dog pancreas microsomes has been well characterized. It contains five polypeptides: SPC12, SPC18, SPC21, SPC22/23 and SPC25 (33) (Fig. 1.2). The catalytic subunits SPC18 and the SPC21 are homologous to bacterial type 1 signal peptidase and carry the catalytic center. They both contain the classic serine protease catalytic triad which consists of serine (S), histidine (H) and aspartate (D). The transmembrane domain at the N-terminus of the subunits translocates the catalytic center into

the luminal side of the ER with the catalytic active serine (S) residue located close to the lipid bilayer (33). Other subunits are probably involved in interacting with the ER translocon and stabilizing the catalytic subunits. They may also play a role in retention the enzyme complex on the lipid bilayer. The information about plant signal peptidase is very limited. The sequence of the catalytic subunits of several plant species is available, including *Arabidopsis thaliana* and *Nicotiana benthamiana*. The structure and processing specificity are currently unknown.

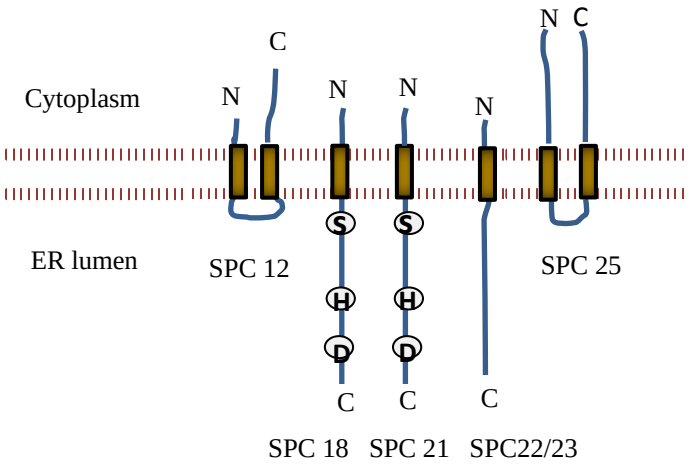


Figure 1.2: Signal peptidase complex and its topology on the ER membrane. Dog pancreatic signal peptidase contains five subunits: SPC12, SPC18, SPC21, SPC22/23 and SPC25. They associate with membrane though their transmembrane domains (shown in brown). The catalytic triad (S, H and D) of subunits SPC18 and SPC21 is located in the luminal side of the ER. Adapted with permission from (Paetzel M, Karla A, Strynadka NC, Dalbey RE. 2002. Signal peptidases. Chem Rev 102:4549-4580). Copyright (2002) American Chemical Society.

The well-known function of signal peptidase is to cleave off the signal peptide from preproteins co-translationally or post-translationally, which allows the release of mature proteins into the cellular secretory pathway. Signal peptides have been extensively studied (40, 57). They vary in length and do not share any sequence similarity. However, they share the same pattern in the arrangement of the signal sequence (Fig. 1.3). The amino-terminal region (n-region) contains positively charged amino acids, followed by a central hydrophobic core region (h-region) and a carboxyl-terminal region (c-region) which contains the signal peptidase cleavage site. The n-region is involved in determining the topology of the protein by following the positive inside rule, whereas the h-region traverses the membrane by forming a α-helix structure and translocates the c-region into the luminal side of the ER (33, 58). Statistical studies indicate that the presence of small and neutral amino acids at the -1 and -3 positions of the cleavage site is required for efficient signal peptidase processing, which is referred as the -3 -1 rule (40). The prevalent motif of the signal peptidase cleavage site is Ala-X-Ala. However, amino acid residues Gly, Cys, and Ser are also observed at the -1 position. Mutation of the -1 position from a preferred amino acid to an unfavorable amino acid completely abolishes signal peptidase processing at the initial site. However, cleavage often occurs at alternative Ala-X-Ala cleavage sites instead (59, 60). Although the majority of proteins have their signal peptide at the N-terminus, there is also evidence that the cleavable signal peptide can be internally located in the protein (61, 62). Sometimes the signal peptide just acts as a signal anchor sequence and no signal peptidase processing occurs (63).

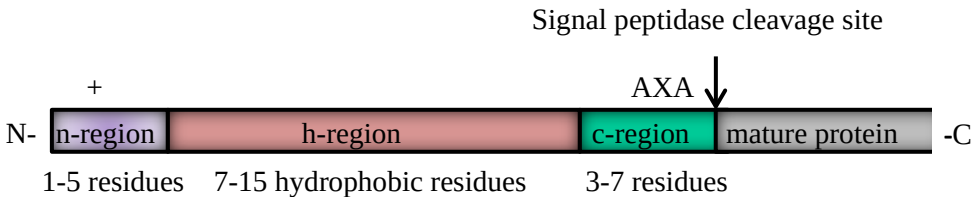


Figure 1.3: Schematic representation of a typical signal peptide. Signal peptide contains three distinct domains: the positively charged n-region (purple), the hydrophobic h-region (pink) and the c-region (green) that contains the signal peptidase cleavage site. The mature protein is indicated by gray color. Cleavage site between the signal peptide and mature protein is indicated with the black arrow. The preferred signal peptidase cleavage site Ala-X-Ala is shown in the figure. Adapted with permission from (Auclair SM, Bhanu MK, Kendall DA. 2012. Signal peptidase I: cleaving the way to mature proteins. Protein Sci 21:13-25). Copyright (2012).

1.2.6.3 Signal peptide peptidase processing

The SPP is an aspartyl proteases with intramembrane proteolytic activity. It is a highly hydrophobic protein that traverses the membrane multiple times with the conserved motif YD and GXGD located on two of the adjacent transmembrane domains (TMDs) (64). Both aspartate residues are located in the catalytic center of the enzyme and mutation of these residues abolishes the enzyme activity (65). The SPP mainly resides in the ER membrane and its homolog SPP-like proteases are found distributed in the Golgi apparatus, plasma membrane and endosomes (66, 67). The function of the SPP is to process the remnant signal peptides which were previously cleaved by signal peptidase (68). SPP is capable of processing both N-terminally located and internally located signal peptides (61, 69), therefore leading to the degradation of the signal peptides and the release of bio-functional protein fragments from the membrane (70).

1.3 Replication cycle of plant positive-strand RNA viruses

The majority of plant viruses contain a positive-strand RNA genome, which acts as an mRNA. Compared with host cells and other intracellular parasites, viral genomes are small and their coding capacity is limited. To establish a successful infection and accomplish their multiplication cycle, viruses take full advantage of their host cells and utilize host factors at every step of their replication cycle. The host translation system, various intracellular organelles, host proteins, host secretory pathway and host degradation pathways are all hijacked by viruses to assist their survival (2, 71-73). Positive-strand RNA viruses that infect plants have been extensively studied. They share a similar life cycle which consists of six key steps (Fig. 1.4):

1. Virus entry: entry of plant viruses into host cells is mostly vector mediated (insects, nematodes, mites and fungi). Some viruses can also be transmitted through pollen and seeds. In the lab, we use mechanical inoculation to help viruses get into the plant cells.

2. Disassembly: upon entry into the cell, virus particles disassembly and release genetic material into the cytosol where it has access to various host cellular machineries.

3. Translation: positive-strand RNA viruses directly recruit host translation machinery to synthesize viral replication proteins as well as structural proteins (74). Viruses also evolved to adopt different strategies to maximize their coding capacity, such as subgenomic RNA, stop codon readthrough, leaky scanning of start codons and ribosomal frameshift. These strategies are well discussed in a number of review papers (75-77). Nepoviruses express their genome through a polyprotein manner and this strategy will be discussed in detail later (Section 1.4).

4. Replication: Virus infection induces membrane proliferation and formation of membrane vesicles, double-membrane vesicles or multivesicular bodies in infected cells (4, 5). These rearranged membrane structures are derived from various intracellular organelles, including the ER, mitochondria, peroxisome, chloroplast, or vacuole (5, 78, 79). The RNA-dependent RNA polymerase, other viral factors, viral genomic RNA and host factors co-localize in these modified intracellular structures and support active viral replication. Formation of this organelle-like microenvironment has several advantages. For example, it locally enriches viral replicases and cellular factors to stimulate replication efficiency, spatially separates and regulates the viral replication step from RNA translation and virion assembly, protects viral RNA and proteins from degradation by the host nuclease and various proteases, supplies an anchor for the viral replication complexes and prevents activation of host defense responses

(2, 80). Viral-encoded membrane proteins play a key role in formation of replication complexes on the membranes; they target to membranes independently and recruit other necessary factors to this replication site, i.e. they act as anchors for the RC. Well-studied membrane anchor proteins include brome mosaic virus (BMV) 1a, tomato bushy stunt virus (TBSV) P33, potyvirus 6K and turnip yellow mosaic virus (TYMV) 140K proteins (2, 5, 81). Once the replication complexes are established, the Pol catalyzes the synthesis of a complementary negative-strand RNA by using the genomic RNA as a template. The positive-strand RNA progenies are produced on this negative-strand RNA template and subsequently released into the cytosol for initiation of more rounds of translation or for assembly of virus particles.

5. Encapsidation: once the progeny RNA and virus structural proteins accumulate in the cell, the virus switches from replication to encapsidation. The newly synthesized viral RNAs are packaged into coat protein (CP) shells to form infectious virus particles.

6. Cell-to-cell and systemic movement: plant viruses spread to adjacent cells by moving through the plasmodesmata (PD) in the form of the virus particles, viral ribonucleoprotein complexes or viral replication complexes (82, 83). The plant picorna-like viruses such as ToRSV, spread to the adjacent cells by using the tubule-guided transportation mechanism (84, 85). In this case, viral-encoded movement proteins (MP) interact with PD-localized proteins and form the tubule structures within the PD, which allows the cell-to-cell transport of virus particles (82).

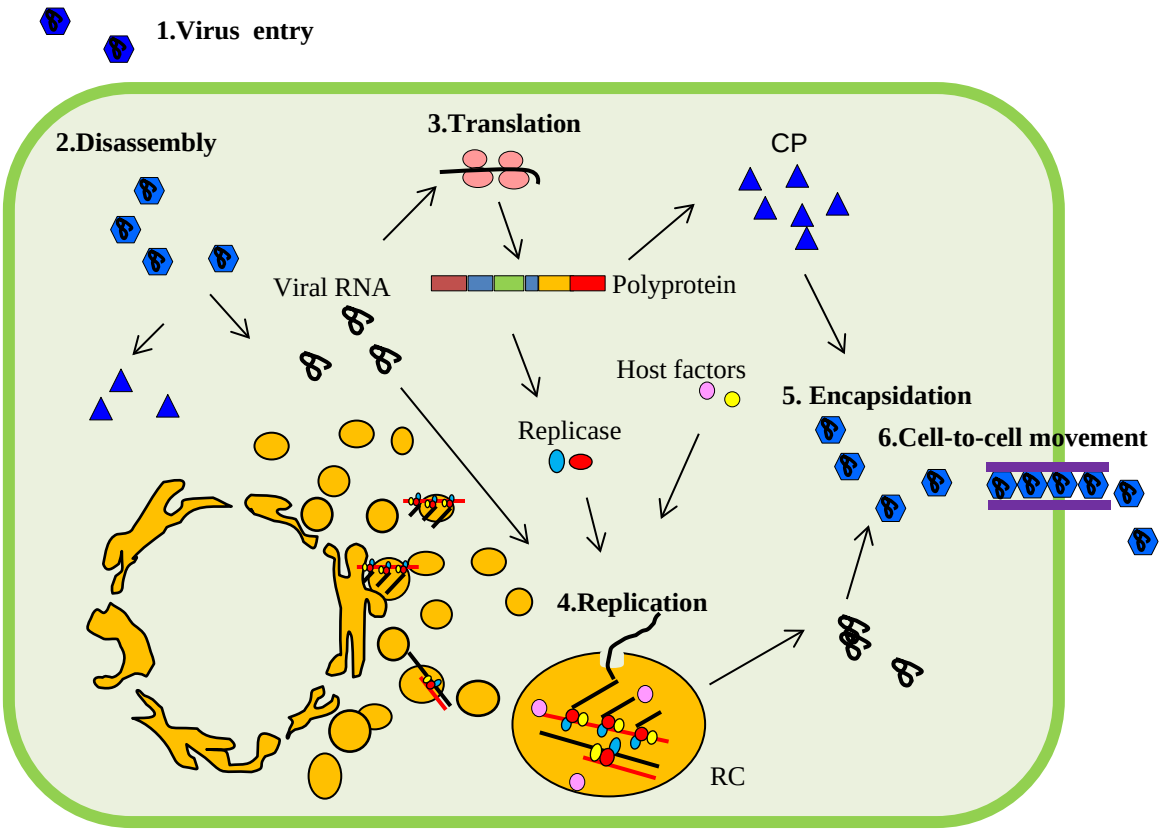


Figure 1.4: Schematic representation of a plant picorna-like virus replication cycle. The plant virus replication cycle usually contains six important steps: 1.Virus entry. 2. Disassembly. 3. Translation. 4. Replication. 5. Encapsidation. 6. Cell-to-cell movement (please see text for details). The host cell is shown by the large green box and virus particles are shown with blue hexagons. The negative-strand RNA is indicated with the red line while the membrane bound replication complexes are shown by orange circles. Adapted with permission from (Sanfacon H, Zhang G, Chisholm J, Jafarpour B, Jovel J. 2006. Molecular biology of Tomato ringspot nepovirus, a pathogen of ornamentals, small fruits and fruit trees, p. 540-546. In Teixeira da Silva J (ed.), Floriculture, Ornamental and Plant Biotechnology: Advances and Topical Issues (1st Edition), vol. III. Global Science Books, London, UK.). Crown copyright (2006).

1.4 Expression of viral proteins using polyprotein strategy

Positive-strand RNA viruses in the families *Picornaviridae, Flaviviridae, Potyviridae,* and *Secoviridae* express their genome by using a polyprotein strategy (1, 86). Translation of viral genomic RNAs produces large polyproteins, which undergo a series of proteolytic processing steps to liberate functional gene products. Cleavage of the polyprotein is conducted by the viral protease and sometimes also by host proteases. Viral-encoded proteases have been extensively studied (86, 87). They hydrolyze the polyprotein peptide bond either *in cis* or *in trans*. When the protease and cleavage sites are present on the same molecule, this cleavage event is called in *in cis* (intramolecular) cleavage. In contrast, when the cleavage sites are present on one molecule and the protease is supplied from a different molecule, the processing event is referred as *in trans* (intermolecular) cleavage. A combination of *cis*- and *trans*-cleavage is observed in viral polyprotein processing. The *cis*-cleavage is usually a rapid process while *trans* cleavage events occur slowly and show more sensitivity to inhibitors and to dilution of the protease concentration (1). Viral-encoded polyprotein processing is a highly specific event. Site-directed mutagenesis indicated that once the cleavage site and the flanking amino acid residues are mutated, it is no longer recognized, or poorly recognized by the viral-encoded protease. *Trans*-cleavage was found to be more sensitive to mutations than *cis*-cleavage (88).

1.4.1 Viral proteases involved in polyprotein processing

All viruses that express their genome in a polyprotein manner encode one or more proteases which actively participate in the release of functional viral proteins (86, 89). Viral proteases are generally classified into four categories based on the conserved amino acid residues that

constitute the active site as well as the three dimensional structure of the protein. These four types of protease are: the serine, the cysteine, the aspartic and the metallo-protease (86). The serine proteases contain a catalytic triad that is usually composed of histidine (His), aspartate (Asp) and serine (Ser) amino acid residues. However, the 3C and 3C-like protease encoded by picorna and picorna-like viruses contain a cysteine (Cys) instead of Ser as a nucleophile, but they fold into a three dimensional structure that is similar to that of serine proteases, therefore these proteases are referred to as serine-like protease (86). The cysteine proteases possess a catalytic dyad that consists of Cys and His while the aspartic proteases are characterized by having a catalytic dyad composed of two Asp residues. These key amino acids are brought together in the three dimensional structure and constitute the catalytic center of the enzyme. Unlike other groups of protease, the metalloproteases require a divalent metal cation at the active site to perform the cleavage event (86). Other than the catalytic center, the substrate-binding pocket, which is the three dimensional structure that recognizes and interacts with the substrate cleavage site, is also crucial for efficient cleavage (1). It specifically binds to the substrate cleavage site and properly exposes the scissile bond to the active site for efficient processing. The substrate binding pocket of the viral protease may show preference for a single amino acid or they may have a very strict requirement of up to seven amino acids that flank the cleavage site. Unlike the conservation that has been observed for the catalytic dyad/triad, the substrate binding pockets exhibit significant sequence variation from one viral protease to another, which ensures the specificity of cleavage.

Viruses encode one or several proteases to accomplish their polyprotein processing. For example, nepoviruses and comoviruses encode only one protease, which is sufficient for cleavage of both RNA1 and RNA2 encoded polyproteins and release of functional mature proteins and precursors required for all the viral activities (1). In contrast, potyvirus encoded polyproteins are processed by three viral-encoded proteases: the virus-encoded proteases P1, the helper-component proteinase HC-Pro and the NIa-Pro (1, 86).

In addition to processing polyproteins, some viral proteases also cleave cellular proteins. For example, picornavirus and coxsackievirus encoded 3C and 2A cleave the translation initiation factor eIF4G and the polyA binding protein, leading to shut down of cellular cap-dependent translation (90-94), Viral gene expression, however, is not affected since they are translated in an internal ribosome entry site (IRES)-dependent manner (95). The leader protein (L^{Pro}) of foot and mouth disease virus also cleaves eIF4G and inhibits cellular protein translation (96). Additionally, the picornavirus 3C also cleaves cellular transcription factors, such as TATA-binding protein, the transcriptional activator cyclic AMP-responsive element-binding protein, the octamer-binding factor and the p53. (97-99). This phenomenon indicates that picornavirus infection not only disrupts cellular gene expression at the translational level, but also at the transcriptional level, therefore benefiting the virus by providing more translation resources for viral genome expression.

1.4.2 Host proteases involved in maturation of viral proteins: signal peptidase, SPP and furin

For some viruses, their own proteases are necessary but not sufficient to release all gene

products and a supplementary protease activity from the host is required. For example, some viruses in the families *Flaviviridae* and *Togaviridae* hijack the ER-associated signal peptidase to release the structural proteins from the viral-encoded polyprotein (100-103). The polyprotein strategy of hepatitis C virus (HCV), which is a member of the family *Flaviviridae,* will be discussed in detail below. An obvious difference between viral-derived signal peptidase substrate and cellular substrate is that the cleavage sites of the viral substrate are usually located internally in the polyprotein. Mutations that block the signal peptidase cleavage remarkably reduce the virulence of the virus (104).

Sometimes, the proteins released after signal peptidase cleavage are not the mature proteins. Further cleavage by SPP is required to generate the functional protein. For example, the core protein of HCV is released from the polyprotein by sequential processing events that are performed by the signal peptidase in the lumen and then by SPP cleavage within the lipid bilayer (105). This sequential cleavage is essential for transport of the core protein to cytosolic lipid droplets and for assembly of virus particles (105-108). The SPP is also involved in releasing the core protein of other flavivirids (109, 110). It is worth noting that some proteins that are not expressed through polyprotein also undergo SPP processing. For example, the envelope protein of foamy virus (a retrovirus) is identified as a substrate of signal peptide peptidase-like 3 (SPPL3), although the biological function of this cleavage still needs to be investigated (111).

Furin is a serine protease that is localized in the trans-Golgi network. A number of viruses are known to take advantage of this host protease for the proteolytic maturation of their

structural proteins. The best studied example is that of the semliki forest virus E3-E2 cleavage site on the sub-genomic RNA encoded polyprotein (112). In addition, the glycoproteins of HIV, influenza virus, respiratory syncytial virus and ebola virus are synthesized as inactive preproteins, which must be cleaved by furin or furin-like proteases to become fully functional (113-116). Evidence suggests that these endoproteolytic cleavage events are important for virus infectivity and pathogenicity.

1.4.3 Polyprotein processing of picornaviruses

The picornavirus genome consists of a single positive-strand RNA molecule, which is linked with the viral genome-linked protein (VPg) at its 5' end and has a 3' poly(A) tail (Fig. 1.5). IRES-dependent translation produces a single large polyprotein (117), which is cleaved by several viral proteases: the leader protein (L, a cysteine protease, only present in aphthoviruses), 2A (a serine-like protease, only present in enterovirusese), 3C (a serine-like protease, encoded by all picornaviruses) and its precursor 3CD (Pro-Pol) (95, 118). The L and 2A proteases cleave at their own C-terminus and N-terminus respectively. 3C and 3CD are responsible for processing at other cleavage sites (86). The primary cleavage occurs rapidly and efficiently in a co-translational manner. It takes place at L/1A,1D/2A and 2C/3A sites and is performed by L, 2A and 3C, respectively (Fig. 1.5). This primary cleavage allows release of the L protein, the P1 (1ABCD) region, which contains the structural proteins, and the P2 (2ABC) and P3 (2ABCD) regions which encode nonstructural proteins. P1, P2 and P3 are further processed by by 3C and 3CD. The polyprotein 3CD , rather than the mature 3C, is responsible for efficient processing of the poliovirus (PV) P1 region (119). In contrast, the 3C of other picornaviruses is capable of cleaving the structural protein precursor P1 (1).

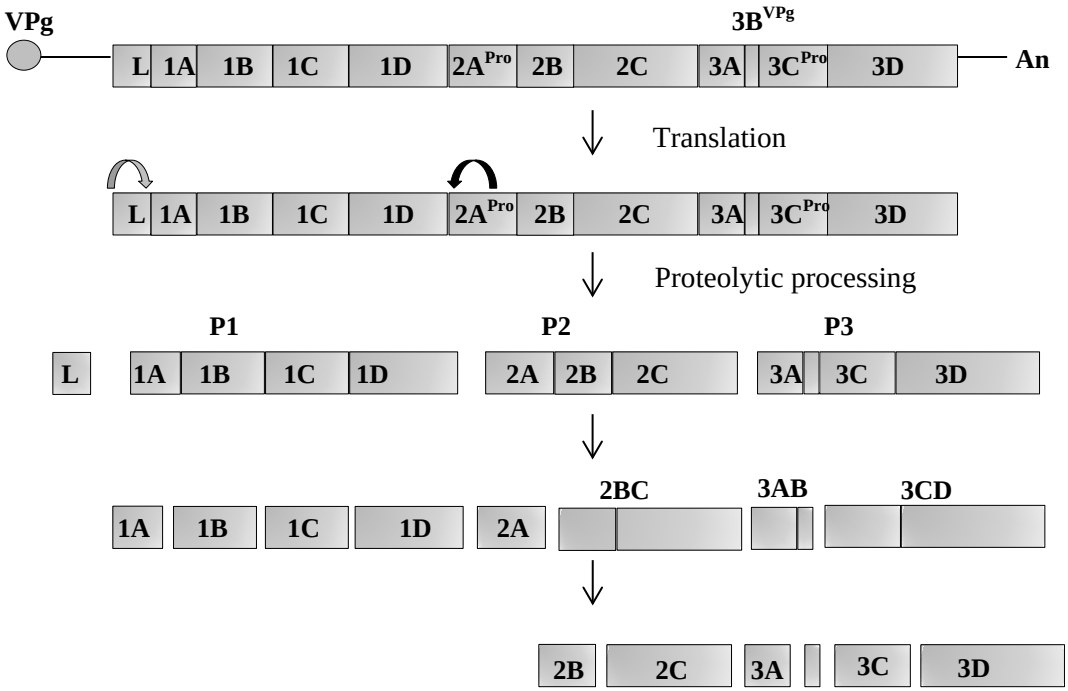


Figure 1.5: Genomic organization and polyprotein processing of picornavirus. The genomic organization of picornavirus is shown at the top of the figure. Different protein domains are indicated with boxes with the vertical lines representing the cleavage sites. Cleavage sites processed by L and 2A proteases are shown with gray and black curved down arrows, respectively. Other cleavage sites are processed by 3C and 3CD. P1, P2, P3, 2BC, 3AB and 3CD are the main polyproteins detected. Adapted from Microbes and Infection 6, Kristin M. Bedard, Bert L. Semler, regulation of picornavirus gene expression, 702–713, Copyright (2004), with permission from Elsevier.

1.4.4 Polyprotein processing of HCV

Viruses in the family *Flaviviridae* also express their positive-strand RNA genome via polyprotein strategy. Polyproteins of these viruses are co- and post-translationally processed by proteases of both viral and host origin. Host proteases process the N-terminal structural protein region. Processing of the non-structural protein region is mainly performed by viral proteases. HCV is a member of the family *Flaviviridae*. Its genome is translated in an IRES-dependent manner to produce a large polyprotein which contains ten protein domains (Fig. 1.6). The signal peptidase cleaves the C/E1 and E1/E2 junctions co-translationally which allows rapid release of C and E1, whereas processing at E2/p7 and p7/NS2 sites are slower and result in the accumulation of the E2-p7-NS2 and E2-p7 polyproteins (120, 121). The released C protein is further cleaved by the SPP to produce bio-functional mature protein for virus assembly (107, 122). Non-structural protein junctions are proteolytically processed by the viral NS2 and NS3 proteases. NS2 is a cysteine protease that cleaves at its own C-terminus to separate NS2-NS3. NS3 is a serine protease that is responsible for processing the remaining junctions. The cofactor NS4A stabilizes NS3 and generally enhance processing efficiency at all cleavage sites (123). Clevage at NS4B -NS5A junction does not occur in the absence of the cofactor NS4A (124).

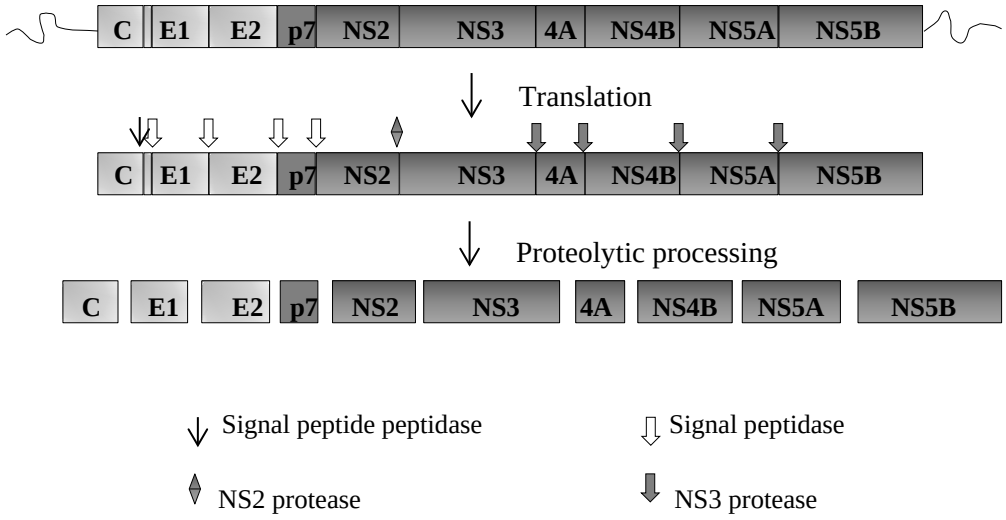


Figure 1.6: Genomic organization and polyprotein processing of HCV. The genomic organization of the HCV is shown at the top the figure. The structural protein region, which is located in the N-terminal region of the polyprotein, is shown in light gray while the non-structural region is indicated with dark gray. Cleavage sites that are processed by viral and host proteases are specified at the bottom of the figure. Adapted with permission from Macmillan Publishers Ltd: [Nature Reviews Microbiology] (Darius Moradpour, François Penin and Charles M. Rice, Replication of hepatitis C virus. 2007 Jun;5(6):453-63), copyright (2007).

1.4.5 Regulation of polyprotein processing

Expression of viral proteins via a polyprotein strategy highly enriches the coding capacity of the viral genome and allows regulating gene expression by sequential, alternative processing. Key features of the polyprotein strategy are listed below:

(1) The initial primary cleavage usually occurs co-translationally when the polypeptides are still growing from the ribosomes. Therefore the full-length genome-encoded polyprotein is never detected in infected cells due to this highly efficient primary cleavage (95).

(2) Unlike other viral gene expression strategies, proteins expressed via polyprotein strategy are produced in equimolar amounts. However, proteolytic processing efficiency at each cleavage site varies, which allows accumulation of various mature proteins as well as intermediate polyproteins. The half-life of these gene products differs from one to another, therefore allowing accumulation of each protein at different concentration (89).

(3) Polyprotein processing occurs in a specific order. As shown in Fig. 1.5, the picornavirus-encoded polyprotein is initially cleaved into three large precursor P1, P2, P3. The 3CD is released from P3 and acts as an active protease that is involved in further processing, followed by release of the active polymerase 3D from the 3CD precursor for viral replication. Cleavage of 1A-1B capsid protein precursor is delayed until the late stage of the picornavirus life cycle, probably preceding the assembly of virus particle (89, 125). Generating viral gene products in this way can perfectly meet the changing requirements at different stage of the virus life cycle.

(4) A single precursor can be processed in an alternative pathway and release different subsets of gene products. For example, the cowpea mosaic virus (CPMV) 170K (NTB-VPg-Pro-Pol) can either be cleaved to produce the 60K (NTB-VPg) and 110K (Pro-Pol) proteins, or the 84K (NTB-VPg-Pro) and 87K (Pol) proteins (126).

(5) The intermediate polyproteins actively participate in the viral replication cycle and may play a different role compared to the mature form of the proteins. The best characterized example is the PV 3CD and 3D proteins. 3CD maintains the protease activity and efficiently releases the capsid proteins from the precursor P1(127). However, it does not display any polymerase activity. The polymerase activity can only be activated when 3D is liberated from the precursor 3CD (128).

(6) As the key enzyme during polyprotein processing, the function and activity of viral-proteases are highly regulated in various ways:

First, different forms of the protease are produced due to alternative processing. Protease activity varies dependent on which form the protease domain is present. For example, the purified recombinant ToRSV Pro is five to ten times more active than the precursor VPg-Pro in processing the RNA2-encoded polyprotein at the X4-MP and MP-CP cleavage sites (129). Similarly, the grapevine fanleaf nepovirus (GFLV) Pro shows higher cleavage efficiency than VPg-Pro in processing the cleavage sites that are located on the RNA2 polyprotein (130). The slow release of the mature Pro from the intermediate precursor VPg-Pro may regulate the maturation of the RNA2-encoded structural proteins and movement protein, which are only required at late stages of the virus replication cycle.

Second, other viral proteins may be involved in regulating the efficient release of the final mature protease from its precursor protein. For example, the purified recombinant 3AB interacts with 3CD and accelerate autoprocessing between 3C and 3D (131). The CPMV Pro-Pol is a stable polyprotein, but adding the upstream 58K-VPg cleavage site and the entire VPg region to this polyprotein significantly enhanced the cleavage between Pro and Pol (132).

Third, different forms of protease may be required to process different cleavage sites on the polyprotein. For example, cleavage of P1 is conducted by PV encoded 3CD while cleaveage of P2 and P3 is performed by 3C (127). The 3D region of P3 is involved in the specific recognition of the P1 substrate and therefore stimulates cleavage by 3C (119, 127).

Finally, sometimes the protease shows a strict requirement for a cofactor for processing at certain cleavage sites. For example, although the CPMV Pro is responsible for processing at all cleavage sites of the B-RNA and M-RNA encoded polyproteins, the 32K protein, which does not exhibit any proteolytic activity, is required for efficient processing at the glutamine-methionine site of the M-RNA encoded polyprotein (133). The HCV encoded NS4A also act as a cofactor that assists cleavage of NS4B-NS5A by NS3 protease (124). These cofactors, probably increase the protease activity by promoting the specific interaction and alignment of the protease with its substrate.

1.5 Replication of picorna and picorna-like viruses

1.5.1 Common features of picorna and picorna-like viruses

Viruses in the family *Picornaviridae* include a number of important pathogens that infect animal and human cells. Plant viruses in the genera *Comovirus* and *Nepovirus* (family *Secoviridae*, order *Picornavirales*) share several common features with picornaviruses in genome organization and gene expression, therefore these viruses are referred to as plant picorna-like viruses (134). These common features are listed as follows: (1) The virus particles are icosahedral and contain positive-strand RNA genome. (2) Viral genome have a small viral- encoded protein VPg (2-3 kDa) covalently linked to the 5’ end and a poly(A) tail at the 3’ end. The viral genome is translated in a cap-independent manner. (3) Genomic

information is expressed via a polyprotein strategy and the viral-encoded protease has a cysteine as nucleophile but folds like a serine protease. (4) Replication proteins, such as helicase, VPg, Pro and Pol are organized in the same order on the genome. In contrast to animal picornaviruses that have a monopartite genome, some plant picorna-like viruses, such as nepoviruses and comoviruses have a bipartite genome with RNA1 encoding replication proteins and RNA2 encoding structural and movement proteins (135).

1.5.2 Replication of picorna and picorna-like virus

Like other positive-strand RNA viruses, replication of picorna- and picorna-like viruses takes place in association with intracellular membranes. Virus infection causes dramatic changes, such as proliferation and rearrangement of the membrane and induces the formation of single or double membrane vesicles that house the replication complexes. Vesicles of picornaviruses are derived from the ER, Golgi apparatus and lysosomes. The host secretory system and autophagy pathway are also involved (136-139). The vesicles of nepo- and comoviruses are mainly derived from the ER membrane (140). As will be described below, viral-encoded integral membrane proteins directly interact and associate with intracellular membranes. Other viral proteins, as well as viral RNA are recruited to the replication complexes via protein-protein and protein-RNA interactions with these membrane proteins. Polyproteins that contain these membrane anchor domains are also directed to the membrane. A variety of cellular factors also actively participate in the replication of picorna- and picorna-like viruses. These host factors are well discussed in a number of review papers (95, 141-143). In the following section, I will focus on viral- encoded proteins and their roles in virus replication.

1.5.2.1 Membrane proteins involved in picorna and picorna-like virus replication

1.5.2.1.1 Picornavirus membrane proteins: 2B, 2C, 3A and their precursors

The best studied picornavirus is PV. Vesicles derived from the ER, Golgi apparatus and lysosomes are observed in PV-infected cells. Replication complexes are assembled on the cytoplasmic face of these vesicles and form rosette-like structures (144, 145). Viral-encoded nonstructural proteins, host factors and viral RNA co-localize in these structures for active virus replication. PV encodes three integral membrane proteins: 2B, 2C and 3A (142). They target to intracellular membranes independently and polyprotein that contain these protein domains also associate with membranes (146).

The 2B protein has an N-terminal amphipathic helix and a C-terminal transmembrane domain which mediate its membrane association. Introducing mutations to either domain lead to defects in viral RNA synthesis (147-149). In addition, PV 2B oligomerizes on the membrane to form tetrameric aqueous pores and therefore increases the membrane permeability. Both the amphipathic helix domain and the transmembrane domain are able to traverse the membrane and are actively involved in pore formation (150). A similar phenomenon is observed in coxsackievirus that 2B is able to enhance the permeability of both the ER membrane and the plasma membrane (151). Later evidence showed that increasing the membrane permeability can facilitate the release of progeny virions at late stages of replication.

Membrane association of 2C is directed by an amphipathic helix which is mapped at the N-terminal of the protein (152-154). Both 2C and its precursor 2BC are responsible for the

formation of vesicles that support active virus replication (155). Transient expression of 2C and 2BC induces massive proliferation of the cytoplasmic membranes. These membrane rearrangements are similar to those observed in virus infection (156). PV 2C mutants show defects in viral RNA replication. However, these defects can be rescued by supplying 2C *in trans* (157). 2C is a multifunctional protein and also specifically binds to 3' UTR of the negative-strand RNA (158, 159). It is hypothesized that this interaction helps retaining the negative-strand RNA in the replication complexes and plays a role in the initiation of positive-strand RNA synthesis (160). 2C is also an NTP-binding protein and helicase (161), it forms homo-oligomers and display ATPase and GTPase activity (162, 163). Additionally, *in vitro* experiments demonstrated that the central and C-terminal region of 2C interact with 3C and negatively regulate the protease activity of 3C and the 3C-induced apoptosis during virus infection (164).

Picornavirus 3A and its precursor 3AB (3B is the VPg) target the ER and Golgi apparatus in the absence of other viral factors (146). The membrane binding motif was mapped to the hydrophobic domain located in the C-terminal region of 3A (165). Both 3A and 3AB are located in the membrane vesicles that support active viral replication (166). Mutations that reduce the hydrophobicity of 3AB impairs PV RNA replication (167). 3AB is a multifunctional protein. It interacts with the 3D polymerase through its 3B region and directs the soluble 3D protein to the membrane fractions (168). 3AB also has RNA binding properties and specifically binds to the viral RNA through the cloverleaf structure which is located in the 5' UTR region (169). 3AB forms a complex with 3CD and this complex is essential for virus replication. Additionally, 3AB stimulates *in cis* proteolytic processing of

3CD and release the functional polymerase 3D for virus replication (131). Both 3A and 3AB get glycosylated when expressed in the *in vitro* translation system in the presence of canine microsomal membranes (165). The glycosylation sequon Asn-Ile-Thr which is located near the middle of the 3A domain likely acts as the glycosylation site. Inhibitor treatment that blocks the glycosylation results in impaired virus replication *in vivo* (165). It was hypothesized that the glycosylated form of 3A and 3AB are required for forming the replication complexes or that uridylylation of VPg might depend on this glycosylation event.

An extensive network of protein-protein interactions are observed among picornavirus encoded membrane proteins. The coxsackievirus 3A protein self-interacts and form dimers. Site-directed mutagenesis that disrupts homodimer formation strongly inhibits intracellular trafficking and virus replication (170). Self-interaction of PV 3AB is also observed and this interaction is directed through the N-terminal region and the C-terminal hydrophobic domain of 3A (171, 172). Yeast two hybrid analysis also revealed interactions among PV P2 and P3 encoded non-structural proteins and their precursors, in particular between 2C and 3AB and between 2B and 3A/3AB (173). Interactions between 2C and 3A and between 2BC and 3A /3AB are also observed (174). This protein-protein interaction network brings all replication factors together on the membrane. However, more work is required to fully understand the functions of these interactions.

1.5.2.1.2 Comovirus membrane anchor proteins: 32K (Co-Pro) and 60K (NTB-VPg)

Infection of CPMV causes drastic morphological changes to the ER membrane and leads to

accumulation of ER-derived membrane vesicles (175, 176). Double-stranded RNA as well as viral replication proteins are found associated with these vesicles (176, 177). B-RNA can replicate without M-RNA while the replication of the M-RNA is dependent on B-RNA encoded replication proteins. Expression vectors that carry the entire B-RNA coding region and express the 200 kDa polyprotein support the replication of M-RNA in protoplast (178). Interestingly, shorter B-RNA encoded proteins that contain the polymerase, such as 170K (NTB-VPg-Pro-Pol), the 110K (Pro-Pol) and the 87K (Pol), are unable to support M-RNA replication (178), indicating that CPMV 32K, which located at the N-terminal end of the B-RNA is indispensable for virus replication.

CPMV 32K (Co-Pro) and the 60K (NTB-VPg) proteins display properties of integral membrane proteins. Both proteins target to the ER membrane in the absence of other viral factors (179). Overexpression of the 32K and 60K using a tobacco rattle virus (TRV) vector induces the formation of local necrotic spots on inoculated *N. benthamiana* leaves (177). Interestingly, this necrosis does not occur in the course of a natural virus infection. The cell death may be due to the cytotoxic properties of these two overexpressed proteins (177). Expression of the 60K protein using the baculovirus expression system induces the formation of membrane vesicles in host cell which confirms that 60K alters membranes (180). Mutation of two conserved amino acids in the NTB domain abolishes CPMV replication in the protoplast, indicating that NTB is an essential factor for viral replication (181). Yeast two hybrid analysis demonstrated that the C-terminal region of NTB interacts with host SNARE-like proteins VAP27-1 and VAP27-2, suggesting that NTB plays a role in hijacking the host transport system for its own benefit (182). However, the function of this interaction still

needs to be clarified. NTB is also a putative helicase that shares conserved amino acid motifs with other helicases (183).

Three stretches of hydrophobic domains are observed in the C-terminal region of the 32K protein (177). In the 60K protein, an amphipathic helix in the N-terminal region and a hydrophobic domain in the distal C-terminal region are conserved among all comoviruses (177). However, the membrane binding capability of these motifs still needs to be confirmed experimentally. Since comoviruses use a polyprotein strategy, various forms of NTB are detected in CPMV-infected protoplasts, including the 170K (NTB-VPg-Pro-Pol), the 84K (NTB-VPg-Pro), the 60K (NTB-VPg) and the 58K (mature NTB) proteins (184). These proteins are believed to maintain the membrane binding ability since they have the integral membrane protein domain. Co-immunoprecipitation indicated that the 32K protein (Co-Pro) interacts with NTB and NTB-containing polyproteins (184). Only the mature form of 32K accumulates in virus infected protoplasts due to the highly efficient primary cleavage of 32K /170K site of RNA-1 encoded polyprotein (184). *In vitro* assay indicated that the 32K remains associated with the 170K polyprotein after this primary cleavage, probably via an interaction with the NTB domain. This interaction plays an important role in regulating polyprotein processing (185). Whether it is also required for forming the replication complexes is still unknown.

Membrane anchor proteins of other comoviruses have also been characterized. Studies with chimeric RNA transcripts indicated that the Co-pro and the C-terminal half of the NTB of the bean pod mottle virus (BPMV) are involved in inducing severe symptoms on soybean plants

(186). Individual expression of these two coding regions via a potato virus X (PVX) vector induces necrosis on *N. benthamiana* plants, suggesting that both Co-Pro and NTB act as symptom determinants (186). In the case of radish mosaic virus, an N-terminal transmembrane helix and amphipathic helix as well as a C-terminal transmembrane helix are identified in the helicase (NTB) domain (187). When transiently overexpressed in *N. benthamiana* leaves, the NTB targets to the ER and induces hypersensitive response -like cell death, which is similar to that observed in a natural virus infection (187). The N-terminal amphipathic helix of the NTB is identified as the elicitor of the plant cell death and this cell death is tightly correlated with accumulation of perinuclear aggregates and alteration of the ER network (187). Furthermore, the cell death-inducing activity is also observed when the N-terminal amphipathic helix of tobacco ringspot virus (TRSV, a nepovirus) or CPMV (derived from the NTB domain) is expressed independently in *N. benthamiana* plants (187), indicating that this phenomenon is conserved among these secoviruses.

1.5.2.1.3 Nepovirus membrane proteins

ER-derived aggregation and proliferation of membrane vesicles are observed in plants infected by two nepoviruses: GFLV and ToRSV (11, 188). Some of GFLV-induced membrane vesicles aggregate in the perinuclear region and form rosette-like structures (188). VPg-containing proteins as well as double-stranded RNA replicative intermediate are found associated with these membrane structures, indicating that they are the site for virus replication (188). Similarly to CPMV, the GFLV RNA1 is sufficient to induce morphological change to the intracellular membrane and replicate independently in the protoplast (189, 190). Treatment of the infected cells with cerulenin (an inhibitor of lipid synthesis) and BFA

(an exocytosis inhibitor) significantly reduces the efficiency of virus replication, suggesting that both *de novo* phospholipid biosynthesis and vesicle dependent ER-Golgi transportation are important for GFLV replication (188). The RNA2-encoded protein 2A was also found associated with membrane structures and plays an essential role in RNA2 replication (189). This membrane association of 2A is likely mediated by RNA1-encoded membrane proteins (189). The membrane association domains of GFLV membrane proteins have not been characterized. However, membrane proteins of ToRSV are well-studied (see below section 1.6.3).

1.5.2.2 Genome-linked protein (VPg)

A small protein VPg (approximately 3kDa) is attached to the 5' end of the RNA genome of picorna- and picorna-like viruses. It is the only viral protein that is present inside the virus particle (191, 192). Evidence has shown that PV VPg is linked to the viral genome by forming a phosphodiester bond between the third amino acid tyrosine of VPg and the 5' terminal UTP of viral RNA (193). The linkage occurs during viral RNA synthesis when VPg serves as a protein primer (194). The viral-encoded polymerase 3D Pol directly interacts with VPg and stimulates the attachment of UMP to the VPg, yielding VPgpU as well as VPgpUpU (195), which is present in the PV replication complex and serves as a primer for both positive and negative-strand viral RNA synthesis (196). Generation of VPgpUpU occurs in a template dependent manner and the *cis*-acting RNA element (CRE) prompts the VPg uridylylation (197, 198). Further evidence indicates CRE-dependent VPg uridylylation is required for generation of PV positive-strand RNA, but not negative-strand RNA (199, 200). The cloverleaf structure in the 5' UTR of PV is essential for forming pre-initiation

replication complexes that are capable of synthesis of VPgpUpU and initiation of negative-strand RNA synthesis (201). Intact membranes are also required for uridylylation of VPg and initiation of RNA synthesis at least *in vitro* (202).

Unlike PV, the VPg proteins of the comovirus CPMV and the nepoviruses TRSV and GFLV are covalently linked to the viral genome through a serine residue instead of a tyrosine (203-205). In CPMV, this serine is located at the very N-terminus of the VPg. Substitution of the serine residue with tyrosine or threonine, or relocation of the serine to position 2 or 3 is lethal to virus infectivity (205). Replacement of the CPMV VPg with the cowpea severe mosaic virus (another comovirus) VPg abolishes virus replication (205). Proteinase K treatment of purified virion RNA to remove the VPg abolishes ToRSV infectivity whereas the infectivity of arabis mosaic virus is less affected (192), indicating that the VPg may protect the viral RNA from degradation. For CPMV, the viral infectivity is not influenced by degradation of the VPg (192). The VPg-containing proteins of GFLV, CPMV and ToRSV are detected in close association with the membrane structures active in viral replication (13, 179, 188). It is likely that the como- and nepoviruses VPg proteins are also uridylylated and act as a primer during viral genome replication, although this remains to be determined.

1.5.2.3 Protease

The function of the viral-encoded protease in polyprotein processing was already discussed in section 1.4. In addition, the protease and its precursor are also involved in virus replication. For example, PV 3C and its precursor 3CD interact with the stem loop I of the 5' cloverleaf secondary structure of the genomic RNA and promote RNA synthesis (206, 207).

This RNA binding is mediated by the 3C region and amino acid substitutions in the S1-specificity pocket of 3C result in alteration of both cleavage sites recognition and RNA binding (208). A precursor form of the ToRSV protease that contains both the VPg and protease domains (VPg-Pro) is able to interact with the eukaryotic initiation factor eIF(iso)4E via the N-terminal region of the Pro domain (209). The presence of VPg on the precursor greatly enhanced the binding affinity between the Pro and the eIF(iso)4E. However, the biological function of this interaction is not known (209).

1.5.2.4 Polymerase

All positive-strand RNA viruses encode an RNA-dependent RNA polymerase. It is the key enzyme located in the viral replication complexes and catalyzes the formation of phosphodiester bonds between ribonucleotides using the viral RNA as a template. The three-dimensional structure of a number of viral RNA-dependent RNA polymerases has been determined (210-212). Although they are diverse in their amino acid sequence, except for a few conserved motifs, they share a right hand structure which contains palm, thumb and finger domains. The palm domain is particularly conserved and is important for selection of ribonucleoside triphosphates over dNTPs.

The Pols share several conserved amino acid motifs (213). In the three-dimensional structure, these motifs are spatially juxtaposed to form the catalytic center, which is important for ribonucleoside triphosphates selection, substrate binding and metal ion binding. The Gly-Asp-Asp (GDD) motif, which is involved in Mg^{2+} or Mn^{2+} binding that is required for polymerase activity, is present in the palm domain (214). Structural comparisons showed that

the Pols from plant, animal and bacterial RNA viruses all contain the GDD motif (215). Mutation of this motif in the classical swine fever virus NS5B (Pol) reduces its enzymatic activity (216). Substitution of glycine by alanine in the GDD motif of rubella virus P90-Pol results in impaired virus infectivity while alteration of either aspartic residue completely abolishes virus replication (217). It has also been reported that the GDD mutation of tobacco etch virus polymerase NIb is lethal and can be rescued *in trans* by wild type NIb expression in transgenic plants (218). Pol is an indispensable enzyme not only because of its function in selecting the RNA template, initiating the replication and elongating the template, but also because it regulates the viral genome variability and recombination, and therefore RNA virus evolution (219, 220).

As a key enzyme in virus replication, Pol usually functions in concert with other host and viral factors in the replication complexes. Most viral Pols are soluble, but viral replication takes place in the membrane-bound replication complex (221). To reach the replication site, polymerases directly or indirectly interact with membrane-associated proteins. For example, the BMV 2a (Pol) is recruited to the ER membrane by the viral-encoded multifunctional membrane protein 1a (222). Similarly, the TYMV 140K protein, a membrane protein, directs the 60K (Pol) to the chloroplast through protein-protein interaction (223). There are also examples in which Pol contains the transmembrane domain which directs its association with intracellular membranes. These enzymes include the HCV NS5 and the flock house virus 1A proteins (224, 225). Mutagenesis studies demonstrated that the transmembrane domain is strictly required for HCV genome replication *in vivo* (225).

The three dimensional structure of the PV Pol has been determined. It shows the conserved overall structure with the palm, finger and thumb domain (226). Biochemical and crystal structure studies indicated that extensive polymerase-polymerase interactions take place and form highly ordered oligomeric structures (227, 228). These interactions are mediated by two important regions, named interface I and interface II. Interface I is essential for substrate RNA binding, while interface II is involved in the formation of the catalytic center (227, 229, 230). As mentioned above, 3D catalyzes the incorporation of UMP to VPg and forms the primer VPgpUpU for viral replication (195). Viral membrane protein 3AB plays an important role in modulating the 3D activity. 3AB not only stimulates the release of active polymerase from the precursor, but also stimulates the 3D polymerase activity on the elongation of the template (131, 231). Moreover, interaction between 3AB and 3D plays an important role in recruiting 3D to membrane vesicles (232). Although 3CD does not possess polymerase activity, it actively participates in virus replication in other ways. For example, PV 3CD binds to the 5' cloverleaf structure of the viral RNA and forms a ribonucleoprotein complex with cellular protein PCBP2. This ternary complex was found to be important for viral RNA replication (233, 234). Mutations introduced in the cloverleaf structure that abolish the 3CD interaction greatly impair RNA replication without influencing RNA stability and genome translation (235, 236).

Although the PV polymerase precursor 3CD does not possess polymerase activity (127), the precursor form of other viral polymerases may have enzymatic activity. For example, the feline calicivirus (FCV) Pro-Pol precursor is a predominant protein with both protease and polymerase activity (237). In the case of norovirus (NV), the Pro-Pol precursor and the

mature Pol are both active (238). It has been suggested that different forms of the polymerase may play different roles in viral replication. An *in vitro* assay indicated that the precursor form of polymerase 3CD of the rabbit hemorrhagic disease virus is more efficient in VPg uridylylation, whereas the mature polymerase 3D shows higher activity in elongation of the RNA template (239).

Como- and nepovirus encoded polymerases (85-95 kDa) are bigger than the PV polymerase (53 kDa) (13). However, the PV polymerase still shares all the conserved motifs with como- and nepoviruses since these motifs are present in the N-terminal region (13). Different forms of the polymerase are detected in como- and nepovirus infected plants (13, 176). In CPMV infected plants, the polyprotein Pro-Pol associates with the replication complexes (240). For tomato black ring nepovirus, the VPg-Pro-Pol polyprotein is the only polymerase form detected *in vivo* (241). Further cleavage of this polyprotein was not observed *in vitro* or *in vivo* (241). For another nepovirus ToRSV, the mature Pol (81 kDa) and a truncated version of VPg-Pro-Pol, named VPg-Pro-Pol' (95 kDa, lacking the C-terminal 15 kDa of Pol) were both detected in infected plants (13). Interestingly, this truncation did not occur when processing the full-length VPg-Pro-Pol *in vitro*, indicating that the truncation of the ToRSV VPg-Pro-Pol is a plant specific event (13). In the case of GFLV, only the mature Pol is detected *in planta*. No truncation of Pol was observed either in infected plants or in agroinfiltated plants that transiently express Pol (242). The C-terminal region (the last 136 aa) of GFLV Pol shows significant variability between different strains and this region is involved in eliciting strain-specific symptoms in *N. benthamiana* and *N. clevelandii* plants (242). Interestingly, symptom severity was not correlated with virus titer (242). A putative

RNA recognition motif as well as zinc fingers are located in the C-terminal region of GFLV Pol, indicating that Pol plays multiple roles during virus infection (242) .

1.6 Overview of ToRSV

ToRSV infects small fruits and fruit trees, resulting in significant economic losses in fruit production (243, 244). In nature, ToRSV also infects a number of herbaceous plants and ornamental plants. Infected plants display various symptoms, including necrosis, ringspot, chlorosis, yellowing and dwarfing. Transmission of ToRSV depends on soil-borne nematode vectors in the genus *Xiphinema* (245, 246). However, it can also be disseminated through seeds or pollen. ToRSV is mostly distributed in North America in the regions around the Great Lakes and along the Pacific Coast since the nematode vectors are restricted to these regions and is considered as a quarantined plant virus. However, ToRSV has also been reported in other regions of the world, probably due to the long distance transport of infected plant material.

1.6.1 Classification of ToRSV

As mentioned above, ToRSV belongs to the genus *Nepovirus*, family *Secoviridae* and order *Picornavirales* (10). Most nepoviruses are transmitted by nematode vectors in a semi-persistent manner and the transmission specificity is solely determined by the viral-encoded coat protein (247). Nepoviruses are also transmitted by seeds and /or pollen (10) Nepoviruses have a bipartite positive-sense RNA genome. The two RNA molecules are encapsidated separately into icosahedral particles. Based on the length of RNA2 and serologic characterization, nepoviruses have being divided into three subgroups (10). Subgroup C

nepoviruses (RNA2 6.4-7.3 kb, e.g., ToRSV) have larger RNA2 than both subgroup A (RNA2 3.7 kb-4.0 kb, e.g., GFLV) and subgroup B (RNA2 4.4 kb-4.7 kb e.g., *tomato black ring virus*, TBRV) nepoviruses.

1.6.2 Genomic organization and gene expression of ToRSV

ToRSV is the best characterized subgroup C nepovirus. RNA1 is approximately 8.2 kb in length and RNA2 is about 7.2 kb in length (248, 249). The length of RNA1 and RNA2 slightly varies among isolates (250). Extensive sequence similarity is observed in the 5'UTR and 3' UTR between RNA1 and RNA2, which is probably due to recombination events during virus replication (250, 251). Both RNAs are translated into large polyproteins that are cleaved by the RNA1-encoded protease.

The translation mechanism of ToRSV is currently unknown. However, an IRES was identified in the 5' UTR of blackcurrant reversion virus (BRV, a subgroup C nepovirus). Based on base-pair complementarity with 18S rRNA, it was suggested that the IRES interacts directly with the 18S rRNA. Mutations that disrupt the base pair complementarity drastically decrease translation (252). In the ToRSV 5' UTR, there are also stretches of possible complementarity to 18S rRNA, suggesting that ToRSV RNAs are translated by an IRES-dependent mechanism that requires viral RNA-rRNA interaction between the 5' IRES and 18S rRNA (252).

RNA1 is translated into a 244 kDa polyprotein, and RNA2 is translated to produce the 207 kDa polyprotein. Both large polyproteins are processed by the viral serine-like protease

which is related to picornavirus 3C protease (253). The ToRSV Pro specifically cleaves at either Q/G (Gln /Gly) or Q/S (Gln/Ser) dipeptides. The histidine in the substrate binding pocket directly interacts with Gln (Q) in the cleavage site and determines the specificity of the cleavage. Mutation of the His inactivate the protease activity (253).

In vitro processing assays identified five cleavage sites in RNA1 encoded polyprotein and three cleavage sites in RNA2 encoded polyprotein. Therefore, six distinct protein domains are delineated on RNA1: X1, X2, NTB, VPg, Pro and Pol (254) (Fig. 1.7). The function of X1 is currently not clear. All other RNA1-encoded protein domains are involved in replication of the ToRSV genome (254). Four protein domains are identified on RNA2: X3, X4, MP and CP (Fig. 1.7). The function of X3 has not been investigated yet. However, in analogy with the corresponding protein of GFLV 2A (a subgroup A nepovirus), it is likely that X3 plays a role in replication of the RNA2 (189). X4 is a unique protein that does not share any similarity to other known proteins in the databases (255). The CP is involved in virus particle assembly and the MP is responsible for forming the tubular structures that assist virus particles in traversing the plasmodesmata (85, 254). The GFLV RNA1 can replicate independently in the protoplast. However, successful plant infection requires the presence of both RNA1 and RNA2, probably because the newly formed RNAs are unable to assemble and spread to adjacent cells without RNA2-encoded proteins (190).

In ToRSV infected plants, the mature NTB and Pol are detected along with the NTB-VPg, X2-NTB-VPg and VPg-Pro-Pol′ polyproteins (11, 13) (Fig. 1.7). Although different forms of NTB are detected, the mature NTB is unlikely to be released from the NTB-VPg and X2-

NTB-VPg polyproteins since RNA1 encoded polyprotein can only be processed *in cis.* Therefore these proteins are probably produced by alternative processing pathways (88, 256). NTB containing proteins are localized in membrane-enriched fractions while the mature Pol is only present in soluble fractions (11, 13). It is interesting to note that two populations of VPg-Pro-Pol′ are present in ToRSV infected plants. One population is in the soluble fraction and the other population is peripherally associated with the ER membranes (13).

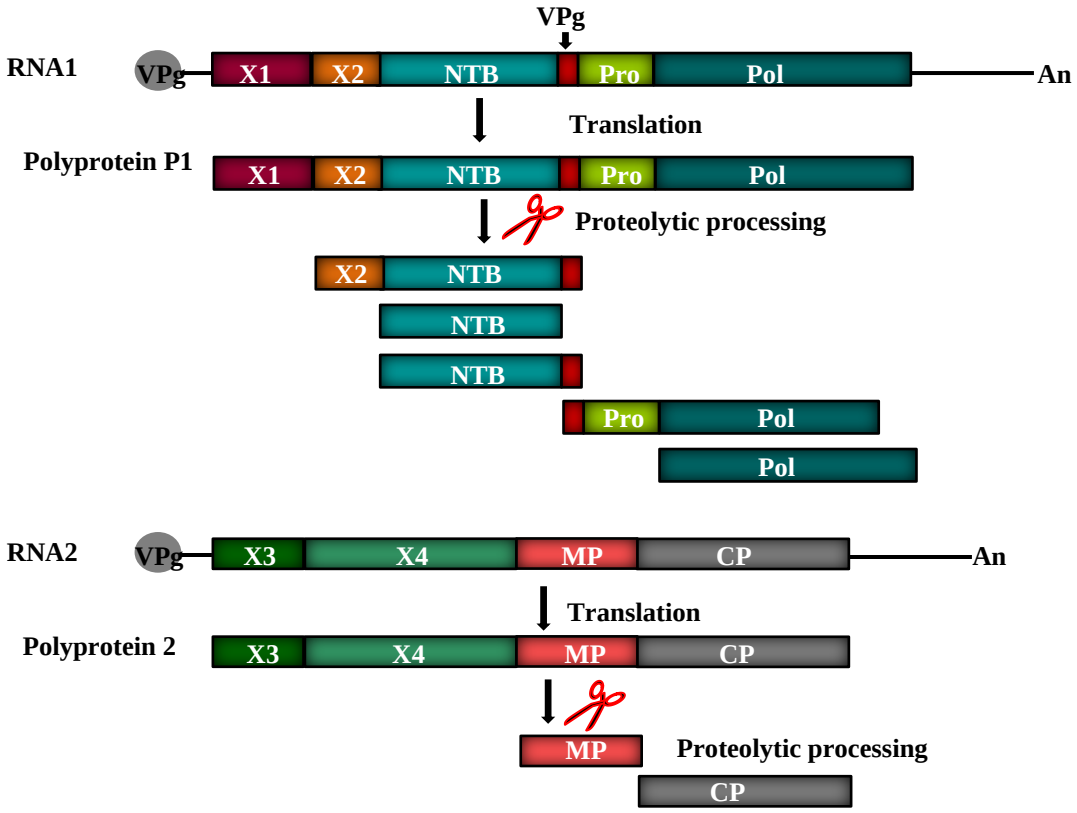


Figure 1.7: Genomic organization of ToRSV and viral proteins detected in ToRSV infected plants. RNA1 encodes six proteins: X1, X2, nucleoside triphosphate binding protein (NTB), viral protein genome-linked (VPg), protease (Pro), polymerase (Pol). RNA2 encodes four proteins: X3, X4, movement protein (MP) and coat protein (CP). Mature proteins and polyproteins detected in infected plants using antibodies against NTB, VPg, Pro, Pol, MP and CP domains are shown below each polyprotein. Adapted with permission from (Sanfacon H, Zhang G, Chisholm J, Jafarpour B, Jovel J. 2006. Molecular biology of Tomato ringspot nepovirus, a pathogen of ornamentals, small fruits and fruit trees, p. 540-546. In Teixeira da Silva J (ed.), Floriculture, Ornamental and Plant Biotechnology: Advances and Topical Issues (1st Edition), vol. III. Global Science Books, London, UK.). Crown copyright (2006).

1.6.3 Replication of ToRSV

Replication of comovirus and nepovirus is associated with ER-derived membrane vesicles (11, 140, 176, 177, 188). Extensive membrane proliferation and perinuclear ER aggregates are observed in ToRSV-infected cells. NTB, NTB-VPg and X2-NTB-VPg are present in membrane fractions derived from the rough ER (11). Detergent treatment is required to dissociate the proteins from the membranes, suggesting that they are integral membrane proteins. Active RNA synthesis co-fractionationates with NTB-containing proteins and the ER membrane (11). Both X2 and NTB directly interact with the ER membrane when expressed ectopically and have been proposed to act as membrane anchors of the replication complexes (12, 140, 257, 258).

1.6.3.1 Two membrane proteins: X2 and NTB-VPg

X2 is a highly hydrophobic protein with an N-terminal amphipathic helix and two C-terminal transmembrane domains (258). These motifs are also observed in the X2 proteins of other nepoviruses (140). Both C-terminal transmembrane domains can direct the fused GFP to ER membrane independently (258). *In vitro* glycosylation mapping revealed that the two transmembrane domains traverse the membrane and form a hairpin-like structure, leading to a cytosolic localization of the C-terminus of the X2 protein (258) (Fig. 1.8). The N-terminal amphipathic helix inserts parallel to the ER membrane initially with its hydrophobic side associated with the lipid bilayer and its hydrophilic side facing the cytosol. However, glycosylation occurs when an N-glycosylation site is introduced at the N-terminus of X2, indicating that the N-terminus of the X2 protein could also be translocated into the luminal side of the ER by the amphipathic helix (258). Oligomerization of X2 is also observed, suggesting strong hydrophobic interactions occurring among X2 molecules (258). Therefore,

it is possible that the N-terminal amphipathic helices of different X2 molecules traverse the membrane and form an aqueous pore, directing the N-terminus of X2 into the ER lumen, whereas the C-terminal transmembrane domains traverse the membrane in a hairpin-like structure and translocate the C-terminus of the protein back into the cytosol (140) (Fig. 1.8).

NTB-VPg associates with ER-derived membranes *in vitro* and *in vivo* (12, 257). This membrane association is mediated by three domains: an N-terminal amphipathic helix and two C-terminal predicted hydrophobic helices (Fig. 1.8). The N-terminal amphipathic helix is predicted to orient itself horizontally on the membrane with the hydrophobic side facing the membrane and the hydrophilic side towards the cytosol (257). However, *in vivo* glycosylation mapping and mutational analysis indicated that the N-terminal amphipathic helix can also traverse the membrane and direct the N-terminus of NTB into the luminal side of ER (257). Proteinase K treatment indicated that the C-terminal region of NTB as well as the entire VPg domain is embedded in the ER lumen (11). An N-glycosylation site, which is located in the VPg region, is glycosylated when the C-terminal region of NTB-VPg (cNTB-VPg, including the two C-terminal transmembrane domains and the VPg domain) was expressed *in vitro* or *in vivo* (12, 257). This result is consistent with the proteinase K protection assay and therefore confirms the luminal localization of the VPg. Glycosylation mapping indicated that the first C-terminal transmembrane domain is highly hydrophobic and traverses the membrane efficiently while the second only shows weak membrane traversing capability (257). Therefore, two different membrane topologies were predicted for NTB-VPg (Fig. 1.8). In the first topology, the N-terminal amphipathic helix associates with the cytosolic side of the membrane via its hydrophobic side. The first C-terminal

transmembrane domain transverses the membrane and directs the downstream hydrophobic domain and the VPg into the ER lumen, where glycosylation occurs. This is considered as the predominant topology. In the second topology, both C-terminal transmembrane domains traverse the membrane and form a hairpin-like structure, translocating the VPg back into the cytosol (140). However, this topology still needs to be confirmed experimentally. Oligomers of NTB-VPg are also observed. In this case, it is likely that the amphipathic helices from different NTB-VPg molecules interact with each other and traverse the lipid bilayer by forming an aqueous pore (140).

Interestingly, a signal peptidase cleavage occurs when the C-terminal half of the NTB-VPg protein (cNTB-VPg) is expressed *in vitro* in the presence of canine microsomal membranes (ER-enriched membranes) (12). The putative cleavage site is between the first C-terminal transmembrane domain and the NTB-VPg junction based on the size of the N-terminal and the C-terminal cleavage products (12) (Fig. 1.8B). The exact cleavage site and the biological function of this cleavage are still unknown. The cleavage may influence the conformation of the protein, the architecture of the replication complex and/or the rate of virus replication. Interestingly, a signal peptidase cleavage site was also predicted in the C-terminal region of NTB from other nepoviruses, such as TBRV, GFLV and grapevine chrome mosaic virus using the SignalP program (a software that is designed to predict signal peptidase cleavage) (12), indicating that this membrane-associated cleavage might be conserved among nepoviruses.

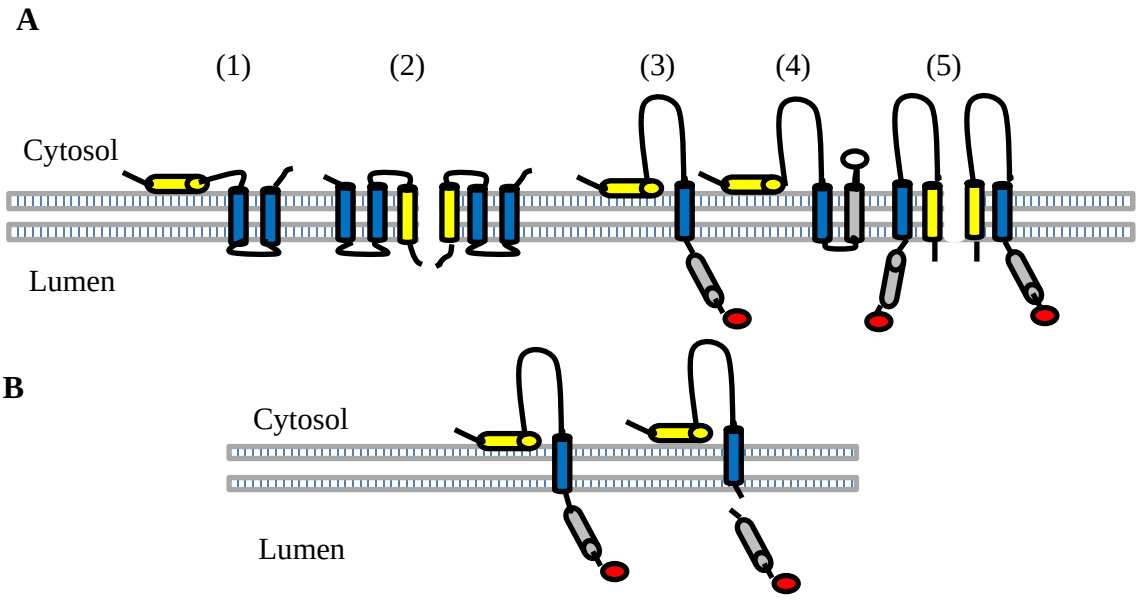


Figure 1.8: Topological models of ToRSV X2 and NTB-VPg protein. A. Topology of X2 and NTB-VPg. (1) Topology of X2 monomers with the amphipathic helix shown in yellow and the transmembrane helices shown in blue. (2) Topology of X2 oligomer. Only two molecules are shown in the figure for simplicity, however, at least four molecules would be required to form an aqueous pore (3) Dominant topology of NTB-VPg monomers. Luminal orientation of the VPg results in its glycosylation (shown with the red circle). (4) Alternative topology of NTB-VPg. Absence of glycosylation in the cytosolic oriented VPg domain is depicted by the empty circle. (5) Topology of the NTB-VPg oligomers. As for X2, only two molecules are shown. B. Signal peptidase processing of NTB-VPg in the luminal side of the ER membrane. This cleavage is inefficient and both uncleaved and cleaved proteins are detected *in vitro*. See text for more details. Adapted with permission from (Sanfacon H. 2013. Investigating the role of viral integral membrane proteins in promoting the assembly of nepovirus and comovirus replication factories. Frontiers in Plant Science 3:313.). Crown copyright (2013).

1.6.3.2 Polymerase and VPg

Two forms of polymerase are detected in ToRSV infected plants: the mature Pol is detected in low amounts if at all, indicating the instability of the protein, whereas a truncated polyprotein VPg-Pro-Pol′ is a stable intermediate that accumulates in plants (13). Whether one or both forms of the polymerase catalyze genome replication still needs to be investigated. It is also possible that the two forms of the polymerase play different roles in virus replication, or function at different stages of the genome replication. As mentioned previously, a sub-population of VPg-Pro-Pol′ is peripherally associated with ER-derived membranes in ToRSV infected plants (13). The fact that the VPg, Pro and Pol mature proteins are soluble when expressed individually suggests that the VPg-Pro-Pol′ is probably directed to the ER through direct or indirect interaction with viral membrane proteins (13). Another possibility is that the VPg-Pro-Pol′ is targeted to membranes as a polyprotein that contains the NTB and /or X2 domain and is released from the polyprotein after membrane association. Active viral replication was detected in membrane fractions that are rich in viral membrane proteins (11).

Similar to picornaviruses, the replication of ToRSV genomic RNA is likely primed by uridylylated VPg (254). In comoviruses, the VPg-Pro-Pol polyprotein acts as a VPg donor *in vivo* (259). For ToRSV, two polyproteins (NTB-VPg and VPg-Pro-Pol′) that contain the VPg domain are detected in infected plants (13). For NTB-VPg, the predominant topology predicted implies that the VPg is translocated into the ER lumen (Fig. 1.8), in addition, NTB-VPg is not cleaved *in trans* by the protease. Therefore it is unlikely to act as a donor for the VPg primer, since the replication takes place on the cytosolic side of the ER membrane. The

VPg is also present in the form of VPg-Pro-Pol′. This polyprotein is present in the cytosol or is peripherally located on the cytosolic side of the ER membrane (13), and likely acts as donor of VPg primer during virus replication.

1.6.4 ToRSV isolates

Several ToRSV isolates from different geographical areas and host plants have been reported. The Rasp2 isolate is a raspberry isolate from the Lower Mainland in B.C. Canada. The Rasp1 isolate is also originally from raspberries, cultivated in Washington State. The PYB isolate (peach yellow bud) is originally from infected peach orchards in California, whereas the GYV isolate (grape yellow vein) is a mild isolate that was discovered in California. Symptom and virulence varies among the isolates. At lower temperature (21°C), Rasp1 infected *N. benthamiana* plants show systemic necrosis leading to death at about 20 days. However, at a higher temperature (27 °C), the newly emerged upper leaves are symptom free and plants recover from infection due to active RNA silencing at this temperature (260). GYV is a mild isolate and plants recover from infection at both 21°C and 27 °C. At the time this thesis was initiated, the entire genome sequence was only available for the Rasp2 isolate (248, 249).

1.7 Thesis objectives

The molecular biology of the ToRSV has been well studied in the last decade. The genome organization, polyprotein processing and subcellular localization of viral proteins have been characterized (254). Viral-encoded membrane anchor proteins have been identified and a putative signal peptidase cleavage has been detected in the C-terminal region of NTB-VPg from the Rasp2 isolate. (12, 140, 257, 258). However, the molecular mechanisms of virus

replication and the biological function of the putative signal peptidase cleavage of the membrane anchor protein are unknown. In addition, the function of various forms of ToRSV polymerase and their relative stability are still not well understood. My starting hypotheses were as follows: (1) Signal peptidase cleavage in the viral encoded membrane protein is likely a highly regulated processing event that may result in changes in the architecture and stability of replication complexes and could affect the viral replication rate. (2) Regulating the stability and activity of polymerase is a sophisticated way for viruses to modulate their replication and prevent catastrophic damage to their hosts. The overall goal of this thesis was to investigate the stability of various forms of viral replication proteins and their role in virus replication. The specific objectives of this thesis were as follows:

(1) Determine the sequence of replication proteins from different ToRSV isolates (see Chapter 2).

(2) Examine the specificity of signal peptidase cleavage in the NTB-VPg protein of various ToRSV isolates and map the cleavage site (see Chapter 2).

(3) Investigate the stability of the various forms of the polymerase (Pol and VPg-Pro-Pol') (see Chapter 3)

(4) Identify the active form of the ToRSV polymerase (see Chapter 4)

Chapter 2

Membrane-associated glycosylation and signal peptidase processing of the cNTB-VPg protein of ToRSV

2.1. Introduction

The ToRSV NTB and the NTB-VPg polyprotein are integral membrane proteins associated with the ER-bound replication complexes (11). N-linked glycosylation and signal peptidase processing of an N-terminally truncated version of the NTB-VPg (cNTB-VPg, cNV) protein derived from the Rasp2 isolate have been observed *in vitro* in the presence of canine microsomal membranes but were not well characterized (12). In this current study, I investigated *in vitro* membrane-associated modification and processing of the cNV truncated protein derived from different ToRSV isolates. I also used a smaller fragment of cNV that is truncated in the N-terminal region (cNV2) to examine membrane-associated modifications *in vivo*. N-linked glycosylation and signal peptidase cleavage were detected for the Rasp2, PYB and GYV isolates, but not for the Rasp1 isolate. Inhibitor treatment confirmed the signal peptidase processing. Mutational analysis allowed identification of the putative signal peptidase cleavage site. Amino acid residues that influence the efficiency of this cleavage were also identified. The potential biological significance of these modifications is discussed.

2.2. Materials and methods

2.2.1 Sequencing of ToRSV NTB-VPg

The NTB-VPg region of the Rasp1 and GYV isolates were sequenced as described (250). The nucleotide sequences were deposited into the GenBank under accession numbers KM083894 and KM083892, respectively.

2.2.2 Computer-based analysis of NTB-VPg

NTB-VPg from different ToRSV isolates was analyzed by computer-based software. Multiple sequence alignment was performed using the online ClustalW2 software. Prediction of transmembrane domains and N-linked glycosylation were performed using the HMMTOP software (261) and the NetNGlyc 1.0 Server, respectively. Secondary structure prediction of cNV was performed using softwares DSC (262), MLRC (263) and PHD (264). The NTB-VPg sequences used in this chapter can be found under the following GenBank accession numbers: Rasp1 (KM083894), Rasp2 (NC_003840), PYB (KM083890) and GYV (KM083892).

2.2.3 Plasmid constructions

The plasmids pCITE-cNV, pCITE-cNV (T_{610}-A) and pCITE-cNV (ΔTM3) derived from the Rasp2 isolate have been described previously (12). The corresponding cNV fragment from the GYV and PYB isolates were amplified using oligos ToRSV295 and ToRSV300, ToRSV345 and ToRSV346, respectively (Table 2.1). The PCR products were digested with *BamH*I and inserted into the *BamH*I site of pCITE-4a (+) (Novagen) for *in vitro* expression. The Rasp1 cNV fragment was amplified using oligos ToRSV302 and ToRSV307 and the amplified fragment was digested with *Bgl*II and inserted into the *BamH*I site of pCITE-4a (+).

The cNV2 fragments (Fig. 2.1) from different ToRSV isolates were cloned into pBIN (+) for expression in *N. benthamiana*. The above cNV constructs were used as templates for PCR amplification. The cNV2 fragments of Rasp2, GYV and Rasp1 were amplified using primers

ToRSV320 and ToRSV318, ToRSV352 and ToRSV353, ToRSV315 and ToRSV319, respectively. An N-terminal Flag tag and a C-terminal HA tag were fused in frame to the cNV2 fragments to allow easier detection of full-length and the modified proteins. All fragments were digested with *Nco*I and *Xba*I and subsequently inserted into the corresponding sites of the intermediate vector pBBI525 (265). For Rasp2 and GYV, the cassettes including the duplicated 35S promoter, cNV2 sequence and the NOS terminator were digested with *EcoR*I and *Hind*III and inserted into the corresponding sites of pBIN (+). For Rasp1, *EcoR*I and *Xho*I sites were used to digest the fragment from the intermediate vector and then were cloned into pBIN (+). FLAG-cNV2-HA constructs were also digested with *Nco*I and *BamH*I from the pBBI525 vector and transferred into pCITE-4a (+) for expression *in vitro*. All constructs were confirmed by sequencing the entire inserted region.

Table 2.1 Primers used in cNTB-VPg plasmid constructions

Primer name	Sequence of the Primer
ToRSV295R	5’-TCTCGGGGATCCTACTGTACAGATTGTGGGCGGAAAACGCGTG-3’
ToRSV300F	5’-TGCGTTGGATCCGAATTAAGTGCTGAGTTGTTGCTGC-3’
ToRSV302R	5’-TCTCGGAGATCTACTGTACAGATTGCGGCCTGAAAACGCGAG-3’
ToRSV307F	5’-TGCGTTAGATCTGAGATGAGTGCTGAGTTATTGCTTAGG-3’
ToRSV315F	5’-CTAGTTCCATG**GACTACAAGGACGATGATGACAAG** GAAGCCCTTTCAAAGGATTCCTTGGAG-3’
ToRSV318R	5’-TGCTTTCTAGATTA**GGCATAGTCAGGAACATCGTATGGGTA** CTGTACAGATTGTGGGCGGA-3’
ToRSV319R	5’-TGCTTTCTAGATTA**GGCATAGTCAGGAACATCGTATGGGTA**CTG TACAGATTGCGGCCTGA-3’
ToRSV320F	5’-CTAGTTCCATG**GACTACAAGGACGATGATGACAAG**GAGGTTTT ATCTAAGGACTCCCTCGAG-3’
ToRSV345F	5’-TGCGTTGGATCCGAACTTAGTGCCGAGCTCA-3’
ToRSV346R	5’-TCTCGGGATCCTACTGCACAGACTGAGGCCT-3’
ToRSV352F	5’-CTAGTTCCATG**GACTACAAGGACGATGATGACAAG**GAAGCTCTT TCCAAGGAATCTTTG-3’
ToRSV353R	5’-TGCTTTCTAGATTA**GGCATAGTCAGGAACATCGTATGGGTA**CTGC ACAGACTGAGGCCT-3’

The restriction enzyme sites in primers are underlined, the start codon and the stop codon are shown in red. The nucleotide sequences of the Flag and HA tags are shown in bold.

2.2.4 Site-directed mutagenesis

A series of mutations were performed according to the protocol of QuickChange II-E Site-Directed Mutagenesis (Stratagene, La Jolla, Ca, USA.). Complementary primers that contain the desired mutations were designed. The wild type Rasp2 pCITE-cNV and Rasp1 pCITE-cNV were used as templates for initial mutations. The Rasp1 pCITE-cNV (A_{608}-T) was used as template for introducing double mutations to the Rasp1 construct. All constructs were verified by sequencing to make sure that there are no unintended site mutations. Primers used for mutagenesis were PAGE purified and the sequences of the primers are listed in Table 2.2.

2.2.5 *In vitro* translation assays

In vitro transcription and translation assays were performed in the presence or in the absence of canine pancreas microsomal membranes (Promega) as previously described (12).

2.2.6 Agroinfiltration of *N. benthamiana* plants

The pBIN (+) constructs containing wild type cNV2 and glycosylation site mutants were transformed into *Agrobacterium tumefaciens* LBA4404. Agroinfiltration of *N. benthamiana* plants was conducted as previously described (257). Samples were collected 3 days post-agroinfiltration (dpa) for protein extraction.

Table 2.2 Primers used for mutagenesis of cNTB-VPg

Primer name	Sequence of the Primer
P108F (Rasp2 K_{567}-T)	5'- CTTTGGGTCTGCTTGTATAACGTTGATGCAGGCCATTTTTTGTG-3'
P109R (Rasp2 K_{567}-T)	5'- CACAAAAAATGGCCTGCATCAACGTTATACAAGCAGACCCAAAG-3'
P110F (Rasp2 L_{568}-A)	5'- GGGTCTGCTTGTATAAAGGCGATGCAGGCCATTTTTTGTGG-3'
P111R (Rasp2 L_{568}-A)	5'-CCACAAAAAATGGCCTGCATCGCCTTTATACAAGCAGACCC -3'
P112F (Rasp2 Q_{570}-R)	5'-GCTTGTATAAAGTTGATGCGGGCCATTTTTTGTGGTGCC-3'
P113R (Rasp2 Q_{570}-R)	5'-GGCACCACAAAAAATGGCCCGCATCAACTTTATACAAGC-3'
P114F (Rasp2 G_{578}-A)	5'-TTTTTTGTGGTGCCGCAGCTGGTACTGTCAGTATGGCTG -3'
P115R (Rasp2 G_{578}-A)	5'-CAGCCATACTGACAGTACCAGCTGCGGCACCACAAAAAA -3'
P116F (Rasp2 $\Delta TV_{580-581}$)	5'-GTGGTGCCGCAGGTGGTA--GTATGGCTGCTGTCGGG-3'
P117R (Rasp2 $\Delta TV_{580-581}$)	5'-CCCGACAGCAGCCATAC--TACCACCTGCGGCACCAC -3'
P120F (Rasp2 I_{572}-V)	5'-GTATAAAGTTGATGCAGGCCGTTTTTTGTGGTGCCGCAGG-3'
P121R (Rasp2 I_{572}-V)	5'-CCTGCGGCACCACAAAAAACGGCCTGCATCAACTTTATAC-3'
P122F (Rasp2 A_{576}-S)	5'-GCAGGCCATTTTTTGTGGTTCCGCAGGTGGTACTGTCAGTA-3'
P123R (Rasp2 A_{576}-S)	5'-TACTGACAGTACCACCTGCGGAACCACAAAAAATGGCCTGC-3'
P124F [Rasp2 A_{576}-S (G_{578}-A)]	5'-GCAGGCCATTTTTTGTGGTTCCGCAGCTGGTACTGTCAG -3'
P125R [Rasp2 A_{576}-S (G_{578}-A)]	5'-CTGACAGTACCAGCTGCGGAACCACAAAAAATGGCCTGC-3'
P126F (Rasp2 M_{583}-A)	5'-CGCAGGTGGTACTGTCAGTGCGGCTGCTGTCGGGAAAATG -3'
P127R (Rasp2 M_{583}-A)	5'-CATTTTCCCGACAGCAGCCGCACTGACAGTACCACCTGCG -3'
P128F (Rasp2 A_{584}-V)	5'-CAGGTGGTACTGTCAGTATGGTTGCTGTCGGGAAAATGACCG -3'
P129R (Rasp2 A_{584}-V)	5'-CGGTCATTTTCCCGACAGCAACCATACTGACAGTACCACCTG -3'
P75F (Rasp1 A_{608}-T)	5'-TAATGCTCGTAACATGACACGCGTTTTCAGGCCGC-3'
P76R (Rasp1 A_{608}-T)	5'-GCGGCCTGAAAACGCGTGTCATGTTACGAGCATTA-3'
P134F (Rasp1 R_{570}-Q)	5'-GTGCGTGTGTTACTGCAATGCAGGCAGTTTTCTGTGGCTCTGC-3'
P135R (Rasp1 R_{570}-Q)	5'-GCAGAGCCACAGAAAACTGCCTGCATTGCAGTAACACACGCAC-3'
P136F (Rasp1 $+TV_{580-581}$)	5'-GTGGCTCTGCTGCAGGAACTGTCAGTGCAGTCGCTGTCGG-3'
P137R (Rasp1 $+TV_{580-581}$)	5'-CCGACAGCGACTGCACTGACAGTTCCTGCAGCAGAGCCAC -3'
P138F (Rasp1 S_{576}-A)	5'-GCGTGCAGTTTTCTGTGGCGCTGCTGCAGGAAGTGCAG-3'
P139R (Rasp1 S_{576}-A)	5'-CTGCACTTCCTGCAGCAGCGCCACAGAAAACTGCACGC-3'

Nucleotides that encode the mutated amino acids are underlined.

2.2.7 Protein extraction and western blot

Protein extraction and western blot analyses were performed as described (11). Monoclonal ANTI-FLAG® M2 antibody (Sigma-Aldrich) and monoclonal Anti-HA−Peroxidase antibody (Sigma-Aldrich) were used as described by the manufacturer.

2.2.8 Inhibitor treatment

1.2 mM signal peptidase inhibitor (MeOSuc-Ala-Ala-Pro-Val chloromethyl ketone; Sigma) was added to the *in vitro* transcription and translation assay in the presence of microsomal membranes. Reactions were stopped at two hours and four hours by adding 2 × protein loading buffer. Cycloheximide treatment (30 ug/ul) of agroinfiltrated leaves was performed as previously described (266). Samples were collected two hours after treatment and used for protein analysis.

2.2.9 Virus-induced gene silencing (VIGS) of *N. benthamiana* signal peptidase

A 356 bp fragment of *N. benthamiana* signal peptidase catalytic subunit was amplified by RT-PCR with primers VIGSSPase-1 (5'-CTATGGATCCAAGTATGGAACCTGGC - 3', corresponds to nts 238-256 of *Solanum lycopersicum* signal peptidase complex catalytic subunit, XM_004251517) and VIGSSPase-2 (5'-CTATGAATTCCCAACAATCCCAGCGC ACC-3', corresponds to nts 595-577 of *Solanum lycopersicum*, XM_004251517). The internal *BamH*1 site was mutated using primers VIGSSPase-3 (5'-CATATGAGTAAGGCT CCTATTCGTGCAG-3') and VIGSSPase-4 (5'-CTGCACGAATAGGAGCCTTACTCAT ATG-3'). The fragment was digested with *EcoR*1 and *BamH*1 and inserted into the tobacco rattle virus vector TRV:00 that corresponds to TRV RNA2 (267). The construct TRV-SPase

was confirmed by sequencing, and the plasmid was transformed into *A. tumefaciens* strain GV3101 for expression in *N. benthamiana* plants in combination with TRV RNA1 construct pBINTRA6 (267).

2.3 Results

2.3.1 Sequence alignment of NTB-VPg from different ToRSV isolates

The NTB-VPg region of ToRSV Rasp1 and GYV isolates was sequenced and the deduced amino acid sequences were aligned with available sequences from the Rasp2 and PYB isolates using the online ClustalW2 software. Transmembrane domains and N-linked glycosylation sites were predicted as described in Materials and Methods. NTB binding motifs, which are thought to play an important role in the helicase activity, were identical among all isolates (Fig. 2.1). Three membrane association domains have been identified in the ToRSV NTB regions (11, 12, 257). The N-terminal putative amphipathic helix (TM1) was strongly conserved among all isolates. The strongly predicted C-terminal hydrophobic domain (TM2) was also relatively conserved except for an I_{558} to V (the number indicates the amino acid position in the NTB-VPg polyprotein with 1 being the first amino acid of the NTB domain) substitution in the PYB isolate and an F_{561} to L substitution in the Rasp1 isolate. These mutations are not predicted to change the overall topology of this transmembrane domain. In contrast, the downstream region of Rasp1 NTB-VPg exhibited remarkable differences when compared to other isolates. I_{566} to V, K_{567} to T and L_{568} to A substitutions were observed. In addition, several mutations (G_{578}-A, M_{583}-A and A_{584}-V) and deletions ($\Delta TV_{580\text{-}581}$) in the weakly predicted C-terminal hydrophobic domain (TM3) of Rasp1 decreased the hydrophobicity of this domain and as a result, it was not predicted as a

transmembrane domain anymore (Fig. 2.1). The N-linked glycosylation site (NMT) in the VPg region was conserved in the Rasp2, PYB and GYV isolates, but was replaced by NMA in the Rasp1 isolate, which is not predicted as an N-glycosylation site. In summary, a number of amino acid substitutions were observed in the C-terminal region of Rasp1 cNTB-VPg, which may alter the conformation of NTB-VPg in the membrane and influence the membrane-associated modifications of the protein.

Putative amphipathic helix (TM1)

```
Rasp2           GLTDVFGVPLSIMNAIGDGLVHHSLDTLTLMGKFGAAMDNVRKGITCMRSFVSWLMEHLA 60
PYB             GLTDVFGVPLSIMNAIGDGLVHHSLDTLTLMGKFGAAMDNVRKGITCMRSFVSWLMEHLA 60
Rasp1           GLTDVFGVPLSIMNALGDGLVHHSLDTLQLMGKFGAAMDNVRKGITCMRSFVSWLMEHLA 60
GYV             GINDVFGVPLSIMNALGDGLVHHSLDTLQMMGKFGAAMDNVRKGITCMRSFVSWLMEHLA 60
                *:.***********:*********** :******************************

Rasp2           LALDKITGKRTSFFRELATLINFDVEKWVRDSQQYLLAAEIYVDGDTVVMDTCRHLLDKG 120
PYB             LALDKITGKRTSFFRELATLINFDVEKWVRDSQQYLLAAEIYVDGDTVVMDTCRHLLDKG 120
Rasp1           LALDKITGKRTAFFRELATLINFDVEKWVRDSQQYLLAAEIYVDGDTVVMDTCRHLLDKG 120
GYV             LALDKITGKRTAFFRELATLINFDVEKWVRDSQQYLLTAEIYADGDAIIMDTCRHLLDKG 120
                ***********:*************************.***:::***********

                                                        NTB-binding motifs
Rasp2           LKLQRMMVSAKSGCSFNYGRLVGDLVKRLSDLHKRYCASGRRVHYRLAPFWVYLYGGPRC 180
PYB             LKLQRMMVSSKSGCSFNYGRLVGDLVKRLSDLHKRYCASGRRVHYRLAPFWVYLYGGPRC 180
Rasp1           LKLQRMMVSSKSGTSFNYGRLVGDLVKRLSDLHKRYCASGRRVHYRLAPYWVYLYGGPRC 180
GYV             TKLQRMMVSSKSGTSFNYVRLVSDLVKRLGDLHKRYQVTGRRVHYRLAPFWVYLYGGPGC 180
                 ********:*** **** ***.******.****** .:**********:******** *

Rasp2           GKSLFAQSFMNAAVDFMGTTVDNCYFKNARDDFWSGYRQEAICCVDDLSSCETQPSIESE 240
PYB             GKSLFAQSFMNAAVDFMGTTVDNCYFKNARDDFWSGYRQEAICCVDDLSSCETQPSIESE 240
Rasp1           GKSLFAQSFMNTAVDFMGTTTDNCYFKNARDDFWSGYRQEAICCVDDLSSCETQPSIESE 240
GYV             GKSIFAQSFMNAAVDFMGTTVDNCYFKNARDDFWSGYRQEAICCVDDLSACETQPSIESE 240
                ***:*******:********.*****************************:**********

Rasp2           FIQLITTMRYGLNMAGVEEKGASFDSKMVITTSNFFTAPTTAKIASKAAYNDRRHACILV 300
PYB             FIQLITTMRYGLNMAGVEEKGASFDSKMVITTSNFFTAPTTAKIASKAAYNDRRHACILV 300
Rasp1           FIQLITTMRYGLNMAGVEEKGAQFNSKMVITTSNFFTAPTTAKIADKAAYNHRRHACVLV 300
GYV             FIQLITTMRYGLNMAGVEEKGAQFDSKMVITTSNFFTAPTTAKIADMSAYNRRRHACVLV 300
                **********************.*:*********************. :*** *****:**

                                                    ↓Start of cNTB-VPg (cNV)
Rasp2           QRKEGVAYNPSDPAAAAEAMFVDSTTQHPLSEWMSMQELSAELLLRYQQHREAQHAEYSY 360
PYB             QRKEGVVYNPSDPAASAEAMFVDSTTQHPLSEWMSMQELSAELLLRYQQHREAQHAEYSY 360
Rasp1           QRKEGVKYDPSNPAAAAEAMFVDNETQHPLSEWMNMQEMSAELLLRYQQHRETQHAEYSY 360
GYV             QKKKGVEYDPSNPSAAAEAMFVDSHTQHPLSEWMSMSELSAELILKYQQHREKQYAEYKY 360
                *:*:** *:**:*:*:*******. *********.*.*:****:*:****** *:***.*

Rasp2           WKSTSRTSHDVFDILQKCVNGDTQWLSLPVDVIPPSIRQKHKGNRVFAIDGRIFMFDYMT 420
PYB             WKSTSRTSHDVFDILQKCVNGDTQWLSLPVDVIPPSIRQKHKGNRVFAIDGRIFMFDYMT 420
Rasp1           WKSTSRSSHDVFDILQKCVDGDVHWLSLPIDVIPPTIRLKYKGNRVFAIDGRTFIFDYMT 420
GYV             WKSTTCVSHDVFDILRKCVDGDTSWLSLPVDVIPPSIRQKYKGNRVFAIDGRLFIFDYLT 420
                ****:  ********:***:**. *****:*****:** *:*********** *:***:*

Rasp2           LEYDEIKEKENLDARHLEARILEKYGDTRLLLEKWGANGVVAQFIEQLLEGPSNVASLEV 480
PYB             LEYDEIKEKENLDARHLEARILEKYGDTRLLLEKWGANGVVAQFIEQLLEGPSNVASLEV 480
Rasp1           LECEEIKEKSEIDARHLESRILEKYGDTRLLLEKWGANGVVAQFIEQLVEGPSNVASMEA 480
GYV             LECEEVFERENLDVRHLESRILEKYGDTRLLLEKWGANGIVAQFIEQLVEGPSNVTSMEA 480
                ** :*: *:.::*.****:*********************:********:******:*:*.

                    ↓ start of cNV2 constructs
Rasp2           LSKDSLESHKEFFSTLGLIERATLRAVQKKIDAAREDLMHLSGLKPGRSLTELFVEAYDW 540
PYB             LSKDSLESHKEFFSTLGLIERATLRAVQKKIDAAREDLMHLSGLKPGRSLTELFVEAYDW 540
Rasp1           LSKDSLESHKEFFSTLGLIERATLRAVQKKIDSAREDLSSFAGLKPGRSLAELFVEAYDW 540
GYV             LSKESLASHKEFFSTLGVIERATLRAVQKKIDIAREDLSSFSGIKPGRSLAELFVEAYDW 540
                ***:** **********:************** *****  ::*:******:*********
```

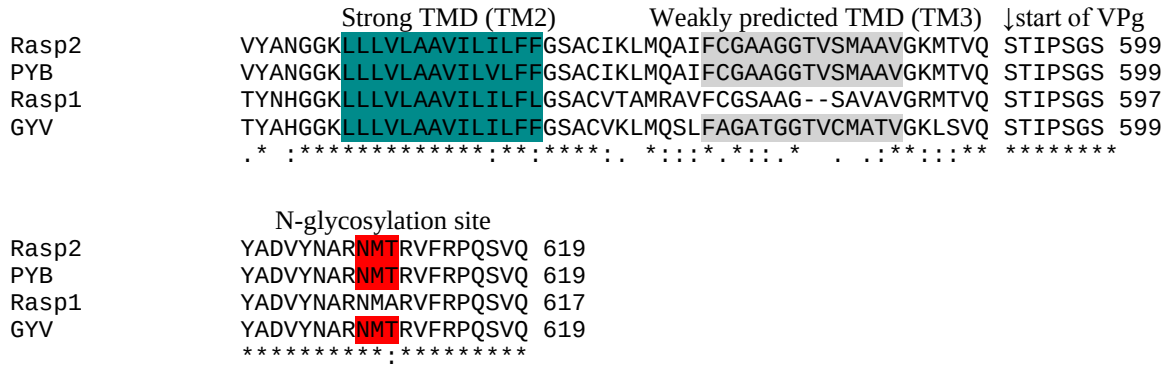


Figure 2.1: Amino acid alignment of NTB-VPg from different ToRSV isolates. Starting amino acids for cNV and cNV2 fragments are indicated with arrows. An asterisk (*) indicates positions which have a fully conserved residue. A colon (:) indicates conservation between groups of strongly similar properties. A period (.) indicates conservation between groups of weakly similar properties

2.3.2 *In vitro* membrane association of cNTB-VPg constructs: evidence for glycosylation and signal peptidase cleavage in the Rasp2, PYB and GYV isolates, but not in the Rasp1 isolate

The coding regions of the truncated cNTB-VPg protein (cNV) (Fig. 2.1) of different ToRSV isolates were cloned into pCITE-4a (+) and expressed in a coupled transcription and translation system. The presence of [^{35}S] methionine in the reactions allowed detection of the translation products by autoradiography. The cNV of all ToRSV isolates were expressed with an expected molecular weight of 34.7 kDa (Fig. 2.2). A protein with slower mobility (about 3 kDa larger) was detected for the wild type (WT) PYB, GYV and Rasp2 constructs in the presence of dog pancreas canine microsomal membrane (a commercial preparation of ER-derived membranes) (Fig. 2.2). However, this band was not detected when a T_{610} to A (T_{610}-A) mutation was introduced in the naturally occurring NMT glycosylation site in the Rasp2 cNV construct ((12) and data not shown), indicating that it is a glycosylated form of cNV. The N-linked glycosylation, which occurs on the luminal side of the ER, is commonly used as a convenient marker of luminal orientation of the proteins (56). Detection of the glycosylated form of cNV indicated that the protein is associated with membranes *in vitro* and that the VPg region is translocated into the ER lumen by the upstream transmembrane domain TM2 located in the NTB region. It was noted that cNV was only partly glycosylated, probably due to the limiting amounts of membranes in the reactions. A glycosylated band was not detected for the Rasp1 cNV protein but was observed for an A to T (A_{608}-T) mutant that introduced the NMT glycosylation site in the VPg domain. This result indicates that the Rasp1 cNV protein adopted a similar topology as Rasp2, PYB and GYV with the VPg translocated into the lumen of the membrane.

Smaller proteins were also detected for PYB, GYV and Rasp2 isolates in the presence of microsomal membranes (Fig. 2.2). The accumulation level of these small proteins increased over time in all PYB, GYV and Rasp2 isolates, indicating that it is a post-translational event, at least *in vitro*. Post-translational signal peptidase cleavage of Rasp2 cNV *in vitro* has been previously reported (12). Based on their size, the smaller fragments observed in PYB and GYV were also likely due to signal peptidase processing. In contrast, cleavage products were not detected in the WT Rasp1 cNV or the A_{608}-T mutant derivative. As mentioned above, detection of the glycosylated form of cNV in the Rasp1 A_{608}-T mutant confirmed that the protein was properly associated with membranes although no cleavage occurred. Taken together, the results indicate that the cNV proteins of all ToRSV isolates associate with microsomal membranes *in vitro* and that the C-terminal region of NTB and the VPg domain is translocated into the luminal side of the membrane, where glycosylation and signal peptidase cleavage occurred in the Rasp2, PYB and GYV isolates but not in the Rasp1 isolate.

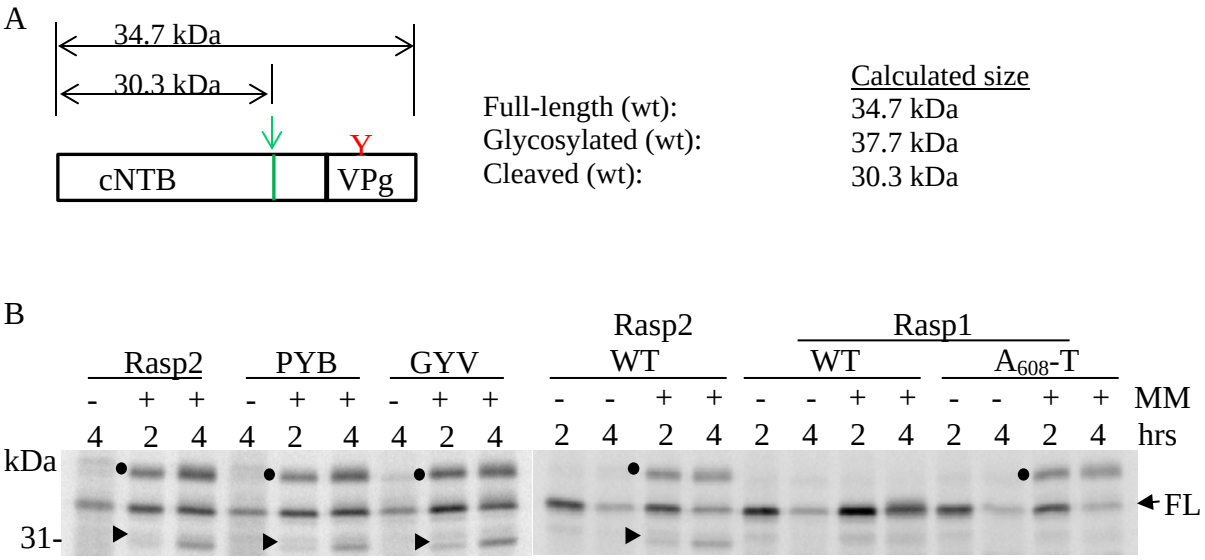


Figure 2.2: *In vitro* translation of the cNV proteins from different ToRSV isolates. A. Schematic representation of the cNTB-VPg construct with the predicted size of the full-length protein and membrane modified products. The arrow indicates the predicted signal peptidase cleavage site and Y represents the predicted glycosylation site. B. The cNV proteins of different ToRSV isolates were expressed *in vitro* in the presence (+) or absence (-) of canine microsomal membrane (MM). Reactions were performed at room temperature for two or four hours (hrs) as indicated. Translation products and membrane-associated modified products were separated by 12% SDS-PAGE and detected by autoradiography. The glycosylated forms of cNV are indicated with black dots (●) and the signal peptidase cleavage products are shown with black triangles (▶). The unglycosylated full-length (FL) proteins are indicated with the arrow.

2.3.3 A signal peptidase inhibitor inhibits the membrane-dependent cleavage of Rasp2 cNTB-VPg *in vitro*

To further confirm that the smaller products observed above resulted from membrane-associated signal peptidase cleavage; a signal peptidase inhibitor (MeOSuc-Ala-Ala-Pro-Val chloromethyl ketone, Sigma) was added to the *in vitro* translation system. The WT Rasp2 cNV and a mutant (G_{578}-A) with increased signal peptidase cleavage (see Fig. 2.6) were used as substrates in this inhibitor assay. Both Rasp2 WT cNV and the G_{578}-A mutant showed glycosylation and signal peptidase processing when the microsomal membranes were added to the reactions (Fig. 2.3). Addition of the inhibitor to the reactions at the final concentration of 1.2 mM resulted in reduced accumulation of the cleaved product in both constructs, demonstrating that the cleavage was due to the activity of membrane-associated signal peptidase. Interestingly, the accumulation of signal peptidase cleavage product over time in the G_{578}-A mutant in the absence of the inhibitor was accompanied with a decrease in the amount of glycosylated protein. The accumulation level of the glycosylated protein in the G_{578}-A mutant was significantly increased in the presence of the inhibitor, suggesting that the glycosylated form of cNV is the main substrate for the signal peptidase. However, signal peptidase processing of the unglycosylated Rasp2 cNV (T_{610}-A) has also been observed previously (12).

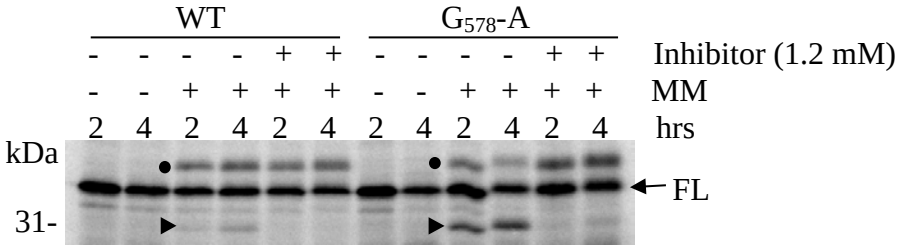
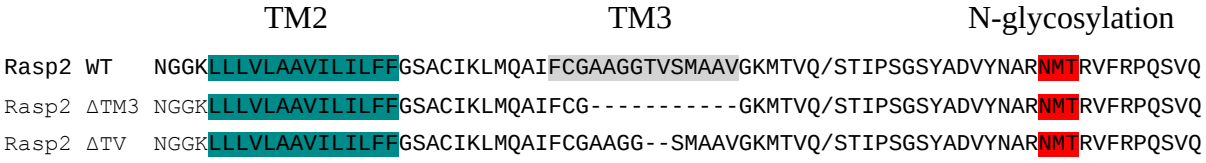


Figure 2.3: Signal peptidase inhibitor treatment of the Rasp2 cNV and its G_{578}-A mutant. The WT Rasp2 cNV and its G_{578}-A mutant were expressed *in vitro* in the presence (+) or absence (-) of microsomal membranes (MM). Signal peptidase inhibitor (MeOSuc-Ala-Ala-Pro-Val chloromethyl ketone, Sigma) was added to the reactions at the final concentration of 1.2 mM. FL indicates the full-length unglycosylated protein. The (●) and (▶) indicate the glycosylated full-length proteins and the N-terminal fragments of the signal peptidase cleavage products, respectively.

2.3.4 Complete or partial deletion of a weak hydrophobic domain (TM3) reduces the efficiency of signal peptidase cleavage

Significant sequence divergence was observed in the C-terminal region of Rasp1 NTB-VPg, which is not a signal peptidase substrate *in vitro*, when compared to other isolates which are processed by the signal peptidase (Fig. 2.4A). The most obvious differences are several amino acid substitutions and a TV deletion in the TM3 domain that drastically reduces its hydrophobicity. Therefore, the TM3 region may influence the efficiency of the signal peptidase processing. As previously shown (12), a Rasp2 cNV mutant with a deletion of almost the entire TM3 (Rasp2 ΔTM3) was glycosylated in the presence of microsomal membranes, indicating that the deletion of TM3 did not prevent the membrane association of this protein and the translocation of the VPg in the lumen (Fig. 2.4B). However, the signal peptidase cleavage was significantly decreased. In addition, deletion of the two amino acids TV ($\Delta TV_{580\text{-}581}$) in the TM3 region of the Rasp2 cNV (corresponding to the normal mutation observed in Rasp1) reduced the efficiency of signal peptidase cleavage although it was not completely eliminated. These results strongly indicate that the TM3 region plays an important role in the signal peptidase processing.

A



B

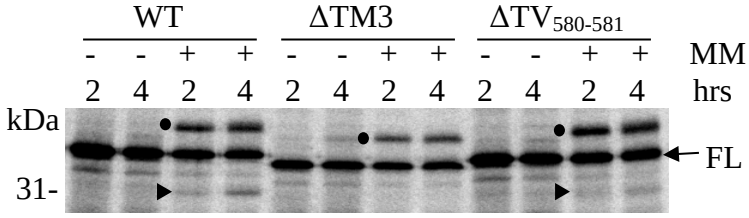
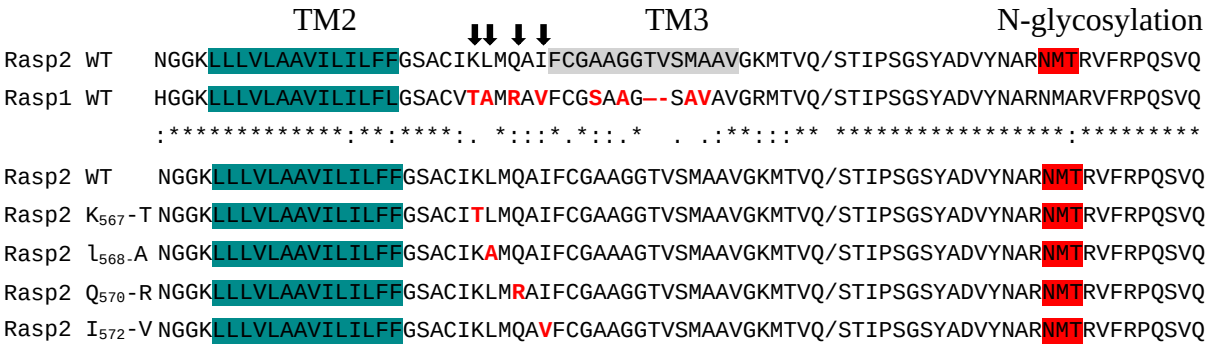


Figure 2.4: *In vitro* translation of Rasp2 cNV and the ΔTM3 and ΔTV mutants. A. Amino acid sequence of the C-terminal region of the Rasp2 cNV protein and its ΔTM3 and $\Delta TV_{580-581}$ mutants. B. *In vitro* translation of Rasp2 WT cNV and the two mutants (ΔTM3 and $\Delta TV_{580-581}$) with (+) or without (-) microsomal membranes.

2.3.5 Single mutations in the upstream N-terminal region of the TM3 domain modestly influence the efficiency of signal peptidase cleavage

Several amino acid substitutions were observed immediately upstream of the TM3 region in the Rasp1 isolate, including a K_{567} to T, L_{568} to A, Q_{570} to R and I_{572} to V mutation (Fig. 2.5A). To investigate whether these amino acid substitutions have an impact on the signal peptidase processing, individual amino acid mutations were introduced to Rasp2 cNV to restore the sequence of the Rasp1 protein (Fig. 2.5A). The L_{568}-A and I_{572}-V mutations did not significantly influence the efficiency of signal peptidase cleavage when compared to the wild type construct (Fig. 2.5B). The K_{567}-T mutation caused a slight increase in the efficiency of cleavage while the Q_{570}-R mutation slightly reduced but did not eliminate the signal peptidase cleavage. In conclusion, single mutations introduced into the immediate upstream region of TM3 of Rasp2 cNV protein generally had modest effects on the efficiency of signal peptidase cleavage. This does not exclude the possibility that double or triple mutations could have a greater impact.

A



B

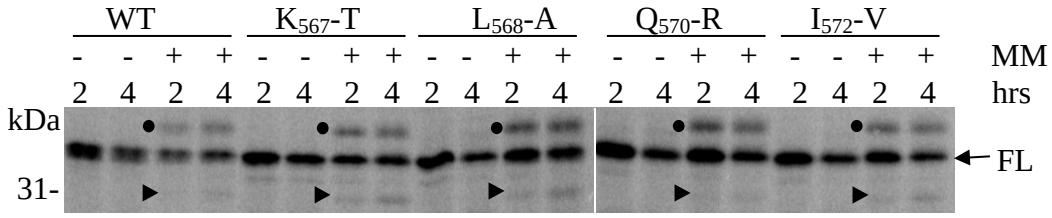


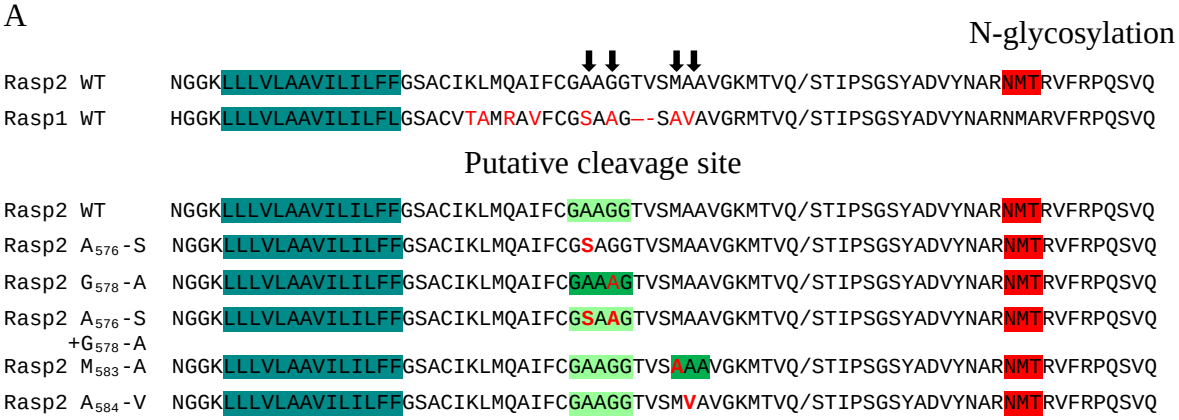
Figure 2.5: *In vitro* signal peptidase processing of Rasp2 cNV and mutants at positions 567-572. A. Amino acid sequence of the Rasp2 cNV mutants used in the *in vitro* translation. Mutations tested in panel B are shown with arrows. B. *In vitro* translation of Rasp2 cNV and the mutants.

2.3.6 Evidence that the sequence AAA corresponds to an optimal signal peptidase cleavage site *in vitro*

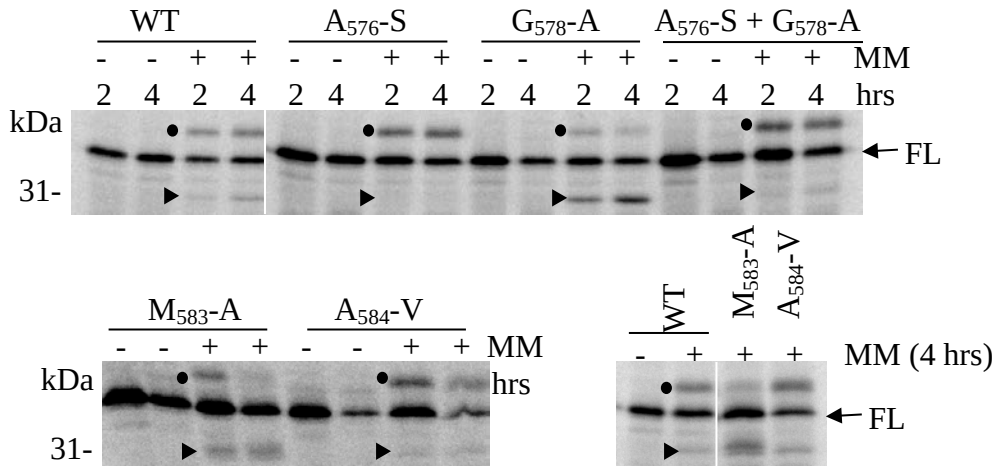
The above results indicated that the TM3 region is critical for signal peptidase processing (Fig. 2.4). A new series of single and double amino acid mutations were introduced at positions 576-583 into the Rasp2 cNV to restore the sequence of Rasp1 (Fig. 2.6A). The signal peptidase cleavage efficiency of these mutants was examined. M_{583} to A (M_{583}-A) and G_{578} to A (G_{578}-A) mutations dramatically increased the efficiency of signal peptidase processing. In both cases, the mutations resulted in the introduction of an AAA sequence, which was not present in the wild type Rasp2 or Rasp1 cNV sequence. Preference of A at the -3 and -1 position of signal peptidase cleavage site has been previously observed in both prokaryotic and eukaryotic cells (57). Based on the similar size of the cleaved products in the WT cNV and G_{578}-A mutant, it was likely that the GAAGG sequence in the wild type Rasp2 protein corresponded to the original signal peptidase cleavage site. Typical signal peptidase cleavage site generally follow the -3,-1 rule (with a preference of small amino acids such as A and G at the -3 and -1 position relative to the cleavage site) (40). Therefore, the GAA, AAG or AGG sequence may have been recognized in the WT constructs. Edman degradation sequencing of cleaved products would be necessary to determine the exact cleavage site but is technically challenging due to the low amounts and small size (8.6 kDa if glycosylated) of the C-terminal cleavage products. Introducing a less favorable amino acid residue S (A_{576}-S mutant) to create the GSAGG sequence eliminated the signal peptidase cleavage completely. A Rasp2 cNV double mutant that contained an A_{576}-S and G_{578}-A mutation (A_{576}-S+G_{578}-A mutant) created a GSAAG sequence that was cleaved by signal peptidase although inefficiently. Interestingly, two cleavage products were detected in the M_{578}-A mutant. The

smaller product was likely due to the cleavage at the original GAAGG sequence while the larger product may result from cleavage at the newly introduced AAA sequence further downstream (Fig. 2.6 A). The A to V (A_{584}-V) mutant did not seem to influence the signal peptidase processing. The entire series of mutations were also introduced into the smaller Rasp2 cNV2 truncated protein (Fig. 2.1). The same trend of results was obtained upon expression of the wild type and mutant cNV2 *in vitro* in the presence of microsomal membrane (data not shown). Taken together, these results indicate that the GAAGG sequence in the wild type Rasp2 cNV is recognized and cleaved by the signal peptidase *in vitro*. The putative -3, -1 position of the cleavage site as well as flanking amino acids can influence the efficiency of the signal peptidase cleavage.

A



B



C

Rasp2 WT	G	A	A	G	G
Rasp2 A_{576}-S	G	S	A	G	G
Rasp2 G_{578}-A	G	A	A	A	G
Rasp2 G_{578}-A+A_{576}-S	G	S	A	A	G

Figure 2.6: *In vitro* signal peptidase processing of Rasp2 cNV and mutants at positions 576-583. A. Amino acid sequence of the Rasp2 cNV mutants. The approximate deduced location of the signal peptidase cleavage sites are shown in green with the darkness of the

green corresponding to the efficiency of signal peptidase processing. Amino acids that are mutated are indicated with arrow. B. *In vitro* translation of WT Rasp2 cNV and its mutants. C. Deduced signal peptidase cleavage sites of the wild type Rasp2 cNV and its mutants. Putative cleavage sites in each construct are indicated with arrow.

2.3.7 Mutations of Rasp1 cNTB-VPg to recreate the Rasp2 sequence restores signal peptidase cleavage

Mutations were also introduced in the Rasp1 cNV A_{608}-T mutant to reconstitute the Rasp2 sequence (Fig. 2.7A). The A_{608}-T mutation recreates the VPg N-glycosylation site, allowing the monitoring of membrane insertion and signal peptidase cleavage simultaneously. Glycosylation was detected in all Rasp1 cNV mutants in the presence of microsomal membranes, confirming membrane association and translocation of the VPg in the membrane lumen (Fig. 2.7B). The R to Q (R_{570}-Q) mutation was able to restore the signal peptidase cleavage, although inefficiently (Fig. 2.7). The S to A (S_{576}-A) mutation, which created a favorable AAA sequence, was efficiently cleaved by the signal peptidase. The size of the cleaved protein was similar to that observed in the R_{570}-Q mutant, suggesting that either the SAA sequence or the adjacent AAG sequence was recognized inefficiently when the Q was introduced. Introduction of the absent amino acids T and V (referred as $+TV_{580\text{-}581}$ mutant) to the Rasp1 cNV (A_{608}-T) also restored efficient signal peptidase cleavage. However, the size of the cleaved protein was slightly larger suggesting that the downstream AVA sequence was recognized as a signal peptidase cleavage site in this mutant. It is interesting to note that the AVA sequence was not recognized as a signal peptidase cleavage site in the absence of the TV residues, suggesting that other factors, possibly the conformation of the protein in the membrane may influence the signal peptidase cleavage.

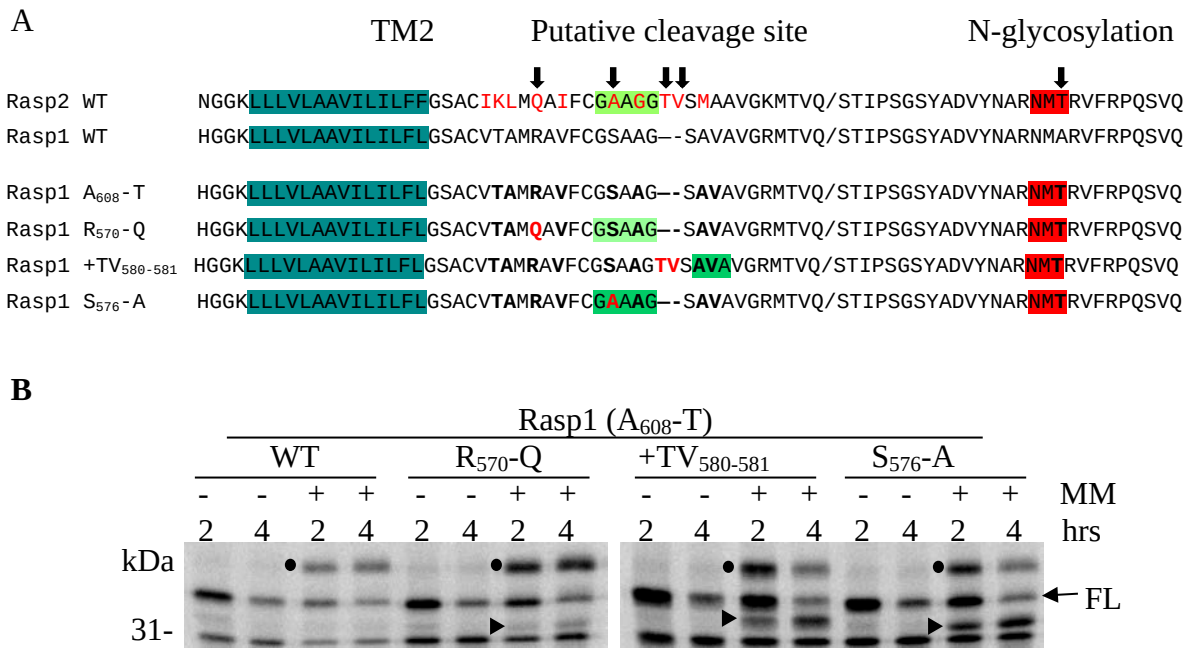


Figure 2.7: *In vitro* translation of Rasp1 cNV (A_{608}-T) and mutant derivatives. A. Amino acid sequence of the Rasp1 cNV (A_{608}-T) and mutants (R_{570}-Q, $+TV_{580-581}$ and S_{576}-A) that mimic the Rasp2 sequence, Please note that the R_{570}-Q, $+TV_{580-581}$ and S_{576}-A mutants have the introduced glycosylation site (A_{608}-T) in the VPg domain in addition to the mutations in the C-terminal region of the NTB domain. The deduced signal peptidase cleavage sites (based on the migration of the cleaved products) are shown in green, with darker color corresponding to higher efficiency of signal peptidase cleavage. B. Translation of Rasp1 cNV (A_{608}-T) and mutant derivatives *in vitro*.

2.3.8 *In vivo* expression of cNV2 and its mutants from different ToRSV isolates

To further investigate whether the N-linked glycosylation and signal peptidase cleavage observed *in vitro* also occurs *in planta*, the cNV2 fragments derived from different ToRSV isolates were fused to an N-terminal Flag tag and a C-terminal HA tag. Fragments were inserted into pBIN (+) and expressed in *N. benthamiana* plants by agroinfiltration. Samples were collected at 3 dpa. Proteins were separated by 16% SDS-PAGE and detected with both Flag and HA antibodies. The glycosylated form of cNV2 of Rasp2 WT, ΔTM3, Rasp1 A_{608}-T and GYV WT constructs and the unglycosylated form of the full-length cNV2 of the Rasp2 T_{610}-A and the Rasp1 WT constructs were detected with Flag antibody (Fig. 2.8). This result is consistent with the *in vitro* result. Interestingly, additional smaller fragments of about 15 kDa were also detected in Rasp2 WT, T_{610}-A, ΔTM3 and GYV WT constructs with the Flag antibody. In contrast, no smaller products were detected in both Rasp1 WT and A_{608}-T mutant. These fragments may correspond to the N-terminal fragment released after signal peptidase cleavage, although they migrated at a position slightly higher than predicted for their calculated size (12.0 kDa, based on the predicted AAG cleavage site). The ratio of the cleaved fragment relative to the unprocessed form of the protein varied. It was very low for the ΔTM3 mutant which is consistent with the inefficient signal peptidase cleavage observed *in vitro*. Using the HA antibody, the full-length glycosylated and unglycosylated forms of the cNV2 proteins were detected, which corresponded to those detected with the Flag antibody. For constructs that included a glycosylation site, small amounts of the unglycosylated proteins were also detected, likely because of the higher detection efficiency of this antibody. In addition, smaller proteins of about 11 kDa were detected in Rasp2 and GYV cNV2 constructs but not in the Rasp2 T_{610}-A and ΔTM3 mutants. It is possible that

these proteins were the C-terminal products of the signal peptidase cleavage event (predicted size 8.6 kDa if glycosylated). Other than the above mentioned bands, an extra band of about 14 kDa was also detected in Rasp2 WT, T_{610}-A, ΔTM3 and GYV WT using the HA antibody. The origin of this band is not known, but it is likely due to internal translation initiation or an uncharacterized protein degradation event. Taken together, the results suggested that the cNV2 fragment of Rasp2 and GYV might be processed by signal peptidase *in vivo*. However, this interpretation must be made carefully because the size of the possible cleavage products migrated slightly higher than expected. Future experiments will be required to make definitive conclusions.

A

(Flag-cNV2 -HA)

	Calculated size
Full-length (wt):	17.6 kDa
Glycosylated (wt):	20.6 kDa
SPase cleaved (wt) (Flag Abs):	12.0 kDa
SPase cleaved (wt) (HA Abs):	5.6 kDa (without glycosylation)
	8.6 kDa (with glycosylation)
Full-length ΔTM3	16.7 kDa (without glycosylation)
	19.7 kDa (with glycosylation)

B

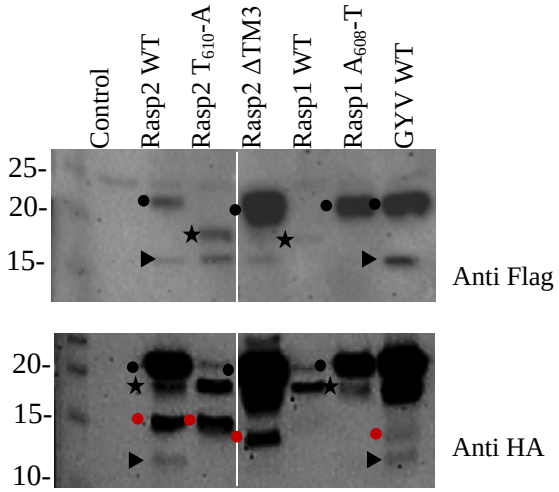


Figure 2.8: Expression of ToRSV-derived Flag-cNV2-HA in *N. benthamiana*. A. Schematic representation of ToRSV cNV2 constructs. The Flag and HA tags were fused to the N-and C-terminus of the cNV2 fragment, respectively. The predicted glycosylation is indicated with Y and the signal peptidase cleavage is indicated with the green vertical line and the green arrow. The calculated size of the glycosylated cNV2, unglycosylated cNV2 and the cleavage products are shown on the right of the figure. B.Western blot analysis of Flag-cNV2-HA. The glycosylated and unglycosylated full-length cNV2 proteins are indicated with black dots and stars, respectively. The putative signal peptidase cleavage products are indicated with black triangles. Bands of unknown origin are indicated with red dots.

2.3.9 Silencing of the signal peptidase causes death of *N. benthamiana*

Viral-induced gene silencing (VIGS) is a powerful tool for down-regulation of host genes. It has been widely used to study gene function. Therefore, I tried to knockdown the signal peptidase of *N. benthamiana* and examine whether this has an impact effect on virus replication or cleavage of Flag-cNV2-HA protein. To do this, a 356 bp fragment of the signal peptidase catalytic subunit was cloned into the TRV00. This construct (named TRV-SPase) was agroinfiltrated into *N. benthamiana* plants as described in Materials and Methods. Systemic silencing was established at 10 dpa. The signal peptidase silenced plants showed obvious stunting, yellowing and vein clearing symptoms and all plants died soon after (Fig. 2.9), suggesting that signal peptidase might be essential for the survival of *N. benthamiana* plants. Silencing of the signal peptidase probably disrupted the cellular secretory system, which in turn caused physiological disorder and malfunctioning of the plant cell and eventually resulted in plant death. Alternatively or additionally, it is also possible that the signal peptidase silenced plants are more susceptible to TRV infection. Because of this severe phenotype, I was unable to inoculate ToRSV or test the Flag-cNV2-HA constructs using these plants.

TRV00 TRV-SPase

Figure 2.9: Symptoms of signal peptidase silenced *N. benthamiana* plants. Pictures were taken 12 days after VIGS treatment. Signal peptidase silenced plants (TRV-SPase) showed lower growing rate, stunting and necrosis when compared to the control plants (TRV: 00).

2.3.10 The glycosylated form of cNV2 is more stable than the unglycosylated protein

N-glycosylation plays an important role in assisting protein folding and increasing protein stability (268-270). The *in vivo* results described above indicated that the full-length unglycosylated cNV2 protein of Rasp2 T_{610}-A and Rasp1 WT accumulated in lesser amounts compared to the corresponding glycosylated forms of Rasp2 WT and Rasp1 A_{608}-T (Fig. 2.8). Therefore, I further investigated whether the glycosylated cNV2 is more stable than the unglycosylated form. To perform this experiment, *N. benthamiana* leaves that transiently express cNV2 of the Rasp2 WT, Rasp2 T_{610}-A, Rasp1 WT and Rasp1 A_{608}-T were collected after 3 dpa and the detached leaves were treated with 30ug/ul cycloheximide (an inhibitor that freezes translational elongation and stops protein synthesis) for two hours as previously described (266). Accumulation levels of the cNV2 proteins were detected by western blots using the HA antibody (Fig. 2.10). According to the result, the cycloheximide treated glycosylated cNV2 of Rasp2 WT and Rasp1 A_{608}-T accumulated at similar levels when compared to the corresponding control plants. In contrast, the unglycosylated cNV2 of Rasp2 T_{610}-A and Rasp1 WT disappeared after cycloheximide treatment. Taken together, the results suggest that adding a glycan to the cNV2 protein significantly enchanced its stability, probably because the glycan helps the protein fold in an optimal conformation.

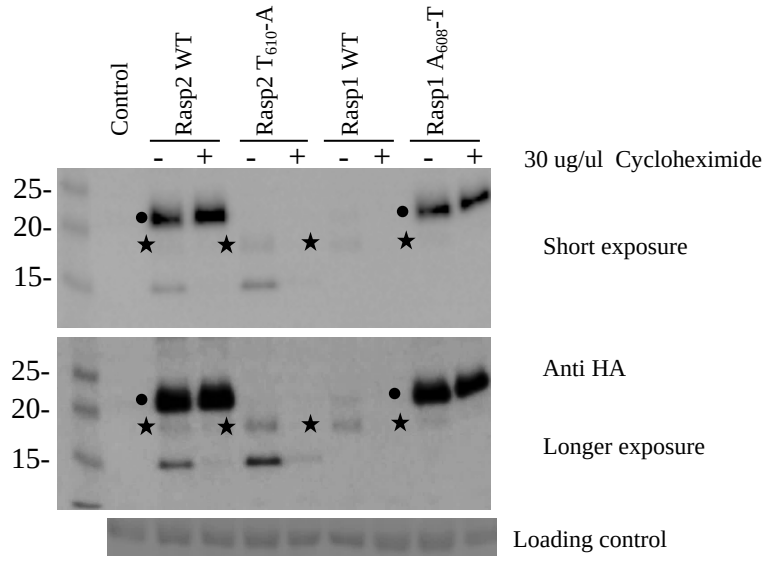


Figure 2.10: Cycloheximide treatment of glycosylated and unglycosylated ToRSV cNV2. Agroinfiltrated leaves were treated with control solution (-) or 30 ug/ul cycloheximide solution (+) (a protein synthesis inhibitor) for two hours. Accumulation levels of proteins were detected using western blot analysis and the HA antibody. The glycosylated proteins are indicated with dots and the unglycosylated proteins are indicated with stars.

2.4 Discussion

Signal peptidase processing of ToRSV-Rasp2 cNV has previously been observed *in vitro* (12). However, the exact cleavage site was not identified and the signal peptidase processing has not been investigated *in vivo*. In this study, the signal peptidase was shown to cleave the cNV truncated protein from the GYV and PYB isolates *in vitro* (Fig. 2.2). These two isolates shared high sequence similarity with the Rasp2 isolate in the C-terminal region of the cNV (Fig. 2.1). In contrast, the Rasp1 cNV protein that exhibited more variations in the C-terminal region of cNV did not serve as a signal peptidase substrate, although it associated with membranes and adopted a similar topology in the membrane with the VPg domain translocated in the lumen (Fig. 2.1 and 2.2). Deletion and mutational studies indicated that the TM3 region is critical for signal peptidase processing (Fig. 2.4 and 2.5). Single amino acid mutations were introduced in the TM3 region of Rasp2 cNV to mimic the Rasp1 sequence, which allowed identification of the GAAGG sequence as a sub-optimal signal peptidase cleavage site in Rasp2 with the AAG sequence as the most likely -3 to -1 position (Fig 2.6). Replacing the G in the putative -1 position with a more favorable amino acid A (G_{578}-A mutant) enhanced the signal peptidase processing while changing the putative -3 position from A to a less favorable amino acid S (A-S mutant) prohibited the cleavage (Fig. 2.6). Additionally, amino acid residues upstream and downstream of the cleavage site also influenced signal peptidase processing. Mutations were also introduced in the Rasp1 cNV A_{608}-T mutant to mimic the Rasp2 sequence. Surprisingly, the signal peptidase cleavage was restored when the R_{570}-Q, $+TV_{580\text{-}581}$ and S_{576}-A substitutions were introduced (Fig. 2.7). The AVA sequence (located downstream of the position of the Rasp2 cleavage site) was recognized as a cleavage site in the $+TV_{580\text{-}581}$ mutant but not in the wild-type sequence,

suggesting that protein conformation plays an important role in determining the efficiency of signal peptidase processing. *In vivo* expression of cNV2 allowed detection of the possible signal peptidase cleavage products in both Rasp2 and GYV wild type constructs but not in Rasp1 (Fig. 2.8), which was consistent with the *in vitro* results.

As discussed in Chapter 1, canonical signal peptides are normally located in the N-terminal region of the protein (40). However, it has been shown that signal peptidase is able to cleave C-terminally located signal peptidase cleavage sites (271). Some viruses take advantage of this host protease for releasing the structural proteins from viral polyproteins (as discussed in Chapter 1). Here, I report that the signal peptidase is involved in processing a C-terminal signal peptidase cleavage site in a ToRSV membrane-associated replication protein. Interestingly, this processing occurred in the PYB, GYV and Rasp2 isolates at least *in vitro*, but not in the Rasp1 isolate (Fig. 2.2). Although the clevage site in the ToRSV cNTB-VPg is C-terminally located, the context of this cleavage site shares certain common features with the typical signal peptides (40, 57). For example, it has a positively charged amino acid K upstream of the transmembrane domain, the core hydrophobic region is rich in L and the -3 and -1 position of the putative cleavage site is occupied with small amino acids (Fig. 2.11).

The mutagenesis and inhibitor studies confirmed that the processing in cNV was executed by the signal peptidase. However, one feature did not fit the scheme of a typical signal peptide. Indeed, the distance between the end of the hydrophobic domain TM2 and the putative cleavage site (approximately 18 residues) exceeds that known for regular signal peptides (3-7 residues) (Fig 2.11A) (40). As mentioned in Chapter 1, signal peptidase is a membrane

protein complex. The catalytic subunits have a transmembrane domain leading to the orientation of the active site on the luminal side of the membrane (33). This forces specific positioning of the enzyme active site in relation to the lipid bilayer. For efficient cleavage, a suitable substrate cleavage site needs to be correctly presented to the active site of the enzyme which is close to the surface of the membrane (272). It has been shown that the optimal distance between the end of the h-region (transmembrane domain) and the signal peptidase cleavage site is 4-5 residues (273). This raises the question of how the ToRSV cNV cleavage site is exposed to and recognized by the active site of the signal peptidase. Amino acid sequence analysis revealed two possible explanations. Secondary structure prediction softwares (DSC, MLRC and PHD) implied a long α-helix in the Rasp2 cNV sequence that encompasses not only the hydrophobic TM2 domain (highlighted in blue in Fig. 2.11A) but also the 12 amino acids further downstream. Projection of the downstream region of the predicted long helix (in yellow) reveals its amphipathic helix property (Fig. 2.11A and B). Therefore, it is possible that the Rasp2 cNV has an unusual h-region with its N-terminal region traversing the membrane and the C-terminal region forming an amphipathic helix that orients itself horizontally on the luminal side of the membrane (Fig. 2.11C). A bend in the transmembrane domain, facilitated by the glycine residue (between the blue and yellow boxes), would be necessary to allow this topology to occur (274). By adopting this topology, the cleavage site, which is located downstream of the putative amphipathic helix, is brought closer to the active site of the signal peptidase for processing. Signal peptidase cleavage after an amphipathic helix was previously observed in classic swine fever virus E^{rns}-E1 protein (275). Alternatively, it is possible that the h-region could form a longer transmembrane domain to traverse the membrane (Fig. 2.11C), in this case, the

helix may be tilted to allow the entire region to be buried in the lipid membrane. Tilting of long hydrophobic regions to accommodate the width of the membrane has been observed in some membrane proteins (274). Adopting this topology can also ensure the contact of the substrate cleavage site and the signal peptidase catalytic site.

The sequence of the Rasp1 cNV was also analyzed. The amino acid sequence downstream of the TM2 hydrophobic domain was also predicted to form an amphipathic helix although with lesser degree of confidence than the Rasp2 sequence (Fig. 2.11B). Cleavage of the Rasp1 S_{576}-A mutant which recreates the AAA sequence indicates that recognition of a signal peptidase cleavage site at this position is also possible in the Rasp1 sequence, indicating that the Rasp1 NTB-VPg may adopt a similar topology as Rasp2.

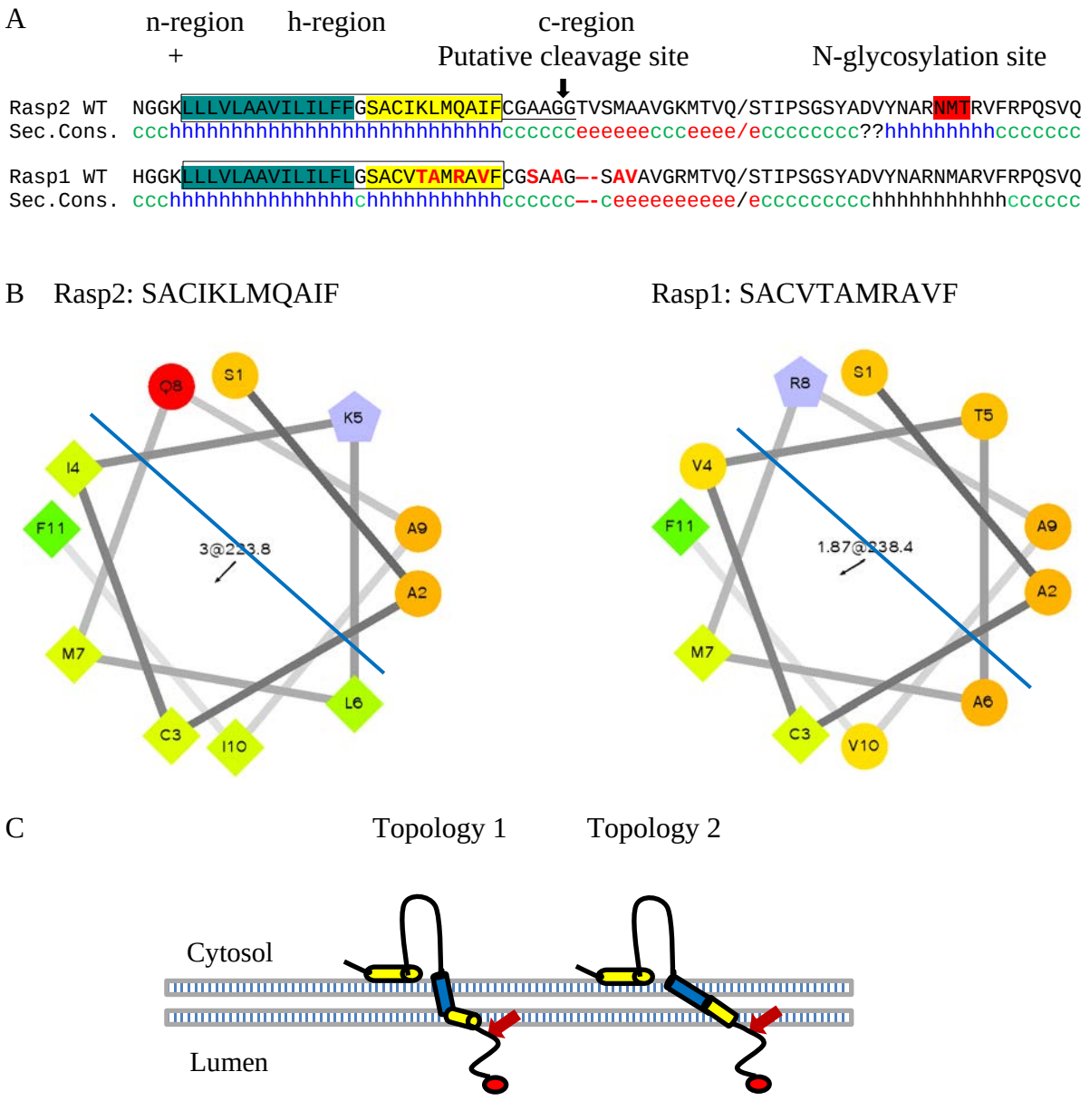


Figure 2.11: Sequence analysis of ToRSV cNV and possible topologies of ToRSV NTB-VPg in the membrane. A. Sequence analysis of Rasp1 and Rasp2 cNV. The n-region includes a positively charged (+) amino acid residue K, the h-region (shown in box) including the TM2 transmembrane domain (highlighted in blue) and a putative amphipathic helix (highlighted in yellow), the c-region (underlined) and the most likely putative cleavage

site (indicated with arrow) are also specified. Secondary structure of this region was predicted with three computer based softwares (DSC, MLRC and PHD); the consensus secondary structures are shown in the figure with h, c and e representing alpha helices, random coils and extended strands, respectively. B. Helical projections of the putative amphipathic helix. Amino acid residues used for helical projections are indicated on the top of the figure. The hydrophilic residues are shown as circles, and hydrophobic residues are shown as diamonds. A blue line dividing the hydrophobic face of the helix (left half) from the hydrophilic face of the helix (right half) is shown. C. Models of possible topology of the NTB-VPg in the ER membrane. Two possible topologies in the membrane are shown. In the first topology, the TM2 hydrophobic domain (blue) traverses the membrane and a kink in this transmembrane domain allows the downstream amphipathic helix (yellow) to associate with the surface of the membrane with its hydrophilic side. The downstream region of NTB and the VPg is directed into the ER lumen where signal peptidase processing occur (red arrow). In the second model, the h-region (including the TM2 and the following 12 amino acid residues)forms a long helix structure traversing the membrane and orients the downstream NTB-VPg and cleavage site into lumen where it gets cleaved by signal peptidase (red arrow).

Interestingly, signal peptidase cleavage sites were also predicted in C-terminal NTB of other nepoviruses, including grapevine chrome mosaic nepovirus, TBRV and GFLV using online SignalP program (12), indicating that this membrane-associated cleavage might be conserved among nepoviruses. As mentioned previously, the ToRSV NTB-VPg protein associates with membrane and is involved in the assembly of virus replication complexes (11). I attempted to test the biological relevance of this signal peptidase cleavage by investigating whether or not down-regulation of signal peptidase would influence ToRSV accumulation. Unfortunately, I was unable to perform the experiment because down-regulation of the signal peptidase caused plant death (Fig. 2.9). It is possible that signal peptidase processing disrupts luminal hydrophobic interactions among TM3 region of different NTB-VPg molecules, which may influence the architecture of the replication complexes and impact virus replication. Future studies will be aimed at obtaining infectious transcripts to introduce mutations that either promote or decrease signal peptidase cleavage. The effect of these mutations on ToRSV replication could then be studied. It is also interesting to note that the signal peptidase cleavage is sub-optimal, it occurs over time and cleavage is not efficient in the wild type Rasp2, PYB and GYV sequence, which probably allows efficient virus replication and accumulation at the early stages of virus infection. Cleavage of NTB-VPg at later stages of infection may slow down the replication and stimulate the transition from replication to the next stage of the virus life cycle.

Glycosylation plays an important role in protein folding and stability by adding sugar chains to the peptide backbone (270, 276). In this study, N-linked glycosylation was observed in the VPg region of the PYB, GYV and Rasp2 isolates (Fig 2.2). No glycosylation was observed

in the Rasp1 cNV or cNV2 since the Rasp1 lacked the glycosylation acceptor sequence (Fig. 2.1 and 2.8). However, glycosylation can be restored once the glycosylation site was introduced to the Rasp1 constructs (Fig. 2.2 and 2.6). Adding the glycan molecule to ToRSV cNV2 strongly enhanced its stability *in vivo* (Fig. 2.10). N-linked glycosylation is commonly seen in viral structural proteins (277, 278). However, it is also observed in viral nonstructural proteins. Mutation that abolishes glycosylation of dengue virus NS1 delays viral RNA synthesis and causes milder cytoplasmic effect (279). Similarly, mutants that lack glycan moieties in the yellow fever virus NS1 lead to reduction in virus yield and deficiency in viral RNA accumulation (280). Additionally, elimination of the glycosylation site in murine hepatitis virus nsp4, which is a viral-encoded multi-transmembrane protein, impairs viral RNA accumulation and this effect correlates directly with the disruption of viral-induced double membrane vesicle formation (281). As mentioned above, N-linked glycosylation is important for the stability of ToRSV derived membrane protein cNV2 (Fig. 2.10). However, it is also possible that glycosylation promotes proper folding of the protein in the membrane and play a role in ER-membrane modification during virus infection.

It is evident that glycosylation can increase the concentration of cNV2 by enhancing its stability (Fig. 2.10). This may lead to efficient assembly of replication complexes and increase in viral replication. The results also suggest that cNTB-VPg of PYB, GYV and Rasp2 is cleaved by signal peptidase. This cleavage may influence the stability of the protein or the architecture of the replication complexes and possibly regulate the stability of the replication complexes. It is interesting that only the PYB, GYV and Rasp2 proteins, which are stabilized by glycosylation, are susceptible to signal peptidase cleavage at least *in vitro*.

These isolates may have evolved to take advantage of signal peptidase processing to destabilize these proteins and regulate the replication activity to achieve optimal virus titer for their survival. In the case of Rasp1, the regulated degradation of the protein in the absence of glycosylation may serve the same function. However, this hypothesis still needs to be experimentally tested.

Chapter 3

Investigation into the stability of ToRSV RNA-dependent RNA polymerase

3.1 Introduction

As discussed in Chapter 1, two forms of Pol are detected in ToRSV infected plants: the mature 81 kDa Pol and the 95 kDa VPg-Pro-Pol′ (including the VPg, Pro and Pol domain but lacking the C-terminal 15 kDa of Pol) (13). Interestingly, the full-length VPg-Pro-Pol (111 kDa) is never detected even when a broad range protease inhibitor is added to the extraction buffer. The full-length Pol has been detected in low amounts if at all. In addition, smaller fragments with molecular masses of 53 kDa and 28 kDa are also detected in infected plants with the Pol antibody, raising the possibility that these fragments are cleavage products derived from the Pol (13). The truncation of the VPg-Pro-Pol, which allows accumulation of the stable polyprotein VPg-Pro-Pol′ *in vivo*, is not observed using *in vitro* processing assays with NTB-VPg-Pro-Pol or VPg-Pro-Pol as a substrate (13, 282). This suggests that the full-length VPg-Pro-Pol and Pol are intrinsically unstable in infected plants but what causes this instability is still unknown. In this chapter, I show that both ToRSV VPg-Pro-Pol and Pol are unstable when ectopically expressed in *N. benthamiana* plants individually. However, the stability of these two proteins can be significantly enhanced when the C-terminal 15-20 kDa of Pol is truncated. In contrast, the accumulated levels of VPg-Pro-Pol and Pol are similar to that of their truncated forms in *E. coli*, suggesting that the instability of VPg-Pro-Pol and Pol is plant specific.

3.2 Materials and methods

3.2.1 Plasmid constructions

To express different forms of polymerase in plants, the coding region of the full-length VPg-Pro-Pol (nts 3710-6664) and two truncated forms: VPg-Pro-Pol′ (nts 3710-6121) and VPg-Pro-Pol″(nts 3710-6226) of the ToRSV PYB isolate were amplified by RT-PCR using primer pairs P16F and P17R, P16F and P18R, P16F and P19R, respectively. The mature full-length Pol (nts 4532-6664) and two truncated forms: Pol′ (nts 4532-6126) and Pol″ (nts 4532-6226), as well as the C-terminal 15 kDa of Pol (nts 6126-6668) were amplified with primer pairs P20F and P17R, P20F and P18R, P20F and P19R, P86F and P17R, respectively (Table 3.1). All reverse primers included an HA coding sequence and allowed fusion of the HA tag at the C-terminus of the proteins. The amplified products were digested with *Xba*I and *Kpn*I and inserted into the modified plant expression vector pBIN (+) that already contains the double 35S promoter and NOS terminator. The HA tag was also fused to the N-terminus of various forms of the polymerase and cloned into pBIN (+) using the same strategy as above. The primer pairs P71F and P72R, P71F and P73R, P74F and P72R, P74F and P73R were used to clone the HA-VPg-Pro-Pol, HA-VPg-Pro-Pol′, HA-Pol and HA-Pol′, respectively (Table 1). The pBIN-p19 and the pBIN-GFP clones have been described previously (257). The pBIN-X2 construct was from Helene Sanfaçon.

Different forms of polymerase were also cloned into pET-21d(+) (Novagen). The coding regions of VPg-Pro-Pol, VPg-Pro-Pol′, VPg-Pro-Pol″, Pol, Pol′ and Pol″ were amplified with primer pairs P37F and P38R, P37F and P39R, P37F and P40R, P41F and P38R, P41F and P39R, P41F and P40R, respectively (Table 3.1). The amplified fragments were digested with

*Xba*I and *Xho*I and subsequently inserted into the corresponding sites of pET-21d(+) , which allows in frame fusion of six histidine residues at the C-terminus of proteins.

Table 3.1 Primers used in polymerase plasmid constructions

Primer name	Sequence of the Primer	Comments
P16F	5'- GCTCTAGATGTCGACGATTCCCTCCGGTAGTTAC-3'	*Xba*1+ ATG+ nts 3710-3733
P17R	5'- GGGGTACCTTA **GGCATAGTCAGGAACATCGTATGGGTA** GCTCGCAATTACACGAGGCTGA-3'	nts 6664–6643+ HA +TAA+*Kpn*I
P18R	5'- GGGGTACCTTA**GGCATAGTCAGGAACATCGTATGGGTA** ATATTGTTCATCATGAAACGAAG-3'	nts 6121-6099– HA +TAA+*Kpn*I
P19R	5'- GGGGTACCTTA**GGCATAGTCAGGAACATCGTATGGGTA** AGCAACAATCGAGAGGGCAGTAGCT-3'	nts 6226–6202+ HA +TAA+*Kpn*I
P20F	5'- GCTCTAG**A**TGTCTAGCGTCATTAAGTCTCT-3'	*Xba*1+ ATG+ nts 4532-4551
P37F	5'- ACGCTCTAGA*AATAATTTTGTTTAACTTTAAGAAGGAGATATACAT* ATGTCGACGATTCCCTCCGGTAGTTAC-3'	*Xba*1+vector sequence + ATG+ nts 3710-3733
P38R	5'- ACCGCTCGAGGCTCGCAATTACACGAGGCTGA-3'	nts 6664–6643+ *Xho*I
P39R	5'- ACCGCTCGAGATATTGTTCATCATGAAACGAAG-3'	nts 6121–6099+*Xho*I
P40R	5'- ACCGCTCGAGAGCAACAATCGAGAGGGCAGTAGCT-3'	nts 6226–6202+*Xho*I
P41F	5'- ACGCTCTAGA*AATAATTTTGTTTAACTTTAAGAAGGAGATATACAT* ATGTCTAGCGTCATTAAGTCTCT-3'	*Xba*1+vector sequence + ATG+ nts 4532-4551
P71F	5'- GCTCTAGATG**TACCCATACGATGTTCCTGACTATGCC** TCGACGATTCCCTCCGGTAGTTAC-3'	*Xba*1+ ATG+HA+nts 3710-3733
P72R	5'- GGGGTACCTTAGCTCGCAATTACACGAGGCTGA-3'	nts 6664–6643+TAA+*Kpn*I
P73R	5'- GGGGTACCTTAATATTGTTCATCATGAAACGAAG-3'	nts 6121–6099+TAA+*Kpn*I
P74F	5'- GCTCTAGATG**TACCCATACGATGTTCCTGACTATGCC** TCTAGCGTCATTAAGTCTCT-3' -3'	*Xba*1+ ATG+HA+ nts 4532-4551
P86F	5'- GCTCTAGATGTCTCAGTGGAAGCCGTGGTCTC-3'	*Xba*1+ ATG+ nts 6122-6143

The restriction enzyme sites in the primers are underlined, the start codon and the stop codon are highlighted in red and the HA tag is shown in bold. Vector sequences are shown in italic. The corresponding nucleotide sequence of primer in ToRSV RNA1 is specified.

3.2.2 Site-directed mutagenesis

For construction of pBIN HA-VPg-Pro^{HD}-Pol and pBIN HA-VPg-Pro^{HD}-Pol′, the wild type fragments were first introduced into pCR-Blunt (Life Technologies) and used as template for

mutagenesis. The site-directed mutagenesis was performed as described in Chapter 2. Primers P60F and P61R (Table 3.2) were used to introduce the H to D mutation in the catalytic triad of the protease. Plasmids with the correct mutation of the HA-VPg-Pro^{HD}-Pol and HA-VPg-Pro^{HD}-Pol′ were confirmed by sequencing the entire inserted region. The fragments were digested with *Xba*I and *Kpn*I for insertion into final vector pBIN (+). The GW to GA double mutation were introduced to the HA-Pol, Pol-HA and the C-terminal 15 kDa fragment of Pol as described above. Primer pairs P88F and P89R, P92F and P93R (Table 3.2) were used sequentially to introduce the first and second GW-GA mutations, respectively.

For construction of the pET VPg-Pro^{HD}-Pol, pET VPg-Pro^{HD}-Pol′ and pET VPg-Pro^{HD}-Pol″, site-directed mutagenesis was conducted using the pET-21d(+) plasmid that contained the wild type fragments as a template, primers P60F and P61R were used for introducing the H to D mutation. All constructs were confirmed by sequencing. The primers used for mutagenesis are shown in Table 3.2. All primers used in mutagenesis were PAGE purified (IDT).

Table 3.2 Primers used for mutagenesis of polymerase

Primer name	Sequence of the Primer
P60F	5’--GTCTTTGGCTTTGACTAAAGATCAGGCCTTAACCATCCCG-3’
P61R	5’- CGGGATGGTTAAGGCCTGATCTTTAGTCAAAGCCAAAGAC-3’
P88F	5’-GCTTGCTGTTGCAGGACCTGGCGCGCGTAATAAAGACCCAGACAGG-3’
P89R	5’-CCTGTCTGGGTCTTTATTACGCGCGCCAGGTCCTGCAACAGCAAGC-3’
P92F	5’-CGAAGGTCCCTTTGTCTCGGGAGCGGCAGCTGCCATTTCCTTCGG-3’
P93R	5’-CCGAAGGAAATGGCAGCTGCCGCTCCCGAGACAAAGGGACCTTCG-3’

Nucleotides that encode the mutated amino acids were underlined.

3.3.3 Agroinfiltration of *N. benthamiana*

The binary vector pBIN (+) containing different forms of ToRSV polymerase were transformed into *A. tumefaciens* LBA4404. Agroinfiltration of *N. benthamiana* leaves were conducted as previously described (283). The plants were kept at 21˚C and the infiltrated leaves were collected 3 dpa for protein and RNA analysis. The transgenic *N. benthamiana* line SGT662 was a gift from Peter Moffett (University of Sherbrooke). Each experiment was repeated at least three times and a representative experiment is shown.

3.2.4 Western blot analysis of proteins

Western blots were performed as described (11). The commercially available monoclonal anti-HA-peroxidase antibody (Sigma-Aldrich), anti-His(C-term)-HRP antibody (Invitrogen) and GFP monoclonal antibody (Clontech) were used as described by the manufacturer. Ponceau S staining of the large RUBISCO subunit was used as a loading control for the western blots.

3.2.5 Isolation and detection of mRNAs

Total RNAs were extracted from the agroinfiltrated leaf samples and subjected to Northern blotting as described (260). The probe was designed against the N-terminal region of the ToRSV polymerase domain and was amplified using the primers 5'-TTTATCGCATCACCG TACCACGC-3' and 5'-GAATCGCGCATATGGCGTACG-3'. Ethidium bromide staining of rRNAs was used as loading control for the Northern blot.

3.2.6 Expression and purification of proteins from *E. coli*

Plasmids that encode different forms of the polymerase were expressed in *E. coli* BL21 (DE3) as previously described (129). The cells were collected by centrifugation. The pellets were resuspended in resuspension buffer (50 mM Tris, pH 8; 0.3 M NaCl, 0.1% Triton X-100, 2 mM EDTA) and sonicated. The cell lysates were mixed with 2 X PLB and proteins were detected by Western blotting using the His antibody.

3.2.7 Inhibitor treatment

Inhibitor treatments were conducted as previously described with minor changes (266). Agroinfiltrated leaves were detached at 2 dpa and the petioles were placed in various inhibitor solutions for 24 hours. Inhibitor E-64 (Sigma-Aldrich), MG123 (Sigma-Aldrich), caspase inhibitor (Z-VAD-FMK, Promega), E-64d (Sigma-Aldrich) were used at final concentrations of 10 µM, 100 µM, 100 µM and 20 µM, respectively. For inhibitors that needed to be resuspended in DMSO, a corresponding diluted DMSO solution was used as a negative control.

3.3 Results

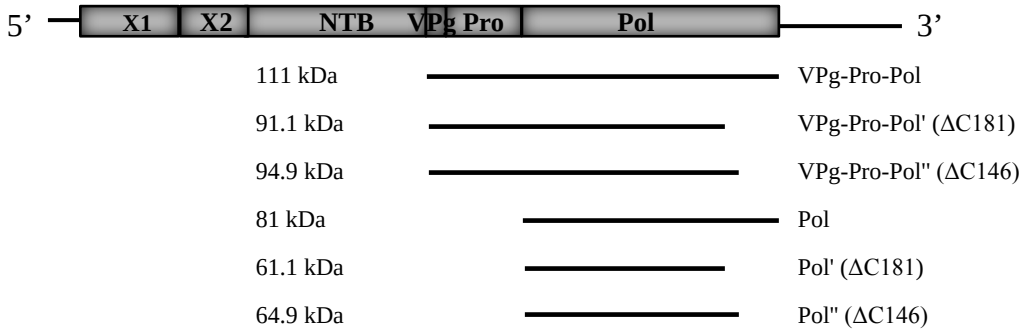
3.3.1 Expression of different forms of the ToRSV polymerase in *N. benthamiana*

As discussed in Chapter 1, the full-length Pol and the C-terminally truncated VPg-Pro-Pol′ are detected in ToRSV infected plants. However, the exact site of truncation is still unknown. Alignment with related viral polymerases shows that the ToRSV Pol has a 15-20 kDa C-terminal extension when compared to FCV and PV polymerase (13). To investigate the stability of the ToRSV polymerase, various forms of the ToRSV Pol were ectopically expressed in *N. benthamiana* plants. The full-length VPg-Pro-Pol (110 kDa) and Pol (81

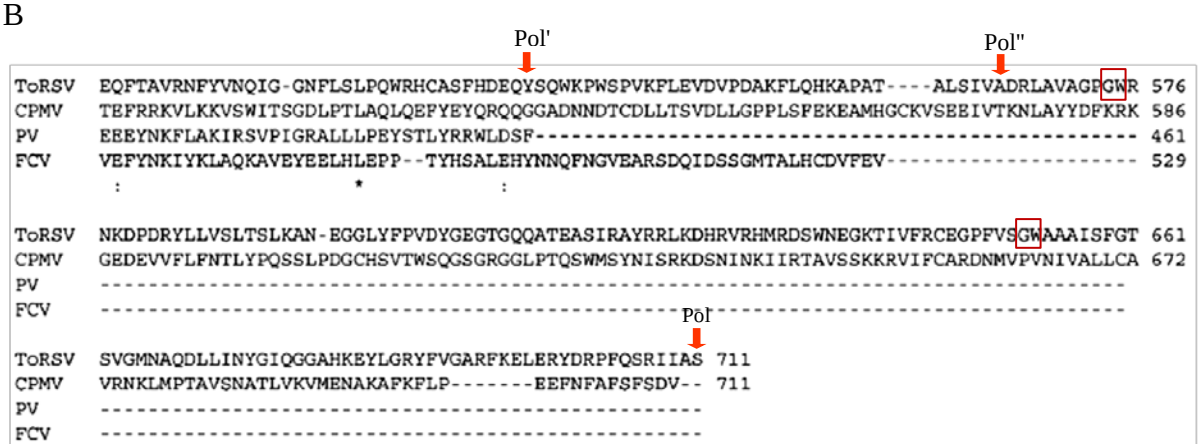
kDa) contain the natural C-terminal end of the Pol (Fig. 3.1A and B), whereas VPg-Pro-Pol′ and Pol′ were designed with their C-terminal end corresponding to the end of the PV polymerase in the sequence alignment (Fig. 3.1A and B) (13). These two truncated proteins lack 181 amino acid residues (approximately 20 kDa) from the C-terminus of the ToRSV Pol (Fig. 3.1A). Similarly, the Pol″ and VPg-Pro-Pol″ lack 146 amino acids (approximately 15 kDa) from the C-terminus of the ToRSV Pol. This truncationwas designed based on the estimated size of the VPg-Pro-Pol′ polyprotein that has been detected *in vivo* in infected plants. (Fig. 3.1A and B). Agroinfiltration of these constructs to express proteins in *N. benthamiana* leaves was performed as described in Material and Methods. The TBSV p19, which is a gene silencing suppressor (284), was co-agroinfiltrated with the polymerases to prevent the onset of RNA silencing and thereby enhance protein expression. Leaf samples were collected at 3 dpa and the steady-state levels of proteins and mRNAs were examined (Fig. 3.1C). The full-length VPg-Pro-Pol-HA was usually not detected (lane 2), although it was occasionally detected in very low amount (Fig. 3.3). In contrast, the VPg-Pro-Pol′-HA and VPg-Pro-Pol″-HA were easily detected in plants (lane 3 and 4). Similarly, the full-length Pol-HA protein (lane 5) was detected at a much lower level than Pol′-HA and Pol″-HA (lane 6 and 7). To eliminate the possibility that the C-terminal HA tag influences the stability of the ToRSV Pol, versions of the polymerase fused with an N-terminal HA tag were also expressed in plants. Similar to what was observed above with the C-terminal HA tag, the full-length HA-VPg-Pro-Pol and HA-VPg-Pro^{HD}-Pol (with the H to D mutation in the catalytic triad to abolish the protease activity) did not accumulate to detectable levels (lane 8 and 9). However, the truncated protein HA-VPg-Pro-Pol′ and HA-VPg-Pro ^{HD}-Pol′ were detected in high amounts (lane 10 and 11). The VPg-Pro cleavage products were also

detected in extracts derived from plants expressing HA-VPg-Pro-Pol and HA-VPg-Pro-Pol′ (lane 8 and 10, see also Fig. 3.4, lane 4), but not HD mutants which include a mutation of the histidine in the catalytic triad of the protease that inactivate the protease (253) (lane 9 and 11). This result strongly suggests that the full-length HA-VPg-Pro-Pol protein was initially expressed at least to some extent but did not accumulate (lane 8). The release of the VPg-Pro product was incomplete and accumulation of the polyprotein HA-VPg-Pro-Pol′ was evident (lane 10). In contrast, the full-length HA-VPg-Pro-Pol was not detected. Truncated product corresponding to HA-VPg-Pro-Pol′ was not detected either (lane 8). To eliminate the possibility that the variation in protein accumulation levels was due to different accumulation of the mRNAs in plants, I also conducted northern blots of full-length and truncated forms of the polymerase mRNAs. A probe was designed against the N-terminal region of the polymerase domain that is present in all constructs. Although resolution of the bands was somewhat obscured by the co-migration of rRNAs, preliminary results suggest that the mRNA levels were similar for all expressed proteins in spite of the significant differences observed in protein levels except lane 3 and 9 where the levels of RNA are quite low along with the protein products (Fig. 3.1 B). Taken together, the results suggest that the full-length ToRSV Pol and VPg-Pro-Pol are not stable when expressed in *N. benthamiana*. However, truncation of 15 or 20 kDa at the C-terminus of Pol significantly stabilized the proteins.

A


5’
X1
X2
NTB
VPg
Pro
Pol
3’
111 kDa
VPg-Pro-Pol
91.1 kDa
VPg-Pro-Pol' (ΔC181)
94.9 kDa
VPg-Pro-Pol" (ΔC146)
81 kDa
Pol
61.1 kDa
Pol' (ΔC181)
64.9 kDa
Pol" (ΔC146)

B


Pol'
Pol"
ToRSV EQFTAVRNFYVNQIG-GNFLSLPQWRHCASFHDEQVSQWKPWSPVKFLEVDVPDAKFLQHKAPAT----ALSIVADRLAVAGPGWR 576
CPMV TEFRRKVLKKVSWITSGDLPTLAQLQEPYEYQRQQGGADNNDTCDLLTSVDLLGPPLSFEKEAMHGCKVSEEIVTKNLAYYDFKRK 586
PV EEEYNKFLAKIRSVPIGRALLLPEYSTLYRRWLDSF---------- 461
FCV VEFYNKIYKLAQKAVEYEELHLEPP--TYHSALEHYNNQFNGVEARSDQIDSSGMTALHCDVFEV---------- 529
ToRSV NKDPDRYLLVSLTSLKAN-EGGLYFPVDYGEGTGQQATEASIRAYRRLKDHRVRHMRDSWNEGKTIVFRCEGPFVSGWAAAISFGT 661
CPMV GEDEVVFLFNTLYPQSSLPDGCHSVTWSQGSGRGGLPTQSWMSYNISRKDSNINKIIRTAVSSKKRVIFCARDNMVPVNIVALLCA 672
PV ----------
FCV ----------
Pol
ToRSV SVGMNAQDLLINYGIQGGAHKEYLGRYFVGARFKELERYDRPFQSRIIAS 711
CPMV VRNKLMPTAVSNATLVKVMENAKAFKFLP-------EEFNFAFSFSDV-- 711
PV ----------
FCV ----------

C

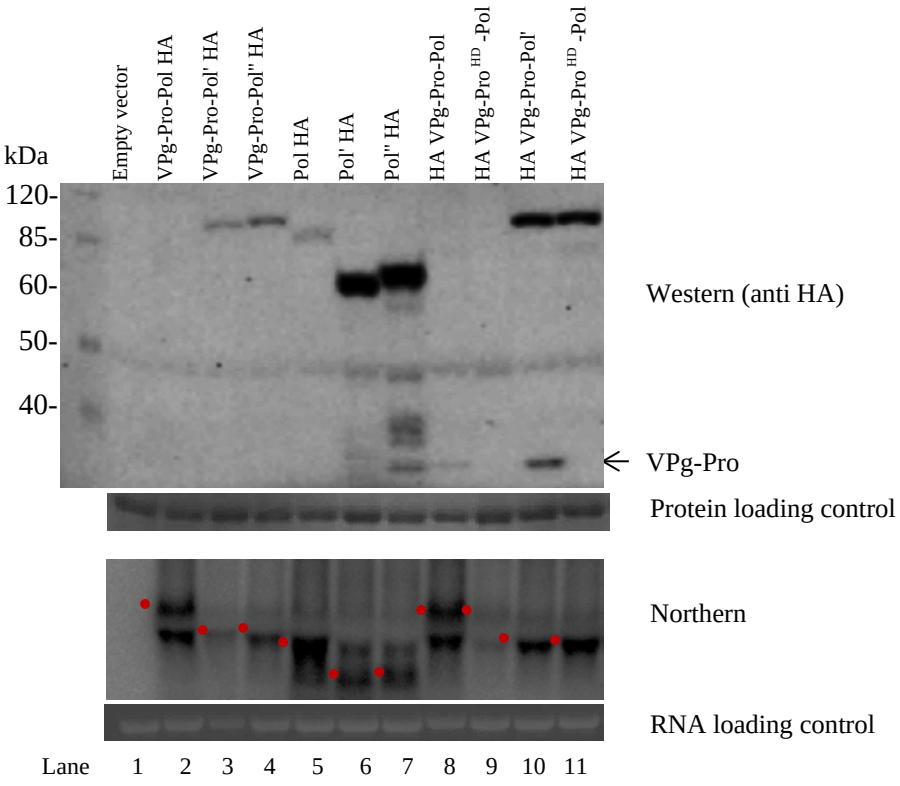

kDa
Empty vector
VPg-Pro-Pol HA
VPg-Pro-Pol' HA
VPg-Pro-Pol" HA
Pol HA
Pol' HA
Pol" HA
HA VPg-Pro-Pol
HA VPg-Pro HD -Pol
HA VPg-Pro-Pol'
HA VPg-Pro HD -Pol
120-
85-
60-
50-
40-
Western (anti HA)
VPg-Pro
Protein loading control
Northern
RNA loading control
Lane 1 2 3 4 5 6 7 8 9 10 11

Figure 3.1: Ectopic expression of different forms of the ToRSV polymerase in *N. benthamiana*. A. Schematic representation of the ToRSV RNA1 and the coding region of various forms of the polymerase used in this study. B. Amino acid alignment of the polymerase from ToRSV, CPMV, FCV and NV. Please note that only the C-terminal region is shown. The C-terminal end of Pol, Pol' and Pol" are indicated with red arrows. Two GW motifs of the ToRSV Pol are shown in boxes. C. Western and northern blots of ectopically expressed forms of the ToRSV polymerase in *N. benthamiana*. The HA tag was fused to the C-terminus (lane 2-7) or N-terminus (lanes 8-11) of the proteins. Expression levels of these proteins in *N. benthamiana* were detected with HA antibody (Sigma). A sample from plants infiltrated with the empty vector was used as a control (lane 1). The VPg-Pro cleavage product released from the HA-VPg-Pro-Pol and HA-VPg-Pro-Pol′ is shown with an arrow. In the northern blot, expected migration of the mRNAs is shown with the red dots. Migration of the bands was partially hindered by co-migration with an abundant rRNA band (white shadow).

3.3.2 Expression of all forms of ToRSV polymerase in *E. coli*

VPg-Pro-Pol, VPg-Pro-Pol′, VPg-Pro-Pol″, Pol, Pol′ and Pol″ were fused to a C-terminal His tag and expressed in *E. coli* BL21 (DE3). Proteins were detected with the His antibody (Fig. 3.2). All forms of the polymerase were successfully expressed with the expected size, although some differences of expression were observed between the full-length Pol (lane 8) and the truncated Pol′ and Pol″ (lane 9 10). The full-length VPg-Pro-Pol (lane 2) and the VPg-Pro^{HD}-Pol (lane 3) were expressed at similar levels when compared with VPg-Pro-Pol′ (lane 4) and VPg-Pro-Pol″ (lane 6) as well as their HD mutants (lane 5 and 7), suggesting that the full-length VPg-Pro-Pol is as stable as the C-terminally truncated forms in *E. coli*. Together with the above results, I conclude that the instability of ToRSV VPg-Pro-Pol and Pol is specific to plant expression, and this instability might be caused by a plant protease(s) or by one of the plant degradation pathways.

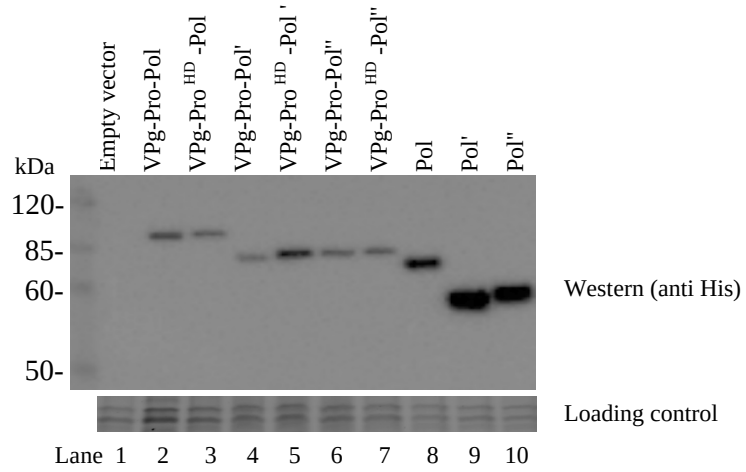


Figure 3.2: Expression of different forms of ToRSV polymerase in *E. coli*. Different forms of the His tagged polymerases and their HD mutants were expressed in *E. coli* BL21 (DE3). Proteins were analyzed using SDS-PAGE and detected with His antibody. The empty vector control is shown in lane 1.

3.3.3 Treatment of agroinfiltrated leaves with inhibitors of protease or of the proteasome did not enhance the expression levels of full-length Pol or VPg-Pro-Pol

To investigate what causes the instability of the full-length polymerase, I tested a number of inhibitors of plant proteolytic pathways. According to my experience, transiently expressed proteins become detectable to low levels at 2 dpa, with the highest expression level achieved at 3 dpa. The protein accumulation level decreases later on. Therefore, for the inhibitor treatments, agroinfiltrated leaves were detached at 2 dpa, the petioles of the detached leaves were treated in the inhibitor solution for 24 hours and leaf samples were collected for protein analysis. This method has been successfully used in our lab for *in vivo* labelling of newly synthesized ToRSV proteins (260), or to demonstrate that the Ago1 protein is protected from degradation by the E-64d inhibitor (266). Inhibitor treatments were performed in the same manner using the inhibitor E-64 (cysteine proteases inhibitor), MG123 (proteasome pathway inhibitor), Z-VAD-FMK (caspase inhibitor) and the E-64d (thiol protease inhibitor). The western result indicated that the accumulation levels of Pol-HA, HA-Pol or VPg-Pro-Pol-HA were not increased in the presence of any of the inhibitors (Fig. 3.3).

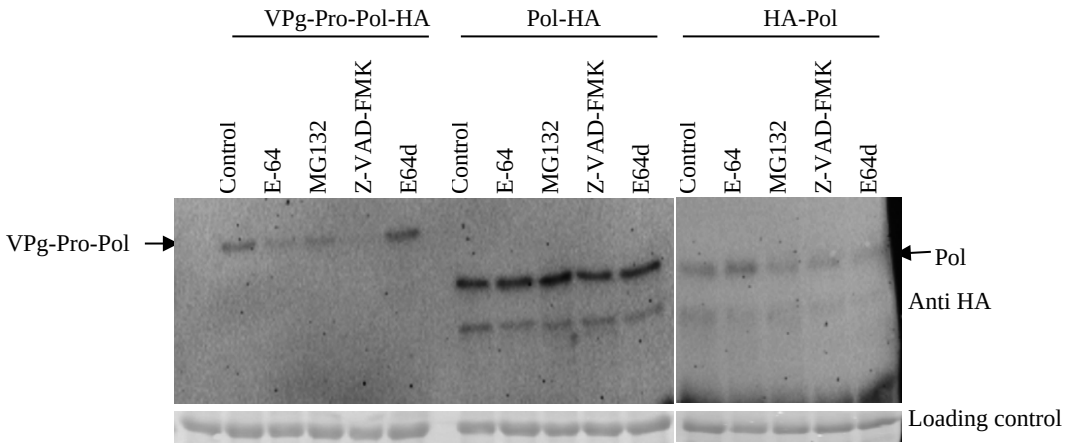


Figure 3.3: Inhibitor treatment of VPg-Pro-Pol and Pol. The E-64, MG123, Z-VAD-FMK and the E-64d were used in the inhibitor assay. H_2O and DMSO were used as a control. The VPg-Pro-Pol and the Pol were indicated with arrow.

3.3.4 Expression of ToRSV Pol is not enhanced in transgenic plants that overexpress the SGT1 co-chaperone

SGT1 (suppressor of G-two allele of SKP1) is a highly conserved protein in yeast, plants and mammalian cells. It has co-chaperone properties and plays an important role in protein folding and stability (285). Recent studies show that SGT1 interacts with molecular chaperones Hsp70 and Hsp90 to stabilize protein and protein complexes (286-289). SGT1 is upregulated in plants infected with PVX or after ectopic expression of the PVX TGBp3 protein (290). Silencing of SGT1 reduces PVX accumulation in *N. benthamiana* plants while overexpression of SGT1 greatly enhances the PVX accumulation (290). Plants silenced for SGT1 showed reduced ToRSV accumulation (Basudev Ghoshal and Helene Sanfaçon, unpublished results). Based on this, I hypothesized that SGT1 may stabilize one or several replication proteins to increase virus replication and viral RNA accumulation. To investigate whether SGT1 could stabilize the ToRSV polymerase, a transgenic *N. benthamiana* line SGT662 that overexpresses SGT1 was used (291). The Pol-HA and HA-VPg-Pro-Pol were agroinfiltrated into the transgenic and the wild type *N. benthamiana* leaves. Samples were collected at 3 dpa and protein levels were detected using the HA antibody (Fig. 3.4). The results showed that accumulation levels of Pol-HA or HA-VPg-Pro-Pol were similar in wild type plants or SGT1 transgenic lines. Therefore, overexpression of the SGT1 co-chaperone did not stabilize the ToRSV VPg-Pro-Pol or Pol proteins.

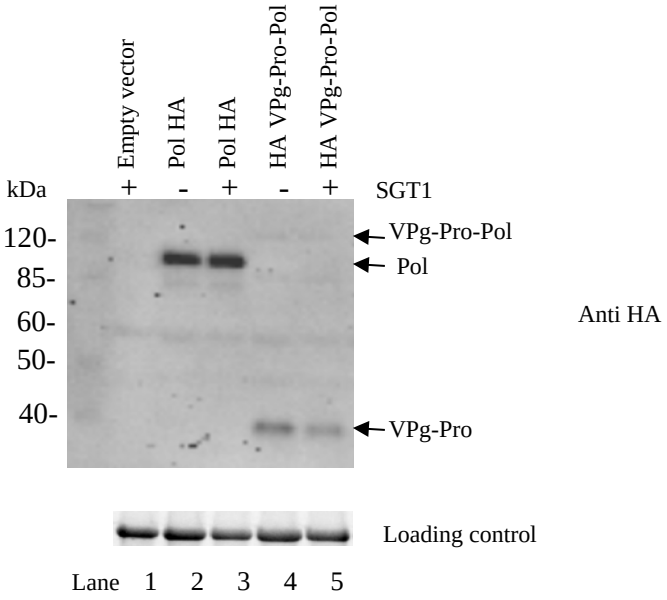


Figure 3.4: Expression of full-length VPg-Pro-Pol and Pol in SGT1 overexpressed plants. The wild type *N. benthamiana* and the SGT1 transgenic line are indicated with (–) or (+) respectively. The expected migration for VPg-Pro-Pol, Pol and the released VPg-Pro are indicated with arrows.

3.3.5 Is the ToRSV polymerase a suppressor of RNA silencing?

RNA silencing is a common host defense mechanism against virus infection. Viral double-stranded RNA replication intermediates or highly structured single strand RNA region in viral RNAs are cleaved by host Dicer-like RNAse III enzyme (DCL) into 21-24 nt small interfering RNA (siRNA) (292). The siRNAs are recruited by Argonaute (AGO) proteins and other host factors to form RNA-induced silencing complexes (RISC), which specifically recognize complementary viral RNAs and lead to their degradation or translation repression (293, 294). To overcome RNA silencing, viruses encode suppressors of RNA silencing (VSRs). VSRs can target and block different steps of the RNA silencing pathway to suppress the plant anti-viral defense (295). The best studied example is the TBSV p19 which binds to siRNA and sequesters these siRNAs to prevent the assembly of the RISC (296). Some VSRs, for example, the turnip crinkle virus P38 protein, directly interacts with AGO protein using Glycine/Tryptophan (GW) motifs (297, 298). This interaction prevents the assembly of AGO into RISC. Interestingly, the C-terminal 15 kDa of ToRSV Pol contains two GW motifs (Fig. 3.1B). This raises the possibility that the full-length Pol, VPg-Pro-Pol or the released 15 kDa fragment obtained after truncation of VPg-Pro-Pol may act as a VSR. To test this hypothesis, the green fluorescence protein (GFP) was co-agroinfiltrated with different forms of ToRSV polymerase in *N. benthamiana* plants. The TBSV p19 was used as a positive control, an empty vector and the ToRSV X2 protein, which is unlikely to be a VSR, were used as negative controls. The expression levels of GFP were monitored at both 3 and 5 dpa (Fig. 3.5). As expected, p19 VSR significantly enhanced the GFP expression level at both 3 dpa and 5 dpa when compared with the negative control (lane 1 and 11). The full-length VPg-Pro-Pol-HA and the HA-VPg-Pro-Pol did not accumulate to detectable levels and they did

not enhance the GFP accumulation level at both time points (lane 2 and 6). The Pol-HA expressed at low levels, but did not increase the GFP accumulation either (lane 4). Truncated proteins VPg-Pro-Pol′-HA, Pol′-HA and HA-VPg-Pro-Pol′ were all expressed to detectable levels but they did not enhance the GFP expression (lane 3, 5 and 7). This is not surprising since these proteins do not contain any GW motifs. The C-terminal 15 kDa fragment of Pol (with two GW motifs) (lane 8) and its mutant GA1+GA2 (with both GW motifs mutated to GA) (lane 9) were also co-infiltrated with GFP. None of them increased GFP accumulation level when compared with the negative control (lane 1 and 10). Taken together, these results indicate that although the ToRSV Pol and the C-terminal 15 kDa contain two GW motifs, they are unlikely to be a VSR like the GW motif containing protein P38. Interestingly, it is noted that mutation of the two GW motifs on the 15 kDa fragment of Pol significantly increased accumulation of this protein (lane 8 and 9).

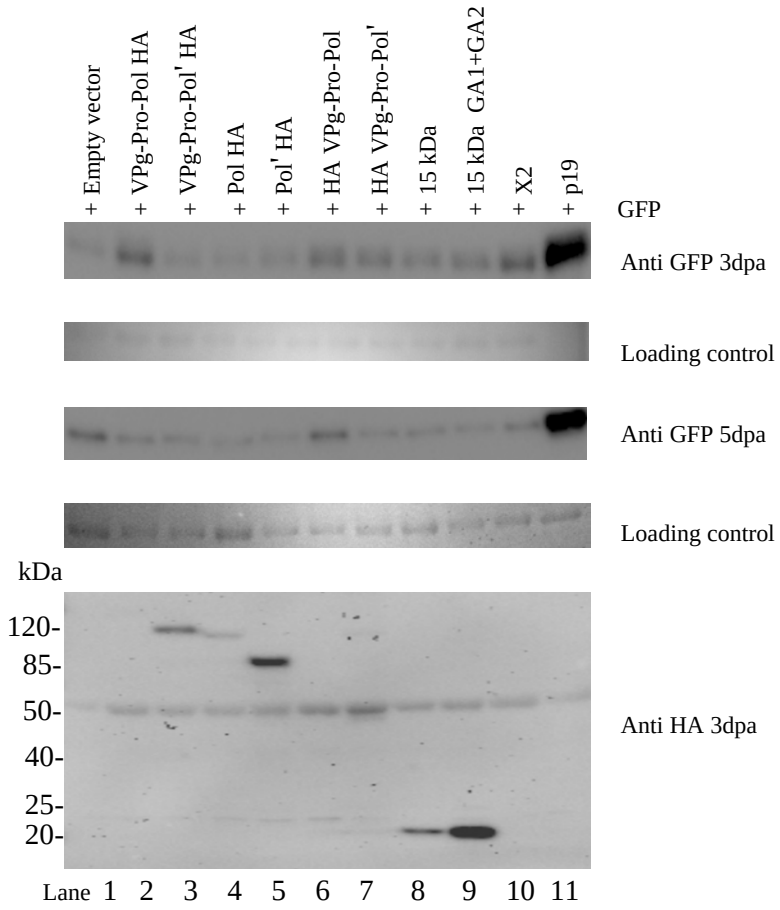


Figure 3.5: Co-expression of ToRSV Pols with GFP in *N. benthamiana* plants. Various forms of the ToRSV polymerase as indicated were co-expressed with GFP. Samples were collected at 3 and 5 dpa. Western blots were developed using GFP antibody or HA antibody as indicated. The empty vector and the X2 protein were co-expressed with GFP to act as negative controls, while the TBSV p19 was used as a positive control.

3.3.6 Mutation of the two GW motifs did not stabilize the full-length Pol

As mentioned above, mutation of the two GW motifs to GA seemed to stabilize the C-terminal 15 kDa fragment of Pol. This was reproducibly observed in three repeats of the experiment. Therefore, I decided to test whether these mutations would also stabilize the full-length Pol. The GW to GA mutation was introduced to both Pol-HA and HA-Pol. The wild type proteins and the GA mutant derivatives were transiently expressed in *N. benthamiana* plants. Samples were collected at 3 dpa and expression level of the Pol and the GA mutants were detected with HA antibody. Surprisingly, the results were opposite to those observed with the small fragment. Instead of stabilizing the Pol, the GA mutations enhanced the degradation of the full-length protein and caused the accumulation of degradation fragments.

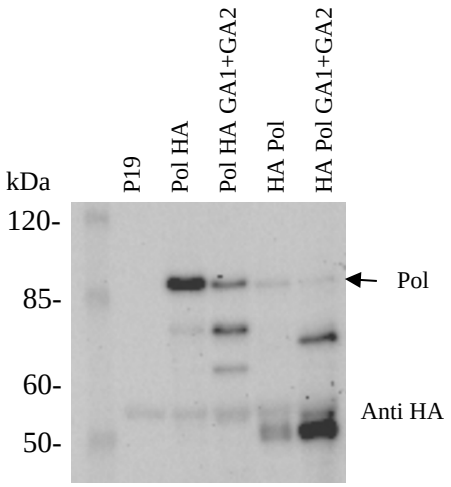


Figure 3.6: Expression of the full-length Pol and its GW motif mutants. The two GW motifs were mutated to GA on the full-length Pol-HA and HA-Pol. The wild type proteins and the GA mutants were expressed in *N. benthamiana* plants. Proteins were extracted and detected with HA antibody. The expected migration for the full-length Pol is indicated with the arrow.

3.4 Discussion

In this study, I transiently expressed various forms of the ToRSV polymerase in *N. benthamiana* plants (Fig. 3.1). The detected mRNA levels of these constructs were similar in plants. However, accumulation of the proteins varied from one to another. The full-length VPg-Pro-Pol and Pol were detected in low amounts, if at all, whereas C-terminally truncated forms of these proteins stably accumulated in plants (Fig. 3.1). This phenomenon is also observed in ToRSV infected plants in which only the VPg-Pro-Pol′ is consistently detected (13). Detection of the released VPg-Pro from the wild type HA-VPg-Pro-Pol in plants indicated that this full-length protein was initially expressed but did not accumulate. Mutation of the protease catalytic triad (HD mutant) of VPg-Pro-Pol did not stabilize the full-length protein, indicating that the instability was not caused by the viral-encoded protease, at least in the transient expression experiments. In contrast with what was observed *in vivo*, the full-length VPg-Pro-Pol and Pol accumulated to similar amounts as their truncated forms when expressed in *E. coli* (Fig. 3.2). Taken together, these results indicate that the full-length ToRSV VPg-Pro-Pol and Pol are unstable when compared to the C-terminally truncated forms of polymerase. This instability is plant specific and probably caused by a plant-derived protease or plant degradation pathway.

A number of protease inhibitors, which target different types of proteasesor the ubiquitin proteasome pathway, were used to investigate the stability of VPg-Pro-Pol and Pol. Unfortunately, none of the inhibitors stabilized these proteins (Fig. 3.3). It is possible that the instability of VPg-Pro-Pol and Pol was caused by an alternative degradation pathway, which was not hindered by the inhibitors tested. I cannot completely exclude the possibility that the

inhibitors were not taken up properly in the detached leaf assay, although this assay has been used successfully in our lab (260, 266) (see also cycloheximide treatments in Chapter 2). As well, overexpression of the co-chaperone SGT1 did not help to stabilize the VPg-Pro-Pol (Fig. 3.4). Therefore, the host factors or host protease(s) that are involved in the instability of VPg-Pro-Pol and Pol are still unclear.

I also tried to investigate the biological functions of the full-length VPg-Pro-Pol, Pol and the C-terminal 15 kDa fragment. Results indicated that although these fragments have two GW motifs, they do not have detectable silencing suppression activity. However, I can not eliminate the possibility that the lack of silencing suppression activity is due to the very low expression levels of the full-length proteins (Fig. 3.5).

The presence of different forms of the polymerase (Pol or VPg-Pro-Pol) is common in picorna- and picorna-like viruses, but no C-terminally truncated forms of polymerase have been reported in these viruses, including the closely related GFLV (242). The fact that no C-terminal truncation was observed on VPg-Pro-Pol and NTB-VPg-Pro-Pol *in vitro* suggests it is probably caused by plant factors (13). I did not detect C-terminally truncated fragments of HA-VPg-Pro-Pol after transient expression of this protein *in vivo* (Fig. 3.1). It is possible that the full-length protein is immediately degraded under these conditions, preventing the detection of the truncated fragments. Alternatively, it is possible that a plant protease responsible for the C-terminal truncation of ToRSV VPg-Pro-Pol is only induced, or activated in the context of a virus infection but not after ectopic expression of the protein.

As the key enzyme in virus replication, the concentration of the polymerase is highly regulated either at the translational level or post-translational level (77, 299). For example, various gene expression strategies are used by viruses to produce and regulate the accumulation of their polymerase (76, 300). Tobamoviruses use the ribosome termination/ read-through mechanism to produce the 126K and 183K (polymerase) proteins at the ratio of 10:1 (301). The same strategy is used by tombusvirus and the read-through efficiency is 5%, leading to lesser amounts of $P92^{Pol}$ than the P33 (302). Bromoviruses have a multicomponent RNA genome, the RNA2-encoded $2a^{Pol}$ expressed 25 fold less efficiently than the RNA1-encoded 1a protein (303). In addition, barley yellow dwarf virus Pol is expressed via a -1 frameshift with an efficiency of about 1-2% (304). Post-translational down-regulation has been observed in a number of viruses, the best-studied example is TYMV Pol, which is targeted and degraded by the host ubiquitin-proteasome pathway (299). The HAV and HCV polymerases are also poly-ubiquitinated and are rapidly degraded by the 26S proteasome (305, 306). Reduced accumulation of the polymerase generally results in a lower copy number of the polymerase within the replication complexes compared to the copy number of viral membrane proteins. The ToRSV polymerase is produced by a polyprotein strategy and is initially expressed at the same level as other viral-encoded proteins (such as the membrane replication proteins). The instability of ToRSV VPg-Pro-Pol and Pol may provide an efficient way of regulating the polymerase concentration. Regulated viral replication rates can benefit viruses by keeping the viral genome integrity (i.e., limiting mutation and recombination) and reducing the chance of being recognized by host defense response (299, 307).

Chapter 4

Investigation into the activity of different forms of the ToRSV polymerase

4.1 Introduction

The presence of more than one form of polymerase is common in viruses that use a polyprotein strategy. The active form of the polymerase could be the mature polymerase, a polyprotein that contains the polymerase domain or both forms of the polymerase (as discussed in section 1.5.2.4). For ToRSV, whether VPg-Pro-Pol′ and/or Pol are the active form of the polymerase and whether truncation of the C-terminal 15 kDa modulates the activity of the full-length VPg-Pro-Pol are still unknown. In this study, I examined the activity of various forms of the ToRSV polymerase using the *E. coli* purified recombinant polymerases and *in vitro* polymerase activity assays. Unfortunately, I was unable to detect any polymerase activity with this approach. The possible reasons for this are discussed in detail.

4.2 Materials and methods

4.2.1 Plasmid constructs

The plasmids pET VPg-Pro^{HD}-Pol, pET VPg-Pro^{HD}-Pol′, pET VPg-Pro^{HD}-Pol″ and pET Pol were already described in Chapter 3. I also introduced the GDD-GAD mutation to the pET Pol by using primers P84F (5’- GGAGGTGTGTTTGATTGTTTATGGTGCTGATATTTAA TTTCTATTAAGCCGG-3’) and P85R (5’- CCGGCTTAATAGAAATTAAATTATCAGC ACCATAAACAATCAAACACACCTCC-3’). The mutagenesis was performed as described in Chapter 3.

4.2.2 Expression and purification of recombinant polymerases from *E. coli*

Plasmids were transformed into *E. coli* BL21 (DE3). Cells were grown at 37 °C to an OD600 of 0.3 and then induced with 1 mM IPTG at 12 °C for overnight. Cells were harvested by centrifugation at 4000g for 10 minutes at 4 °C. For purification, the pellets were dissolved in lysis buffer (50 mM phosphate buffer, PH 8.0, 300 mM NaCl, 5 mM imidazole, 10% glycerol) and cells were lysed using the French Press. The disrupted cells were centrifuged at 20000g for 20 minutes to eliminate the insoluble fractions. The supernatant was filtered (0.2 µm) and added to the Talon (Clontech, USA) resin for purification as described by manufacturer. The purified proteins were dialyzed and concentrated in the storage buffer (20 mM potassium phosphate buffer (pH 7.7), 1 mM EDTA, 10 mM DTT, 0.1 M NaCl, 0.1% Triton® X-100 and 50% (v/v) glycerol.). All purified proteins were aliquot and stored at -80 °C.

4.2.3 Preparation of RNA templates and primers

The RNA oligonucleotides 5'- UCUCUCUCUCUCUCUCUCUCAUAACUUAAUCUCA CAUAGC-3' and 5' GCUAUGUGAGAUUAAGUUAU-3', as well as the (U)12 were ordered from Invitrogen. The poly(A) oligomer was purchased from Roche. ToRSV viral RNA was purified as previously described (see Chapter 3).

4.2.4 End-labelling of primer and template

The primer and template were end-labeled with [γ-^{32}P]-ATP using the 5' End Labelling Kit (GE healthcare) as described by manufacturer.

4.2.5 Polymerase activity assay

To test the activity of different forms of ToRSV Pol, 250 ng of purified polymerase were added to the polymerase assays which contained 1 µM template, 0.2 µM primer, 0.17 µM α-^{32}P ATP (10 µCi/µl), 50 µM ATP, 50 µM GTP, 20 mM Tris-HCl (pH 8.0), 10 mM KCl, 6 mM $MgCl_2$ and/ or 6 mM $MnCl_2$, 0.01% Triton X-100, 1 mM DTT, 1 mM rifampin and 0.5 U RNAguard (Amersham). Experiments were carried out at room temperature for two hours and the reactions were stopped by adding an equal volume of 2×siRNA loading buffer (90% formamide, 50 mM EDTA pH 8.0, 0.01% bromophenol blue and xylene cyanol). Polymerase activity products were analyzed using 8 M urea, 15 % PAGE gel, and detected autoradiographed.

For activity assays using poly(A) and (U) 12 as primer, experiments were performed as above with minor changes. UTP and [α-^{32}P]-UTP were added to the reactions instead of ATP and [α-^{32}P]-ATP. For activity assays using viral RNA template, reactions were performed in the presence of ATP, UTP, GTP, CTP as well as [α-^{32}P]-UTP, reactions were stopped by adding 2×RNA loading buffer (95% formamide, 0.025% SDS, 0.025% bromophenol blue, 0.025% xylene cyanol, 0.5 mM EDTA). The RNA products were analyzed on denaturing 1% agarose gels containing formaldehyde.

4.3 Results

4.3.1 Expression and purification of ToRSV polymerases in *E. coli*

Constructs encoding ToRSV VPg-Pro^{HD}-Pol, VPg-Pro^{HD}-Pol′ VPg-Pro^{HD}-Pol″ and Pol were transformed and expressed in *E. coli* BL21 (DE3). Each fragment contained a hexahistidine

tag at the C-terminus for affinity purification of the protein. The cultures were grown at low temperature (12 °C) to avoid protein aggregation. Collected cells were disrupted by French Press in the presence of protease inhibitor. Only the soluble fractions of the cell lysate were used for protein purification. The purified proteins were dialyzed into storage buffer and then analyzed by SDS-PAGE and Coomassie brilliant blue staining. The result indicated that the VPg-Pro^{HD}-Pol, VPg-Pro^{HD}-Pol′ and Pol were successfully expressed at the expected sizes and these proteins were purified to near homogeneity (Fig. 4.1). In addition to the above mentioned proteins, the Pol-GAD, which contains the GDD-GAD mutation in the conserved GDD motif, was also expressed and purified (Fig. 4.1). It has been demonstrated that the first aspartic acid of the GDD triplet is strictly required for polymerase activity (308), mutation of this aspartic acid totally abolishes the enzymatic activity of the polymerase (309). Therefore this mutant was used as a negative control for the activity assays. All purified proteins were stored at -80 °C in aliquotes to avoid repetitive freezing and thawing, which may cause inactivation of the polymerases.

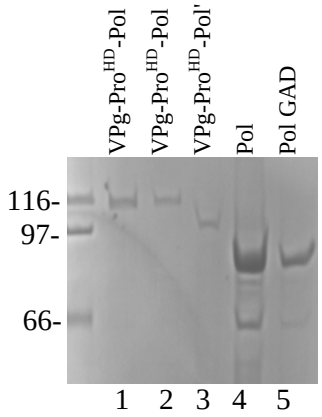


Figure 4.1: SDS-PAGE analysis of purified ToRSV polymerases from *E. coli* . Different forms of purified recombinant ToRSV polymerase were analyzed with 10% SDS-PAGE and Coomassie brilliant blue staining. For VPg-Pro^{HD}-Pol, two different batches of purified protein were analyzed (lane 1 and 2).

4.3.2 Activity assay of purified ToRSV polymerases

As the key enzyme for virus replication, biochemical properties of viral polymerases have been extensively studied. In many cases, the Pol preparations expressed in heterologous systems share identical properties with those of polymerases purified from infected cells (310, 311). Using heterologously expressed and purified Pols and *in vitro* activity assays is a well-established approach for studying the activity and biochemical properties of viral polymerases. In this approach, viral Pols are predicted to catalyze RNA synthesis on a given RNA template in the appropriate environment (Fig. 4.2A). RNA synthesis initiation can be *de novo* (312) or primer dependent (311). Adding [α-^{32}P]-NTP to the reactions allow the newly synthesized products to be radioactively-labeled and easily detected by autoradiography. Unlike the *in vivo* replication, *in vitro* replication assays lack host factors that are often important for specific selection of RNA templates (72, 313, 314). Therefore, the purified polymerases generally do not exhibit viral RNA template-specificity *in vitro* and are capable of copying exogenously added non-viral RNA templates (315).

To identify the active form of the ToRSV polymerase, the above purified polymerases were tested in the *in vitro* activity assays as described in Materials and Methods. An exogenous RNA template was used in the initial assays (316) (Fig. 4.2B). A primer was also added to the reaction since the related PV Pol catalyzes the synthesis of RNA *in vitro* in a primer-dependent manner (311). In addition, by analogy with other picorna-like viruses, *in vivo* replication of ToRSV is likely to require the uridylylated VPg as a primer. Activity assays were performed at room temperature for 2 hours and ^{32}P- labeled products were analyzed on 8 M-urea denatured 15 % PAGE gel (Fig. 4.2B). According to this initial assay, a

radioactively-labeled band was detected when using the purified full-length VPg-Pro^{HD}-Pol (lane 1) and Pol (lane 3) in the assay, whereas no band was observed when the purified VPg-Pro^{HD}-Pol′ ((lane 2) or Pol-GAD (lane 4) mutant was tested. Further experiments were conducted to determine whether this band is an extension product resulting from the polymerase activity of the purified ToRSV VPg-Pro-Pol and Pol.

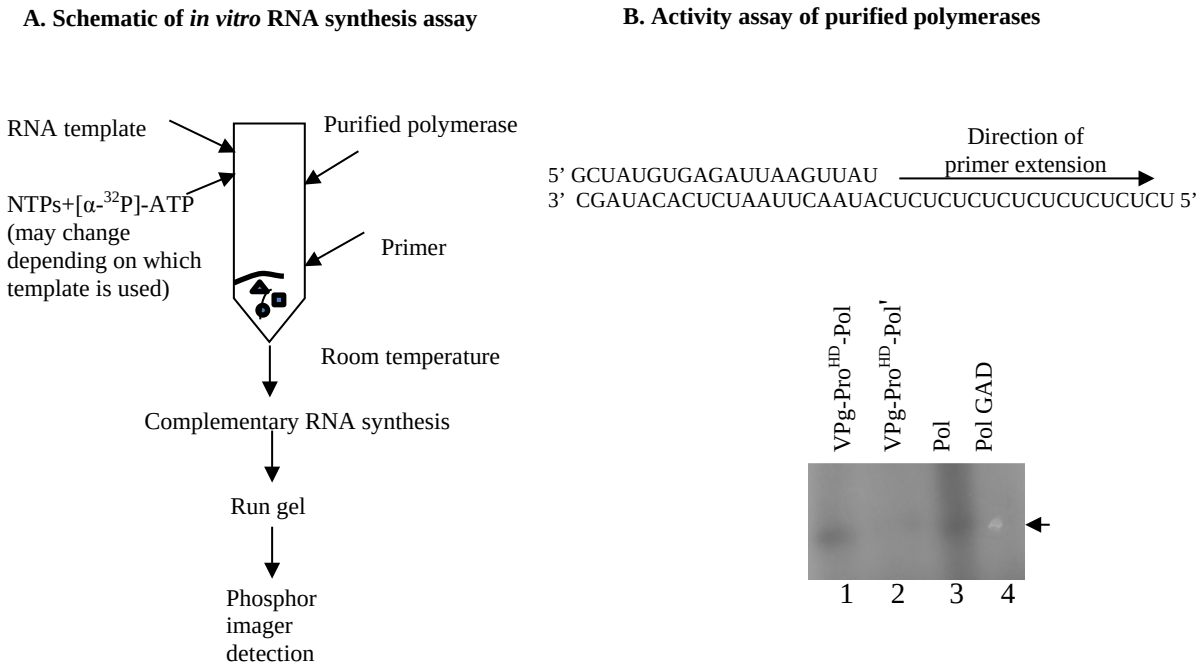


Figure 4.2: Activity asssay of purified ToRSV polymerases and the active-site mutant. A. Schematic of the *in vitro* RNA activity assay. B. RNA products synthesized in the presence of various ToRSV polymerase. Primer and template used in this assay are indicated. Products (shown with arrow) are analyzed on 8M -urea denatured 15 % PAGE gel and detected by autoradiography.

The primer and template were end-labeled with [γ-^{32}P]-ATP and used as markers to help identify the radioactively-labeled bands observed in the initial assays. The expected polymerase extension product should be of the same size as the labeled template, although some purified polymerase have been shown to extend products slightly longer than the template due to terminal transferase activity (317, 318). The commercially available T7 polymerase (a DNA-dependent RNA polymerase, Promega) was used as a positive control. Similar to what has been observed in the initial assays, a product was observed when different concentration of the ToRSV Pol were added to the activity assay (Fig. 4.3A). The size of the product was similar to the template as well as the radioactively-labeled product synthesized in the presence of the T7 polymerase (Fig. 4.3A and B). In the absence of template and primer, both purified Pol or the Pol-GAD mutant failed to synthesize a product as expected (Fig. 4.3B, lane 2, 5). To further test whether the product observed after incubation with the ToRSV Pol is the result of extension due to the polymerase activity, or to a terminal transferase activity, the experiment was repeated as above but excluding GTP (Fig. 4.3B). Because the extension product is expected to include GA repeats (Fig. 4.2B), the absence of GTP should prevent elongation of a product synthesized from the template. However, the radioactive band was detected in the presence or in the absence of GTP (Fig. 4.3B, lane 3). In addition, the same band was observed when a new batch of purified Pol-GAD was added to the reaction (lane 4), which differs from my initial results (Fig. 4.2). I repeated this experiment several times with different batches of purified Pol-GAD (data not shown) and was convinced that the radioactively-labeled band was produced in the presence of Pol-GAD. Therefore, the band observed was not due to the polymerase activity. However, it is also unlikely caused by a terminal transferase activity from the purified Pol either, since

both activities should be abolished in the GAD mutant (312). Taken together, the band detected was more likely a product of a contaminating activity present in the *E. coli* extracts.

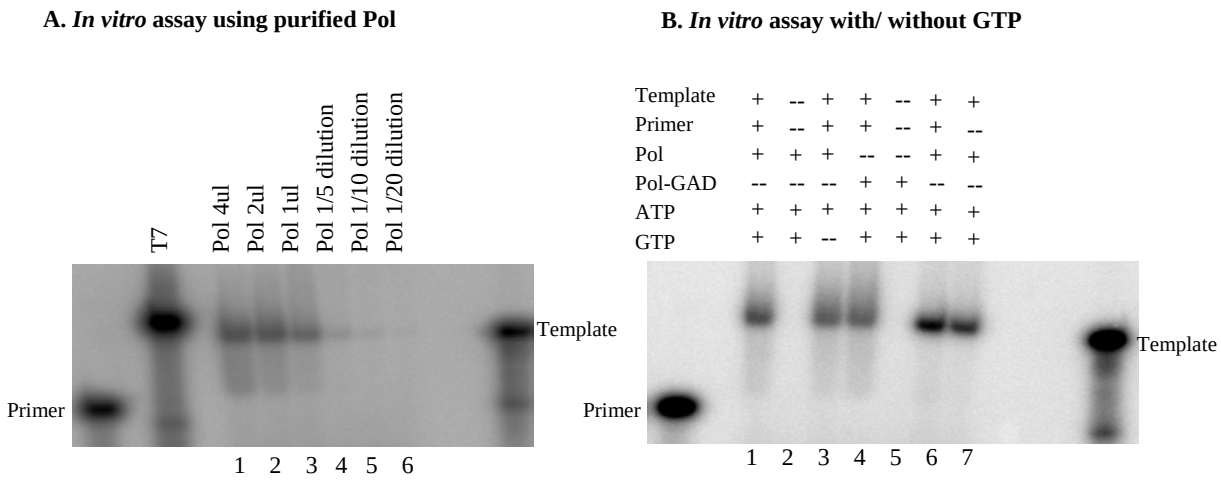


Figure 4.3: Polymerase activity asssay using purified ToRSV polymerase. A. *In vitro* polymerase activity assay using different concentation of purified Pol. B. *In vitro* polymerase activity assay in the presence or absence of GTP. T7 polymerase is used as a positive control, radioactively-labeled primer and template are shown as markers. Products were analyzed on a 8 M-urea denatured 15 % PAGE gel and auto radiographed using phosphor screen.

4.3.3 Activity assay using crude *E. coli* extracts

Since no polymerase activity was detected above and to eliminate the possibility that the ToRSV polymerase activity is lost during purification, I performed activity assays using crude *E. coli* extracts (the filtered supernatant after the French Press step). Polymerase activity has been detected in crude *E. coli* extracts for various viral polymerases (315, 319). However, no synthesized products were detected after incubation with crude extracts of *E. coli* cells expressing various forms of ToRSV Pol (Fig. 4.4). Only background levels were observed which were similar to those observed for cells transformed with the empty vector or expressing the Pol-GAD mutant (lane 3-8). To determine whether the crude extracts contain factors that inhibit polymerase activity or degrade the synthesized products, I also mixed the T7 polymerase with the crude extracts derived from cells that express the empty vector. Interestingly, the amount of radioactivity-labeled products synthesized by the T7 polymerase was significantly reduced, suggesting that something in the crude extracts did inhibit the T7 polymerase activity (lane 2 and 3).

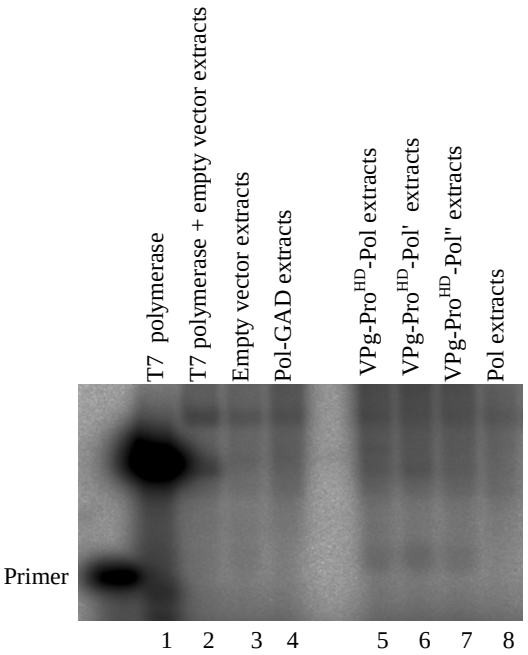


Figure 4.4: Activity asssay using crude extracts of ToRSV polymerase. Products are analyzed on 8M-urea denatured 15 % PAGE gel. T7 polymerase is used as a positive control.

4.3.4 Activity assay using poly(A) and (U)12

The template used in the above experiments was not originally derived from the ToRSV genome. Therefore it is possible that the absence of polymerase activity was caused by an inappropriate template in the activity assay. The ToRSV genomic RNAs are polyadenylated and replication of genomic RNA is likely primed by uridylated VPg. In addition, a number of viral polymerases, including the related PV are capable of copying poly(A) template in the presence of oligoU primer (320, 321). Therefore, activity assays were performed using poly(A) template and (U)12 primer in the presence of purified ToRSV polymerase (Fig. 4.4A) or the crude extracts (Fig. 4.4B). Unfortunately, no detectable polymerase activity was observed. I attempted different reaction conditions (pH, divalent ion, reaction temperature, concentration of polymerase) but with no success (data not shown). Experiments were also performed using purified ToRSV genomic RNA as a template and (U)12 as a primer. Similarly, none of the purified polymerases or crude *E. coli* extracts displayed any activity (data not shown).

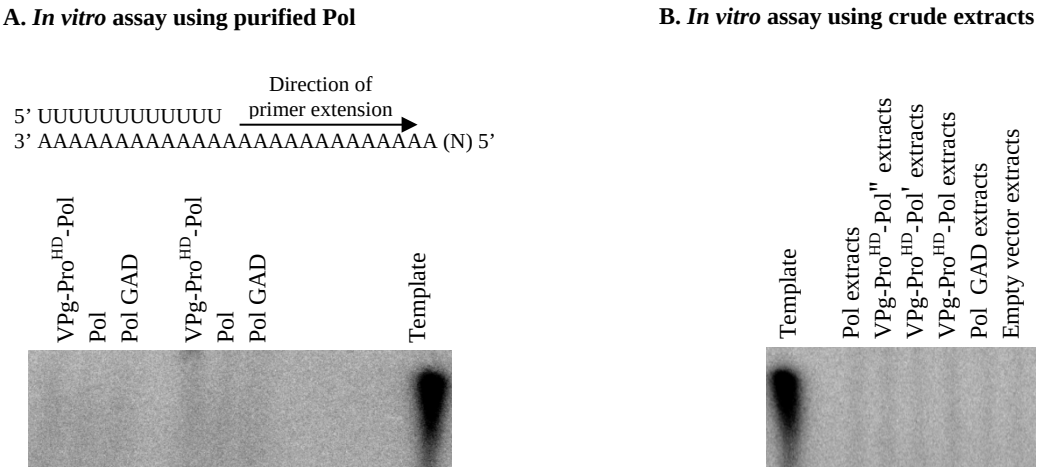


Figure 4.5: Activity assay of ToRSV polymerase using poly(A) as template and (U)12 as primer. A. *In vitro* RNA synthesis assay using purified Pol. B. *In vitro* RNA synthesis assay using crude extracts derived from *E. coli* cells expressing different forms of the polymerase. Synthesized products were separated in 1% agarose/ formaldehyde gel and detected by autoradiography.

4.4 Discussion

In this work, I tried to identify the active form of the ToRSV polymerase using various forms of recombinant polymerase and *in vitro* activity assays. This approach has been successfully used in investigation of the biochemical properties of a number of viral polymerases, including those from PV (322), NV (238, 318) and FCV (237). Unfortunately, I did not detect any ToRSV polymerase activity using either purified proteins or crude *E. coli* extracts (Fig. 4.3, 4.4 and 4.5). The lack of polymerase activity is unlikely due to the inappropriate RNA template because a number of templates were tried, including a non-viral template, a poly(A) template and the purified ToRSV genomic RNA template (Fig. 4.2, 4.3, 4.4 and 4.5). It is also unlikely that reaction conditions are responsible for the lack of the ToRSV polymerase activity, because different conditions (including different pH, polymerase concentration, divalent ions, reaction temperature and reaction time) were tested (results not shown). Detection of a RNA product with T7 polymerase in the same reaction condition strongly suggests that the reaction conditions were within the acceptable range for polymerase activities (Fig. 4.3 and 4.4).

It is possible that the purified polymerases added to the reactions were inactive either because they are not properly folded in *E. coli* or because they may require post-translational modifications, which occur in plants but not in *E. coli*. I also cannot exclude the possibility that the ToRSV polymerase is very sensitive and loses its activity rapidly during purification. Additionally, there is evidence indicating that having the authentic N- and C-terminus is important for polymerase activity (322). Therefore, adding an N-terminal methionine (from the AUG start codon) and the C-terminal hexahistidine tag to the ToRSV polymerase may

have disrupted the spatial positions of both ends in the three dimensional structure and caused inactivation of the polymerase.

I also cannot exclude the possibility that the primer used in the activity assay is not optimal. As mentioned before, *in vivo* ToRSV replication is likely primed by the uridylated VPg protein as was demonstrated for PV. However, the purified PV Pol is capable of catalyzing RNA synthesis using an RNA oligo as primer (311). The ToRSV Pol may have a more strict requirement for using VPg-pU-pU as primer instead of an RNA oligo. In addition, other viral proteins might be required in the assay to stimulate the ToRSV polymerase activity. It has been shown that the *E. coli* purified sindbis virus nsP4 only exhibits polymerase activity in the presence of viral protein P123 (312). Finally, host factors and membranes are also important for polymerase activity. For example, the heat shock protein 70 (Hsp70), which is a chaperone protein, can activate the TBSV polymerase *in vitro* and an inhibitor that blocks the Hsp70 function also inhibits RNA synthesis (323). Furthermore, neutral phospholipids, such as phosphatidylethanolamine and phosphatidylcholine, were shown to enhance polymerase activation *in vitro* by stimulating the polymerase and viral RNA interaction. In contrast, the acidic phosphatidylglycerol has an inhibitory effect on polymerase activation by interfering with the binding of the viral RNA to the polymerase (323). Preliminary yeast two hybrid results indicated that the ToRSV polymerase interacts with a number of host factors (results not shown). It is possible that these host factors are also required for ToRSV polymerase activity *in vitro*. Although the closely related picornavirus Pols have been extensively studied using heterologously expressed proteins and *in vitro* activity assays (315, 322, 324), the only study of viral polymerase from members of the family *Secoviridae*

reported to date is that of CPMV. As mentioned previously, the Pols of ToRSV and CPMV have a C-terminal 15-20 kDa extension when compared to the Pol of PV (13). Detection of a sub-population of VPg-Pro-Pol′ in membrane-enriched fractions active in ToRSV replication using endogenous templates indicates that the VPg-Pro-Pol′ might be the active form of polymerase during virus replication (13). However, I was unable to detect any polymerase activity with the *E. coli* purified VPg-Pro-Pol′ *in vitro* (Fig. 4.2, 4.4 and 4.5). Similarly, a sub-population of CPMV 110K protein (Pro-Pol) co-purifies with viral replication complexes from infected cowpea leaves and is capable of stimulating RNA chain elongation (240). Although the CPMV 110K acts as an active form of the polymerase *in vivo*, no detectable polymerase activity was observed when recombinant CPMV 110K protein was prepared either from *E. coli* or from the baculovirus vector expression system (315, 325). In addition, the 87K (Pol) and 170K (NTB-VPg-Pro-Pol) are also unable to support RNA synthesis *in vitro* (315, 325). Interestingly, protoplasts transiently expressing 200K (the entire RNA1 polyprotein) but not 87K, 110K or 170K are capable of supporting the replication of co-inoculated M-RNA (178). However, addition of exogenous template RNA (poly(A)-oligo(U) or other template/primer combinations) to extracts derived from protoplasts expressing the 200 K polyprotein did not result in a detectable polymerase activity (178). Together, this information suggests that the mechanism of nepovirus and comovirus polymerase activity is more complex than that of picornaviruses, and that a simple *in vitro* system cannot support polymerase activity. Additionally, the heterologous system, which is capable of producing active polymerase of picornaviruses, may fail to produce active nepovirus and comovirus polymerases.

Chapter 5

General discussion

As a key step in the virus multiplication cycle, replication of positive-strand RNA viruses takes place on membrane-associated replication complexes (4). During this process, viral-encoded integral membrane proteins target to the intracellular membranes independently and act as membrane anchors for the replication complexes (5). The viral polymerases, which catalyze the synthesis of the progeny RNAs, are recruited to the replication complexes by directly or indirectly interacting with membrane anchor proteins (5). In addition, various host factors are also involved in virus replication (72). In this thesis, I investigated the stability of two ToRSV encoded replication proteins: the membrane protein NTB-VPg and the polymerase. My results indicated that the C-terminal region of the NTB-VPg polyprotein (cNTB-VPg) of several ToRSV isolates is glycosylated by the membrane-associated OST, and cleaved by the signal peptidase complex. In addition, I show that the full-length Pol and the polyprotein VPg-Pro-Pol are unstable in *N. benthamiana* plants and this instability is likely caused by a plant protease, and/or a plant degradation pathway. The possible biological relevance of the instability and membrane-associated modification of viral replication proteins are discussed below.

5.1 Membrane-associated glycosylation and processing of C-terminal NTB-VPg

The ToRSV encoded NTB-VPg is an integral membrane protein that associates with ER enriched fractions that support active replication of the ToRSV genome (11). In this thesis, the C-terminal region of NTB-VPg (cNTB-VPg) from different ToRSV isolates was expressed *in vitro* and *in vivo*. These results indicate that (1) cNTB-VPg derived from all

ToRSV isolates is able to associate with ER-derived membranes. The transmembrane domain at the C-terminal region of NTB directs the translocation of the downstream NTB region and the entire VPg region into the ER lumen. (2) Membrane-associated N-linked glycosylation and signal peptidase processing occur in cNTB-VPg of PYB, GYV and Rasp2 isolates, but not in that of the Rasp1 isolate. (3) Glycosylation of cNTB-VPg is an efficient event and it significantly enhances the protein stability. (4) The signal peptidase processing event is suboptimal and inefficient. (5) The signal peptidase cleavage site is atypical, as it is located in the C-terminal region of the NTB protein. However, it still follows the -3, -1 rule similar to typical N-terminally located signal peptidase cleavage sites.

As a common protein modification process, N-linked glycosylation plays a role in various aspects of the virus life cycle. It is mostly observed in viral-encoded structural proteins presented on the surface of the virus particle (278). Glycosylation of viral structural proteins plays an important role in virus survival and virulence. For example, glycosylation in the premembrane (prM) and envelope (E) protein of West Nile Virus influences virus particle assembly, virus release and virus infectivity (326). Similarly, glycosylation of influenza virus hemagglutinin protein has been linked to receptor binding, infectivity, and virus release (278, 327, 328). Depending on the virus considered, glycosylation of viral-encoded non-structural proteins is also observed but less common. Adding a glycan to the replication proteins could delay the virus infection; change the morphology of the replication complexes and influence virus replication and RNA accumulation (279-281). My results provide evidence that N-linked glycosylation occurs on a truncated form of the ToRSV membrane protein NTB-VPg and this glycosylation stabilizes the protein *in planta*. Since NTB-VPg is actively involved in

virus replication complex assembly, it is possible that glycosylation of the membrane protein influences ToRSV replication. However, this hypothesis still needs to be confirmed with further experiments.

Eukaryotic signal peptidase is a multi-subunit membrane protein complex with the active site on the luminal side of the ER (33). It is an important enzyme that is involved in cleaving N-terminally located signal peptides of cellular proteins (33). Viruses hijack this host machinery for processing internally located signal peptides and liberating the structural proteins from the viral-encoded polyprotein (101, 105, 120). In this thesis, the signal peptidase was shown to be responsible for cleaving a truncated protein derived from the ToRSV NTB-VPg polyprotein and likely also the full-length NTB-VPg. Unlike other viruses, the signal peptidase did not cleave at the boundary between two viral protein domains, but instead cleaved in the C-terminal region of NTB which is upstream of the NTB-VPg cleavage site. Mutational analysis allowed identification of the putative cleavage site, which strictly follows the -3, -1 rule as other typical signal peptidase cleavage site. Mutations that introduced favorable amino acids in the cleavage site resulted in increased signal peptidase processing, while mutations that introduced unfavorable ones lead to inefficient or abolished cleavage.

Interestingly, both N-linked glycosylation and signal peptidase processing occurs on the cNTB-VPg of PYB, GYV and Rasp2. In contrast, these membrane-associated modifications were not detected on the cNTB-VPg of the Rasp1 isolate, at least *in vitro*. A number of amino acid differences were observed in the C-terminal region of NTB and VPg in the Rasp1

isolate. The mutagenesis study identified several key amino acids differences that account for the absence of glycosylation and signal peptidase processing in the Rasp1 cNTB-VPg. As far as I know, the ToRSV cNTB-VPg is the first plant virus replication protein suggested to be cleaved by the signal peptidase.

Based on the above results, a model of how membrane-associated enzymes modify the ToRSV replication protein NTB-VPg is proposed: (1) For PYB, GYV and Rasp2 isolates, the NTB-VPg is released from the polyprotein by the ToRSV protease early in infection and efficiently targets to the ER membrane via its hydrophobic domains. The C-terminal transmembrane domain translocates the downstream region of NTB-VPg into the ER lumen, where the consensus glycosylation site is exposed to the active site of the OST and is glycosylated. This glycosylation enhances the stability of this membrane protein and possibly also of the replication complexes. The signal peptidase slowly cleaves off the C-terminal region of NTB-VPg, which may change the architecture of the viral replication complexes. It is possible that this change modulates the activity of the replication complexes and stimulates the transition from replication to the next stage of the virus multiplication cycle. (2) In the case of the Rasp1 isolate, the lack of glycosylation causes instability of the NTB-VPg protein and only smaller amounts of the NTB-VPg accumulate in the cell. This may result in formation of fewer replication complexes compared with other ToRSV isolates. However, signal peptidase processing is inefficient or absent in the NTB-VPg and the replication complexes may remain active until later stages of infection. Comparing the kinetics of replication of the different ToRSV isolates in protoplasts would be helpful to test this model.

5.2 Instability of ToRSV polymerase

As mentioned previously, the mature Pol and the VPg-Pro-Pol' are detected in ToRSV infected plants (13). The full-length VPg-Pro-Pol is undetectable, and the Pol is detected in low amounts, if at all. Therefore, I investigated the stability of the ToRSV polymerase by expressing the full-length Pol and VPg-Pro-Pol as well as C-terminally truncated proteins in *N. benthamiana* plants and in *E. coli*. My results indicated that: (1) the full-length Pol and VPg-Pro-Pol show intrinsic instability in *N. benthamiana* plants, but not in *E. coli*; (2) mutations that inactivate the viral protease did not stabilize the transiently expressed full-length VPg-Pro-Pol in *N. benthamiana* plants; (3) the C-terminal 15-20 kDa truncated forms of VPg-Pro-Pol and Pol show significantly enhanced stability compared to the corresponding full-length proteins.

Truncation of the VPg-Pro-Pol in infected plants is evident but the exact truncation site is still unclear (13). It is estimated that the full-length polyprotein loses approximately 15 kDa in the C-terminal region of Pol based on the migration of VPg-Pro-Pol' in SDS-PAGE (13). The cleaved C-terminal 15 kDa fragment of Pol was never detected probably due to the instability of this fragment. Therefore it was not possible to sequence this fragment using the Edman degradation method. The C-terminal truncation was not detected when the VPg-Pro-Pol was transiently expressed in plants. Several hypotheses can explain this result. (1) The truncation may only occur on the RNA1-encoded large polyprotein. In this case, the transiently expressed VPg–Pro-Pol is folded in a conformation that differs from that adopted when it is present within the large polyprotein, and as a result it is not truncated. (2) Truncation in the C-terminal region of Pol has a strict requirement for the presence of other

viral proteins, although an active viral protease may not be required. (3) The protease or other host factors that are involved in the truncation may only be activated during viral infection.

Truncation of the polymerase was not reported in other related nepoviruses or in picornaviruses as far as I know, indicating that it is unique in ToRSV. Interestingly, the CPMV and ToRSV polymerases have a C-terminal 15-20 kDa extension compared to other picorna-like viruses (13). Therefore, truncation of the C-terminal 15 kDa of ToRSV Pol makes it similar in size to the picornavirus polymerase. What causes this truncation is still unknown, but interestingly, when I expressed the VPg-Pro-Pol mutants that lacked the C-terminal 15 or 20 kDa fragment of Pol in plants, the stability of the protein was significantly increased when compared to the full-length VPg-Pro-Pol. As mentioned previously, although the VPg-Pro-Pol' is truncated, it still maintains all the conserved polymerase motifs upstream of this truncation site (13). Therefore it is possible that this truncated protein has polymerase activity and that truncation of the polymerase in *N. benthamiana* plants plays a role in modulating the polymerase activity. The fact that VPg-Pro-Pol' peripherally associates with the membrane-bound replication complexes indicates that it might be involved in catalyzing the synthesis of positive- and negative-strand RNAs during replication.

The instability of the full-length Pol and VPg-Pro-Pol is observed in both transient expression and in natural virus infection of *N. benthamiana* plants. The mechanism of this instability is still unknown. However, it is worth mentioning that down-regulation of viral polymerase concentration has been observed in a number of viruses. As mentioned

previously, it could occur at the translational level. For example, polymerase of barley yellow dwarf virus is expressed via inefficient -1 ribosomal frameshift (304), the tobacco mosaic virus and TBSV polymerase are produced via inefficient translational readthrough (302, 329), whereas the BMV polymerase is encoded on a separate RNA molecule (303). For TYMV, HAV and HCV, polymerases are post-translationally targeted and degraded by ubiquitin-proteasome pathway (299, 305, 306). The ToRSV polymerase is initially produced in amounts equal to those of other RNA1-encoded proteins due to the polyprotein strategy. Regulated degradation of Pol and VPg-Pro-Pol during virus infection may play a role in regulating the polymerase concentration. This may allow the replication proteins to accumulate in an optimal ratio for virus replication.

Based on all the results, a model of how polymerase concentration and activity is modulated during ToRSV infection is suggested (Fig 5.1). Alternative processing of the ToRSV RNA1 encoded polyprotein allows release of two forms of the polymerase: Pol and VPg-Pro-Pol. These full-length polymerases are readily degraded by unknown host degradation machinery through interacts with the C-terminal 15 kDa of Pol. To escape from this host degradation, the virus has evolved to use a host protease that specifically cleaves the C-terminal 15 kDa from the full-length VPg-Pro-Pol and significantly enhances the protein stability, resulting in accumulation of VPg-Pro-Pol' in the cell. A sub-population of the VPg-Pro-Pol' is recruited to the ER membrane-associated replication complexes via direct or indirect interaction with the viral membrane proteins, and this protein catalyzes the synthesis of progeny viral RNA.

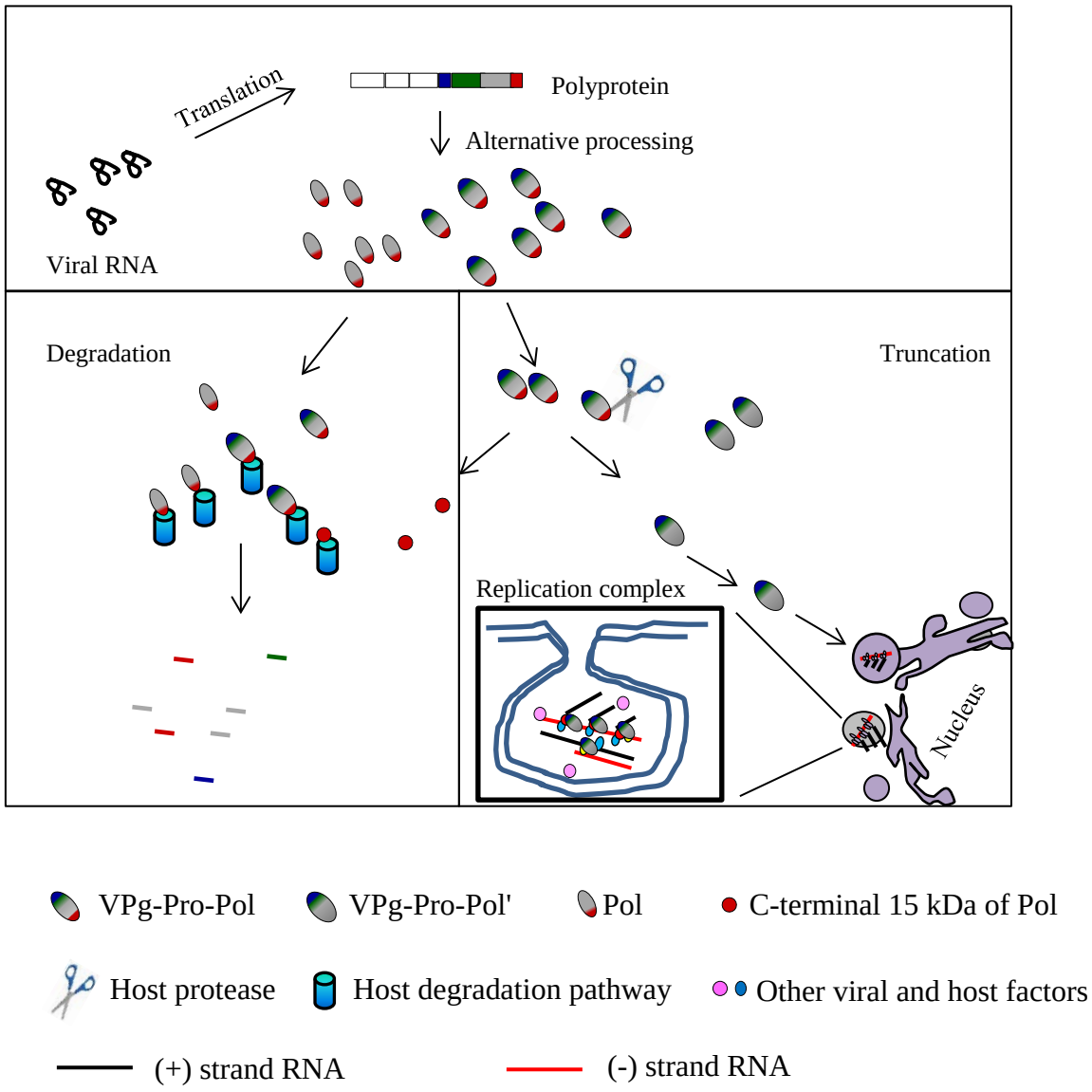


Figure 5.1: Model for the regulation of ToRSV polymerase during virus replication. The full-length VPg-Pro-Pol and Pol are unstable and are degraded by an unknown protein degradation pathway. VPg-Pro-Pol molecules are specifically cleaved by a host protease and thus lose the C-terminal 15 kDa of Pol, which allows accumulation of VPg-Pro-Pol'. A sub-population of the truncated VPg-Pro-Pol' protein associates with ER-bound replication complexes and catalyzes the synthesis of positive and negative-strand RNAs.

5.3 The role of regulating virus replication proteins by host proteases and degradation machineries

My results provide evidence that host proteases and degradation machineries are involved in regulating the stability and truncation of the ToRSV encoded NTB-VPg and polymerase. Regulation of viral protein stability is an important mechanism to regulate viral infection and has been observed in a number of viruses (299, 306). In this thesis, both proteins studied are key components of the ToRSV replication complexes. Therefore, regulation of the stability of these proteins may directly impact virus replication. This could be advantageous for the virus in a number of ways. (1) Degradation of polymerases can down-regulate virus replication which in turn reduces the rate of RNA mutation and keeps the integrity of the viral genome. (2) Less virus replication also reduces the chance of double strand replication intermediates or highly structured double strand regions of single strand RNA molecules being recognized by host anti-viral defenses such as RNA silencing. (3) Signal peptidase processing of NTB-VPg in the late stages of virus replication might also play a role in disassembly of the virus replication complexes and stimulate the switch from replication to assembly of the viral RNA into virus particles. (4) The instability of the polymerase and the signal peptidase processing of the NTB-VPg may contribute to a host cell defense pathway against virus infection. (5) Viruses utilize the host machineries to optimize viral protein concentration and virus replication, which may help enhance its infectivity. (6) Regulation of viral proteins by host machineries may help keep an optimal balance between plants and viruses that ensure virus persistence without killing the plant cell. In conclusion, the susceptibility of plant viruses to host proteases and degradation machineries is probably a reflection of the co-evolution between the host plant and its pathogens.

5.4 Future experiments

A number of experiments could be performed in the future to help further our understanding of the biological relevance of membrane-associated modification of the cNTB-VPg. First, once infectious transcripts of ToRSV become available, mutations increasing or decreasing signal peptidase processing could be introduced in these transcripts. The effect of these mutations could be tested by protoplast transfection and RNA accumulation level could be monitored. This would help confirm the above hypothesis that signal peptidase processing plays a role in regulating virus replication. Similarly, mutations in the consensus glycosylation site could be introduced to test whether or not glycosylation enhances virus replication. In addition, it would also be interesting to investigate whether the glycosylation and signal peptidase processing play a role in the morphological change of membrane-associated replication complexes. To do this, membrane modification could be monitored in plant infected with wild type or mutated transcripts using electron microscopy.

Although my results provide important evidence that the full-length Pol and VPg-Pro-Pol of ToRSV are unstable *in planta*, there are still important questions that need to be addressed. A number of experiments could be performed in the future to further understand the mechanism of ToRSV replication. First, purified viral RNAs could be used to transfect *N. benthamiana* protoplasts and inhibitors against different protein degradation enzymes or pathways could be used to examine whether or not truncation and degradation of full-length polymerase occurs. These experiments may help identify what causes the instability of the ToRSV polymerase. Second, since I was unable to identify the active form of the polymerase using *in vitro* polymerase activity assay, it would be interesting to delete the C-terminal 15 kDa

region of the polymerase in infectious transcripts to test whether or not the truncated polymerase is able to direct virus replication. Viral RNA accumulation levels could be compared between wild type and mutated transcripts to identify whether this truncation activates or modulates the polymerase activities. Third, it would also be interesting to investigate how the VPg-Pro-Pol' is brought to the ER-associated replication complexes. To do this, yeast two hybrid experiments and/or co- immunoprecipitations could be performed to investigate whether VPg-Pro-Pol' interacts with viral membrane proteins, such as NTB, NTB-VPg and/or X2-NTB-VPg. At the same time, the yeast two hybrid screen of VPg-Pro-Pol' against the *N. benthamiana* library may provide us with more information about what host factors are involved in the assembly of virus replication complexes.

Bibliography

1. **Ryan MD, Flint M.** 1997. Virus-encoded proteinases of the picornavirus super-group. J Gen Virol **78 (Pt 4):**699-723.
2. **Laliberte JF, Sanfacon H.** 2010. Cellular remodeling during plant virus infection. Annu Rev Phytopathol **48:**69-91.
3. **Novoa RR, Calderita G, Arranz R, Fontana J, Granzow H, Risco C.** 2005. Virus factories: associations of cell organelles for viral replication and morphogenesis. Biol Cell **97:**147-172.
4. **den Boon JA, Ahlquist P.** 2010. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. Annu Rev Microbiol **64:**241-256.
5. **Sanfacon H.** 2005. Replication of positive-strand RNA viruses in plants: Contact points between plant and virus components. Can J Bot **83:**1529-1549.
6. **Wang A.** 2015. Dissecting the Molecular Network of Virus-Plant Interactions: The Complex Roles of Host Factors. Annu Rev Phytopathol **53:**45-66.
7. **Reid CR, Airo AM, Hobman TC.** 2015. The Virus-Host Interplay: Biogenesis of +RNA Replication Complexes. Viruses **7:**4385-4413.
8. **Nagy PD, Barajas D, Pogany J.** 2012. Host factors with regulatory roles in tombusvirus replication. Curr Opin Virol **2:**691-698.
9. **Sanfacon H, Iwanami T, Karasev A, Van der Vlugt R, Wellink J, Wetzel T, Yoshikawa N.** 2011. Family Secoviridae, p. 881-899. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed.), Virus Taxonomy: Classification and Nomenclature of Viruses. Ninth Report of the International Committee on the Taxonomy of Viruses. Elseviers, San Diego.
10. **Sanfacon H.** 2008. Nepovirus, p. 405-413. *In* Mahy BWJ, Van Regenmortel MH (ed.), Encyclopedia of Virology, Third Edition, vol. 3. Elsevier, Oxford.
11. **Han S, Sanfacon H.** 2003. Tomato ringspot virus proteins containing the nucleoside triphosphate binding domain are transmembrane proteins that associate with the endoplasmic reticulum and cofractionate with replication complexes. J Virol **77:**523-534.
12. **Wang A, Han S, Sanfacon H.** 2004. Topogenesis in membranes of the NTB-VPg protein of Tomato ringspot nepovirus: definition of the C-terminal transmembrane domain. J Gen Virol **85:**535-545.
13. **Chisholm J, Zhang G, Wang A, Sanfacon H.** 2007. Peripheral association of a polyprotein precursor form of the RNA-dependent RNA polymerase of Tomato ringspot virus with the membrane-bound viral replication complex. Virology **368:**133-144.
14. **Singer SJ.** 1972. A fluid lipid-globular protein mosaic model of membrane structure. Ann N Y Acad Sci **195:**16-23.
15. **Cheung AY, de Vries SC.** 2008. Membrane trafficking: intracellular highways and country roads. Plant Physiol **147:**1451-1453.
16. **Sanderfoot AA, Raikhel NV.** 1999. The specificity of vesicle trafficking: coat proteins and SNAREs. Plant Cell **11:**629-642.
17. **Voeltz GK, Rolls MM, Rapoport TA.** 2002. Structural organization of the endoplasmic reticulum. EMBO Rep **3:**944-950.

18. **Vitale A, Denecke J.** 1999. The endoplasmic reticulum-gateway of the secretory pathway. Plant Cell **11:**615-628.
19. **Jennings ML.** 1989. Topography of membrane proteins. Annu Rev Biochem **58:**999-1027.
20. **Segrest JP, Jackson RL, Morrisett JD, Gotto AM, Jr.** 1974. A molecular theory of lipid-protein interactions in the plasma lipoproteins. FEBS Lett **38:**247-258.
21. **Cornell RB, Taneva SG.** 2006. Amphipathic helices as mediators of the membrane interaction of amphitropic proteins, and as modulators of bilayer physical properties. Curr Protein Pept Sci **7:**539-552.
22. **Segrest JP, De Loof H, Dohlman JG, Brouillette CG, Anantharamaiah GM.** 1990. Amphipathic helix motif: classes and properties. Proteins **8:**103-117.
23. **Zemel A, Ben-Shaul A, May S.** 2005. Perturbation of a lipid membrane by amphipathic peptides and its role in pore formation. Eur Biophys J **34:**230-242.
24. **Drin G, Antonny B.** 2010. Amphipathic helices and membrane curvature. FEBS Lett **584:**1840-1847.
25. **McMahon HT, Gallop JL.** 2005. Membrane curvature and mechanisms of dynamic cell membrane remodelling. Nature **438:**590-596.
26. **Ulmschneider MB, Sansom MS.** 2001. Amino acid distributions in integral membrane protein structures. Biochim Biophys Acta **1512:**1-14.
27. **DeGrado WF, Gratkowski H, Lear JD.** 2003. How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles. Protein Sci **12:**647-665.
28. **Fink A, Sal-Man N, Gerber D, Shai Y.** 2012. Transmembrane domains interactions within the membrane milieu: principles, advances and challenges. Biochim Biophys Acta **1818:**974-983.
29. **Akopian D, Shen K, Zhang X, Shan SO.** 2013. Signal recognition particle: an essential protein-targeting machine. Annu Rev Biochem **82:**693-721.
30. **Lutcke H, Dobberstein B.** 1993. Structure and function of signal recognition particle (SRP). Mol Biol Rep **18:**143-147.
31. **Schaffitzel C, Oswald M, Berger I, Ishikawa T, Abrahams JP, Koerten HK, Koning RI, Ban N.** 2006. Structure of the E. coli signal recognition particle bound to a translating ribosome. Nature **444:**503-506.
32. **Denks K, Vogt A, Sachelaru I, Petriman NA, Kudva R, Koch HG.** 2014. The Sec translocon mediated protein transport in prokaryotes and eukaryotes. Mol Membr Biol **31:**58-84.
33. **Paetzel M, Karla A, Strynadka NC, Dalbey RE.** 2002. Signal peptidases. Chem Rev **102:**4549-4580.
34. **Higy M, Junne T, Spiess M.** 2004. Topogenesis of membrane proteins at the endoplasmic reticulum. Biochemistry **43:**12716-12722.
35. **Jung SJ, Kim JE, Reithinger JH, Kim H.** 2014. The Sec62-Sec63 translocon facilitates translocation of the C-terminus of membrane proteins. J Cell Sci **127:**4270-4278.
36. **Lyman SK, Schekman R.** 1997. Binding of secretory precursor polypeptides to a translocon subcomplex is regulated by BiP. Cell **88:**85-96.
37. **Nilsson J, Persson B, von Heijne G.** 2005. Comparative analysis of amino acid distributions in integral membrane proteins from 107 genomes. Proteins **60:**606-616.

38. **Denzer AJ, Nabholz CE, Spiess M.** 1995. Transmembrane orientation of signal-anchor proteins is affected by the folding state but not the size of the N-terminal domain. EMBO J **14:**6311-6317.
39. **Wahlberg JM, Spiess M.** 1997. Multiple determinants direct the orientation of signal-anchor proteins: the topogenic role of the hydrophobic signal domain. J Cell Biol **137:**555-562.
40. **von Heijne G.** 1990. The signal peptide. J Membr Biol **115:**195-201.
41. **von Heijne G.** 1989. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. Nature **341:**456-458.
42. **Ota K, Sakaguchi M, Hamasaki N, Mihara K.** 2000. Membrane integration of the second transmembrane segment of band 3 requires a closely apposed preceding signal-anchor sequence. J Biol Chem **275:**29743-29748.
43. **Goder V, Bieri C, Spiess M.** 1999. Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. J Cell Biol **147:**257-266.
44. **Heinrich SU, Rapoport TA.** 2003. Cooperation of transmembrane segments during the integration of a double-spanning protein into the ER membrane. EMBO J **22:**3654-3663.
45. **Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC, Hernandez M, Kelly JF, Wacker M, Aebi M.** 2006. Definition of the bacterial N-glycosylation site consensus sequence. EMBO J **25:**1957-1966.
46. **Chavan M, Lennarz W.** 2006. The molecular basis of coupling of translocation and N-glycosylation. Trends Biochem Sci **31:**17-20.
47. **Helenius A, Aebi M.** 2004. Roles of N-linked glycans in the endoplasmic reticulum. Annu Rev Biochem **73:**1019-1049.
48. **Kelleher DJ, Gilmore R.** 2006. An evolving view of the eukaryotic oligosaccharyltransferase. Glycobiology **16:**47R-62R.
49. **Yan A, Lennarz WJ.** 2005. Unraveling the mechanism of protein N-glycosylation. J Biol Chem **280:**3121-3124.
50. **Nilsson IM, von Heijne G.** 1993. Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. J Biol Chem **268:**5798-5801.
51. **Chavan M, Yan A, Lennarz WJ.** 2005. Subunits of the translocon interact with components of the oligosaccharyl transferase complex. J Biol Chem **280:**22917-22924.
52. **Harada Y, Li H, Li H, Lennarz WJ.** 2009. Oligosaccharyltransferase directly binds to ribosome at a location near the translocon-binding site. Proc Natl Acad Sci U S A **106:**6945-6949.
53. **Kasturi L, Eshleman JR, Wunner WH, Shakin-Eshleman SH.** 1995. The hydroxy amino acid in an Asn-X-Ser/Thr sequon can influence N-linked core glycosylation efficiency and the level of expression of a cell surface glycoprotein. J Biol Chem **270:**14756-14761.
54. **Kasturi L, Chen H, Shakin-Eshleman SH.** 1997. Regulation of N-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequons that are poor oligosaccharide acceptors. Biochem J **323 (Pt 2):**415-419.

55. **Ruiz-Canada C, Kelleher DJ, Gilmore R.** 2009. Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. Cell **136:**272-283.
56. **Nilsson IM, von Heijne G.** 1993. Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. J Biol Chem **268:**5798-5801.
57. **Choo KH, Ranganathan S.** 2008. Flanking signal and mature peptide residues influence signal peptide cleavage. BMC Bioinformatics **9 Suppl 12:**S15.
58. **Auclair SM, Bhanu MK, Kendall DA.** 2012. Signal peptidase I: cleaving the way to mature proteins. Protein Sci **21:**13-25.
59. **Jain RG, Rusch SL, Kendall DA.** 1994. Signal peptide cleavage regions. Functional limits on length and topological implications. J Biol Chem **269:**16305-16310.
60. **Fikes JD, Barkocy-Gallagher GA, Klapper DG, Bassford PJ, Jr.** 1990. Maturation of Escherichia coli maltose-binding protein by signal peptidase I in vivo. Sequence requirements for efficient processing and demonstration of an alternate cleavage site. J Biol Chem **265:**3417-3423.
61. **Robakis T, Bak B, Lin SH, Bernard DJ, Scheiffele P.** 2008. An internal signal sequence directs intramembrane proteolysis of a cellular immunoglobulin domain protein. J Biol Chem **283:**36369-36376.
62. **Guo KK, Tang QH, Zhang YM, Kang K, He L.** 2011. Identification of two internal signal peptide sequences: critical for classical swine fever virus non-structural protein 2 to trans-localize to the endoplasmic reticulum. Virol J **8:**236.
63. **Ouzzine M, Magdalou J, Burchell B, Fournel-Gigleux S.** 1999. An internal signal sequence mediates the targeting and retention of the human UDP-glucuronosyltransferase 1A6 to the endoplasmic reticulum. J Biol Chem **274:**31401-31409.
64. **Fluhrer R, Steiner H, Haass C.** 2009. Intramembrane proteolysis by signal peptide peptidases: a comparative discussion of GXGD-type aspartyl proteases. J Biol Chem **284:**13975-13979.
65. **Friedmann E, Lemberg MK, Weihofen A, Dev KK, Dengler U, Rovelli G, Martoglio B.** 2004. Consensus analysis of signal peptide peptidase and homologous human aspartic proteases reveals opposite topology of catalytic domains compared with presenilins. J Biol Chem **279:**50790-50798.
66. **Krawitz P, Haffner C, Fluhrer R, Steiner H, Schmid B, Haass C.** 2005. Differential localization and identification of a critical aspartate suggest non-redundant proteolytic functions of the presenilin in homologues SPPL2b and SPPL3. J Biol Chem.
67. **Friedmann E, Hauben E, Maylandt K, Schleeger S, Vreugde S, Lichtenthaler SF, Kuhn PH, Stauffer D, Rovelli G, Martoglio B.** 2006. SPPL2a and SPPL2b promote intramembrane proteolysis of TNFalpha in activated dendritic cells to trigger IL-12 production. Nat Cell Biol **8:**843-848.
68. **Weihofen A, Binns K, Lemberg MK, Ashman K, Martoglio B.** 2002. Identification of signal peptide peptidase, a presenilin-type aspartic protease. Science **296:**2215-2218.

69. **Voss M, Schroder B, Fluhrer R.** 2013. Mechanism, specificity, and physiology of signal peptide peptidase (SPP) and SPP-like proteases. Biochim Biophys Acta **1828:**2828-2839.
70. **Weihofen A, Lemberg MK, Ploegh HL, Bogyo M, Martoglio B.** 2000. Release of signal peptide fragments into the cytosol requires cleavage in the transmembrane region by a protease activity that is specifically blocked by a novel cysteine protease inhibitor. J Biol Chem **275:**30951-30956.
71. **Mine A, Okuno T.** 2012. Composition of plant virus RNA replicase complexes. Curr Opin Virol **2:**669-675.
72. **Ahlquist P, Noueiry AO, Lee WM, Kushner DB, Dye BT.** 2003. Host factors in positive-strand RNA virus genome replication. J Virol **77:**8181-8186.
73. **Patarroyo C, Laliberte JF, Zheng H.** 2012. Hijack it, change it: how do plant viruses utilize the host secretory pathway for efficient viral replication and spread? Front Plant Sci **3:**308.
74. **Bushell M, Sarnow P.** 2002. Hijacking the translation apparatus by RNA viruses. J Cell Biol **158:**395-399.
75. **Miller WA, Koev G.** 2000. Synthesis of subgenomic RNAs by positive-strand RNA viruses. Virology **273:**1-8.
76. **Dreher TW, Miller WA.** 2006. Translational control in positive strand RNA plant viruses. Virology **344:**185-197.
77. **Firth AE, Brierley I.** 2012. Non-canonical translation in RNA viruses. J Gen Virol.
78. **Verchot J.** 2011. Wrapping membranes around plant virus infection. Curr Opin Virol **1:**388-395.
79. **Belov GA, van Kuppeveld FJ.** 2012. (+)RNA viruses rewire cellular pathways to build replication organelles. Curr Opin Virol **2:**740-747.
80. **Paul D, Bartenschlager R.** 2013. Architecture and biogenesis of plus-strand RNA virus replication factories. World journal of virology **2:**32-48.
81. **White KA, Nagy PD.** 2004. Advances in the molecular biology of tombusviruses: gene expression, genome replication, and recombination. Prog Nucleic Acid Res Mol Biol **78:**187-226.
82. **Niehl A, Heinlein M.** 2011. Cellular pathways for viral transport through plasmodesmata. Protoplasma **248:**75-99.
83. **Kawakami S, Watanabe Y, Beachy RN.** 2004. Tobacco mosaic virus infection spreads cell to cell as intact replication complexes. Proc Natl Acad Sci U S A **101:**6291-6296.
84. **Ritzenthaler C, Schmit AC, Michler P, Stussi-Garaud C, Pinck L.** 1995. Grapevine fanleaf nepovirus P38 putative movement protein is located on tubules *in vivo*. Mol Plant Microbe Interact **8:**379-387.
85. **Wieczorek A, Sanfacon H.** 1993. Characterization and subcellular localization of tomato ringspot nepovirus putative movement protein. Virology **194:**734-742.
86. **Tong L.** 2002. Viral proteases. Chem Rev **102:**4609-4626.
87. **Dougherty WG, Semler BL.** 1993. Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. Microbiol Rev **57:**781-822.

88. **Carrier K, Hans F, Sanfacon H.** 1999. Mutagenesis of amino acids at two tomato ringspot nepovirus cleavage sites: effect on proteolytic processing in cis and in trans by the 3C-like protease. Virology **258:**161-175.
89. **Wellink J, van Kammen A.** 1988. Proteases involved in the processing of viral polyproteins. Brief review. Archive of Virology **98:**1-26.
90. **Lloyd RE, Jense HG, Ehrenfeld E.** 1987. Restriction of translation of capped mRNA in vitro as a model for poliovirus-induced inhibition of host cell protein synthesis: relationship to p220 cleavage. J Virol **61:**2480-2488.
91. **Krausslich HG, Nicklin MJ, Toyoda H, Etchison D, Wimmer E.** 1987. Poliovirus proteinase 2A induces cleavage of eucaryotic initiation factor 4F polypeptide p220. J Virol **61:**2711-2718.
92. **Joachims M, Van Breugel PC, Lloyd RE.** 1999. Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. J Virol **73:**718-727.
93. **Kerekatte V, Keiper BD, Badorff C, Cai A, Knowlton KU, Rhoads RE.** 1999. Cleavage of Poly(A)-binding protein by coxsackievirus 2A protease in vitro and in vivo: another mechanism for host protein synthesis shutoff? J Virol **73:**709-717.
94. **Etchison D, Milburn SC, Edery I, Sonenberg N, Hershey JW.** 1982. Inhibition of HeLa cell protein synthesis following poliovirus infection correlates with the proteolysis of a 220,000-dalton polypeptide associated with eucaryotic initiation factor 3 and a cap binding protein complex. J Biol Chem **257:**14806-14810.
95. **Bedard KM, Semler BL.** 2004. Regulation of picornavirus gene expression. Microbes Infect **6:**702-713.
96. **Devaney MA, Vakharia VN, Lloyd RE, Ehrenfeld E, Grubman MJ.** 1988. Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. J Virol **62:**4407-4409.
97. **Yalamanchili P, Harris K, Wimmer E, Dasgupta A.** 1996. Inhibition of basal transcription by poliovirus: a virus- encoded protease (3Cpro) inhibits formation of TBP-TATA box complex in vitro. J Virol **70:**2922-2929.
98. **Clark ME, Lieberman PM, Berk AJ, Dasgupta A.** 1993. Direct cleavage of human TATA-binding protein by poliovirus protease 3C in vivo and in vitro. Mol Cell Biol **13:**1232-1237.
99. **Yalamanchili P, Weidman K, Dasgupta A.** 1997. Cleavage of transcriptional activator Oct-1 by poliovirus encoded protease 3Cpro. Virology **239:**176-185.
100. **Stocks CE, Lobigs M.** 1995. Posttranslational signal peptidase cleavage at the flavivirus C-prM junction in vitro. J Virol **69:**8123-8126.
101. **Stocks CE, Lobigs M.** 1998. Signal peptidase cleavage at the flavivirus C-prM junction: dependence on the viral NS2B-3 protease for efficient processing requires determinants in C, the signal peptide, and prM. J Virol **72:**2141-2149.
102. **Wu JZ.** 2001. Internally located signal peptides direct hepatitis C virus polyprotein processing in the ER membrane. IUBMB Life **51:**19-23.
103. **Liljestrom P, Garoff H.** 1991. Internally located cleavable signal sequences direct the formation of Semliki Forest virus membrane proteins from a polyprotein precursor. J Virol **65:**147-154.

104. **Lee E, Stocks CE, Amberg SM, Rice CM, Lobigs M.** 2000. Mutagenesis of the signal sequence of yellow fever virus prM protein: enhancement of signalase cleavage In vitro is lethal for virus production. J Virol **74:**24-32.
105. **Ait-Goughoulte M, Hourioux C, Patient R, Trassard S, Brand D, Roingeard P.** 2006. Core protein cleavage by signal peptide peptidase is required for hepatitis C virus-like particle assembly. J Gen Virol **87:**855-860.
106. **McLauchlan J, Lemberg MK, Hope G, Martoglio B.** 2002. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. EMBO J **21:**3980-3988.
107. **Pene V, Hernandez C, Vauloup-Fellous C, Garaud-Aunis J, Rosenberg AR.** 2009. Sequential processing of hepatitis C virus core protein by host cell signal peptidase and signal peptide peptidase: a reassessment. J Viral Hepat **16:**705-715.
108. **Okamoto K, Moriishi K, Miyamura T, Matsuura Y.** 2004. Intramembrane proteolysis and endoplasmic reticulum retention of hepatitis C virus core protein. J Virol **78:**6370-6380.
109. **Heimann M, Roman-Sosa G, Martoglio B, Thiel HJ, Rumenapf T.** 2006. Core protein of pestiviruses is processed at the C terminus by signal peptide peptidase. J Virol **80:**1915-1921.
110. **Targett-Adams P, Schaller T, Hope G, Lanford RE, Lemon SM, Martin A, McLauchlan J.** 2006. Signal peptide peptidase cleavage of GB virus B core protein is required for productive infection in vivo. J Biol Chem **281:**29221-29227.
111. **Voss M, Fukumori A, Kuhn PH, Kunzel U, Klier B, Grammer G, Haug-Kroper M, Kremmer E, Lichtenthaler SF, Steiner H, Schroder B, Haass C, Fluhrer R.** 2012. Foamy virus envelope protein is a substrate for signal peptide peptidase-like 3 (SPPL3). J Biol Chem **287:**43401-43409.
112. **Zhang X, Fugere M, Day R, Kielian M.** 2003. Furin processing and proteolytic activation of Semliki Forest virus. J Virol **77:**2981-2989.
113. **Moulard M, Hallenberger S, Garten W, Klenk HD.** 1999. Processing and routage of HIV glycoproteins by furin to the cell surface. Virus Res **60:**55-65.
114. **Stieneke-Grober A, Vey M, Angliker H, Shaw E, Thomas G, Roberts C, Klenk HD, Garten W.** 1992. Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease. EMBO J **11:**2407-2414.
115. **Zimmer G, Budz L, Herrler G.** 2001. Proteolytic activation of respiratory syncytial virus fusion protein. Cleavage at two furin consensus sequences. J Biol Chem **276:**31642-31650.
116. **Volchkov VE, Feldmann H, Volchkova VA, Klenk HD.** 1998. Processing of the Ebola virus glycoprotein by the proprotein convertase furin. Proc Natl Acad Sci U S A **95:**5762-5767.
117. **Martinez-Salas E, Fernandez-Miragall O.** 2004. Picornavirus IRES: structure function relationship. Curr Pharm Des **10:**3757-3767.
118. **Allaire M, Chernaia MM, Malcolm BA, James MN.** 1994. Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. Nature **369:**72-76.
119. **Ypma-Wong MF, Dewalt PG, Johnson VH, Lamb JG, Semler BL.** 1988. Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. Virology **166:**265-270.

120. **Carrere-Kremer S, Montpellier C, Lorenzo L, Brulin B, Cocquerel L, Belouzard S, Penin F, Dubuisson J.** 2004. Regulation of hepatitis C virus polyprotein processing by signal peptidase involves structural determinants at the p7 sequence junctions. J Biol Chem **279:**41384-41392.
121. **Brazzoli M, Helenius A, Foung SK, Houghton M, Abrignani S, Merola M.** 2005. Folding and dimerization of hepatitis C virus E1 and E2 glycoproteins in stably transfected CHO cells. Virology **332:**438-453.
122. **Targett-Adams P, Hope G, Boulant S, McLauchlan J.** 2008. Maturation of hepatitis C virus core protein by signal peptide peptidase is required for virus production. J Biol Chem **283:**16850-16859.
123. **Failla C, Tomei L, De Francesco R.** 1994. Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. J Virol **68:**3753-3760.
124. **Koch JO, Lohmann V, Herian U, Bartenschlager R.** 1996. In vitro studies on the activation of the hepatitis C virus NS3 proteinase by the NS4A cofactor. Virology **221:**54-66.
125. **Arnold E, Luo M, Vriend G, Rossmann MG, Palmenberg AC, Parks GD, Nicklin MJ, Wimmer E.** 1987. Implications of the picornavirus capsid structure for polyprotein processing. Proc Natl Acad Sci U S A **84:**21-25.
126. **van Bokhoven H, van Lent JW, Custers R, Vlak JM, Wellink J, van Kammen A.** 1992. Synthesis of the complete 200K polyprotein encoded by cowpea mosaic virus B-RNA in insect cells. J Gen Virol **73 (Pt 11):**2775-2784.
127. **Jore J, De Geus B, Jackson RJ, Pouwels PH, Enger-Valk BE.** 1988. Poliovirus protein 3CD is the active protease for processing of the precursor protein P1 in vitro. J Gen Virol **69 (Pt 7):**1627-1636.
128. **Harris KS, Reddigari SR, Nicklin MJ, Hammerle T, Wimmer E.** 1992. Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. J Virol **66:**7481-7489.
129. **Chisholm J, Wieczorek A, Sanfacon H.** 2001. Expression and partial purification of recombinant tomato ringspot nepovirus 3C-like proteinase: comparison of the activity of the mature proteinase and the VPg-proteinase precursor. Virus Res **79:**153-164.
130. **Margis R, Viry M, Pinck M, Bardonnet N, Pinck L.** 1994. Differential proteolytic activities of precursor and mature forms of the 24K proteinase of grapevine fanleaf nepovirus. Virology **200:**79-86.
131. **Molla A, Harris KS, Paul AV, Shin SH, Mugavero J, Wimmer E.** 1994. Stimulation of poliovirus proteinase 3Cpro-related proteolysis by the genome-linked protein VPg and its precursor 3AB. J Biol Chem **269:**27015-27020.
132. **Dessens JT, Lomonossoff GP.** 1992. Sequence upstream of the 24K protease enhances cleavage of the cowpea mosaic virus B RNA-encoded polyprotein at the junction between the 24K and 87K proteins. Virology **189:**225-232.
133. **Vos P, Verver J, Jaegle M, Wellink J, van Kammen A, Goldbach R.** 1988. Two viral proteins involved in the proteolytic processing of the cowpea mosaic virus polyproteins. Nucleic Acids Res **16:**1967-1985.
134. **Sanfacon H, Gorbalenya AE, Knowles NJ, Chen Y.** 2011. Order Picornavirales, p. 835-839. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed.), Virus

Taxonomy: Classification and Nomenclature of Viruses. Ninth Report of the International Committee on the Taxonomy of Viruses. Elseviers, San Diego.
135. **Sanfacon H.** 2009. Secoviridae: The amalgamation of the families Sequiviridae and Comoviridae, Encyclopedia of Life Sciences (ELS). John Wiley & Sons Ltd, Chichester.
136. **Taylor MP, Kirkegaard K.** 2008. Potential subversion of autophagosomal pathway by picornaviruses. Autophagy **4:**286-289.
137. **Klein KA, Jackson WT.** 2011. Picornavirus subversion of the autophagy pathway. Viruses **3:**1549-1561.
138. **Wang J, Wu Z, Jin Q.** 2012. COPI is required for enterovirus 71 replication. PLoS One **7:**e38035.
139. **Schlegel A, Giddings TH, Jr., Ladinsky MS, Kirkegaard K.** 1996. Cellular origin and ultrastructure of membranes induced during poliovirus infection. J Virol **70:**6576-6588.
140. **Sanfacon H.** 2013. Investigating the role of viral integral membrane proteins in promoting the assembly of nepovirus and comovirus replication factories. Frontiers in Plant Science **3:**313.
141. **Chase AJ, Semler BL.** 2012. Viral subversion of host functions for picornavirus translation and RNA replication. Future virology **7:**179-191.
142. **Lin JY, Chen TC, Weng KF, Chang SC, Chen LL, Shih SR.** 2009. Viral and host proteins involved in picornavirus life cycle. Journal of biomedical science **16:**103.
143. **Belov GA, Sztul E.** 2014. Rewiring of cellular membrane homeostasis by picornaviruses. J Virol **88:**9478-9489.
144. **Egger D, Bienz K.** 2002. Recombination of poliovirus RNA proceeds in mixed replication complexes originating from distinct replication start sites. J Virol **76:**10960-10971.
145. **Bienz K, Egger D, Pfister T, Troxler M.** 1992. Structural and functional characterization of the poliovirus replication complex. J Virol **66:**2740-2747.
146. **Fujita K, Krishnakumar SS, Franco D, Paul AV, London E, Wimmer E.** 2007. Membrane topography of the hydrophobic anchor sequence of poliovirus 3A and 3AB proteins and the functional effect of 3A/3AB membrane association upon RNA replication. Biochemistry **46:**5185-5199.
147. **Johnson KL, Sarnow P.** 1991. Three poliovirus 2B mutants exhibit noncomplementable defects in viral RNA amplification and display dosage-dependent dominance over wild-type poliovirus. J Virol **65:**4341-4349.
148. **van Kuppeveld FJ, Galama JM, Zoll J, Melchers WJ.** 1995. Genetic analysis of a hydrophobic domain of coxsackie B3 virus protein 2B: a moderate degree of hydrophobicity is required for a cis-acting function in viral RNA synthesis. J Virol **69:**7782-7790.
149. **Van kuppeveld FJ, Melchers WJ, Kirkegaard K, Doedens JR.** 1997. Structure-function analysis of coxsackie B3 virus protein 2B. Virology **227:**111-118.
150. **Agirre A, Barco A, Carrasco L, Nieva JL.** 2002. Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. J Biol Chem **277:**40434-40441.
151. **van Kuppeveld FJ, Hoenderop JG, Smeets RL, Willems PH, Dijkman HB, Galama JM, Melchers WJ.** 1997. Coxsackievirus protein 2B modifies endoplasmic

reticulum membrane and plasma membrane permeability and facilitates virus release. EMBO J **16:**3519-3532.
152. **Echeverri AC, Dasgupta A.** 1995. Amino terminal regions of poliovirus 2C protein mediate membrane binding. Virology **208:**540-553.
153. **Paul AV, Molla A, Wimmer E.** 1994. Studies of a putative amphipathic helix in the N-terminus of poliovirus protein 2C. Virology **199:**188-199.
154. **Echeverri A, Banerjee R, Dasgupta A.** 1998. Amino-terminal region of poliovirus 2C protein is sufficient for membrane binding. Virus Res **54:**217-223.
155. **Cho MW, Teterina N, Egger D, Bienz K, Ehrenfeld E.** 1994. Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells. Virology **202:**129-145.
156. **Aldabe R, Carrasco L.** 1995. Induction of membrane proliferation by poliovirus proteins 2C and 2BC. Biochem Biophys Res Commun **206:**64-76.
157. **Li JP, Baltimore D.** 1988. Isolation of poliovirus 2C mutants defective in viral RNA synthesis. J Virol **62:**4016-4021.
158. **Banerjee R, Echeverri A, Dasgupta A.** 1997. Poliovirus-encoded 2C polypeptide specifically binds to the 3'-terminal sequences of viral negative-strand RNA. J Virol **71:**9570-9578.
159. **Banerjee R, Tsai W, Kim W, Dasgupta A.** 2001. Interaction of poliovirus-encoded 2C/2BC polypeptides with the 3' terminus negative-strand cloverleaf requires an intact stem-loop b. Virology **280:**41-51.
160. **Banerjee R, Dasgupta A.** 2001. Interaction of picornavirus 2C polypeptide with the viral negative-strand RNA. J Gen Virol **82:**2621-2627.
161. **Teterina NL, Kean KM, Gorbalenya AE, Agol VI, Girard M.** 1992. Analysis of the functional significance of amino acid residues in the putative NTP-binding pattern of the poliovirus 2C protein. J Gen Virol **73 (Pt 8):**1977-1986.
162. **Rodriguez PL, Carrasco L.** 1993. Poliovirus protein 2C has ATPase and GTPase activities. J Biol Chem **268:**8105-8110.
163. **Adams P, Kandiah E, Effantin G, Steven AC, Ehrenfeld E.** 2009. Poliovirus 2C protein forms homo-oligomeric structures required for ATPase activity. The Journal of biological chemistry **284:**22012-22021.
164. **Banerjee R, Weidman MK, Echeverri A, Kundu P, Dasgupta A.** 2004. Regulation of poliovirus 3C protease by the 2C polypeptide. J Virol **78:**9243-9256.
165. **Datta U, Dasgupta A.** 1994. Expression and subcellular localization of poliovirus VPg-precursor protein 3AB in eukaryotic cells: evidence for glycosylation in vitro. J Virol **68:**4468-4477.
166. **Towner JS, Ho TV, Semler BL.** 1996. Determinants of membrane association for poliovirus protein 3AB. J Biol Chem **271:**26810-26818.
167. **Giachetti C, Hwang SS, Semler BL.** 1992. cis-acting lesions targeted to the hydrophobic domain of a poliovirus membrane protein involved in RNA replication. J Virol **66:**6045-6057.
168. **Strauss DM, Wuttke DS.** 2007. Characterization of protein-protein interactions critical for poliovirus replication: analysis of 3AB and VPg binding to the RNA-dependent RNA polymerase. J Virol **81:**6369-6378.

169. **Xiang W, Harris KS, Alexander L, Wimmer E.** 1995. Interaction between the 5'-terminal cloverleaf and 3AB/3CDpro of poliovirus is essential for RNA replication. J Virol **69:**3658-3667.
170. **Wessels E, Duijsings D, Lanke KH, van Dooren SH, Jackson CL, Melchers WJ, van Kuppeveld FJ.** 2006. Effects of picornavirus 3A Proteins on Protein Transport and GBF1-dependent COP-I recruitment. J Virol **80:**11852-11860.
171. **Xiang W, Cuconati A, Hope D, Kirkegaard K, Wimmer E.** 1998. Complete protein linkage map of poliovirus P3 proteins: interaction of polymerase 3Dpol with VPg and with genetic variants of 3AB. J Virol **72:**6732-6741.
172. **Strauss DM, Glustrom LW, Wuttke DS.** 2003. Towards an understanding of the poliovirus replication complex: the solution structure of the soluble domain of the poliovirus 3A protein. J Mol Biol **330:**225-234.
173. **Yin J, Liu Y, Wimmer E, Paul AV.** 2007. Complete protein linkage map between the P2 and P3 non-structural proteins of poliovirus. J Gen Virol **88:**2259-2267.
174. **Teterina NL, Levenson E, Rinaudo MS, Egger D, Bienz K, Gorbalenya AE, Ehrenfeld E.** 2006. Evidence for functional protein interactions required for poliovirus RNA replication. J Virol **80:**5327-5337.
175. **de Zoeten GA, Assink AM, van Kammen A.** 1974. Association of cowpea mosaic virus-induced double-stranded RNA with a cytopathological structure in infected cells. Virology **59:**341-355.
176. **Carette JE, Stuiver M, Van Lent J, Wellink J, Van Kammen A.** 2000. Cowpea mosaic virus infection induces a massive proliferation of endoplasmic reticulum but not Golgi membranes and is dependent on de novo membrane synthesis. J Virol **74:**6556-6563.
177. **Carette JE, van Lent J, MacFarlane SA, Wellink J, van Kammen A.** 2002. Cowpea mosaic virus 32- and 60-kilodalton replication proteins target and change the morphology of endoplasmic reticulum membranes. J Virol **76:**6293-6301.
178. **van Bokhoven H, Verver J, Wellink J, van Kammen A.** 1993. Protoplasts transiently expressing the 200K coding sequence of cowpea mosaic virus B-RNA support replication of M-RNA. J Gen Virol **74 (Pt 10):**2233-2241.
179. **Carette JE, Guhl K, Wellink J, Van Kammen A.** 2002. Coalescence of the sites of cowpea mosaic virus RNA replication into a cytopathic structure. J Virol **76:**6235-6243.
180. **van Bokhoven H, Wellink J, Usmany M, Vlak JM, Goldbach R, van Kammen A.** 1990. Expression of plant virus genes in animal cells: high-level synthesis of cowpea mosaic virus B-RNA-encoded proteins with baculovirus expression vectors. J Gen Virol **71 (Pt 11):**2509-2517.
181. **Peters SA, Verver J, Nollen EA, van Lent JW, Wellink J, van Kammen A.** 1994. The NTP-binding motif in cowpea mosaic virus B polyprotein is essential for viral replication. J Gen Virol **75 (Pt 11):**3167-3176.
182. **Carette JE, Verver J, Martens J, van Kampen T, Wellink J, van Kammen A.** 2002. Characterization of plant proteins that interact with cowpea mosaic virus '60K' protein in the yeast two-hybrid system. J Gen Virol **83:**885-893.
183. **Lain S, Riechmann JL, Martin MT, Garcia JA.** 1989. Homologous potyvirus and flavivirus proteins belonging to a superfamily of helicase-like proteins. Gene **82:**357-362.

184. **Wellink J, Rezelman G, Goldbach R, Beyreuther K.** 1986. Determination of the proteolytic processing sites in the polyprotein encoded by the bottom-component RNA of Cowpea mosaic virus. J Virol **59:**50-58.
185. **Peters SA, Voorhorst WG, Wery J, Wellink J, van Kammen A.** 1992. A regulatory role for the 32K protein in proteolytic processing of cowpea mosaic virus polyproteins. Virology **191:**81-89.
186. **Gu H, Ghabrial SA.** 2005. The Bean pod mottle virus proteinase cofactor and putative helicase are symptom severity determinants. Virology **333:**271-283.
187. **Hashimoto M, Komatsu K, Iwai R, Keima T, Maejima K, Shiraishi T, Ishikawa K, Yoshida T, Kitazawa Y, Okano Y, Yamaji Y, Namba S.** 2015. Cell Death Triggered by a Putative Amphipathic Helix of Radish mosaic virus Helicase Protein is Tightly Correlated with Host Membrane Modification. Mol Plant Microbe Interact.
188. **Ritzenthaler C, Laporte C, Gaire F, Dunoyer P, Schmitt C, Duval S, Piequet A, Loudes AM, Rohfritsch O, Stussi-Garaud C, Pfeiffer P.** 2002. Grapevine fanleaf virus replication occurs on endoplasmic reticulum-derived membranes. J Virol **76:**8808-8819.
189. **Gaire F, Schmitt C, Stussi-Garaud C, Pinck L, Ritzenthaler C.** 1999. Protein 2A of grapevine fanleaf nepovirus is implicated in RNA2 replication and colocalizes to the replication site. Virology **264:**25-36.
190. **Viry M, Serghini MA, Hans F, Ritzenthaler C, Pinck M, Pinck L.** 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. J Gen Virol **74:**169-174.
191. **Lee YF, Nomoto A, Detjen BM, Wimmer E.** 1977. A protein covalently linked to poliovirus genome RNA. Proc Natl Acad Sci U S A **74:**59-63.
192. **Mayo MA, Barker H, Harrison BD.** 1982. Specificity and properties of the genome-linked proteins of nepoviruses. J Gen Virol **59:**149-162.
193. **Tobin GJ, Young DC, Flanegan JB.** 1989. Self-catalyzed linkage of poliovirus terminal protein VPg to poliovirus RNA. Cell **59:**511-519.
194. **Goodfellow I.** 2011. The genome-linked protein VPg of vertebrate viruses - a multifaceted protein. Curr Opin Virol **1:**355-362.
195. **Steil BP, Barton DJ.** 2009. Conversion of VPg into VPgpUpUOH before and during poliovirus negative-strand RNA synthesis. J Virol **83:**12660-12670.
196. **Paul AV, van Boom JH, Filippov D, Wimmer E.** 1998. Protein-primed RNA synthesis by purified poliovirus RNA polymerase. Nature **393:**280-284.
197. **Paul AV, Rieder E, Kim DW, van Boom JH, Wimmer E.** 2000. Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. J Virol **74:**10359-10370.
198. **Goodfellow IG, Polacek C, Andino R, Evans DJ.** 2003. The poliovirus 2C cis-acting replication element-mediated uridylylation of VPg is not required for synthesis of negative-sense genomes. J Gen Virol **84:**2359-2363.
199. **Murray KE, Barton DJ.** 2003. Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. J Virol **77:**4739-4750.
200. **Morasco BJ, Sharma N, Parilla J, Flanegan JB.** 2003. Poliovirus cre(2C)-dependent synthesis of VPgpUpU is required for positive- but not negative-strand RNA synthesis. J Virol **77:**5136-5144.

201. **Lyons T, Murray KE, Roberts AW, Barton DJ.** 2001. Poliovirus 5'-terminal cloverleaf RNA is required in cis for VPg uridylylation and the initiation of negative-strand RNA synthesis. J Virol **75:**10696-10708.
202. **Fogg MH, Teterina NL, Ehrenfeld E.** 2003. Membrane Requirements for Uridylylation of the Poliovirus VPg Protein and Viral RNA Synthesis In Vitro. J Virol **77:**11408-11416.
203. **Pinck M, Reinbolt J, Loudes AM, Le Ret M, Pinck L.** 1991. Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus. FEBS Lett **284:**117-119.
204. **Zalloua PA, Buzayan JM, Bruening G.** 1996. Chemical cleavage of 5'-linked protein from tobacco ringspot virus genomic RNAs and characterization of the protein-RNA linkage. Virology **219:**1-8.
205. **Carette JE, Kujawa A, Guhl K, Verver J, Wellink J, Van Kammen A.** 2001. Mutational analysis of the genome-linked protein of cowpea mosaic virus. Virology **290:**21-29.
206. **Cornell CT, Semler BL.** 2002. Subdomain specific functions of the RNA polymerase region of poliovirus 3CD polypeptide. Virology **298:**200-213.
207. **Blair WS, Parsley TB, Bogerd HP, Towner JS, Semler BL, Cullen BR.** 1998. Utilization of a mammalian cell-based RNA binding assay to characterize the RNA binding properties of picornavirus 3C proteinases. Rna **4:**215-225.
208. **Blair WS, Nguyen JH, Parsley TB, Semler BL.** 1996. Mutations in the poliovirus 3CD proteinase S1-specificity pocket affect substrate recognition and RNA binding. Virology **218:**1-13.
209. **Leonard S, Chisholm J, Laliberte JF, Sanfacon H.** 2002. Interaction in vitro between the proteinase of Tomato ringspot virus (genus Nepovirus) and the eukaryotic translation initiation factor iso4E from Arabidopsis thaliana. J Gen Virol **83:**2085-2089.
210. **Lee JH, Alam I, Han KR, Cho S, Shin S, Kang S, Yang JM, Kim KH.** 2011. Crystal structures of murine norovirus-1 RNA-dependent RNA polymerase. J Gen Virol **92:**1607-1616.
211. **Yap TL, Xu T, Chen YL, Malet H, Egloff MP, Canard B, Vasudevan SG, Lescar J.** 2007. Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. J Virol **81:**4753-4765.
212. **Malet H, Egloff MP, Selisko B, Butcher RE, Wright PJ, Roberts M, Gruez A, Sulzenbacher G, Vonrhein C, Bricogne G, Mackenzie JM, Khromykh AA, Davidson AD, Canard B.** 2007. Crystal structure of the RNA polymerase domain of the West Nile virus non-structural protein 5. J Biol Chem **282:**10678-10689.
213. **Boonrod K, Chotewutmontri S, Galetzka D, Krczal G.** 2005. Analysis of tombusvirus revertants to identify essential amino acid residues within RNA-dependent RNA polymerase motifs. J Gen Virol **86:**823-826.
214. **Gohara DW, Crotty S, Arnold JJ, Yoder JD, Andino R, Cameron CE.** 2000. Poliovirus RNA-dependent RNA polymerase (3Dpol): structural, biochemical, and biological analysis of conserved structural motifs A and B. J Biol Chem **275:**25523-25532.

215. **Kamer G, Argos P.** 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. Nucleic Acids Res **12:**7269-7282.
216. **Wang Y, Xiao M, Chen J, Zhang W, Luo J, Bao K, Nie M, Li B.** 2007. Mutational analysis of the GDD sequence motif of classical swine fever virus RNA-dependent RNA polymerases. Virus Genes **34:**63-65.
217. **Wang X, Gillam S.** 2001. Mutations in the GDD motif of rubella virus putative RNA-dependent RNA polymerase affect virus replication. Virology **285:**322-331.
218. **Li XH, Carrington JC.** 1995. Complementation of tobacco etch potyvirus mutants by active RNA polymerase expressed in transgenic cells. Proc Natl Acad Sci U S A **92:**457-461.
219. **Ahlquist P.** 2002. RNA-dependent RNA polymerases, viruses, and RNA silencing. Science **296:**1270-1273.
220. **Perez-Losada M, Arenas M, Galan JC, Palero F, Gonzalez-Candelas F.** 2015. Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences. Infect Genet Evol **30C:**296-307.
221. **Mackenzie J.** 2005. Wrapping things up about virus RNA replication. Traffic **6:**967-977.
222. **Restrepo-Hartwig MA, Ahlquist P.** 1996. Brome mosaic virus helicase- and polymerase-like proteins colocalize on the endoplasmic reticulum at sites of viral RNA synthesis. J Virol **70:**8908-8916.
223. **Jakubiec A, Notaise J, Tournier V, Hericourt F, Block MA, Drugeon G, van Aelst L, Jupin I.** 2004. Assembly of turnip yellow mosaic virus replication complexes: interaction between the proteinase and polymerase domains of the replication proteins. J Virol **78:**7945-7957.
224. **Miller DJ, Ahlquist P.** 2002. Flock house virus RNA polymerase is a transmembrane protein with amino-terminal sequences sufficient for mitochondrial localization and membrane insertion. J Virol **76:**9856-9867.
225. **Lee KJ, Choi J, Ou JH, Lai MM.** 2004. The C-terminal transmembrane domain of hepatitis C virus (HCV) RNA polymerase is essential for HCV replication in vivo. J Virol **78:**3797-3802.
226. **Hansen JL, Long AM, Schultz SC.** 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. Structure **5:**1109-1122.
227. **Hobson SD, Rosenblum ES, Richards OC, Richmond K, Kirkegaard K, Schultz SC.** 2001. Oligomeric structures of poliovirus polymerase are important for function. EMBO J **20:**1153-1163.
228. **Pata JD, Schultz SC, Kirkegaard K.** 1995. Functional oligomerization of poliovirus RNA-dependent RNA polymerase. Rna **1:**466-477.
229. **Tellez AB, Wang J, Tanner EJ, Spagnolo JF, Kirkegaard K, Bullitt E.** 2011. Interstitial contacts in an RNA-dependent RNA polymerase lattice. J Mol Biol **412:**737-750.
230. **Lyle JM, Bullitt E, Bienz K, Kirkegaard K.** 2002. Visualization and functional analysis of RNA-dependent RNA polymerase lattices. Science **296:**2218-2222.
231. **Paul AV, Cao X, Harris KS, Lama J, Wimmer E.** 1994. Studies with poliovirus polymerase 3Dpol. Stimulation of poly(U) synthesis in vitro by purified poliovirus protein 3AB. J Biol Chem **269:**29173-29181.

232. **Hope DA, Diamond SE, Kirkegaard K.** 1997. Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3AB. J Virol **71:**9490-9498.
233. **Andino R, Rieckhof GE, Achacoso PL, Baltimore D.** 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. EMBO J **12:**3587-3598.
234. **Parsley TB, Towner JS, Blyn LB, Ehrenfeld E, Semler BL.** 1997. Poly (rC) binding protein 2 forms a ternary complex with the 5'-terminal sequences of poliovirus RNA and the viral 3CD proteinase. Rna **3:**1124-1134.
235. **Rieder E, Xiang W, Paul A, Wimmer E.** 2003. Analysis of the cloverleaf element in a human rhinovirus type 14/poliovirus chimera: correlation of subdomain D structure, ternary protein complex formation and virus replication. J Gen Virol **84:**2203-2216.
236. **Andino R, Rieckhof GE, Baltimore D.** 1990. A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. Cell **63:**369-380.
237. **Wei L, Huhn JS, Mory A, Pathak HB, Sosnovtsev SV, Green KY, Cameron CE.** 2001. Proteinase-polymerase precursor as the active form of feline calicivirus RNA-dependent RNA polymerase. J Virol **75:**1211-1219.
238. **Belliot G, Sosnovtsev SV, Chang KO, Babu V, Uche U, Arnold JJ, Cameron CE, Green KY.** 2005. Norovirus proteinase-polymerase and polymerase are both active forms of RNA-dependent RNA polymerase. J Virol **79:**2393-2403.
239. **Machin A, Martin Alonso JM, Dalton KP, Parra F.** 2009. Functional differences between precursor and mature forms of the RNA-dependent RNA polymerase from rabbit hemorrhagic disease virus. J Gen Virol **90:**2114-2118.
240. **Dorssers L, Van der Kroll S, Van der Meer J, Van Kammen A, Zabel P.** 1984. Purification of cowpea mosaic virus RNA replication complex: identification of a virus-encoded 110,000-dalton polypeptide responsible for RNA chain elongation. Proc Natl Acad Sci U S A **81:**1951-1955.
241. **Demangeat G, Hemmer O, Reinbolt J, Mayo MA, Fritsch C.** 1992. Virus-specific proteins in cells infected with tomato black ring nepovirus: evidence for proteolytic processing in vivo. J Gen Virol **73:**1609-1614.
242. **Vigne E, Gottula J, Schmitt-Keichinger C, Komar V, Ackerer L, Belval L, Rakotomalala L, Lemaire O, Ritzenthaler C, Fuchs M.** 2013. A strain-specific segment of the RNA-dependent RNA polymerase of Grapevine fanleaf virus determines symptoms in Nicotiana species. J Gen Virol.
243. **Brunt AA, Crabtree K, Dallwitz MJ, Gibbs AJ, Watson L, Zurcher EJ** 1996, posting date. Plant Viruses Online: Descriptions and Lists from the VIDE Database. [Online.]
244. **Stace-Smith R.** 1984. Tomato ringspot virus. CMI/AAB Descriptions of Plant Viruses **290**.
245. **Brown DJF, Halbrendt JM, Robbins RT, Vrain TC.** 1993. Transmission of nepoviruses by Xiphinema americanum-group nematodes. J Nematol **25:**349-354.
246. **Brown DJF, Trudgill DL, Robertson WM.** 1996. Nepoviruses: Transmission by nematodes, p. 187-209. *In* Harrison BD, Murant AF (ed.), The plant viruses, Volume 5: Polyhedral virions and bipartite RNA genomes. Plenum Press, New York.

247. **Andret-Link P, Schmitt-Keichinger C, Demangeat G, Komar V, Fuchs M.** 2004. The specific transmission of Grapevine fanleaf virus by its nematode vector Xiphinema index is solely determined by the viral coat protein. Virology **320:**12-22.
248. **Rott ME, Gilchrist A, Lee L, Rochon D.** 1995. Nucleotide sequence of tomato ringspot virus RNA1. J Gen Virol **76:**465-473.
249. **Rott ME, Tremaine JH, Rochon DM.** 1991. Nucleotide sequence of tomato ringspot virus RNA-2. J Gen Virol **72:**1505-1514.
250. **Walker M, Chisholm J, Wei T, Ghoshal B, Saeed H, Rott M, Sanfacon H.** 2015. Complete genome sequence of three tomato ringspot virus isolates: evidence for reassortment and recombination. Arch Virol **160:**543-547.
251. **Rott ME, Tremaine JH, Rochon DM.** 1991. Comparison of the 5' and 3' termini of tomato ringspot virus RNA1 and RNA2: evidence for RNA recombination. Virology **185:**468-472.
252. **Karetnikov A, Lehto K.** 2007. The RNA2 5' leader of Blackcurrant reversion virus mediates efficient in vivo translation through an internal ribosomal entry site mechanism. J Gen Virol **88:**286-297.
253. **Hans F, Sanfacon H.** 1995. Tomato ringspot nepovirus protease: characterization and cleavage site specificity. J Gen Virol **76:**917-927.
254. **Sanfacon H, Zhang G, Chisholm J, Jafarpour B, Jovel J.** 2006. Molecular biology of Tomato ringspot nepovirus, a pathogen of ornamentals, small fruits and fruit trees, p. 540-546. *In* Teixeira da Silva J (ed.), Floriculture, Ornamental and Plant Biotechnology: Advances and Topical Issues (1st Edition), vol. III. Global Science Books, London, UK.
255. **Jafarpour B, Sanfacon H.** 2009. Insertion of large amino acid repeats and point mutations contribute to a high degree of sequence diversity in the X4 protein of tomato ringspot virus (genus Nepovirus). Arch Virol **154:**1713-1717.
256. **Wang A, Sanfacon H.** 2000. Proteolytic processing at a novel cleavage site in the N-terminal region of the tomato ringspot nepovirus RNA-1-encoded polyprotein in vitro. J Gen Virol **81:**2771-2781.
257. **Zhang SC, Zhang G, Yang L, Chisholm J, Sanfacon H.** 2005. Evidence that insertion of Tomato ringspot nepovirus NTB-VPg protein in endoplasmic reticulum membranes is directed by two domains: a C-terminal transmembrane helix and an N-terminal amphipathic helix. J Virol **79:**11752-11765.
258. **Zhang G, Sanfacon H.** 2006. Characterization of membrane-association domains within the Tomato ringspot nepovirus X2 protein, an endoplasmic reticulum-targeted polytopic membrane protein. J Virol **80:**10847-10857.
259. **Peters SA, Mesnard JM, Kooter IM, Verver J, Wellink J, van Kammen A.** 1995. The cowpea mosaic virus RNA 1-encoded 112 kDa protein may function as a VPg precursor in vivo. J Gen Virol **76 (Pt 7):**1807-1813.
260. **Ghoshal B, Sanfacon H.** 2014. Temperature-dependent symptom recovery in Nicotiana benthamiana plants infected with tomato ringspot virus is associated with reduced translation of viral RNA2 and requires ARGONAUTE 1. Virology **456-457:**188-197.
261. **Tusnady GE, Simon I.** 1998. Principles governing amino acid composition of integral membrane proteins: Applications to topology prediction. J Mol Biol **283:**489-506.

262. **King RD, Saqi M, Sayle R, Sternberg MJ.** 1997. DSC: public domain protein secondary structure predication. Comput Appl Biosci **13:**473-474.
263. **Guermeur Y, Geourjon C, Gallinari P, Deleage G.** 1999. Improved performance in protein secondary structure prediction by inhomogeneous score combination. Bioinformatics **15:**413-421.
264. **Rost B.** 1996. PHD: predicting one-dimensional protein structure by profile-based neural networks. Methods Enzymol **266:**525-539.
265. **Sun F, Xiang Y, Sanfacon H.** 2001. Homology-dependent resistance to tomato ringspot nepovirus in plants transformed with the VPg-protease coding region. Can J Plant Pathol **23:**292-299.
266. **Karran RA, Sanfacon H.** 2014. Tomato ringspot virus Coat Protein Binds to ARGONAUTE 1 and Suppresses the Translation Repression of a Reporter Gene. Mol Plant Microbe Interact **27:**933-943.
267. **Ratcliff F, Martin-Hernandez AM, Baulcombe DC.** 2001. Technical Advance. Tobacco rattle virus as a vector for analysis of gene function by silencing. Plant J **25:**237-245.
268. **Hanson SR, Culyba EK, Hsu TL, Wong CH, Kelly JW, Powers ET.** 2009. The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. Proc Natl Acad Sci U S A **106:**3131-3136.
269. **Wormald MR, Dwek RA.** 1999. Glycoproteins: glycan presentation and protein-fold stability. Structure **7:**R155-160.
270. **Rudd PM, Joao HC, Coghill E, Fiten P, Saunders MR, Opdenakker G, Dwek RA.** 1994. Glycoforms modify the dynamic stability and functional activity of an enzyme. Biochemistry **33:**17-22.
271. **Nilsson I, Johnson AE, von Heijne G.** 2002. Cleavage of a tail-anchored protein by signal peptidase. FEBS Lett **516:**106-108.
272. **Nilsson I, Whitley P, von Heijne G.** 1994. The COOH-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase. J Cell Biol **126:**1127-1132.
273. **Nothwehr SF, Folz RJ, Gordon JI.** 1989. Uncoupling of co-translational translocation from signal peptidase processing in a mutant rat preapolipoprotein-A-IV with a deletion that includes the COOH-terminal region of its signal peptide. J Biol Chem **264:**4642-4647.
274. **Yeagle PL, Bennett M, Lemaitre V, Watts A.** 2007. Transmembrane helices of membrane proteins may flex to satisfy hydrophobic mismatch. Biochim Biophys Acta **1768:**530-537.
275. **Bintintan I, Meyers G.** 2010. A new type of signal peptidase cleavage site identified in an RNA virus polyprotein. J Biol Chem **285:**8572-8584.
276. **Price JL, Culyba EK, Chen W, Murray AN, Hanson SR, Wong CH, Powers ET, Kelly JW.** 2012. N-glycosylation of enhanced aromatic sequons to increase glycoprotein stability. Biopolymers **98:**195-211.
277. **Dubuisson J, Duvet S, Meunier JC, Op De Beeck A, Cacan R, Wychowski C, Cocquerel L.** 2000. Glycosylation of the hepatitis C virus envelope protein E1 is dependent on the presence of a downstream sequence on the viral polyprotein. J Biol Chem **275:**30605-30609.

278. **Vigerust DJ, Shepherd VL.** 2007. Virus glycosylation: role in virulence and immune interactions. Trends Microbiol **15:**211-218.
279. **Fan J, Liu Y, Yuan Z.** 2014. Critical role of Dengue Virus NS1 protein in viral replication. Virol Sin **29:**162-169.
280. **Muylaert IR, Chambers TJ, Galler R, Rice CM.** 1996. Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. Virology **222:**159-168.
281. **Gadlage MJ, Sparks JS, Beachboard DC, Cox RG, Doyle JD, Stobart CC, Denison MR.** 2010. Murine hepatitis virus nonstructural protein 4 regulates virus-induced membrane modifications and replication complex function. J Virol **84:**280-290.
282. **Wang A, Carrier K, Chisholm J, Wieczorek A, Huguenot C, Sanfacon H.** 1999. Proteolytic processing of tomato ringspot nepovirus 3C-like protease precursors: definition of the domains for the VPg, protease and putative RNA-dependent RNA polymerase. J Gen Virol **80:**799-809.
283. **Voinnet O, Rivas S, Mestre P, Baulcombe D.** 2003. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J **33:**949-956.
284. **Xia Z, Zhu Z, Zhu J, Zhou R.** 2009. Recognition mechanism of siRNA by viral p19 suppressor of RNA silencing: a molecular dynamics study. Biophys J **96:**1761-1769.
285. **Zabka M, Lesniak W, Prus W, Kuznicki J, Filipek A.** 2008. Sgt1 has co-chaperone properties and is up-regulated by heat shock. Biochemical and biophysical research communications **370:**179-183.
286. **Kadota Y, Amigues B, Ducassou L, Madaoui H, Ochsenbein F, Guerois R, Shirasu K.** 2008. Structural and functional analysis of SGT1-HSP90 core complex required for innate immunity in plants. EMBO reports **9:**1209-1215.
287. **Mayor A, Martinon F, De Smedt T, Petrilli V, Tschopp J.** 2007. A crucial function of SGT1 and HSP90 in inflammasome activity links mammalian and plant innate immune responses. Nature immunology **8:**497-503.
288. **Shirasu K.** 2009. The HSP90-SGT1 chaperone complex for NLR immune sensors. Annu Rev Plant Biol **60:**139-164.
289. **Spiechowicz M, Zylicz A, Bieganowski P, Kuznicki J, Filipek A.** 2007. Hsp70 is a new target of Sgt1--an interaction modulated by S100A6. Biochemical and biophysical research communications **357:**1148-1153.
290. **Ye CM, Kelly V, Payton M, Dickman MB, Verchot J.** 2012. SGT1 Is Induced by the Potato virus X TGBp3 and Enhances Virus Accumulation in Nicotiana benthamiana. Mol Plant.
291. **Wang K, Uppalapati SR, Zhu X, Dinesh-Kumar SP, Mysore KS.** 2010. SGT1 positively regulates the process of plant cell death during both compatible and incompatible plant-pathogen interactions. Mol Plant Pathol **11:**597-611.
292. **Qi Y, Hannon GJ.** 2005. Uncovering RNAi mechanisms in plants: biochemistry enters the foray. FEBS Lett **579:**5899-5903.
293. **Bagasra O, Prilliman KR.** 2004. RNA interference: the molecular immune system. J Mol Histol **35:**545-553.
294. **Iwasaki S, Tomari Y.** 2009. Argonaute-mediated translational repression (and activation). Fly (Austin) **3:**204-206.

295. **Burgyan J, Havelda Z.** 2011. Viral suppressors of RNA silencing. Trends Plant Sci **16:**265-272.
296. **Qiu W, Park JW, Scholthof HB.** 2002. Tombusvirus P19-mediated suppression of virus-induced gene silencing is controlled by genetic and dosage features that influence pathogenicity. Mol Plant Microbe Interact **15:**269-280.
297. **Jin H, Zhu JK.** 2010. A viral suppressor protein inhibits host RNA silencing by hooking up with Argonautes. Genes Dev **24:**853-856.
298. **Bivalkar-Mehla S, Vakharia J, Mehla R, Abreha M, Kanwar JR, Tikoo A, Chauhan A.** 2010. Viral RNA silencing suppressors (RSS): Novel strategy of viruses to ablate the host RNA interference (RNAi) defense system. Virus Res.
299. **Camborde L, Planchais S, Tournier V, Jakubiec A, Drugeon G, Lacassagne E, Pflieger S, Chenon M, Jupin I.** 2010. The ubiquitin-proteasome system regulates the accumulation of Turnip yellow mosaic virus RNA-dependent RNA polymerase during viral infection. Plant Cell **22:**3142-3152.
300. **Miller WA, White KA.** 2006. Long-Distance RNA-RNA Interactions in Plant Virus Gene Expression and Replication. Annu Rev Phytopathol.
301. **Watanabe T, Honda A, Iwata A, Ueda S, Hibi T, Ishihama A.** 1999. Isolation from tobacco mosaic virus-infected tobacco of a solubilized template-specific RNA-dependent RNA polymerase containing a 126K/183K protein heterodimer. J Virol **73:**2633-2640.
302. **Scholthof KB, Scholthof HB, Jackson AO.** 1995. The tomato bushy stunt virus replicase proteins are coordinately expressed and membrane associated. Virology **208:**365-369.
303. **Schwartz M, Chen J, Janda M, Sullivan M, den Boon J, Ahlquist P.** 2002. A positive-strand RNA virus replication complex parallels form and function of retrovirus capsids. Mol Cell **9:**505-514.
304. **Barry JK, Miller WA.** 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 U S A **99:**11133-11138.
305. **Losick VP, Schlax PE, Emmons RA, Lawson TG.** 2003. Signals in hepatitis A virus P3 region proteins recognized by the ubiquitin-mediated proteolytic system. Virology **309:**306-319.
306. **Gao L, Tu H, Shi ST, Lee KJ, Asanaka M, Hwang SB, Lai MM.** 2003. Interaction with a ubiquitin-like protein enhances the ubiquitination and degradation of hepatitis C virus RNA-dependent RNA polymerase. J Virol **77:**4149-4159.
307. **Jaag HM, Stork J, Nagy PD.** 2007. Host transcription factor Rpb11p affects tombusvirus replication and recombination via regulating the accumulation of viral replication proteins. Virology **368:**388-404.
308. **Molinari P, Marusic C, Lucioli A, Tavazza R, Tavazza M.** 1998. Identification of artichoke mottled crinkle virus (AMCV) proteins required for virus replication: complementation of AMCV p33 and p92 replication-defective mutants. J Gen Virol **79 (Pt 3):**639-647.
309. **Lohmann V, Korner F, Herian U, Bartenschlager R.** 1997. Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. J Virol **71:**8416-8428.

310. **Rajendran KS, Pogany J, Nagy PD.** 2002. Comparison of turnip crinkle virus RNA-dependent RNA polymerase preparations expressed in Escherichia coli or derived from infected plants. J Virol **76:**1707-1717.
311. **Neufeld KL, Richards OC, Ehrenfeld E.** 1991. Purification, characterization, and comparison of poliovirus RNA polymerase from native and recombinant sources. J Biol Chem **266:**24212-24219.
312. **Rubach JK, Wasik BR, Rupp JC, Kuhn RJ, Hardy RW, Smith JL.** 2009. Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. Virology **384:**201-208.
313. **Lai MM.** 1998. Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. Virology **244:**1-12.
314. **Strauss JH, Strauss EG.** 1999. Viral RNA replication. With a little help from the host. Science **283:**802-804.
315. **van Bokhoven H, Mulders M, Wellink J, Vlak JM, Goldbach R, van Kammen A.** 1991. Evidence for dissimilar properties of comoviral and picornaviral RNA polymerases. J Gen Virol **72 (Pt 3):**567-572.
316. **te Velthuis AJ, Arnold JJ, Cameron CE, van den Worm SH, Snijder EJ.** 2010. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. Nucleic Acids Res **38:**203-214.
317. **Ranjith-Kumar CT, Gajewski J, Gutshall L, Maley D, Sarisky RT, Kao CC.** 2001. Terminal nucleotidyl transferase activity of recombinant Flaviviridae RNA-dependent RNA polymerases: implication for viral RNA synthesis. J Virol **75:**8615-8623.
318. **Rohayem J, Jager K, Robel I, Scheffler U, Temme A, Rudolph W.** 2006. Characterization of norovirus 3Dpol RNA-dependent RNA polymerase activity and initiation of RNA synthesis. J Gen Virol **87:**2621-2630.
319. **Morrow CD, Warren B, Lentz MR.** 1987. Expression of enzymatically active poliovirus RNA-dependent RNA polymerase in Escherichia coli. Proc Natl Acad Sci U S A **84:**6050-6054.
320. **Flanegan JB, Baltimore D.** 1977. Poliovirus-specific primer-dependent RNA polymerase able to copy poly(A). Proc Natl Acad Sci U S A **74:**3677-3680.
321. **Ishii K, Tanaka Y, Yap CC, Aizaki H, Matsuura Y, Miyamura T.** 1999. Expression of hepatitis C virus NS5B protein: characterization of its RNA polymerase activity and RNA binding. Hepatology **29:**1227-1235.
322. **Gohara DW, Ha CS, Kumar S, Ghosh B, Arnold JJ, Wisniewski TJ, Cameron CE.** 1999. Production of "authentic" poliovirus RNA-dependent RNA polymerase (3D(pol)) by ubiquitin-protease-mediated cleavage in Escherichia coli. Protein Expr Purif **17:**128-138.
323. **Pogany J, Nagy PD.** 2015. Activation of Tomato Bushy Stunt Virus RNA-Dependent RNA Polymerase by Cellular Heat Shock Protein 70 Is Enhanced by Phospholipids In Vitro. J Virol **89:**5714-5723.
324. **Shatskaya GS, Drutsa VL, Koroleva ON, Osterman IA, Dmitrieva TM.** 2013. Investigation of Activity of Recombinant Mengovirus RNA-Dependent RNA Polymerase and Its Mutants. Biochemistry (Mosc) **78:**96-101.

325. **Richards OC, Eggen R, Goldbach R, van Kammen A.** 1989. High-level synthesis of cowpea mosaic virus RNA polymerase and protease in Escherichia coli. Gene **78:**135-146.
326. **Hanna SL, Pierson TC, Sanchez MD, Ahmed AA, Murtadha MM, Doms RW.** 2005. N-linked glycosylation of west nile virus envelope proteins influences particle assembly and infectivity. J Virol **79:**13262-13274.
327. **Bottcher-Friebertshauser E, Garten W, Matrosovich M, Klenk HD.** 2014. The hemagglutinin: a determinant of pathogenicity. Curr Top Microbiol Immunol **385:**3-34.
328. **Klenk HD, Wagner R, Heuer D, Wolff T.** 2002. Importance of hemagglutinin glycosylation for the biological functions of influenza virus. Virus Res **82:**73-75.
329. **Osman TA, Buck KW.** 1996. Complete replication in vitro of tobacco mosaic virus RNA by a template-dependent, membrane-bound RNA polymerase. J Virol **70:**6227-6234.