Characterization of the X4 protein of *Tomato ringspot virus* and analysis of its variability among virus isolates

by

Bita Jafarpour

M.Sc. Ferdowsi University of Mashhad, College of Agriculture, 2000
B.Sc. Ferdowsi University of Mashhad, College of Agriculture, 1997

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

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies
(Botany)

The University of British Columbia
(Vancouver)
March 2010

©Bita Jafarpour 2010
Abstract

The genome of *Tomato ringspot virus* (nepovirus) encodes a unique X4 protein of unknown function. X4 is absent from the genome of many other nepoviruses and does not have significant sequence homology with other proteins in available databases.

I have studied the possible function(s) of the ToRSV-X4 protein. I used three ToRSV isolates, Rasp2, Rasp1 and PYB. Sequence analysis of the X4 region from these isolates showed a surprising level of sequence variability. The central region of the X4 protein contains two series of amino acid sequence tandem repeats. The number of these repeats varies among ToRSV isolates. *N. benthamiana* plants inoculated with the ToRSV-Rasp1 recover from infection at temperatures equal to or above 27°C. At 21°C, recovery does not occur and plants eventually die. In ToRSV-PYB1 infection, plants recover from infection at all temperature tested. Since RNA silencing is a ubiquitous plant defence response and was observed in nepovirus infected plants, I used agroinfiltration assays to test the effect of the X4 protein on the induction of RNA silencing directed at a green fluorescent protein reporter gene (GFP). Co-expression of X4-Rasp1 or X4-Rasp2 proteins with GFP in *N. benthamiana* plants transiently enhanced the expression of GFP. To detect the ToRSV X4 protein in virus infected plants, I used polyclonal antibodies raised against the C-terminal region of the X4-Rasp2 protein. This antibody detected a 60 kDa protein in ToRSV-Rasp1 infected plants. The cross-reactivity of this X4 (64-65) antibody with the CP suggested that the 60 kDa protein corresponds to the CP. However, X4(3743-44) and X4 (3741-42) antibodies derived from the X4 sequence, did not detect this 60 kDa protein. The predicted 82 kDa X4 full-length protein was not detected by any of the anti-X4 antibodies. This suggests that X4 may be an unstable protein.

In conclusion, the variability of the X4 protein among ToRSV isolates combined with different symptomatology and the preliminary evidence that X4 may act as a suppressor of silencing in plants, suggest that X4 may be a multifunctional protein that is involved in protein–protein interaction, host specificity, symptomatology and/or interaction with host defence responses.
# Table of Contents

Abstract ....................................................................................................................................ii  
Table of Contents ....................................................................................................................iii  
List of Tables ..........................................................................................................................vii  
List of Figures ........................................................................................................................viii  
List of Abbreviations .................................................................................................................x  
Dedications ............................................................................................................................xvi  
Co-Authorship Statement ...................................................................................................xvii  
Chapter 1 ..................................................................................................................................1  
1. Introduction .........................................................................................................................1  
1.1 Virus structure ................................................................................................................2  
1.2 Nepoviruses ...................................................................................................................3  
1.2.1 Nepovirus structure...................................................................................................4  
1.3 Replication cycle of plant viruses within the cell .........................................................8  
1.3.1 Translation of viral proteins......................................................................................9  
1.3.2 Polyprotein processing............................................................................................10  
1.3.3 Assembly of the viral replication complex .............................................................11  
1.3.3.1 Replication and translation of nepoviruses......................................................12  
1.3.3.2 Replication of RNA2 in nepoviruses...............................................................13  
1.4 Movement of plant viruses .........................................................................................14  
1.4.1 Plasmodesmata .......................................................................................................15  
1.4.2 Types of movement from cell to cell by plant viruses............................................16  
1.4.2.1 Movement of virus-like particles.....................................................................16  
1.4.2.1.1 Cell-to-cell movement of comoviruses, nepoviruses and related viruses. 16  
1.5 Virus transmission .........................................................................................................17  
1.5.1 Nematode transmission...........................................................................................18  
1.5.1.1 Tobraviruses ....................................................................................................18  
1.5.1.2 Nepoviruses .....................................................................................................18  
1.6 ToRSV isolates ..............................................................................................................19  
1.6.1 Virus evolution, diversity and sources of variation ...............................................20  
1.6.1.1 Mutation, insertion and deletion ....................................................................20  
1.6.1.2 Recombination.................................................................................................20
3.1.1 Research objective and hypothesis ........................................................................71
3.2 Materials and methods ..................................................................................................72
  3.2.1 Production of antibodies against the C-terminal and the N-terminal region of X4 72
  3.2.2 Construction of plasmids used in the expression of full-length or truncated forms of X4, CP and MP in vitro translations or in E. coli .................................................................73
  3.2.3 Preparation of plant extracts and subcellular fractionations ..................................76
  3.2.4 Purification of ToRSV particles .............................................................................76
  3.2.5 Immunoblotting ..................................................................................................76
  3.2.6 Immuno-electron microscopy .............................................................................77
  3.2.7 Trypsin digestion of ToRSV particles ..................................................................77
  3.2.8 Liquid chromatography-mass spectrometry with peptide mass fingerprinting (LC-MS/MS) ......................................................................................................................78
3.3 Results .......................................................................................................................78
  3.3.1 Preparation of antibodies against the X4 protein .................................................78
  3.3.2 Detection of a 60 kDa protein by the X4(64-65) antibodies in ToRSV-infected plant extracts .................................................................................................................................81
  3.3.3 Co-fractionation of the 60 kDa protein detected by the X4(64-65) antibodies with empty and full virus particles in sucrose gradient fractionation assay ..................83
  3.3.4 Co-purification of the 60 kDa protein detected by the X4(64-65) antibodies with the virus particles ..................................................................................................................85
  3.3.5 Immunogold labelling of purified virus particles with X4(64-65) antibodies ......87
  3.3.6 Susceptibility of the 60 kDa protein recognized by the X4(64-65) antibodies to trypsin digestion ......................................................................................................................88
  3.3.7 The X4 protein is not detected in association with virions using liquid chromatography-mass spectrometry with peptide mass fingerprinting (LC-MS/MS) ....89
  3.3.8 Re-examination of the specificity of the X4(64-65) antibodies and the α-virion antibodies ......................................................................................................................89
  3.3.9 In vitro co-immunoprecipitation of X4 and CP proteins with the X4(64-65) antibodies ......................................................................................................................92
    3.3.9.1 The MP did not co-immunoprecipitate with the CP in vitro ..........................93
  3.3.10. Production of antibodies from different regions of the X4 and the CP ...........94
3.4 Discussion .................................................................................................................100
  3.4.1. Summary and conclusion ...............................................................................106
3.5 Bibliography .............................................................................................................107
Chapter 4 ..............................................................................................................................111

Preliminary evidence that the X4 protein may act as a silencing suppressor in plants 111

4.1. Introduction................................................................................................................112
4.1.1. Research hypothesis and objectives.................................................................113

4.2 Materials and methods ............................................................................................114
4.2.1 PCR amplification of a region of the RNA2 open reading frame coding for a
portion of X3, the entire X4 and a portion of MP for the Chickadee isolates ..........114
4.2.2 Symptom evaluation in plants infected with Rasp1 and PYB ToRSV isolates....114
4.2.3. Plasmid construction.......................................................................................115
4.2.4 Agroinfiltration of N. benthamiana plants and immunoblotting ......................115

4.3 Results.......................................................................................................................116
4.3.1 RT-PCR amplification of ToRSV Chickadee isolate and symptom development in
ToRSV–infected Rasp1 and PYB ToRSV isolates ..................................................116
4.3.2 Preliminary evidence that the X4 protein may act as a silencing suppressor in
plants............................................................................................................................118

4.4 Discussion.................................................................................................................120
4.4.1 Summary and conclusion..................................................................................122

4.5 Bibliography ..........................................................................................................124

Chapter 5 ............................................................................................................................127

General discussion and future projects .......................................................................127

5.1 General discussion.................................................................................................128
5.2 Future projects ......................................................................................................131
5.2.1 Potential impact of this research on the future on molecular virology research ..134
5.3 Bibliography .........................................................................................................135
List of Tables

Table 1.1 PEST motifs as potential proteolytic cleavage sites in Rasp2 isolate ....................35
Table 1.2 PEST motifs in the X4 protein of PYB isolate ......................................................36
Table 1.3 PEST motifs in the X4 protein of Rasp1 isolate ......................................................37
Table 1.4 Detail description of the topology prediction of ToRSV isolates .........................41
Table 2.1 Percentage of identity in the deduced amino acid sequence of various protein
    domains among ToRSV isolates ......................................................................................62
List of Figures

Fig. 1.1  A simplified representation of the structure of TRSV. ............................................... 6
Fig. 1.2 Pseudo-T = 3 capsid of Picornavirales. ........................................................................ 8
Fig. 1.3 Schematic representation of the replication cycle of nepoviruses within the cell ....... 9
Fig. 1.4 Genome organisation of ToRSV .................................................................................. 11
Fig. 1.5 Schematic representation of a plasmodesmata transversal section ......................... 16
Fig. 1.6 Functional regions of the helper component proteinase (HC-Pro)............................ 25
Fig. 1.7 The silencing pathway .............................................................................................. 29
Fig. 1.8 Mechanism of the silencing suppression activity of p19. ........................................ 31
Fig. 1.9 Comparison of the genomic structure of tomato ringspot virus and grapevine fanleaf
   virus (nepoviruses) ........................................................................................................... 34
Fig. 1.10 Predicted phosphorylation sites .............................................................................. 39
Fig. 1.11 Topology prediction of ToRSV isolate. ..................................................................... 41
Fig. 1.12 Secondary structure prediction in different ToRSV isolates ..................................... 42
Fig. 2.1 Comparison of the X4 protein domain in the ToRSV Rasp1 and Rasp2 isolates ....... 63
Fig. 3.1 Genome organisation of ToRSV and GFLV and production of the X4(64-65)
   antibodies ......................................................................................................................... 79
Fig. 3.2 Specificity of the X4(64-65) and α-virion antibodies tested by immunoprecipitation
   assays .................................................................................................................................. 81
Fig. 3.3 Detection of a 60 kDa protein by the X4(64-65) antibodies in infected plant extracts.
   .............................................................................................................................................. 82
Fig. 3.4 Immunoblot analysis and electron microscopy of subcellular fractionation of proteins
   by sucrose gradient ............................................................................................................. 84
Fig. 3.5 Co-purification of the 60 kDa protein detected by the X4(64-65) antibodies with
   purified virus particles ........................................................................................................ 86
Fig. 3.6 Immunogold labelling of purified virus particles using the X4(64-65) antibodies ... 88
Fig. 3.7 Susceptibility of the 60 kDa protein to trypsin digestions ......................................... 89
Fig. 3.8 Immunoblot analyses of CP, CP-150, CP-360, X4(64-65), X4-TM (Rasp2) proteins.
   ........................................................................................................................................... 91
Fig. 3.9 Immunoprecipitation of the CP and the X4 protein .................................................... 93
Fig. 3.10 In vitro protein-protein interaction assays ............................................................... 94
Fig. 3.11 X4 protein: Alignment of 3 different isolates of ToRSV (Rasp1, Rasp2 and PYB).
   ........................................................................................................................................... 95
Fig. 3.12 CP protein: Alignment of 3 different isolates of ToRSV (Rasp1, Rasp2 and PYB).
.................................................................................................................................................................96

Fig. 3.13 Immunoblot analysis of the infected (I) and healthy (H) N. benthamiana extracts
using the anti-virion, anti-CP (3739-40) and anti-CP (3737-38) antibodies. .........................98

Fig. 3.14 Immunoblot analysis of the infected (Inf) and healthy (H) N. benthamiana extracts
using the X4-N, X4(3743-44) and X4 (3741-42) antibodies.........................................................99

Fig. 3.15. Sequence alignment between the X4 protein and the CP of ToRSV .....................104

Fig. 4.1   PCR amplification of a region of the RNA2 open reading frame coding for a portion
of X3, the entire X4 and a portion of MP in Rasp1, PYB and Chickadee isolates. ......117

Fig. 4.2 Symptom development in ToRSV–infected Rasp1 and PYB ToRSV isolates........117

Fig. 4.3 Expression levels of GFP in a transient expression assay conducted in the presence
or absence of X4. .................................................................................................................................119

Fig. 4.4 Analysing the expression level of the X4 protein in infiltrated patches by
immunoblotting. .................................................................................................................................120
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-MP</td>
<td>Anti movement protein antibodies</td>
</tr>
<tr>
<td>α-virion</td>
<td>Anti –virion antibodies</td>
</tr>
<tr>
<td>X4-C</td>
<td>Recombinant protein</td>
</tr>
<tr>
<td>a.a</td>
<td>Amino acid</td>
</tr>
<tr>
<td>Abs</td>
<td>Antibodies</td>
</tr>
<tr>
<td>AFM</td>
<td>Atomic force microscopy</td>
</tr>
<tr>
<td>ALSV</td>
<td><em>Apple latent spherical virus</em></td>
</tr>
<tr>
<td>ArMV</td>
<td><em>Arabis mosaic virus</em></td>
</tr>
<tr>
<td>B</td>
<td>Bottom component</td>
</tr>
<tr>
<td>BLCV</td>
<td><em>Beet leaf curl virus</em></td>
</tr>
<tr>
<td>BLMV</td>
<td><em>Blueberry leaf mottle virus</em></td>
</tr>
<tr>
<td>BRAV</td>
<td><em>Blackcurrant reversion associated virus</em></td>
</tr>
<tr>
<td>BRV</td>
<td><em>Blackcurrant reversion virus</em></td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>BYV</td>
<td><em>Beet yellow virus</em></td>
</tr>
<tr>
<td>C. quinoa</td>
<td>Chenopodium quinoa</td>
</tr>
<tr>
<td>C. sativus</td>
<td><em>Cucumis sativus</em></td>
</tr>
<tr>
<td>CaMV</td>
<td><em>Cauliflower mosaic virus</em></td>
</tr>
<tr>
<td>CI</td>
<td>Cylindrical inclusion</td>
</tr>
<tr>
<td>CMV</td>
<td><em>Cucumber mosaic virus</em></td>
</tr>
<tr>
<td>CP</td>
<td>Coat protein</td>
</tr>
<tr>
<td>CPm</td>
<td>Minor capsid protein</td>
</tr>
<tr>
<td>CPMV</td>
<td>Cow pea mosaic virus</td>
</tr>
<tr>
<td>CPS</td>
<td>Small coat protein</td>
</tr>
<tr>
<td>CsCL</td>
<td>Cesium chloride</td>
</tr>
<tr>
<td>CVYV</td>
<td><em>Cucumber vein yellowing virus</em></td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dpi</td>
<td>Days post infiltration</td>
</tr>
<tr>
<td>dpi</td>
<td>Days post inoculation</td>
</tr>
<tr>
<td>dsRNA</td>
<td>Double stranded RNA</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>EM</td>
<td>Electron microscopy</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>Fig</td>
<td>Figure</td>
</tr>
<tr>
<td>GFLV</td>
<td>Grapevine fanleaf virus</td>
</tr>
<tr>
<td>HA</td>
<td>Poly-histidine tail</td>
</tr>
<tr>
<td>HC-Pro</td>
<td>Helper component proteinase</td>
</tr>
<tr>
<td>HCRSV</td>
<td><em>Hibiscus chlorotic ringspot virus</em></td>
</tr>
<tr>
<td>Hsp70</td>
<td>Heat-Shock cognate 70-kDa Protein</td>
</tr>
<tr>
<td>Hsp70h</td>
<td>Hsp70 homologue</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal ribosome binding site</td>
</tr>
<tr>
<td>L</td>
<td>Large</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>Liquid chromatography-mass spectrometry with Peptide mass fingerprinting</td>
</tr>
<tr>
<td>LIYV</td>
<td><em>Lettuce infection yellow virus</em></td>
</tr>
<tr>
<td>LMV</td>
<td><em>Lettuce mosaic virus</em></td>
</tr>
<tr>
<td>M</td>
<td>Morphological subunits</td>
</tr>
<tr>
<td>m7GTP</td>
<td>Methyl-guanidine triphosphate</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>MP</td>
<td>Movement protein</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MSDB</td>
<td>Protein sequence database designed for massspectrometry applications</td>
</tr>
<tr>
<td>N. benthamiana</td>
<td><em>Nicotiana benthamiana</em></td>
</tr>
<tr>
<td>NIa</td>
<td>Nuclear inclusion a</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometre</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotide</td>
</tr>
<tr>
<td>NTB</td>
<td>Nucleotide triphosphate binding protein</td>
</tr>
<tr>
<td>P1</td>
<td>Polyprotein encoded by ToRSV RNA1</td>
</tr>
<tr>
<td>P1</td>
<td>Serine protease in the family potyviridae</td>
</tr>
<tr>
<td>p19</td>
<td>19 kDa protein of <em>Tomato bushy stunt</em> virus</td>
</tr>
<tr>
<td>PAZ</td>
<td>PIWI Argonaut and Zwille</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PEBV</td>
<td><em>Pea early browning virus</em></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PIPO</td>
<td>Pretty interesting potyviridae ORF</td>
</tr>
<tr>
<td>PIWI</td>
<td>P-element induced wimpy testis in Drosophila</td>
</tr>
<tr>
<td>Pol</td>
<td>Polymérase</td>
</tr>
<tr>
<td>PPV</td>
<td>Plum pox virus</td>
</tr>
<tr>
<td>Pro</td>
<td>Proteinase</td>
</tr>
<tr>
<td>PVIP</td>
<td>Potyvirus VPg-interacting protein</td>
</tr>
<tr>
<td>PVX</td>
<td>Potato virus X</td>
</tr>
<tr>
<td>PYB</td>
<td>Peach yellow bud isolate</td>
</tr>
<tr>
<td>RASP</td>
<td>Raspberry isolate</td>
</tr>
<tr>
<td>RdRp</td>
<td>RNA dependent RNA polymerase</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-induced silencing complex</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription-polymerase chain reaction</td>
</tr>
<tr>
<td>S</td>
<td>Small</td>
</tr>
<tr>
<td>S</td>
<td>Sedimentation coefficient</td>
</tr>
<tr>
<td>S3</td>
<td>Subcellular fractionation</td>
</tr>
<tr>
<td>S30</td>
<td>Soluble protein-enriched fraction</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SEL</td>
<td>Size exclusion limit</td>
</tr>
<tr>
<td>siRNA</td>
<td>Short interfering RNA</td>
</tr>
<tr>
<td>SLRSV</td>
<td>Strawberry latent ringspot virus</td>
</tr>
<tr>
<td>ssRNA</td>
<td>Single stranded RNA</td>
</tr>
<tr>
<td>T</td>
<td>Top component</td>
</tr>
<tr>
<td>TBRV</td>
<td>Tomato blackring virus</td>
</tr>
<tr>
<td>TBSV</td>
<td>Tomato bushy stunt virus</td>
</tr>
<tr>
<td>TCV</td>
<td>Turnip crinkle virus</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission electron microscope</td>
</tr>
<tr>
<td>TGB</td>
<td>Tripple gene block</td>
</tr>
<tr>
<td>TM</td>
<td>Tranmembrane domain</td>
</tr>
<tr>
<td>TRSV</td>
<td>Tobacco ringspot virus</td>
</tr>
<tr>
<td>TRV</td>
<td>Tobacco rattle virus</td>
</tr>
<tr>
<td>TSWV</td>
<td>Tomato spotted wilt virus</td>
</tr>
<tr>
<td>TVMV</td>
<td>Tobacco vein mottling virus</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>V</td>
<td>Voltage</td>
</tr>
<tr>
<td>vol/vol</td>
<td>Volume per volume</td>
</tr>
<tr>
<td>VAP</td>
<td>Virion associated protein</td>
</tr>
<tr>
<td>VPG</td>
<td>Genome linked viral protein</td>
</tr>
<tr>
<td>VPG-Pro-Pol</td>
<td>ToRSV precursor protein of VPG,Pro,Pol</td>
</tr>
<tr>
<td>wt:vol</td>
<td>Weight per volume</td>
</tr>
<tr>
<td>X2-NTB-VPG</td>
<td>ToRSV precursor protein of X2, NTB and VPG</td>
</tr>
<tr>
<td>xg</td>
<td>Unit of rotational speed based on gravitational force</td>
</tr>
<tr>
<td>μg</td>
<td>Microgram</td>
</tr>
</tbody>
</table>
I would like to thank my supervisor Dr. Helene Sanfaçon whose expertise, understanding and enthusiasm added considerably to my graduate experience. She spent a lot of time explaining the research projects from the very early stage of this research as well as putting a great effort in helping me with the proposal writing, preparing presentations, preparing for the comprehensive exam and also writing the manuscript. I always enjoyed our scientific discussions throughout my PhD. She gave me the chance to attend several international conferences, which had a great impact on my knowledge and gave me extraordinary experiences through out the work. I would like to thank my co-supervisor Dr. Carl Douglas and my committee members Dr. Francois Jean and Dr. Steven Lund for their encouragement, support and their valuable suggestions and scientific discussions, advices and critical review of this thesis. I am very grateful to have them as my Ph.D. committee members. I would like to thank Dr. Janet K. Chantler for the valuable suggestions and critical review of this thesis. Also I would like to thank the staff members of the department of Botany especially Mrs. Veronica Oxtoby (Graduate Secretary) for her assistance.

I would like to thank our lab members Mrs. Joan Chisholm and Mrs. Melanie Walker for their assistance, friendship and technical help and also for obtaining the required lab supplies. I would like to also thank Dr. Juan Jovel for his support and encouragement during the research.

I wish to thank Dr. D’Ann Rochon, for scientific discussions and help during the research. I also like to thank Mr. Michel Weis for his help with the electron microscopy. I would like to thank Dr. Giuseppe (Joe) Mazza and Dr. Farah Hossenian for their encouragement and support. I was lucky to do the research part of my Ph.D. in Summerland and learn more about the Canadian culture in the beautiful Okanagan Valley. This enriched my life experience. I would like to thank the friendly people of Agriculture and Agri- Food Canada for their kindness and help and also for the generous permission to use their facilities.

I would like to thank Mrs. Bell Kendrick, Mr. and Mrs. Peter and Zhila Schofield and Ms. Julie Boulé for their friendship and kindness during my stay in Penticton. Also I would like to thank all my good friends at UBC Vancouver for their friendship and support.

Ultimately, I would like to express my deepest gratitude to my dear parents, my father Dr. Behrooz Jafarpour who opened my eyes to the wonderful world of virology research and my mother Mrs. Behnaz Jafarpour for her amazing support. I thank my brother Dr. Behnam
Jafarpour for the good times we had in Penticton. I am very grateful for their encouragement and support.
Dedications

I dedicate this thesis to my parents
Co-Authorship Statement

Below is a statement of my contribution to each chapter of this thesis. I wrote each chapter and they were subsequently edited by Dr. H. Sanfaçon. I entered corrections and was responsible for the final version of each chapter. My overall contribution for each chapter is estimated as follows: Chapter 1 (100%), Chapter 2 (75%), Chapter 3 (90%), Chapter 4 (25%) and Chapter 5 (100%).

In Chapter 2, I extracted the RNA and made cDNA for different ToRSV isolates. Also I was responsible for the initial identification of protein repeats, the clones design and production, the primer design and the entire sequencing of the Rasp1 isolate. I also sequenced a portion of the PYB isolate. Dr. H. Sanfaçon sequenced approximately half of the PYB isolate. I was actively involved in the preparation of a research paper describing results presented in this chapter under the guidance of Dr. H. Sanfaçon.

[Jafarpour and Sanfaçon (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). Archives of Virology, in press (DOI: 10.1007/s00705-009-0497-3).]

In Chapter 3, Mrs. Joan Chisholm produced the X4(64-65) antibodies and conducted preliminary sucrose gradient fractionation (shown in Fig. 3.4A). I repeated the sucrose gradient fractionation several times. Fig. 3.8A, Fig.3.13 and Fig. 3.14 are also a courtesy of Mrs. Joan Chisholm. I conducted the rest of the experiments.

In Chapter 4, I extracted the RNA and made cDNA to produce the pCITE-X4-HA plasmids containing full-length or truncated version (deletion mutants) of X4-Rasp1 or X4-Rasp2. The transfer of the X4 fragments contained in these plasmids into the pBIN-X4 vector and the symptomatology analysis was done by Dr. H. Sanfaçon (Fig. 4.2). She also conducted preliminary agroinfiltration and immunoblotting experiments with GFP and HA Abs (Fig. 4.3A). I repeated the agroinfiltration experiments and immunoblotting assays using the GFP (Fig 4.3B) and the HA antibodies. Mrs. Joan Chisholm also repeated immunoblotting with the HA Abs and contributed the Fig. 4.4.
Chapter 1

Literature review
1. Introduction

A virus is a sub-microscopic infectious agent that is unable to grow or reproduce itself outside of a host cell. Viruses depend on cells for every step of their replication cycle. Replication begins with the release of the viral genome within the host cell. Once it is released, the viral genome is used as a template for synthesising viral proteins by the translational machinery of the host. The viral genome is replicated by viral enzymes and the host factors. The newly replicated genome is then encapsidated by the viral coat protein to form the progeny virus particles. The virus can then be transmitted from one cell to another or from one host to another.

In this review, I will begin with a brief overview of virus replication cycle. Viruses have limited genetic information, their protein are often multifunctional and play important roles in key steps of the replication cycle. In this thesis, I have characterized the X4 protein of *Tomato ringspot virus* (ToRSV), a virus belonging to the genus *Nepovirus* in the order Picornavirales (Le Gall *et al.*, 2008). The function of the X4 protein is unknown. I will focus this section of the literature review on the replication cycle of nepoviruses and the function of various viral proteins in these steps. When appropriate, I will also give some examples using well-characterized proteins from other plant viruses. Finally, I will provide some computer prediction based on the protein sequence of the X4 protein of tomato ringspot virus (ToRSV) at the end of this section.

1.1 Virus structure

Viruses come in many shapes and sizes. Their genome consists of one or several molecules of nucleic acid. The genetic material of the virus is surrounded by a capsid shell made up of virus-encoded proteins. The nucleic acid can be deoxyribonucleic acid or ribonucleic acid and can be single or double stranded, linear or circular. The term capsid has been proposed for the closed shell or tube of viruses. The mature virus has been termed the virion (infective virus particle). The coat protein (CP) has an early function in disassembly of parental virus and a late function in assembly of progeny virus. The CP may play a role in many other steps of the infection cycle between the early and the late function such as viral movement in the host and transmission of the virus from host to host (Caspar and Klug, 1962; Knipe *et al.*, 2001).
The capsid provides a protein shell in which the chemically labile viral genome can be maintained in a stable environment. The primary structure of the viral coat proteins and nucleic acids depends on covalent bonds. However, in the final structure of the simple geometric virus, these two major components are held together by non-covalent bonds. It has been suggested that three kinds of interaction can be involved in the assembly or stability of the virions: protein-protein, protein-RNA and RNA-RNA interactions. In addition, small molecules such as divalent metal ions (e.g. Ca\(^{2+}\)) influence or enhance the stability of some virus particles. Caspar and Klug (1962) proposed the theory of quasi-equivalence, which means that not all chemical subunits in the shell need to be arranged in an exact equivalent way (Hull, 2002).

In icosahedral viruses such as nepoviruses, the basic icosahedron has 20 faces with three subunits in identical positions on each face, giving 60 structural subunits in an icosahedron (Hull, 2002).

1.2 Nepoviruses

Nepoviruses are classified in the order Picornavirales and together with other plant viruses of the order have recently been reassigned to the family Secoviridae, sub-family Comovirinae (Le Gall et al., 2007; Sanfacon et al., 2009). All picornavirales have a single strand positive sense RNA genome that may be monopartite or bipartite. They have small icosahedral particles (25-30 nm) with a pseudo T=3 symmetry. The CP is made up of jelly-rolls that can be present in one large CP, or divided among two or three smaller CPs. Each genome encodes a large polyprotein, which is cleaved by the viral protease. The genome contains a replication block that includes a helicase, a 3C like protease and a RNA dependent RNA polymerase. Picornavirales includes viruses that infect vertebrates, arthropods, higher plants, fungi and algae (Le Gall et al., 2007; Sanfacon et al., 2009).

There are currently 32 different species of nepoviruses which makes the genus Nepovirus the largest genus of plant picorna-like viruses (Rochon and Sanfacon, 2001). By definition, nepoviruses are a group of viruses with an icosahedral structure that are transmitted from plant to plant by soil nematodes (Longidoridae) in a semi-persistent manner. However, there are nepoviruses with other types of vectors. For example, blackcurrant reversion virus
is transmitted by mites. Several nepoviruses are transmitted by pollen and/or by seed (Le Gall et al., 2007; Sanfacon et al., 2009).

Tomato ringspot virus (ToRSV, nepovirus) is a major pathogen of small fruit crops and fruit trees in North America. There are three nepovirus subgroups termed A, B and C. Subgroup C nepoviruses have an additional protein domain (the X4 protein) on their RNA2. ToRSV which is the focus of this study belongs to the subgroup C of nepoviruses. ToRSV can be transmitted by the nematode Xiphinema americanum (Dorylamidae). It can also be transmitted by mechanical inoculation, grafting, by seeds (demonstrated in Rubus idaeus, Nicotiana tabacum, Glycine max and Fragaria x ananassa) and also by pollen to seed (Brunt et al., 1996).

1.2.1 Nepovirus structure

As I mentioned above, nepoviruses have isometric particles, which contain 60 molecules of a single coat protein (CP) with a molecular mass of 53-60 kDa (Sanfacon, 2008) and are members of the order Picornavirales (previously referred to as picorna-like viruses or members of the picornavirus-like superfamily or supergroup) that share many common properties. In all members of the order Picornavirales, the capsid is composed of three jelly roll domains (Le Gall et al., 2008). Within the plant picornavirales branch, several lineages are formed based on their hierarchical clustering of the Pro–Pol amino acid sequence. These lineages correspond to the different genera and have been regrouped in the family Secoviridae. The family Secoviridae includes the genera Comovirus, Fabavirus, Nepovirus, Sequivirus, Waikavirus, Cheravirus, Sadwavirus. The genus Torradovirus with the type species Tomato torrado virus is also included in Secoviridae family. Nepoviruses, fabaviruses and comoviruses are closely related to each other and are grouped together in the subfamily Comovirinae within the family Secoviridae.

Nepoviruses have a bipartite single-strand RNA genome. The two genome segments are encapsulated separately into two different icosahedral particles. In nepoviruses, the two RNA molecules of ToRSV are polyadenylated at the 3’ end and are covalently linked to a small viral protein (VPg) at their 5’ end (Sanfacon, 1995). Each RNA codes for a large polyprotein. The polyprotein is cleaved by the viral protease (Pro) to mature and intermediate proteins. In addition to the coding region, each RNA has a long untranslated
region at its 5' and 3' ends. In subgroup A nepoviruses, the 5' and 3' untranslated region of RNA1 and RNA2 share 70-79% sequence identity. In ToRSV, the 5' untranslated region of RNA1 and RNA2 shares 100% identity and this region of sequence identity extends into the coding region. The 3' non-coding region of RNA1 and RNA2 is almost identical. The sequence identity of the 5' and 3' ends of ToRSV RNA might be the cause of a recombination event during the replication of the virus (Rott et al., 1991a). The genus Nepovirus is divided into three sub-groups: A, B and C based on the length of RNA2, serological properties and the similarity of their genome sequence (Wang et al., 2004). Members of the genus Nepovirus vary in the number of processing sites, polyproteins and in the specificity of their proteinase.

In animal picornaviruses and in the plant cheraviruses and torradoviruses, the capsid protein precursor is cleaved at two sites to yield three subunits. Each subunits fold into a single β-barrel domain. In comoviruses, fabaviruses and sadwaviruses, the CP precursor is cleaved at one site to give rise to two subunits of the CP, one which contain one β-barrel domain (CPS, small CP) and the other which is composed of two β-barrels (CPL, large CP). In nepoviruses the capsid protein is a single large protein that contains three covalently linked β-barrels (Chandrasekar and Johnson, 1998). The three β-barrels are named C, B and A domain from the N-terminus to the C- terminus. The B and C domains lay side by side around the three fold axis of symmetry. The A domain lay around the five-fold axes (Fig.1.1 and Fig.1.2) (Chandrasekar and Johnson, 1998).

In nepoviruses, cesium chloride (CsCL) equilibrium centrifugation of purified virus particles typically reveals the presence of three types of viral particles. T-particles are empty virus particles without an RNA component, and sediment at 50S. B-particles contain a single molecule of RNA1 and sediment at 115-134S. M-particles sediment at 86-128S and contain a single molecule of RNA2. In ToRSV, particles separate in two peaks on the gradients. The top component (between 50-55S) consists of empty virus particles. The bottom component (between 115-130S) is composed of two (B1 and B2) nucleoprotein components, which contain either RNA1 or RNA2. Because RNA1 and RNA2 are very similar in size in ToRSV, the two components are difficult to separate (Allen and Dias, 1977).
The structure of tobacco ringspot virus (TRSV, nepovirus of subgroup A) and blackcurrant reversion virus (BRV, nepovirus of subgroup C) has been resolved (Chandrasekar and Johnson, 1998; Seitsonen et al., 2008). Preparations of nepovirus particles purified from indicator plants (C. quinoa or N. benthamiana) are composed of two forms of the unique coat protein with identical N-termini (Lemmetty et al., 1997). The slightly smaller coat protein bands can be separated from the full-length CP on sodium dodecyl sulfate polyacrylamide gel (SDS–PAGE). In the case of BRV particles the two coat protein forms are 54 and 55 kDa in size (Latvala et al., 1998; Lemmetty et al., 1997) with identical N-termini.

The atomic model of TRSV (Chandrasekar and Johnson, 1998) was fitted into the BRV reconstruction and the difference map was calculated. The BRV homology model fitted the cryo EM reconstruction of BRV well. One major difference of BRV and TRSV is a C-terminal extension of 19 amino acids which is not present in TRSV. Based on the homology model of BRV, it is predicted that the C-terminal 14 amino acid residues of BRV projects out of the surface of the virus particle. The model also predicts that the N-terminal domain of the BRV capsid protein extends into the capsid interior to interact with the RNA (Seitsonen et al., 2008). The homology modelling of the BRV capsid was also used to identify potential sequences that could be used for mite interaction. One of the most obvious regions is predicted to be the C-terminal 19 residues extension. The C-terminal 19 residues are some of the least conserved in the sequence alignment among other nepoviruses. Because the C-terminus is extended from the virion surface, thus it is also a suitable region for antibody generation. It has previously been shown that virus preparations containing the shorter form
of the capsid protein are infectious by mechanical inoculation, resulting in symptoms identical to BRV symptoms (Lemmetty et al., 1997). The resulting progeny viruses contain both protein forms (Latvala et al., 1998). The shorter form of CP is due to truncation of the C-terminal extension. Thus, the C-terminus is not important for the infectivity of the virus but may indeed serve as a determinant for mite transmission (Seitsonen et al., 2008). Also, two forms of coat protein (59 and 57 kDa) were detected in tomato black ring virus (TBRV, Nepovirus). The C-terminal extension of the larger coat protein is lost in vivo late in TBRV infection and during virus purification. Proteins were extracted from infected plant extracts (N. clevelandii) at various time lines and were immunoblotted with antiserum raised against purified virus particles. In samples extracted from leaves 3 or 5 days post inoculation, the 59 kDa CP was predominant, but samples taken at later time contained the 57 kDa and 59 kDa CP in approximately equal amounts (Demangeat et al., 1992). In vitro translation of TBRV RNA yields only the 59 kDa CP. Partially purified virus contained both the 57 kDa and 59 kDa CP, while highly purified virus contained only the 57 kDa protein. The 59 and 57 kDa protein shared the same N-terminus suggesting that similarly to BRV, the 57 kDa protein arise by the loss of the C-terminal amino acids (Demangeat et al., 1992). It is known that proteases can remove amino acids from the C-termiini of the coat proteins of tobacco mosaic virus (Harris and Knight, 1952), potato virus X (Koenig et al., 1978) and potyviruses (Shukla et al., 1988). Presumably, as with these viruses, the C-terminal amino acids of the TBRV coat protein protrude from the virus particle surface and can be removed without disrupting the virion. This exposed detachable fragment may play a significant role in TBRV biology as does the protruding N-terminal fragment of potyvirus coat proteins in their transmission by aphids (Atreya et al., 1990). The C-terminal extensions have been reported for comoviruses, strawberry latent ringspot virus (SLRSV), and tomato black ring virus (TBRV-S) (Le Gall et al., 1995). Based on the multiple alignment of nepovirus coat protein, a C-terminal extension of 54 amino acids was also predicted for ToRSV. There is no amino acid sequence similarity between these C-terminal extensions (Latvala et al., 1998).
Fig. 1.2 Pseudo-T = 3 capsid of Picornavirales. The beta-barrel or jelly-roll structure is shown in the left. The three jelly-rolls are assembled into individual capsomers as shown in the middle. Sixty of these protomers are assembled to form the icosahedral capsid (right of the figure). The three jelly-roll are separated in three CP subunits in most taxa, including Enterovirus, Rhinovirus, Cardiovirus, Aphthovirus, Cripavirus (VP1, VP2 and VP3, corresponding to the CP1, CP2 and CP3 domains, respectively, as shown at the top of the figure). The genera Comovirus, Fabavirus and Sadwavirus (S and L) encode two subunits (one large, L, and one small, S) that contain two and one jelly rolls, respectively (shown in the middle). A single large CP subunit contains all three jelly rolls in the genus Nepovirus (CP, shown at the bottom of the figure) (Picture courtesy of Olivier Le Gall, reprinted with permission from Le Gall et al., 2008).

1.3 Replication cycle of plant viruses within the cell

Since the function of the X4 protein of ToRSV is unknown, I will briefly explain the replication cycle of plant viruses and the function of the viral proteins in different steps of this cycle with a focus on ToRSV and nepoviruses. All viruses share the same basic replication cycle which involves the following steps: (1) virus entry into the cell, (2) viral gene expression and genome replication and (3) capsid formation and virion assembly (Fig. 1.3).
Fig. 1. Schematic representation of the replication cycle of nepoviruses within the cell.

The virus (depicted with the yellow hexagons) is released by the nematode vector into the plant cell, the viral RNA (shown by the thick black line) is uncoated and the (+) strand RNA genome is translated to polyprotein using the host translational machinery (ribosomes shown in red). The polyprotein (blue rectangle) is post and co-translationally cleaved by viral encoded proteases (scissors). The next step is assembly of the viral replication complex which is associated with the ER membrane (shown in purple) followed by viral replication. The final stage is capsid formation and virion assembly, the progeny viral RNAs are packaged and translocated from cell to cell.

1.3.1 Translation of viral proteins

Most plant viruses, including nepoviruses, have a positive-sense single-stranded [(+)-strand] RNA genome. The first steps of the replication cycle following cell entry and uncoating of the virus particle are translation and replication of the viral genome. A single strand RNA genome with a positive polarity means that the genomic RNA can act as a messenger RNA (mRNA) immediately as it enters the cell.

RNA viruses code for their own replication protein, the RNA–dependent RNA polymerase (RdRp). Viruses do not encode translation factors or ribosomes. Therefore they must use the host translation machinery to translate their protein. The highjacking of translation factors by viruses often results in host translation shutdown (Sarnow, 2003). Cellular mRNAs contain a cap structure or methyl-guanidine triphosphate \((\text{m}^7\text{GTP})\) at their 5’ end and a poly A tail at
their 3' end. Both elements are essential to recruit host translation factors and form the translation complex (Sonenberg and Dever, 2003). In nepoviruses and other picorna-like viruses, a small protein, termed genome-linked viral protein (VPg), is covalently linked to the 5' end of the RNA, therefore replacing the cap structure. The 5' untranslated region (UTR) in picornaviruses, potyviruses and possibly comoviruses contain internal ribosome entry site (IRES) which are composed of stem loop structures and are necessary for recruiting the host translation factors and allow the ribosomes to initiate translation effectively on these regions in the absence of the cap structure (Gallie, 2001; Hellen and Sarnow, 2001; Verver et al., 1991). In Blackcurrant reversion virus (nepovirus), sequences located at the 5' and 3' untranslated region of both genomic RNA facilitate the translation of these RNAs using internal ribosome entry site (Karetnikov and Lehto, 2008).

1.3.2 Polyprotein processing

This approach allows the synthesis of multiple protein products from a single RNA. The large polyprotein is often not detected in infected cells since it is processed as soon as the protease coding sequence has been translated (Knipe et al., 2001). In picornaviruses, a single polyprotein is first cleaved at three sites to produce P1, P2 and P3 intermediate polyprotein. P1 contain structural proteins, while P2 and P3 contain RNA replication proteins. The polyprotein is cleaved into smaller functional proteins by viral proteases (Knipe et al., 2001).

Nepoviruses encode a single proteinase which is related to the 3CPro of picornaviruses. The catalytic triad contains histidine, aspartic acid and cysteine (Sanfacon, 2008).

The proteinase also has a substrate-binding pocket that determines its cleavage site specificity. The subgroup C nepovirus proteinase recognizes a glutamine, asparagine or aspartate at -1 position of its cleavage site. In contrast to subgroup C, the proteinase for the subgroup A and B nepovirus recognise different cleavage sites which have lysine, cysteine, arginine or glycine at the -1 position. The RNA1 encoded polyprotein is cleaved intramolecularly (cis-cleavage) and the RNA2 encoded polyprotein is cleaved in trans by the viral protease. ToRSV P1 polyprotein contains the domain for the replication proteins which are RNA–dependent RNA polymerase, the proteinase, genome linked viral protein (VPg), putative helicase (also termed NTB) which has a putative nucleoside triphosphate binding
activity and X1 and X2 for which the functions are unknown (Rott et al., 1995; Wang and Sanfacon, 2000b). RNA2 encodes the structural proteins which include the coat protein (CP) and movement protein (MP), as well as the X4 and X3 proteins of unknown function (Carrier et al., 2001; Hans and Sanfacon, 1995). The cleavage sites within the ToRSV polyproteins consist of Q/G (glutamine/glycine) or Q/S (glutamine/serine) dipeptides. Site-directed mutagenesis of two ToRSV cleavage sites showed that efficient processing by the viral protease requires the presence of a Q at position -1 and a S or G at position +1 of the viral cleavage site (Carrier et al., 1999) (Fig. 1.4).

Fig. 1.4 Genome organisation of ToRSV. The polyprotein of nepoviruses are shown (P1 and P2). P1 is cleaved by the viral protease (Pro) at 5 sites to release 6 mature proteins and several intermediate precursors. P2 is cleaved at 3 sites to release 4 protein domain and possible precursors.

1.3.3 Assembly of the viral replication complex

All RNA viruses whether they infect mammalian, insect or plant cells induce the proliferation of membrane vesicles often but not always in the perinuclear area. Electron microscopy (EM) observations made more than 40 years ago described clusters of heterogeneously sized vesicles of 70–400 nm in diameter that were present in the perinuclear regions of poliovirus infected cells (Dales et al., 1965). Other examples of (+)-RNA viruses that are well known for replicating their genomes on intracellular membranes, include members of the Picornaviridae, Flaviviridae, Togaviridae, Coronaviridae and Arteriviridae families, the insect viruses of the Nodaviridae family and many plant viruses, such as Tobacco mosaic virus. As mentioned above, one of the best-documented examples of a virus that induces
membrane alterations is the human pathogen poliovirus, a member of the *Picornaviridae* family. Poliovirus-induced vesicle clusters are probably derived from the endoplasmic reticulum (ER). Other members of the *Picornaviridae* family have also been shown to replicate their RNA genomes on modified membranes that accumulate in the cytosol of infected cells (Wessels et al., 2006). (+)RNA viruses that belong to the *Flaviviridae* family and the *Nidovirales* order typically induce the formation of double-membrane vesicles that are spherical membrane structures 50–400 nm in diameter and composed of two closely apposed membrane bilayers (Mackenzie, 2005; Miller and Krijnse-Locker, 2008).

Viruses probably induce proliferation or reorganization of cellular membranes and vesicles in order to increase the surface area for RNA replication. The viral replication complex is associated with the intracellular membranes and includes viral and host proteins with the template RNA. The formation of this compartment may provide a protecting environment against RNA degradation (Sanfacon, 2005). In picornaviruses, RNA replication complex are associated with clusters of smooth vesicles which accumulate in the cytoplasm (Bienz et al., 1992). Cowpea mosaic virus (*Comovirus*) infection in plants induces the production of vesicles and massive proliferation of the endoplasmic reticulum (ER) membranes (Carette et al., 2000). Viral and host protein are brought to the replication complex via protein-membrane and protein-protein interactions. Some replication proteins are brought to the replication complex as polyprotein precursors which include the domains for the viral membrane anchor (Bedard and Semler, 2004).

### 1.3.3.1 Replication and translation of nepoviruses

As for other related viruses such as picornaviruses and comoviruses, nepovirus infection induces membrane proliferation and morphological changes in the ER membrane (Han and Sanfacon, 2003). Replication proteins and replication intermediates of comoviruses and nepoviruses are found in association with the ER derived membranous vesicles (Carette et al., 2000; Han and Sanfacon, 2003; Ritzenthaler et al., 2002; Schaad et al., 1997). Viral proteins act as membrane anchors for the replication complex. The membrane anchors associate with the ER and other viral and host proteins that are brought to the replication complex through protein–protein interaction with the membrane anchor protein. In ToRSV, the putative nucleoside triphosphate binding protein (NTB) has a stretch of hydrophobic residues at its C-terminal end which is proposed to anchor the replication complex on the
membrane (Han and Sanfaçon 2003). NTB-VPg associates with microsomal membranes in vitro and the VPg domain is translocated in the lumen of the membranes (Wang et al., 2004). In picornaviruses the VPg acts as a primer for viral replication. Since replication takes place in the cytoplasmic side, the ToRSV NTB-VPg is unlikely active in the replication of the viral RNA, suggesting that another form of the VPg may be the active primer for replication (Chisholm et al., 2007; Wang, 2004). The N-terminus of NTB can also be translocated in the lumen at least in vitro. This translocation is dependent on the presence of putative amphipathic helix, suggesting that at least two distinct elements may play a key role in the insertion of NTB-VPg in the membranes: a C-terminal transmembrane helix and an N-terminal amphipathic helix (Zhang et al., 2005).

In the RNA1 ToRSV, X2 protein has a highly hydrophobic protein domain located upstream of the NTB domain in the RNA1-encoded polyprotein. X2 has conserved sequence motifs with the comovirus 32 kDa protein which is an ER-targeted protein implicated in the viral replication complex assembly. Based on mutagenesis studies and confocal microscopy it was suggested that X2 is targeted to the ER membranes. This raises the possibility that it might act as a second membrane anchor for the viral replication complexes (Zhang and Sanfacon, 2006). Also a subpopulation of VPg-Pro-Pol which contains the truncated RNA-dependent RNA polymerase (Pol), the proteinase (Pro) and the VPg was peripherally associated with the ER-derived membranes active in viral replication (Chisholm et al., 2007). This suggests that the peripheral association of the ToRSV VPg-Pro-Pol polyprotein with the ER-derived membranes may be mediated by its interaction with one or several viral membrane proteins (X2 or NTB-VPg) (Han and Sanfacon, 2003; Zhang and Sanfacon, 2006; Zhang et al., 2005).

### 1.3.3.2 Replication of RNA2 in nepoviruses

Using grapevine fanleaf virus, it was shown previously that RNA1 replicates independently of RNA2. RNA2 requires the RNA1 replication machinery for its own replication. The RNA2 contains three protein domains named 2A, 2B (MP) and the 2C (CP). The lack of replication of GFLV RNA2 mutant, in which the 2A-coding sequence was deleted, strongly indicates that 2A or its coding sequence is essential for RNA2 replication. However, protein 2A alone is not enough and additional RNA or protein sequences downstream of 2A are essential (a minimum length of RNA2 exceeding the 2A coding sequence) for RNA2
replication. It was suggested that the requirement for downstream stabilizing sequences is related to the fact that 2A is an unstable protein at least in vitro (Margis et al., 1993). Also in cowpea mosaic virus (CPMV), the B-RNA contains the replication proteins and the M-RNA contains the movement protein and the coat protein. It was shown that the N-terminal domain of the 58 kDa protein encoded by the M-RNA is required for the replication of M-RNA (Van Bokhoven et al., 1993). The RNA2-encoded 58 kDa protein of CPMV contains both the 2A and movement protein domains (Van Bokhoven et al., 1993). It was suggested that the CPMV 58 kDa protein could be involved in binding RNA2 to form a ribonucleoprotein complex that would be recognized, on its own or together with unknown cellular factors, by the RNA1-encoded replicative machinery. In nepoviruses, it was suggested that the 2A protein is not active in replication as a mature protein but rather as a precursor form such as the 2AB or polyprotein P2. Thus the 2AB precursor protein in GFLV would be the functional equivalent of the 58 kDa protein in CPMV (Gaire et al., 1999). The following model is suggested for RNA2 replication in GFLV. After viral inoculation, RNA1 and RNA2 which were encapsidated in separate virus particles are released in the cytoplasm. Both RNAs are then translated into the P1 and P2 polyproteins. Translation of RNA1 and self-processing of the polyprotein P1 provide virus-encoded replication proteins that most likely co-assemble with host factors and membranes to form the viral replication complex, first as punctate structures and afterwards as juxtanuclear aggregated structures. RNA2 depends on RNA1 for its replication and polyprotein processing. Therefore, RNA2 needs to integrate within the replication complex initiated by RNA1-encoded proteins. Since 2A is required for RNA2 replication, it is suggested that the 2A domain within the nascent polyprotein either directs RNA2 to the replication site or interacts with the same cellular structure as RNA1-derived proteins to the juxtanuclear location. Another possibility is that the P2-polyprotein bound to RNA2 would be recruited by the replication complex (Gaire et al., 1999).

1.4 Movement of plant viruses

In animals, the spread of viral infection from cell to cell is by means of endocytosis or by fusion of the viral envelope with plasma membrane. A fundamental difference between plant cell and animal cells is that each plant cell is surrounded by a rigid cell wall. Plant viruses must overcome the barrier of the plant cell wall through cytoplasmic connections between adjacent cells, the plasmodesmata. Plant viruses encode proteins that assist their movement
from cell to cell. These movement proteins interact with the plasmodesmata and modify the plasmodesmal structure and function. Within the cell, the viral genome must be moved from the site of replication to the plasmodesmata. For viruses replicating in the cytoplasm, this often involves association with elements of the cytoskeleton. Cell to cell or short distance movement is defined as the movement of the virus from primarily infected cells (epidermal or mesophyll cells) to vascular bundle. Long distance transport of the virus occurs through vascular tissue, usually the phloem sieve tubes.

1.4.1 Plasmodesmata

Plasmodesmata are tubular extension of plasma membrane (40-50 nm in diameter) that traverse the cell wall and connect the cytoplasm of adjacent cells. The plasmodesmata are specialized channels that allow the intercellular movement of water, sugar, nutrients, and other molecules. They have a complex internal structure that regulates macromolecular traffic from cell to cell. Each plasmodesmata contains a long narrow tubule of the endoplasmic reticulum (ER) called desmotubule which is continuous with the ER of the adjacent cells. The cytoplasmic sleeve is a narrow space between the desmotubule and the plasma membrane. The trafficking of the macromolecules via plasmodesmata occurs in the cytoplasmic sleeve (Aaziz et al., 2001). In the cytoplasmic sleeve both the desmotubule and the plasma membrane globular proteins arrange in helical rows (Fig. 1.5). The size exclusion limit (SEL) of plasmodesmata is about 2.0nm. The SEL is not fixed and can be regulated. Plant virus particles or viral genome are too large to pass through the plasmodesmata. However, it has been suggested that many viral movement proteins increase the plasmodesmata SEL. The mechanism for regulating the SEL is poorly understood (Lincoln Taiz, 2006).
1.4.2 Types of movement from cell to cell by plant viruses

Two types of cell to cell movement of plant viruses with a (+) strand RNA genome have been characterized: movement of viral RNA as a nucleoprotein complex and movement of virus-like particles. Since nepoviruses move from cell to cell as a virus–like particle, I will explain this type of virus movement in more details.

1.4.2.1 Movement of virus-like particles

Plasmodesmata are modified by insertion of tubular structures that allows the transport of virus-like particles through them (tubule-guided virion movement). Such tubular structures have been described for viruses of different genera including Caulimovirus, Nepovirus, Bromovirus and Tospovirus.

1.4.2.1.1 Cell-to-cell movement of comoviruses, nepoviruses and related viruses

In CPMV, cell-to-cell movement is characterized by transport of mature virions through tubules that are assembled inside the plasmodesmal pore. The viral MP is a structural component of the tubular structures (Pouwels et al., 2004). Similar tubular structures are
formed at the surface of CPMV infected protoplasts. In protoplasts and in plant tissue; virus-like particles appear in a single and continuous row within the tubules. Tubule assembly does not depend on the presence of virions or capsid proteins, as expression of MP alone in protoplasts also leads to the formation of (empty) tubules (Angell et al., 1996; Carvalho et al., 2003; Kasteel et al., 1993; Van Lent, 1991; Wellink, 1989). The CPMV MP binds to intact virions and to the large CP subunit (Carvalho et al., 2003). The various steps leading to the formation of CPMV induced tubular structures have been studied. Dimeric or multimeric MP are first targeted to the plasma membrane. At the plasma membrane the MP accumulates in peripheral punctuate spots from which tubule formation begins. The C-terminus of MP interacts with the virus particle in the tubule and the N-terminal region of the MP is involved in tubule formation (Pouwels et al., 2003).

As with comoviruses, nepovirus-infected cells are characterized by the formation of tubular structures containing virus-like particles and traversing the cell wall. The movement protein of two nepoviruses, GFLV and ToRSV, was shown to be a structural component of the tubular structures (Ritzenthaler et al., 1995b; Wieczorek and Sanfacon, 1993). Tubular structures were formed both in plants and protoplasts infected with GFLV (Ritzenthaler et al., 1995a). The MP of GFLV may use the secretory pathway and the cytoskeleton for intracellular targeting and tubule assembly (Laporte et al., 2003; Taliansky et al., 2008). In ToRSV infected plants, MP is associated with the tubular structure containing virus particles. The tubular structures are often found protruding through the cell wall in to the cytoplasm of adjacent cells (Wieczorek and Sanfacon, 1993). The CP and the MP have been detected in infected plants and protoplasts (Sanfacon et al., 1995; Wieczorek and Sanfacon, 1993). In infected ToRSV protoplasts the MP was less stable than the CP (Sanfacon et al., 1995).

1.5 Virus transmission

Most plant viruses depend on a vector for transmission from plant to plant. Insects are the most common vectors (e.g. aphids), although other vectors, such as nematodes and fungi, are also important. The virus–vector interaction is very specific. Nepoviruses are transmitted by nematode therefore in this section I will focus on the transmission of plant viruses by nematodes.
1.5.1 Nematode transmission

There are two genera of plant viruses which are transmitted by nematodes: (1) Tobraviruses which are transmitted by species of the genera Trichodorus and Paratrichodorus and (2) Nepoviruses which are transmitted by species of the genus Xiphinema (Hull, 2002).

1.5.1.1 Tobraviruses

*Pea early browning virus* is a rod shape virus with bipartite (+)-strand RNA genome. Deletion of a gene encoding a 29 kDa protein (2b protein) abolished its nematode transmission without affecting the virus particle formation (MacFarlane, 1996). Comparison of the highly transmissible isolate and a poorly transmissible isolate showed two amino acid substitutions in the 2b protein (Vellios, 2002). In another tobravirus, *Tobacco rattle virus*, a strong interaction was detected between a 40 kDa protein and the CP. The 40 kDa protein is required for transmission of the virus by the nematode vector. This suggested that the 40 kDa protein of Tobacco rattle virus could act as a helper protein in nematode transmission (Visser and Bol, 1999).

1.5.1.2 Nepoviruses

ToRSV and Tobacco ringspot virus (another nepovirus of subgroup A) infect a wide range of fruit crops and woody plants in North America and are transmitted by *X. americanum* sensu strico Cobb, an ectoparasitic nematode. This nematode also transmits two other nepoviruses: Cherry rasp leaf virus and Peach rosette mosaic virus. For nematode transmission to occur, the virus must first dissociate from its retention site, a specific area in the nematode food canal. The dissociated viruses are injected to the plant root during the nematode feeding. Transmission of nepovirus by its nematode species is specific. It has been suggested that during nematode feeding the change of pH induced by the nematode salvation glands, releases the virus from its retention site (Wang, 2002). Based on electron microscopy and immunofluorescent labelling of the nematode *X. americanum*, the retention sites for TRSV are localized at the inner lining of the stylet extension and the esophageal lumen (Wang, 1998; Wang and Gergerich, 1998). In contrast, ToRSV particles are retained in the triradiate lumen of the esophageal bulb. The different retention sites suggest significant differences in the mechanism of virus release from the nematode vector. Immunofluorescent labelling with ToRSV CP antibodies did not detect the virus in the vector nematode. However, TRSV CP
antibodies readily labelled TRSV within the nematode vector. This suggests that ToRSV virions might be coated with other components from the nematode or from the host plant blocking the physical binding between virus particle and the antibody. Based on transmission assays, ToRSV is transmitted more efficiently (~100%) than TRSV (75%), although few virus particles were observed in the nematode vector. This suggests that ToRSV is more readily released from the retention site while TRSV might remain strongly attached to the nematode receptor (Wang, 2002). The determinants for the specificity of the retention site and the efficiency of transmission of these viruses remain to be determined.

Grapevine fanleaf virus is transmitted from plant to plant by *Xiphinema index*. The viral determinants responsible for nematode transmission were studied by engineering chimeric constructs in which the coding region of various proteins of GFLV was replaced with their counterparts from Arabis mosaic virus (ArMV, Nepovirus). All hybrid viruses with the CP of GFLV were transmitted by *X. index*. Other hybrid viruses with the CP of ArMV were not transmitted by this nematode. These results indicated that the CP of GFLV is the sole viral determinants for the specificity of the nematode vector (Andret-Link, 2004).

1.6 ToRSV isolates

ToRSV is a major pathogen of small fruit crops and fruit trees in North America, it also occurs in Australia, Bulgaria, Canada, Chile, China, Cyprus, Germany, Italy, Japan, Korea, New Zealand, Peru, Puerto Rico, Turkey, the United States of America, the USSR (former), Yugoslavia. The susceptible host species are found in the Family *Amaranthaceae, Caryophyllaceae, Chenopodiaceae, Cucurbitaceae, Geraniaceae, Leguminosae, Papilionoideae, Rosaceae, Solanaceae* and *Umbelliferae* (Brunt et al., 1996).

The complete nucleotide sequence of isolate Rasp2, a raspberry isolate from lower mainland, B.C., Canada has been determined (Rott et al., 1995; Rott et al., 1991b). The nucleotide and amino acid sequence for the VPG, Pro, Pol and CP of other ToRSV isolates were analyzed and compared with the published sequence (Rasp2). These isolates were Rasp1 (originally from raspberry plants in Washington state), GYV, grape yellow vein isolate (originally from infected grapevines in California), PYB-1, peach yellow bud-1 (originally from infected peach orchards in California) and T392 isolate from an unknown origin. The VPG sequence
and conserved amino acid motifs within the Pro, Pol and the CP domain were identical in all isolates (Wang and Sanfacon, 2000a).

In Chapter 2, I will explain the insertion of large amino acid repeats and point mutations contribute to a high degree of sequence diversity in the X4 protein of ToRSV (Jafarpour and Sanfaçon, 2009). Therefore, I will briefly explain below the general principles of virus evolution and possible sources of variation in the viral genome.

1.6.1 Virus evolution, diversity and sources of variation

1.6.1.1 Mutation, insertion and deletion

In theory, RNA viruses have a large population of diversity since they have an error-prone replication and a short generation time. The genetically diverse populations of RNA viruses are referred to as quasispecies (Roossinck, 1997). Maintaining diverse quasispecies is advantageous because when the virus is exposed to a new environment or a new selective regiment, the variant in the population that is more fit to the new environment can survive (Roossinck, 2003; Schneider and Roossinck, 2001). In nature, vector transmission can potentially transmit viruses to a variety of hosts and host adaptability is essential for the survival of plant viruses. Deletion and/or insertion may include single or several nucleotides. Some mutations will only have an effect in the nucleotide sequence and will therefore be silent at the amino acid level. But some mutations can result in amino acid change and frame-shift mutations that cause global changes.

1.6.1.2 Recombination

Recombination is a process in which segments of genes are swapped between different genetic variant or different strains during viral replication. It is a general phenomenon and it plays an essential role in the genetic variability and evolution of plant viruses (Roossinck, 2003; Schneider and Roossinck, 2001). In tobraviruses, sequencing of the RNA2 from more than 12 tobravirus isolates showed extensive diversity in the nucleotide sequence which is the result of recombination between nonhomologus regions of RNA1 and RNA2. In 8 of the 12 isolates the 3'-terminal sequence is replaced with those of RNA1 sequence (Vassilakos et al., 2001).
1.6.1.2.1 Protein domain repeats

Repetitive nucleotide sequences are frequently found in eukaryotic genomes. These regions are often hypermutable, rapidly gaining and losing repeats during the course of evolution. They are usually found in non-coding genomic regions but they can also be found within the genes. The repeats might be as short as three amino acid residues tandem repeats to repetition of homologous domains of 100 amino acid or more residues (Marcotte, 1998). It has been suggested that they arise through internal tandem duplication and recombination within a single gene. However the exact mechanism behind this phenomenon is not fully understood (Andrade, 2001; Bjorklund et al., 2006). Domain repeats are often involved in interaction with proteins or other ligands such as DNA or RNA. They are found in proteins with highly diverse functions (Bjorklund et al., 2006).

1.6.1.3 Serial passages of viruses, a source of creating sequence diversity

High level of genetic diversity and rapid evolution of viral RNA genomes are well documented (Kurath and Dodds, 1995). The majority of these changes are seen during serial passaging of viruses. Few studies have characterized the rate and nature of ongoing genetic changes over time under controlled experimental conditions in plant hosts by monitoring multiple serial passage lines of a virus populations for changes in their consensus sequences (Kurath and Dodds, 1995). Evidence for genetic variation of plant viruses was reported as early as 1926 (Agudelo-Romero et al., 2008). It was derived from the possibility of isolating different symptom variants from the same virus source. These variants were first obtained by isolation of symptom mutants from areas that showed atypical symptoms in systemically infected plants (McKinney, 1935). Also, it was soon observed that serial passaging under different conditions could alter viral properties. Most often these experiments involved a host shift (Garcia-Arenal et al., 2001). Viruses adapt to their hosts by evading their defense mechanisms and taking over the host cellular metabolism for their own benefit. Changing the cell metabolism as well as side-effects of antiviral responses contribute to symptoms development and virulence. Sometimes, a virus may move from its usual host species into a new one, or might fail to successfully infect and further transmit the next host. During serial passaging of the virus by mechanical inoculation, some functions such as vector transmission may be completely lost. In cucumber necrosis virus defective interfering particles are generated after at least eight serial passages (Rochon, 2008; Rochon, 1991). Defective interfering particles are virus particles that encapsidate defective genomes that are grossly
altered genetically, usually by significant deletion of essential function. Defective viral RNAs keep their replication origins and packaging signals allowing their encapsidation (Knipe et al., 2001). In potyviruses, most characterized virus proteins are multifunctional, and under different selective constraints may evolve differently corresponding to various functions. For instance, the helper component for the transmission of the potyviruses also has a role in the proteolytic processing of the genome translation product and in systemic movement. The fact that its function in aphid transmission is easily lost upon mechanical passage suggests that trade offs occur for the optimization of its other functions (reviewed by Garcia-Arenal et al., 2001). Below are a few documented examples of virus evolution in response to serial passaging in a new host.

In Tobacco etch virus (Potyvirus) a virus infected plant was allowed to infect native hosts. After 17 serial passages, the viral genome has evolved only five amino acid changes. An amino acid substitution in the viral VPg protein was responsible for the appearance of symptoms, while one substitution in the viral P3 protein contributed to increase in the symptom severity (Agudelo-Romero et al., 2008).

The coat protein (CP) of Hibiscus chlorotic ringspot virus (HCRSV) was identified as a strong RNA-silencing suppressor. CP mutants resulting from serial passage of HCRSV in its local lesion host showed a significantly reduced silencing suppression function, indicating that host-induced mutations lead to avirulence of HCRSV in plant that correlates with its reduced ability to suppress post-transcriptional gene silencing (Meng et al., 2006).

Plum pox virus (PPV), populations from peaches are able to adapt consistently to herbaceous hosts. They are characterized by reduction in time to symptom development, increase in inoculation efficiency and increase in viral titres. Two isolates of PPV from peaches were inoculated and passaged ten times in peas. Sequence analyses of the PPV population lines inoculated (pea-adapted strains of PPV), identified a specific mutation occurring consistently in the nuclear inclusion bodies gene when compared with the same PPV isolates passaged in parallel in peach. The mutation allowed PPV to replicate up to 20 times faster in its new host (Wallis et al., 2007).

Two isolates of plum pox virus, PPV-D and PPV-R were different in their host range. Some isolates lost their ability to infect their natural prunus hosts after repeated passages in
herbaceous hosts. The changes were introduced in PPV-R isolate after long term replication in *N. clevelandii*. Based on pathogenicity experiments of the two PPV isolates, PPV-R and PPV-D, the determinants for host specificity were located at the N-terminal region of the CP (Salvador *et al*., 2008).

In tobacco rattle virus (TRV, a tobravirus) the genome is very flexible and based on selection pressure; nematode transmissibility and encapsidation, which are encoded by RNA2, were abolished from the genome of the PPK20 isolate when this isolate was serially passaged in tobacco by mechanical inoculation (Hernandez *et al*., 1996).

### 1.7 Virion associated proteins

I have briefly discussed virus structure above (section 1.1). Although the main component of the virion is the coat protein(s) (CP), other minor virus-encoded protein components have also been found in association with plant virus particles. In Chapter 3, I will discuss the possibility that the X4 protein of ToRSV may be a virion associated protein (VAP). Therefore in this section, I will explain what virion associated proteins are and bring some examples of well studied VAP in plant viruses.

Virion associated protein (VAP) are proteins associated (directly or indirectly) with the capsid protein in the virion shell. Because they are exposed on the virus particle, they are readily available for interaction with host or vector components. VAP plays an important role in the virus replication cycle, including cell-to-cell or long distance movement in the plant, suppression of RNA silencing, stability of virus particles, regulation of the translation of the viral genome and vector transmission.

#### 1.7.1 CaMV P3 virion associated protein

In cauliflower mosaic virus (*Caulimoviridae*), the virion-associated protein (P3 protein) is a small 15 kDa protein with a bipolar functional structure. The P3 protein is important for aphid transmission of the virus and acts as a bridge between the CP and the aphid transmission factor (P2). Also it has been suggested that the virion bound P3, participates in the cell to cell or long distance movement of the virus inside the host (Plisson *et al*., 2003; Plisson *et al*., 2005). P3 interacts with the virus particle through its C-terminal region. Its N-
terminal region has a high α-helical content and contains a coiled-coil motif. The coiled-coil domain interacts with the aphid transmission factor (P2 protein). The N-terminal region of P2 interacts with the aphid stylet and its C-terminal domain interacts with the virion associated protein (P3). The MP oligomerizes as a trimer through a C-terminally located coiled-coil domain. The coiled-coil domain of MP also interacts with the coiled-coil domain of the VAP located at the virion surface. The MP-VAP interaction is necessary for the spread of virus in the plant but dispensable for replication of the virus (Stavolone et al., 2005).

1.7.2 HC-Pro Helper component proteinase

The potyvirus genome encodes a single polyprotein which is processed by three viral proteinases, P1, HC-Pro (Helper component proteinase) and NIa (nuclear inclusion), to release all viral proteins required for virus infection. P1 and HC-Pro are responsible for two distinct cleavages within the N-terminal region of the polyprotein. NIa carries out the remaining cleavages at the C-terminal two-third of the poly protein (Carrington et al., 1993).

HC-Pro is a virion-associated protein and a multifunctional protein (Kasschau and Carrington, 2001). It was first recognized as a helper factor for vector transmission (Thornbury, 1985). Its proteinase activity was later discovered (Oh and Carrington, 1989). The HC-Pro protein contains three regions: an N-terminal region essential for vector transmission, a C-terminal region which contains the proteinase activity and also overlaps with the cell-to-cell movement function and a central region implicated in all other functions (Plisson et al., 2003). The central region of HC-Pro is responsible for genome amplification (IGN motif), synergism with other viruses, systemic movement within the host plant (CC/SC motif) and silencing suppression (Kasschau et al., 1997) (Fig. 1.6).
Potato virus A (PVA, another potyvirus) produces pinwheel shape cytoplasmic inclusion bodies, which are made of the inclusion protein (Cl). The Cl of several potyviruses anchors to the cell wall or to the close proximity of the plasmodesmata. The ability to form a pinwheel structure and to assist the cell to cell movement of the virus depends on the ability of its self interaction (Lopez, 2001). The Cl of PVA is associated with a sub-population of the virus particles, possibly by interaction with the virion-bound HC-Pro. Both HC-Pro and CI are localized at the protruding tip at the end of the virus particle (Gabrenaite-Verkhovskaya, 2008).

1.7.3 Tobravirus 2b nematode transmission helper protein

As mentioned above, tobraviruses such as tobacco rattle virus (TRV) and pea early browning virus (PEBV) are transmitted by trichodorid vector nematodes. The 2b protein is a non structural protein encoded by RNA2. The 2b protein interacts with the CP, forming a bridge between the virus particle and the internal surface of the nematode vector feeding tube. Immunogold labelling of thin sections of PEBV infected plants probed with 2b specific antibodies showed that 2b is localized with the virus particle (Vellios et al., 2002).
interacting domains were investigated by yeast two-hybrid experiments. A central domain of the CP was identified that promote CP: CP interaction. Deletion of a conserved central coiled-coil domain of the 2b protein prevented its interaction with the CP, while deletion of N-and C-terminal domains of the 2b protein greatly enhanced its interaction with the CP. A C-terminal domain of the CP was shown to be important for interaction with the 2b protein (Holeva and Macfarlane, 2006).

1.7.4 Closterovirus virion associated proteins

Closteroviruses (family Closteroviridae) are long filamentous virus particle with a long body and a short tail. In Beet yellows virus, the virion is composed of two principal components, the minor CP assembles at one end of the virion and the major CP forms the long virion body. In addition there are three VAPs presented at the virion tail: the viral Hsp70 homologue (Hsp70h), a 64 kDa protein and a 20 kDa protein (Peremyslov et al., 2004). Comparative genomic analyses revealed that Hsp70h and p64-like proteins are conserved in all known viruses of the family Closteroviridae (Dolja et al., 2006). Hsp70h and p64 are each required for efficient assembly of the closterovirus virion tail (Alzhanova et al., 2007; Satyanarayana et al., 2000). The tail is narrower than the body (8 nm vs. 12 nm). Thus, the virions exhibit a peculiar three-segment architecture with the pointed tip segment likely formed by p20 (Peremyslov et al., 2004). The p20 is required for long-distance transport of the virus via the phloem. In Closteroviruses, particle assembly is a prerequisite for Beet yellows virus transport of the large RNA genome (Alzhanova, 2007).

Another member of the family Closteroviridae, lettuce infection yellow virus (LIYV) is transmitted from plant to plant by whiteflies (Bemesia tabaci). Virion preparation of LIYV showed that HSP70 and p59 (a protein of unknown function) are associated with the virus particle. LIYV minor CP antibody neutralizes LIYV transmission by B. tabaci. This suggests that the minor CP may be important for virus transmission (Tian, 1999).

1.7.5 Potexvirus virion associated protein

In potexviruses such as PVX, the triple-gene block proteins (TGBp1, TGBp2 and TGBp3) together with the CP are involved in cell to cell movement of PVX. The 25 kDa TGBp1 is multifunctional virion associated protein (VAP) protein involved in many activities including RNA helicase activity and RNA binding activity. TGBp1 has been suggested to increase the
plasmodesmal size exclusion limit and is also a suppressor of RNA silencing (Verchot-Lubicz, 2005). It has been suggested that the binding of the 25-kDa MP to the viral RNA molecule mediates the translation of the viral RNA. The 25 kDa protein binds selectively to one end of the virus particle. Binding of the TGBp1 to the end of the virus particle causes conformational changes in the terminal subunits and alters the stability of the 5'-terminus of the RNA. It has been suggested that binding of TGBp1 renders the viral RNA more accessible to the ribosome and that it triggers co-translation disassembly of PVX RNA (Atabekov et al., 2000; Karpova et al., 2006).

1.8 RNA silencing and viral suppressors of silencing and silencing activity in Nepoviruses

In Chapter 4 I will explain the possible role of the X4 protein as a suppressor of silencing. Therefore, in the section below, I will briefly explain the mechanism of RNA silencing in plants and the mode of action of suppressor of silencing encoded by plant viruses.

1.8.1 Mechanisms of RNA silencing

RNA silencing is an ancient defence mechanism. The first line of research towards recognition of RNA silencing as an important mechanism of defence against viruses led to the discovery of transgene-induced RNA silencing. It was found that over-expressing endogenous genes by introducing additional transgenic copies resulted in turning off the endogenous gene as well as the transgene (Napoli et al., 1990; Roth et al., 2004; Smith et al., 1990). The next piece of information came from studies of pathogen-derived resistance in which RNA silencing directed against a viral transgene provided resistance to the virus carrying the targeted sequence (Baulcombe, 1996; Goodwin et al., 1996). It was later found that RNA silencing is also induced as an antiviral response to virus invasion, even in the absence of transgenes (Ratcliff et al., 1999). Therefore, viruses are targets of RNA silencing. RNA silencing is triggered by double-stranded RNA (dsRNA). The dsRNA trigger can be derived from inverted–repeat transcripts or from complementary RNAs that anneal by base pairing. In the case of viral RNA, it has been proposed that double-stranded RNA formed as a replication intermediate or specific secondary structures in the viral RNA (such as large hairpins) can be recognized by the RNA silencing machinery (Baulcombe, 2005; Roth et al., 2004). Silencing is initiated by the cleavage of dsRNA or extensive folded RNAs by an
RNase III-like protein, called Dicer to 21-28 nucleotides short RNA (siRNA) with a 2-nt overhang at the 3' ends. The production of siRNA by Dicer is ATP-dependent. The two strands of the double stranded siRNA are separated by a helicase and one of the strands is recruited as the guide RNA to the RNA–induced silencing complex (RISC), which contain a member of the Argonaute (Ago) protein family. Agos have a siRNA-binding PAZ domain (PIWI Argonaut and Zwille) and also contain a PIWI domain (homologous to P-element induced wimpy testis in Drosophila)(Cerutti et al., 2000). Argonaute proteins are evolutionary conserved and are composed of two distinguishing domains PAZ and PIWI, while PIWI domain is limited to Argonaute proteins, the PAZ domain is also found in the DICER family of proteins (Song et al., 2003). The PIWI domain provides endonucleolytic activity to RISC complex to cleave the target RNAs and specifically inhibit the expression of the target gene (Baulcombe, 2005; Brodersen and Voinnet, 2006; Lözsa R, 2008) (Fig. 1.7). Host RNA-dependent RNA polymerases (RdRp) are also involved in the silencing pathway (Baulcombe, 2007).

Virus infection starts from entering through a small wound. The virus replicates in the cell and spreads from cell to cell until it reaches the vascular system. At this stage, the virus can rapidly move throughout the plant. In response to the virus, host plants initiate RNA silencing against the viral RNA and produce mobile silencing signals. If the virus moves faster than the silencing signal, a viral infection is established. If the silencing signal enters the cell first, then the infection fails to become systemic (Roth et al., 2004).
**Fig. 1.7 The silencing pathway.** Double-stranded RNA formed as a replication intermediate or specific secondary structures in the viral RNA can be recognized by the RNA silencing machinery. The silencing pathway is initiated by the cleavage of dsRNA or extensive folded RNAs by Dicer into 21-28 nucleotides short RNA (siRNA). The two strands of the double stranded siRNA are separated by the helicase and one of the strands is recruited as the guide RNA to the RNA–induced silencing complex (RISC). Within the activated RISC, single strand siRNAs bring the complex into contact with target RNAs (such as mRNA) in a sequence specific manner and specifically cause their degradation. Therefore, the expression of the target gene is inhibited.

**1.8.2 Viral suppressors of silencing**

To counteract the RNA silencing–mediated defence response of plants, many viruses express proteins with silencing suppressor activity. These proteins are not highly conserved and the mode of action of these proteins against RNA silencing is diverse. Theoretically, viruses can combat RNA-silencing mediated defence in at least three ways: preventing the generation of siRNAs, inhibiting the incorporation of siRNAs into effector complexes and interfering with one of the effector complexes.

Synergistic viral diseases caused by certain pairs of co-infecting viruses like (e.g., co-infection of tobacco plants with Potato virus X (PVX) and any of several members of the genus *Potyvirus*) causes a synergistic disease characterized by a dramatic increase in symptom
severity correlated with a 3 to 10 fold increase in the accumulation of PVX in the first systemically infected leaves. Based on these studies, a viral protein called helper component proteinase (HC-Pro) was suggested to mediate one class of viral synergistic disease (Shi et al., 1997; Vance et al., 1995). Expression of HC-Pro in transgenic plants allowed a broad range of heterologous viruses to accumulate beyond the normal level, suggesting that HC-Pro blocks a general plant defence mechanism (Pruss et al., 1997). The mechanism blocked by HC-Pro was found to be RNA silencing (Anandalakshmi et al., 1998; Brigneti et al., 1998; Kasschau and Carrington, 1998). Since the initial demonstration that HC-Pro blocks RNA silencing (silencing suppressors) many other plant viral suppressors of silencing have been identified.

It has been proposed that the P1 protein, which is a serine protease (Potyviridae), enhance the suppressor activity of HC-Pro in several potyviruses (Rajamaki et al., 2005). Interestingly, the HC-Pro protein is not encoded by cucumber vein yellowing virus (Ipomovirus). In this virus, P1 is duplicated (P1a and P1b, arranged in tandem) and the second copy, P1b, has an RNA silencing suppression activity (Valli et al., 2006).

The 19 kDa protein (p19) of tomato bushy stunt virus (genus tombusvirus) is another well studied plant virus suppressors of silencing. The p19 protein was first identified as a symptom determinant and subsequently, as a suppressor of silencing protein that binds to the siRNAs in vitro (Silhavy et al., 2002). The crystal structures of carnation Italian ringspot virus and tomato bushy stunt virus p19 proteins bound to the siRNA duplexes was elucidated (Vargason, 2003; Ye, 2003). It was suggested that p19 interferes with the assembly of RISC by preventing the RNA helicase from gaining access to the unpaired ends of a siRNA duplex thereby preventing strand separation and RISC assembly (Fig. 1.8).
Fig. 1.8 Mechanism of the silencing suppression activity of p19. The normal silencing pathway is shown on the left of the figure. The mechanism of inhibition of p19 is shown on the right. The tombusvirus p19 protein (shown in green) binds to the double-stranded siRNAs and prevents their loading into the RISC complex.

In turnip crinkle virus (TCV), a small ssRNA icosahedral virus, the CP strongly suppresses gene silencing induced by sense, antisense and double-stranded RNAs. It prevents systemic silencing in *N. benthamiana* plants and prevents the formation of siRNAs (Qu *et al.*, 2003).

As mentioned above, CPMV (*Comovirus*) codes for two viral coat protein: the large (L) and the small (S) CP. The small coat protein is a suppressor of post-transcription gene silencing, although it has a weak activity compared to other suppressors like HC-Pro (Cañizares, 2004; Voinnet *et al.*, 1999). The suppressor activity lies within the C-terminal 24 a.a of the S protein (a.a 198-213). This sequence is exposed on the surface of the virus particle and would be instantly available when the virus enters the cell (Cañizares, 2004).

In potexviruses (family *Flexiviridae*), three overlapping open reading frames, known as triple gene block (TGB), are required for virus cell to cell movement. One of these proteins,
TGBp1, binds to one end of the virion (single-tailed particles) (Atabekov et al., 2000). Although siRNAs were found in inoculated tobacco leaves, the accumulation of these siRNA is apparently controlled by the TGBp1 protein. When the TGBp1 coding region was deleted from the PVX genome, the 25 nt siRNAs accumulated at higher levels and systemic silencing occurred. Therefore TGBp1 is a suppressor of systemic RNA silencing, acting by regulating the accumulation of the 25 nt siRNAs (Hamilton et al., 2002; Verchot-Lubicz et al., 2007; Ye et al., 2009).

Citrus tristeza virus (Closterovirus) encodes at least three suppressor of RNA silencing (p23 and p20 proteins and the CP). The CP targets the intercellular spread of the silencing signal but does not suppress intracellular silencing (Lu et al., 2004).

Cucumber mosaic virus (CMV) 2b protein encoded by the RNA2 of the tripartite viral RNA genome is one of the first identified RNA silencing suppressors (Brigneti et al., 1998). It is believed that the 2b protein inhibits siRNA-directed systemic silencing and local silencing (Goto, 2007; Guo and Ding, 2002; Qi et al., 2004). It has been shown that the 2b protein from a severe CMV strain interacts with the AGO1 protein and blocks the AGO1 slicer activity (Argonaute protein AGO1, is the core component of the RISC complex) (Zhang et al., 2006). Based on RNA sequence data CMV strains can be classified into three subgroups: IA, IB (severe subgroup) and II (mild subgroup). It has been recently reported that the severe subgroup (SD-CMV) has a stronger 2b suppressor activity than the mild subgroup (Q-CMV). Also it has been suggested that the variable domain in SD2b which is absent in Q2b is the major contributor to the stronger activity of SD2b (Ye et al., 2009).

Most plant viruses suppress silencing by binding to the long double strand RNA or to the siRNA. There are two exceptions known so far. One of these exceptions is the 2b protein of Cucumoviruses which binds to the AGO1 effector protein and inactivates its activity. The second example is the P0 RNA silencing suppressor of the poleroviruses (family Luteoviridae). P0 is an F box protein that interacts with Arabidopsis homologs of the S-phase kinase-related protein 1 (SKP1) ASK1 and ASK2. P0 does not have RNA binding activity. SKP1 is a core subunit of multicomponent SCF (SKP1/Cullin/F-box/RBX1) E3 ubiquitin ligase. It was suggested that P0, ASK1 and ASK2, RBX1 and CULLIN1 would form an SCF-like E3 ubiquitin ligase involved in the specific ubiquitination and possibly subsequent degradation of a host protein implicated in RNA silencing.
E3 ubiquitin ligase is present in diverse protein complexes that direct ubiquitination of specific target proteins as a signal for their degradation to the 26S proteasome. The F box proteins are components of E3 ubiquitin ligase complexes that add polyubiquitin tracts on selected lysine residues and mark the protein for degradation (Baumberger et al., 2007). Below I will briefly discuss this degradation pathway and provide an example of plant virus proteins studied so far that uses this pathway.

The most widely used and general degradation pathway in both nucleus and cytoplasm is the 26S proteasome, which is conserved in all eukaryotes. Proteins are targeted to 26S proteasome mostly by covalent linking of ubiquitin chains. This process is catalyzed by ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzymes and an ubiquitin protein ligase (E3). In animals and yeast the ubiquitination/26S proteasome regulate the degradation of short-lived enzymes and regulatory proteins (Karsies et al., 2001). A well-studied plant virus that uses the degradation pathway is tobacco mosaic virus (TMV). The TMV movement protein is targeted through degradation through the ubiquitin-proteasome pathway (Reichel and Beachy, 2000). The CaMV capsid protein precursor expressed in plant protoplasts is also very unstable (Leclerc et al., 1999). Degradation signals within both terminal domains of the cauliflower mosaic virus capsid protein precursor (p44) have been identified, The N-terminus of p44 contains a degradation motif that is characterized by proline, glutamate, aspartate, serine and threonine residues (PEST-rich sequence). PEST–rich sequences are potential degradation signals that direct proteins to the ubiquitin/proteasome pathway. The PEST signal in the CaMV CP protein can be inactivated by mutation of glutamic acid residues to alanines (Karsies et al., 2001). It has been reported that phosphorylation of serine or threonine residues can activate latent PEST sequences (Hericourt et al., 2000).

1.8.3 Host recovery from virus infection and the role of RNA silencing in nepoviruses

After initial symptomatic infection, the upper leaves of plants infected with certain viruses, including nepoviruses, become symptom-free. These leaves are said to be recovered. In tobacco ringspot virus (a nepovirus of subgroup A)-infected plants, recovered leaves are protected against secondary infection with the same virus (Wingard, 1928). In plants infected with another nepovirus (tomato black ring virus), a tobravirus (tobacco rattle virus), a caulimovirus and a geminivirus (pepper golden mosaic virus), the recovery phenotype was
shown to be accompanied by a decrease in the virus titre and induction of virus-specific RNA silencing (Covey et al., 1997; Rodriguez-Negrete et al., 2009). Recovery of N. benthamiana plants from the necrotic symptoms induced by ToRSV is also associated with RNA silencing. However, the virus titre is not significantly reduced in recovered leaves (Jovel et al., 2007). Similarly, in tobacco streak virus, recovery of host can be initiated by minimal genetic changes in a viral genome and it can occur in the absence of virus clearance (Xin and Ding, 2003). The role of RNA silencing in recovery from infection by ToRSV and Tobacco streak virus remains to be clarified.

1.9 The X4 protein of ToRSV

The ToRSV genome consists of two molecule of single-stranded RNA. As I discussed earlier, the RNA2 of ToRSV encodes the coat protein (CP), movement protein (MP), as well as X4 and X3 for which the function is unknown (Carrier et al., 2001; Hans and Sanfacon, 1995). The ToRSV X3 protein shares conserved sequence motifs with the 2a protein of subgroup A and B nepoviruses (Rott et al., 1991c). As discussed above, the 2a protein of grapevine fanleaf virus has been shown to play a role in the replication of RNA2 (Gaire et al., 1999). The ToRSV X4 protein does not have significant sequence identity with any other proteins available in the database and does not have a functional equivalent in the genome of nepoviruses of subgroup A and B (e.g. grapevine fanleaf virus) (Fig 1.9). The objective of this thesis is to provide new information to increase our understanding of the possible role of the X4 protein in ToRSV replication cycle. In order to understand the nature of X4, different computer prediction programmes and the protein sequence of three isolate of ToRSV were used. The Rasp2 isolate was sequenced previously (Rott et al., 1991b). The Rasp1 and PYB isolate were sequenced during this project (Jafarpour and Sanfaçon, 2009).

![Fig.1.9 Comparison of the genomic structure of tomato ringspot virus and grapevine fanleaf virus (nepoviruses).](image-url)
The following sections are computer predictions based on the protein sequence of X4. As I mentioned before, based on blast search, X4 has no significant homology with other proteins in the database. Based on the protein sequence of X4, I have listed a few computer predictions that may help understand the possible function this protein.

1.10 Computer predictions based on ToRSV X4 amino acid sequences.

1.10.1 PEST motifs

As mentioned before, PEST motifs play a role as protein degradation signal (Karsies et al., 2001; Rogers, 1986). The deduced amino acid sequence from the X4- Rasp1 has 13 poor PEST motifs. X4-Rasp2 has 10 PEST motifs and the deduced amino acid sequence from the X4- PYB has 9 PEST motifs (Tables 1.1., 1.2 and 1.3). Below is a more detailed description of this prediction. Phosphorylation of serine or threonine residues can activate latent PEST sequences (Fig.1.10) (Hericourt et al., 2000) (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=epestfind and http://www.cbs.dtu.dk/services/NetPhos/).

Table 1.1 PEST motifs as potential proteolytic cleavage sites in Rasp2 isolate. 10 PEST motifs were identified in X4 (Rasp2) isolate from amino acid positions 1 to 634 and sorted by score.

<table>
<thead>
<tr>
<th>Rasp2</th>
<th>PEST motif amino acids</th>
<th>a.a</th>
<th>PEST Score</th>
<th>PEST motif</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HQEEREPEMVPAVLEAADSVDG</td>
<td>407-458</td>
<td>8.04</td>
<td>Potential</td>
</tr>
<tr>
<td></td>
<td>DITEAFFDDLECESFYSY...AE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>RAVPQIVQQTMMTEEEFEVPS</td>
<td>232-266</td>
<td>-3.97</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>SSSPLPLFANFK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>HEYPEISSAPVQSSLFSR</td>
<td>596 -613</td>
<td>-4.68</td>
<td>Poor</td>
</tr>
<tr>
<td>4</td>
<td>RVVLPDECMRLLSLFEDQPLL</td>
<td>333 - 372</td>
<td>-4.89</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>PEGPLPSFSWSSPLLFASFK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>RVVLPDECMRLLSLFEDQPLL</td>
<td>280-319</td>
<td>-5.76</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>PEGPLPSFSWSSPLLFANFK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>RQAVCPVEMPSCATSVSEK</td>
<td>169-187</td>
<td>-7.79</td>
<td>Poor</td>
</tr>
<tr>
<td>7</td>
<td>RQEACSVVAAPPIVEPVLWVP</td>
<td>84-115</td>
<td>-8.05</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>PVLWVPPLSEYANDFPK</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1.2  PEST motifs in the X4 protein of PYB isolate.  9 PEST motifs were identified in PYB isolate from positions 1 to 635 and sorted by score.

<table>
<thead>
<tr>
<th>8</th>
<th>RLVTLQSFEELPLYSSR</th>
<th>488-504</th>
<th>-12.2</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>KVVSDEFMDVLPLFSPLVSH</td>
<td>387-407</td>
<td>-15.9</td>
<td>Poor</td>
</tr>
<tr>
<td>10</td>
<td>HDYVEGCMASTVLGCAVPVQR</td>
<td>63-84</td>
<td>-21.0</td>
<td>Poor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PYB</th>
<th>PEST motif amino acids</th>
<th>a.a</th>
<th>PEST Score</th>
<th>PEST motif</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RMVPAVLEAANSIDDVTEAFFDDCESFYDSYDEEEAEWAEVPR</td>
<td>413-458</td>
<td>3.40</td>
<td>Poor</td>
</tr>
<tr>
<td>2</td>
<td>RIVQQPTMTEEELDFEVPSSSWSSPLPLFANFK</td>
<td>236-266</td>
<td>-3.69</td>
<td>Poor</td>
</tr>
<tr>
<td>3</td>
<td>HEYPEIPSAPVQSSLFSR</td>
<td>596-613</td>
<td>-4.25</td>
<td>Poor</td>
</tr>
<tr>
<td>4</td>
<td>KESCAVVAVPSGEFTSWVPSCTEYVSNFK</td>
<td>85-115</td>
<td>-4.29</td>
<td>Poor</td>
</tr>
<tr>
<td>5</td>
<td>RVVLPDECMDDLSSLFLFDQLESEGPLPPFSWSSPLPLFASFK</td>
<td>280-319</td>
<td>-4.89</td>
<td>Poor</td>
</tr>
<tr>
<td>6</td>
<td>RVVLPDECMDDLSSLFLFDQLESEGPLPPFSWSSPLPLFASFK</td>
<td>333-372</td>
<td>-4.89</td>
<td>Poor</td>
</tr>
<tr>
<td>7</td>
<td>RVVPDEFMDVSPFLLTTLPLVLH</td>
<td>387-407</td>
<td>-11.28</td>
<td>Poor</td>
</tr>
<tr>
<td>8</td>
<td>HEYVEGSLASTILECAGPVQSK</td>
<td>63-85</td>
<td>-11.48</td>
<td>Poor</td>
</tr>
<tr>
<td>9</td>
<td>REMLPCAVFVSEK</td>
<td>175-187</td>
<td>-22.00</td>
<td>Poor</td>
</tr>
</tbody>
</table>
**Table 1.3 PEST motifs in the X4 protein of Rasp1 isolate.** 13 PEST motifs were identified in X4 protein Rasp1 isolate from positions 1 to 732 and sorted by score.

<table>
<thead>
<tr>
<th>Raspa 1</th>
<th>PEST motif amino acids</th>
<th>a.a</th>
<th>PEST Score</th>
<th>PEST motif</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RMVPAVLEAANSIDDVTEAFFDLECSEFYDSYSCEEEAEWAENPR</td>
<td>510-555</td>
<td>1.52</td>
<td>Poor</td>
</tr>
<tr>
<td>2</td>
<td>HQEEPEMAPAVLEAAATTPLVLH</td>
<td>460-482</td>
<td>-1.34</td>
<td>Poor</td>
</tr>
<tr>
<td>3</td>
<td>HQEEPEMAPAVLEAVTTPLVLH</td>
<td>482-504</td>
<td>-2.25</td>
<td>Poor</td>
</tr>
<tr>
<td>4</td>
<td>RVVLPECMDDLSSLFDQPLPEGPLPPFSWSSPLPLFTSFK</td>
<td>386-425</td>
<td>-3.30</td>
<td>Poor</td>
</tr>
<tr>
<td>5</td>
<td>RIVQQPTMTELEFEVPSWSSPLPLFANFK</td>
<td>236-266</td>
<td>-3.46</td>
<td>Poor</td>
</tr>
<tr>
<td>6</td>
<td>KESCAVVAVPSEELISWVPSCTEYASNPR</td>
<td>85-115</td>
<td>-3.74</td>
<td>Poor</td>
</tr>
<tr>
<td>7</td>
<td>HEYPEIPSAPVQSSLFSR</td>
<td>693-710</td>
<td>-4.25</td>
<td>Poor</td>
</tr>
<tr>
<td>8</td>
<td>RVVLPECMDDLSSLFDQPLPEGPLPPFSWSSPLPLFANFK</td>
<td>280-319</td>
<td>-5.58</td>
<td>Poor</td>
</tr>
<tr>
<td>9</td>
<td>RVVLPECMDDLSSLFDQPLPEGPLPPFSWSSPLPLFANFK</td>
<td>333-372</td>
<td>-5.58</td>
<td>Poor</td>
</tr>
<tr>
<td>10</td>
<td>HEYVGSLASTILECAGPVVQSK</td>
<td>63-85</td>
<td>-11.48</td>
<td>Poor</td>
</tr>
<tr>
<td>11</td>
<td>KSLVPLDEWSFYLR</td>
<td>187-200</td>
<td>-18.63</td>
<td>Poor</td>
</tr>
<tr>
<td>12</td>
<td>RVVLDEFMDVLPFLLSPVNLH</td>
<td>440-460</td>
<td>-18.94</td>
<td>Poor</td>
</tr>
<tr>
<td>13</td>
<td>REMLPCAVFVSEK</td>
<td>175-184</td>
<td>-22.00</td>
<td>Poor</td>
</tr>
</tbody>
</table>
1.10.2 Phosphorylation sites in ToRSV isolates

A

NetPhos 2.0: predicted phosphorylation sites in Sequence

B

NetPhos 2.0: predicted phosphorylation sites in Sequence
Fig. 1.10 Predicted phosphorylation sites. Phosphorylation sites in 732 amino acid sequence of X4-Rasp1 isolate, based on the sequence. There are 22 Serine, 5 Threonine and 7 Tyrosine phosphorylation sites predicted in the sequence (A). Predicted phosphorylation sites in 640 amino acid sequences of X4-PYB isolate, based on the a.a sequence. There are 22 Serine, 4 Threonine and 9 Tyrosine phosphorylation sites predicted in the sequence (B). Predicted phosphorylation sites in 640 amino acid sequence of X4-Rasp2 isolate, based on the sequence. There are 24 Serine, 5 Threonine and 8 Tyrosine phosphorylation sites predicted in the sequence (C).

1.10.3 Topology prediction of membrane proteins.

The topology prediction of membrane proteins (TopPred), based on the amino acid sequence of X4 from various ToRSV isolate is shown below (Fig.1.11 and table 1.4) (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=toppred). There are 4-5 candidate membrane-spanning segments in the X4 protein. The hydrophobicity value is low however further studies are needed to test the possibility of X4 being a membrane protein. The X4 secondary structures as predicted by the Hnn program show high helix probability and high sheet probability in the X4 protein. The secondary structure prediction suggested two helix-rich regions that are separated by a less structured hinge region. The predicted structural domains may facilitate further understanding of the structural and functional characterization of the X4 protein (Fig.1.12).
Fig. 1.11 Topology prediction of ToRSV isolate. (A) Rasp2 isolate has 5 candidate membrane-spanning segments. (B) X4-Rasp1 has 4 candidate membrane-spanning segments and (C) X4-PYB has 4 segments. Score details are listed in the table below.

Table 1.4 Detail description of the topology prediction of ToRSV isolates.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Helix Begin - End</th>
<th>Score</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasp2</td>
<td>64 - 84 a.a</td>
<td>0.884</td>
<td>Putative</td>
</tr>
<tr>
<td></td>
<td>87 - 107 a.a</td>
<td>0.969</td>
<td>Putative</td>
</tr>
<tr>
<td></td>
<td>144 - 164 a.a</td>
<td>1.459</td>
<td>Certain</td>
</tr>
<tr>
<td></td>
<td>300 - 320 a.a</td>
<td>0.934</td>
<td>Putative</td>
</tr>
<tr>
<td></td>
<td>353 - 373 a.a</td>
<td>1.159</td>
<td>Certain</td>
</tr>
</tbody>
</table>
### 1.10.4 Secondary structure prediction in different ToRSV isolates

<table>
<thead>
<tr>
<th></th>
<th>Helix Begin - End</th>
<th>Score</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasp1</td>
<td>140 - 160 a.a.</td>
<td>1.392</td>
<td>Certain</td>
</tr>
<tr>
<td></td>
<td>300 - 320 a.a</td>
<td>0.884</td>
<td>Putative</td>
</tr>
<tr>
<td></td>
<td>353 - 373 a.a</td>
<td>0.884</td>
<td>Putative</td>
</tr>
<tr>
<td></td>
<td>406 – 426 a.a</td>
<td>1.089</td>
<td>Certain</td>
</tr>
<tr>
<td>PYB</td>
<td>87 - 107 a.a</td>
<td>0.829</td>
<td>Putative</td>
</tr>
<tr>
<td></td>
<td>143 - 163 a.a</td>
<td>1.718</td>
<td>Certain</td>
</tr>
<tr>
<td></td>
<td>300 - 320 a.a.</td>
<td>1.109</td>
<td>Certain</td>
</tr>
<tr>
<td></td>
<td>353 - 373 a.a</td>
<td>1.109</td>
<td>Certain</td>
</tr>
</tbody>
</table>

**Fig.1.12 Secondary structure prediction in different ToRSV isolates.**  
A) Hierarchical Neural Network result for X4-Rasp2 isolate. Sequence length is 640 amino acids. Alpha helix (Hh): 244 is 38.49%, 3₁₀ helix (Gg): 0 is 0.00%, Pi helix (Ii): 0 is 0.00%, Beta bridge (Bb): 0 is 0.00%, Extended strand (Ee): 69 is 10.88%, Beta turn (Tt): 0 is 0.00%, Bend region (Ss): 0 is 0.00%, Random coil (Cc): 321 is 50.63%, Ambiguous states (?): 0 is 0.00%, Other states: 0 is 0.00%.  
B) Hierarchical Neural Network result for X4-Rasp1 isolate. Sequence length is 732 amino acids. Alpha helix (Hh): 278 is 37.98%, 3₁₀ helix (Gg): 0 is 0.00%, Pi helix (Ii): 0 is 0.00%, Beta bridge (Bb): 0 is 0.00%, Extended strand (Ee): 104 is 14.21%, Beta turn (Tt): 0 is 0.00%,
1.11 Research hypothesis and objectives

As discussed above, the ToRSV X4 protein does not have significant sequence identity with any other proteins available in the database. Also, it is absent in the genome of other nepoviruses. The X4 protein could be involved in functions such as the specificity of nematode transmission or interaction of the virus with host defence responses. Also it is possible that the X4 plays an accessory role in cell to cell movement or the RNA2 replication. The goal of this thesis is to characterize the X4 protein at the molecular level. The specific objectives of this thesis are as follows:

1- Characterize possible interactions between the X4 protein and other viral proteins.
2- Analyse the variability of X4 among other ToRSV isolate.
3- Investigate the possible function of the X4 protein.
1.12 Bibliography


Chapter 2

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

A version of this chapter has been published. Jafarpour, B. and 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). Archives of Virology, 154, 1713-1717
2.1 Introduction

Tomato ringspot virus (ToRSV) is a serious pathogen of small fruits and fruit trees in North America (Sanfacon et al., 2006). Several isolates have been described that differ in their natural host range and in the intensity of symptoms they induce in herbaceous hosts (Bitterlin and Gonsalves, 1988; Wang and Sanfacon, 2000). *Tomato ringspot virus* is a member of the genus *Nepovirus* (Sanfacon, 2008), which has recently been reassigned to the proposed new family “secoviridae” (sub-family “comovirinae”) within the order *Picornavirales* (Sanfacon et al., 2009). Nepoviruses have a bipartite positive-strand RNA genome. Each RNA encodes a polyprotein that is cleaved by the viral protease (Pro) at specific cleavage sites. RNA1 encodes replication proteins while RNA2 codes for the coat protein (CP), movement protein (MP) and other protein(s) of less defined function. Using *in vitro* processing assays, three protease cleavage sites have been identified in the RNA2-encoded polyprotein of ToRSV (a subgroup C nepovirus), allowing the definition of four protein domains (X3, X4, MP and CP) (Carrier et al., 1999; Carrier et al., 2001). In contrast, only three protein domains are present in the polyprotein of subgroups A and B nepoviruses (2a, MP and CP, see Fig. 2.1A). The ToRSV X3 protein shares conserved sequence motifs with the 2a protein of subgroup A and B nepoviruses. It has been suggested that the 2a protein plays a role in the replication of RNA2 (Gaire et al., 1999; Rott et al., 1991). The ToRSV X4 protein does not have significant sequence identity with proteins available in the database and does not have a functional equivalent in the genome of nepoviruses of subgroups A and B (Fig. 2.1A). The only other nepovirus of subgroup C for which the entire nucleotide sequence is available is blackcurrant reversion virus (BRV). The extent of sequence identity between ToRSV and BRV is very low in the deduced amino acid (a. a.) sequence of the X4 protein (8 %) but higher in the deduced a. a. sequence of the CP (29 %) or the VPg (viral protein linked to the genome), Pro and polymerase (Pol) (36 %). In this study, we examined the extent of sequence diversity in the X4 protein among closely related ToRSV isolates. We show that there is a high degree of sequence diversity in the X4 protein, which is due in part to the insertion of multiple copies of two types of large amino acid repeats.
2.2 Materials and methods

2.2.1 Provenance of the virus material

The complete nucleotide sequence of a ToRSV raspberry isolate from British Columbia (isolate Rasp2) has been previously determined (Genbank accession number NP_60762) and cDNA clones of RNA1 and RNA2 are available, although they are not infectious (Rott et al., 1995; Rott et al., 1991). The original Rasp2 isolate was lost and is not available for inoculations. A different ToRSV raspberry isolate (Rasp1) is available to our laboratory and has been maintained in cucumber for several years. Rasp1 was originally isolated by Dr. R. R. Martin (USDA-ARS, Corvallis) from raspberry fields in Washington State. Sequence analysis of the CP, X2, VPg, Pro and Pol coding regions of Rasp1 revealed that it is closely related to Rasp2 [sequence identity of 97-100 % at the nucleotide (nt) level and 98-100 % at the a. a. level] (Wang and Sanfacon, 2000; Zhang and Sanfacon, 2006). We have previously shown that *N. benthamiana* plants inoculated with Rasp1 recover from infection at temperatures equal to or above 27° C and young leaves become symptom-free after an initial symptomatic systemic infection (Jovel et al., 2007). At lower temperatures (21° C), recovery does not occur and the plants eventually die as a result of the necrosis induced upon ToRSV infection (results not shown). Isolate PYB1 (peach yellow bud) was originally collected from a peach orchard in California and was obtained from Dr. A. Rowhani (UC-Davis). PYB1 is closely related to Rasp1 and Rasp2 as revealed by high level of sequence identity in the CP, VPg, Pro and Pol coding regions (93-98 % identity at the nt level and 97-100 % at the a. a. level) (Wang and Sanfacon, 2000). In *N. benthamiana* plants infected with PYB1, symptoms are milder than those observed in Rasp1-infected plants and PYB1-infected plants recover from infection even when cultivated at 21° C (data not shown).

Leaves from cucumber plants infected with Rasp1 or from *N. benthamiana* plants infected with PYB1 were ground in liquid nitrogen and total RNA was extracted using the TRIzol reagent (Invitrogen). Because of the high degree of sequence conservation previously observed between Rasp1, Rasp2 and PYB1, I designed primers that were based on the published Rasp2 nucleotide sequence. First–strand cDNA was synthesised using 1 μg of total RNA as a template and primer CP-R (5'-GTCAAGCTTGCCACGCCCGAAAGGAT-3', complementary to nts 5723-5707 of the RNA2 of ToRSV isolate Rasp2) (Rott et al., 1991) and Superscript II Reverse Transcriptase (Invitrogen). For the PCR reaction, I used primers F208 (5'-GAGGCCGAATTGGCCTCAAG-3', corresponding to nts 801-822 of Rasp2
RNA2, approximately 180 nts upstream of the X4 coding region) and R207 (5'-GCACCCGCATCAGAGGATC-3', complementary to nts 2998-2980 of Rasp2 RNA2, approximately 120 nts downstream of the X4 coding region) (Fig. 2.1A). As expected, a fragment of approximately 2200 nts was amplified when these primers were used in a direct PCR reaction with a cDNA clone derived from Rasp2 RNA2 or in an RT-PCR reaction with total RNA extracted from PYB1-infected plants (Fig. 2.1B). However, a significantly larger cDNA fragment (approximately 2500 nts in length) was amplified by RT-PCR with total RNA extracted from Rasp1-infected plants (Fig. 2.1B). Sequencing reactions were initially conducted directly on amplified cDNA fragments, using primers R207 and F208. Other primers were designed as the sequence of the Rasp1 and PYB1 isolates were determined. The sequence was confirmed by partial sequencing of several individual cDNA clones. Each nucleotide was sequenced at least twice in each direction. Sequences were assembled using vector NTI advance software (Invitrogen). Multiple sequence alignments were completed using Clustal W2 (Chenna et al., 2003). The sequences for the X4 coding regions of ToRSV-PYB and ToRSV-Rasp1 were deposited in the GenBank database under accession numbers GQ325249 and GQ325250, respectively.

2.3 Results

2.3.1 Sequence properties

The previously identified X3-X4 and X4-MP cleavage sites were conserved in the Rasp1 and PYB1 isolates (Fig. 2.1C). The Rasp1 X4 coding region was 2190 nts in length compared to 1899 nts for the X4 coding region from Rasp2 and PYB1. The predicted size of the Rasp1-X4 protein is 730 a. a. (calculated Mr of 82 kDa) compared to 633 a. a. for the Rasp2 and PYB1 X4 protein (calculated Mr of 71 kDa). The degree of sequence identity in the X4 protein among the three ToRSV isolates was 75-87 % at the a. a. level (78-89 % at the nt level), a value significantly lower than that previously observed for the CP or VPg-Pro-Pol protein domains (Table 2.1). This high degree of diversity was due in part to the insertion of sequences in the X4 protein from Rasp1. However, even when gaps were excluded from the analysis, the degree of sequence identity in the X4 protein was lower than that observed for other protein domains (Table 2.1). Analysis of individual cDNA clones of the X4 coding region of Rasp1 revealed at least two sites of sequence heterogeneity in the virus population. One of these mutations was silent, while the other would cause a significant change in the
deduced a. a. sequence from a glycine (shown with the vertical arrow in the alignment in Fig. 2.1C) to an aspartic acid.

Table 2.1 Percentage of identity in the deduced amino acid sequence of various protein domains among ToRSV isolates.

<table>
<thead>
<tr>
<th>Isolate pair</th>
<th>CP\textsuperscript{1}</th>
<th>VPg-Pro-Pol\textsuperscript{1}</th>
<th>X4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Including gaps</td>
<td>Excluding gaps</td>
<td></td>
</tr>
<tr>
<td>Rasp1-Rasp2</td>
<td>98.4</td>
<td>99.9</td>
<td>75.1</td>
</tr>
<tr>
<td>Rasp1-PYB1</td>
<td>97.0</td>
<td>98.0</td>
<td>80.7</td>
</tr>
<tr>
<td>Rasp2-PYB1</td>
<td>97.9</td>
<td>98.1</td>
<td>86.6</td>
</tr>
</tbody>
</table>

\textsuperscript{1} Values for the CP and VPg-Pro-Pol protein domains were taken from (Wang and Sanfacon, 2000).

Two series of tandem repeats were identified in the deduced a. a. sequence from the Rasp1-X4 protein. The first type of repeat (repeats A1 to A4' in Fig. 2.1C) is a 53 a. a. motif that was previously identified in the deduced a. a. sequence of the ToRSV-Rasp2 RNA2-encoded polyprotein (Rott et al., 1991). However, the number of copies of this repeat varied with the particular ToRSV isolate. While three copies of the repeat (one of which is an imperfect and truncated copy) are present in the Rasp2 and PYB1 sequences, the Rasp1 sequence contains an additional perfect copy of the repeat. The second type of repeat (repeat B1 to B3 in Fig. 2.1C) is 22 a. a. long and was not identified previously. Three copies of the repeat are found in the Rasp1 sequence (two perfect copies and one imperfect copy) while Rasp2 and PYB1 contain a single copy of the repeat. The repeats may represent novel classes of a. a. repeats as they did not share obvious sequence homology with known classes of repeats. Repeats are found in proteins with diverse functions and are often involved in protein-protein interactions or in ligand binding (Andrade et al., 2001; Grove et al., 2008; Main et al., 2003). The role of the a. a. repeats in the ToRSV X4 protein biological function requires further investigation.
Fig. 2.1 Comparison of the X4 protein domain in the ToRSV Rasp1 and Rasp2 isolates

(A) Schematic representation of the X4 protein domain of ToRSV-Rasp1, Rasp2 and PYB1 isolates. The RNA2-encoded polyprotein from nepoviruses of subgroups A and B and of ToRSV is shown at the top of the figure with the vertical lines representing the proteinase cleavage sites. The lower portion of the figure represents the X4 protein domain of Rasp1, Rasp2 and PYB1. The series of repeats are shown with arrows. The approximate position of two primers used for RT-PCR is shown above the schematic representation of the ToRSV RNA2-encoded polyprotein. (B) PCR amplification of a region of the RNA2 open reading frame coding for a portion of X3, the entire X4 and a portion of MP. PCR products were amplified by direct PCR using available cDNA clones for Rasp2, or by RT-PCR using total RNA purified from cucumber or N. benthamiana plants infected with ToRSV Rasp1 or PYB1. Primer pair R207/F208 was used for the amplification. (C) Alignment of the deduced amino acid sequence of the X4 protein from Rasp1, Rasp2 and PYB1 isolates. The sequence of Rasp2 was previously published (Rott et al., 1991). The previously identified X3-X4 and X4-MP cleavage sites (Carrier et al., 1999; Carrier et al., 2001) are shown with the boxes. Amino acid sequence repeats are indicated with the horizontal arrows above the sequence. Repeats A1-A3 and repeats B1-B2 are perfect copies of the first and second sets of repeats, respectively. Repeat A4’ is an imperfect and truncated copy of the first set of repeats and repeat B3 is a full-length imperfect copy of the second set of repeats. The site of sequence heterogeneity in the virus population that would result in an amino acid change (from glycine to aspartic acid) is underlined in the Rasp1 sequence and shown by the vertical arrow.
### A

**Subgroups A and B**

<table>
<thead>
<tr>
<th></th>
<th>Subgroup A</th>
<th>Subgroup B</th>
</tr>
</thead>
<tbody>
<tr>
<td>F208</td>
<td>Rasp2</td>
<td>Rasp1</td>
</tr>
<tr>
<td>R207</td>
<td>Rasp1</td>
<td>PYB</td>
</tr>
</tbody>
</table>

**ToRSV (subgroup C)**

<table>
<thead>
<tr>
<th></th>
<th>Raspl</th>
<th>Rasp2</th>
<th>PYB</th>
</tr>
</thead>
<tbody>
<tr>
<td>X3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Subgroups A and B**

**ToRSV (subgroup C)**

<table>
<thead>
<tr>
<th></th>
<th>Raspl</th>
<th>Rasp2</th>
<th>PYB</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ToRSV (subgroup C)**

<table>
<thead>
<tr>
<th></th>
<th>Raspl</th>
<th>Rasp2</th>
<th>PYB</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### B

**Mq (kb)**

- Raspl: 3.0
- Rasp2: 2.0
- PYB: 1.5

### C

<table>
<thead>
<tr>
<th></th>
<th>Raspl</th>
<th>Rasp2</th>
<th>PYB</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ToRSV (subgroup C)**

<table>
<thead>
<tr>
<th></th>
<th>Raspl</th>
<th>Rasp2</th>
<th>PYB</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.4 Discussion

The high variability of the X4 protein sequence suggests that it is not involved in a core process, such as virus replication. Serial passage of ToRSV isolates by mechanical inoculation would eliminate selection pressure for the maintenance of sequences important for transmission of the virus by the *Xiphinema americanum* nematode, the natural vector for ToRSV (Sanfacon, 2008; Sanfacon et al., 2006). Although the CP has been suggested to be the sole determinant for the specificity of nematode transmission in a subgroup A nepovirus (Andret-Link et al., 2004), it is possible that X4 plays a role in the specificity and/or efficiency of nematode transmission in the case of ToRSV. The X4 protein could also be involved in host specificity, symptomatology and/or interaction with host defence responses. Herbaceous plants infected with nepoviruses often recover from an initial symptomatic infection. Recovery from tomato black ring nepovirus (TBRV, a subgroup B nepovirus) has been associated with the induction of RNA silencing and a concurrent dramatic reduction in virus titer in the recovered leaves (Ratcliff et al., 1997). We have recently shown that ToRSV (Rasp1 isolate) accumulates to high titre in recovered leaves, in spite of active RNA silencing directed at ToRSV sequences (Jovel et al., 2007). One possible interpretation for these results is that ToRSV encodes a suppressor of silencing, thereby allowing the virus to accumulate in recovered leaves. The X4 protein, which is absent from the TBRV genome is an attractive candidate for the silencing suppression function. The availability of two closely related ToRSV isolates with diverging X4 sequences and differences in symptomatology will provide us with an invaluable tool to address this question.
2.5 Bibliography


Chapter 3

Characterization of the X4 protein in ToRSV-infected plants

3.1 Introduction

Most plant viruses, encapsidate their small genome in a non-enveloped virion that has an icosahedral or helical structure. Nepovirus particles are made of 60 copies of the coat protein (CP) and have a pseudo T=3 structure. The capsid protein is made up of three $\beta$-barrel domains. The observation that transgenic plants expressing the CP from Arabis mosaic virus form empty virus particles, similar to those observed in natural infection, suggested that other viral proteins are not required for virion assembly (Bertioli et al., 1991). The CP of ToRSV is a polypeptide of 562 amino acids and shows the highest similarity to Blueberry leaf mottle virus (BLMV) and Blackcurrant reversion virus (BRV), two members of the nepovirus subgroup C. This subgroup is characterized by a larger RNA2 component (Rochon and Sanfacon, 2001).

Preparations of nepovirus particles purified from indicator plants (C. quinoa or N. benthamiana) are composed of two forms of the unique coat protein (Latvala et al., 1998; Lemmetty et al., 1997). Similarly, two forms of the small comovirus CP are also often detected (Taylor et al., 1999). The structure of tobacco ringspot virus (TRSV, a subgroup A nepovirus) and BRV have been resolved. Most preparations of BRV particles purified from indicator plants (C. quinoa or N. benthamiana) are composed of two forms of the single coat protein of 54 and 55 kDa in size with identical N-termini (Latvala et al., 1998; Lemmetty et al., 1997). Using antibodies specific for various regions of the CP, it was determined that the smaller form of the coat protein is a truncated version of the CP, which has lost a C-terminal extension of 19 amino acids. This truncation may be the result of processing occurring during the course of virus infection or during the various steps of virus purification. As for BRV, two forms of the coat protein have also been detected in tomato black ring virus (TBRV, a subgroup B nepovirus). Purified BRV virion containing only the smaller form of the CP (54 kDa) can be used to infect indicator plants (C. quinoa) by mechanical inoculation resulting in symptoms identical to BRV symptoms (Lemmetty et al., 1997). However, the resulting progeny viruses contain both protein forms. Thus, the C-terminus is not important for the infectivity of the virus. It was proposed that the C-terminal extension may serve as a determinant for mite transmission however this needs to be confirmed experimentally. Sequence alignments of the CP from various nepoviruses, revealed that the C-terminal extension is very variable in sequence and varies in length from one nepovirus to another (Seitsonen et al., 2008). As mentioned above the BRV coat protein has a C-terminal
extension of 19 amino acids. In contrast the tobacco ringspot virus coat protein does not include an apparent C-terminal extension. In the case of tomato ringspot virus (ToRSV), a possible C-terminal extension is 54 amino acids long (Seitsonen et al., 2008). Based on homology modeling of the BRV and TRSV structures and also based on sequence alignment of the coat proteins, it has been predicted that the C-terminal 14 residues of BRV project out from the surface of the BRV virus. Since it is extended from the virion surface it is also a potential specific antigenic site (Seitsonen et al., 2008).

Although the main component of the virion is the coat protein(s) (CP), other minor virus-encoded protein components have also been found in association with plant virus particles. These virion-associated proteins (VAP) interact directly or indirectly with the CP and are often loosely associated with the virus particle. Because they are exposed on the surface of the virion, VAP are readily available for interaction with host or vector components. They may play important roles in the virus replication cycle, including cell-to-cell or long distance movement in the plant, suppression of RNA silencing, stability of virus particles, regulation of translation of the viral genome and vector transmission. Several VAP have been identified in association with helical viruses. The potyvirus helper component proteinase (HC-Pro) is a multifunctional protein that was first recognized as an essential factor for aphid transmission, but was later found to act as a proteinase, an enhancer of genome amplification and of cell-to-cell and systemic movement of the virus in the plant, a determinant of symptom formation and a suppressor of silencing (reviewed in Maia et al., 1996; Roth et al., 2004; Urcuqui-Inchima et al., 2001).

The Cauliflower mosaic virus (CaMV) pIII protein is a well studied VAP of a spherical virus. This protein is indispensable for aphid transmission and cell-to-cell movement but is not required for virus encapsidation or replication (Kobayashi et al., 2002; Leh et al., 1999). The pIII protein interacts directly with the CP through its C-terminal domain (Leclerc et al., 2001; Leh et al., 2001). The pIII protein co-purifies with virus particles under certain conditions, although it can easily dissociate from the purified virions (Dautel et al., 1994). Elucidation of the structure of the reconstituted pIII-virion complex showed that pIII is exposed on the surface of the virion with its C-terminus anchored in the inner shell of the virion (Plisson et al., 2005). The N-terminal region of the protein includes a coiled-coiled domain, which may be essential for dimerization and for interaction with other viral proteins. One of these
interactions is with the pII protein. An aphid transmission factor that itself interacts with the aphid stylet (Leh et al., 1999; Uzest et al., 2007). Another interaction is between the coiled-coil domain of pIII and a coiled-coil domain of the movement protein (MP) of CaMV (Stavolone et al., 2005). The viral MP is a structural component of the tubular structures that contain virus-like particles, traverse the cell wall and probably enable cell-to-cell movement of the virus (Perbal et al., 1993 and references therein).

As mentioned above, members of subgroup C have a larger RNA2 than members of subgroup A and B. Analysis of the polyprotein processing of ToRSV, a subgroup C nepovirus, revealed the presence of three proteinase cleavage sites allowing the definition of four protein domains, termed X3, X4, MP and CP (Carrier et al., 2001). In contrast, only three protein domains are present in the polyprotein of subgroups A and B nepoviruses termed 2a, MP (or 2b) and CP (or 2c), see Fig. 3.1A. The ToRSV X3 protein has some conserved sequence motifs with the 2a protein of nepoviruses of subgroup A and B (Rott et al., 1991). The grapevine fanleaf virus (GFLV, a subgroup A nepovirus) 2a protein was suggested to play a role in the replication of RNA2 (Gaire et al., 1999). The MP of ToRSV and GFLV is a structural component of tubular structures that are very similar to those observed in CaMV and cowpea mosaic virus (CPMV, a comovirus) infected plant tissues in that they contain virus-like particles and are found traversing the cell wall (Ritzenthaler et al., 1995; Wieczorek and Sanfacon, 1993 and references therein). In grapevine fanleaf virus (GFLV) analysis of chimeric viruses in which the MP and CP domains were exchanged suggested that the CP and MP interact with each other, either directly or indirectly to promote cell-to-cell movement (Belin et al., 1999). In addition, similar analyses of chimeric viruses suggested that the CP is the sole determinant for the specificity of transmission of GFLV by its nematode vector (Andret-Link, 2004).

3.1.1 Research objective and hypothesis

The ToRSV X4 protein does not have significant sequence identity with any other proteins available in the database and does not have a functional equivalent in the genome of nepoviruses of subgroup A and B (Fig. 3.1A). In this chapter, I will describe the detection of a 60 kDa protein in ToRSV-Rasp1 infected plants using a polyclonal antibody X4(64-65), raised against a recombinant protein corresponding to the C-terminal portion of the X4 protein. Subsequent analysis revealed that this protein is associated with empty and full virus
particles. The initial hypothesis was that this protein (60 kDa) may be a virion-associated protein corresponding to a truncated form of X4. I conducted a series of experiments to investigate the nature of the 60 kDa protein recognized by the X4(64-65) antibodies and its possible interaction with the virus particle. The results indicated that the epitope recognized by the X4(64-65) antibodies is exposed at the surface of the viral particle. It was later found that the X4(64-65) antibodies cross reacted with the CP. Therefore the 60 kDa protein detected by this antibody was the full length ToRSV CP. Two additional antibodies raised against the X4 protein did not detect this 60 kDa protein confirming the suggestion that the 60 kDa protein is not derived from the X4 coding region. A protein corresponding to the full length X4 (82 kDa in Rasp1) was not detected in plant extracts by any of the X4 antibodies suggesting that X4 may be an unstable protein or that the concentration of the protein was lower than the detection level of the antibodies. Although smaller (e.g. 45 kDa) or larger ~120 kDa proteins were detected in infected plant extracts by the X4 antibodies, the nature of these proteins needs to be further investigated.

### 3.2 Materials and methods

#### 3.2.1 Production of antibodies against the C-terminal and the N-terminal region of X4

When this part of my research was initiated, the sequence of Rasp1 and PYB isolates were not available to our laboratory. The sequences of CP, VPg, Pro and Pol region were highly conserved among closely related ToRSV isolates. Therefore, recombinant proteins or peptides for antibody production were initially designed based on the available Rasp2 sequence. Later after sequencing the X4 region from a few isolates, we observed a high degree of sequence diversity in the X4 protein (Chapter 2).

The C-terminal portion of the X4-Rasp2 coding region was amplified by PCR from the cDNA clone of ToRSV RNA2 (Rott et al., 1991) using primers 64 (5’-AACGCCATGGGGCTTGTGACTTACCTTCAG-3’, corresponding to nts 2440-2459 of ToRSV-Rasp2 RNA2) and 65 (5’-AACCCTCGAGCTGTACTGGGGCAGACG-3’, complementary to nts 2798-2782 of ToRSV-Rasp2 RNA2). The amplified fragment was cleaved by NeoI and XhoI and inserted into the corresponding site of the pET21d (Novagen)
expression vector to create plasmid pET-X4-C (Rasp2). This plasmid was transformed into *E. coli* BL21 (DE3). Bacteria were grown until the OD600 of the culture reached approximately 0.6. Protein expression was induced by adding 1 mM IPTG for 2 h at 37°C. Cells were harvested and purification of the recombinant X4-C His-tagged 15 kDa X4 protein was carried out under denaturing conditions (8 M urea) using a Ni+2 resin column as recommended by the manufacturer (Qiagen). After purification, the urea was eliminated by gradual dialysis against sodium phosphate buffer [50 mM NaPO4 (pH 7.5), 0.3 M NaCl] containing decreasing concentration of urea. Rabbits were injected with the purified protein for antibody production.

Another antibody was raised against the N-terminal end of the X4 protein using a synthetic peptide KSKSQNWFRSMGIAHDYVEG corresponding to the coding region of Rasp2 sequence (nts 1119-1182 of RNA2). The antibody was produced commercially (Bioworld Molecular Tools and Laboratory Essentials, Dublin, OH, USA).

### 3.2.2 Construction of plasmids used in the expression of full-length or truncated forms of X4, CP and MP in *in vitro* translations or in *E. coli*

To express proteins *in vitro* I used coupled *in vitro* transcription/translation system (Invitrogen). Plasmid pCITE-X4-Rasp2 and pCITE-CP-Rasp2 were constructed as follows. A cDNA fragment containing the entire X4 coding region from the Rasp2 isolate was amplified using primers FX4-Rasp2 [5’-GATTCCATGGGCTTCTCCCTCCA-3’, corresponding to nts 981-999 of ToRSV-Rasp2 RNA2 and containing an *Nco*I site (underlined)] and RX4-Rasp2 [5’-AATCCTCGAGTCAGTTAGAGCG-3’, complementary to nts 2865-2879 of ToRSV-Rasp2 RNA2 and containing an *Xho*I site (underlined)]. The amplified cDNA fragment was digested by *Nco*I and *Xho*I and inserted into the corresponding sites of plasmid pCITE-4a (+) (Novagen) to create pCITE-X4-Rasp2. Plasmid pCITE-CP-Rasp2 was constructed in a similar manner using primers FCP [5’-GATTCCATGGGCGGGTCCTGCAAGAAG-3’, corresponding to nts 4338-4356 of ToRSV-Rasp2 RNA2 and containing an *Nco*I site (underlined)] and RCP-2 [5’-TATCCTCGAGTCAGCCACCGGAAAGG-3’, complementary to nts 5708-5723 of ToRSV-Rasp2 RNA2 and containing an *Xho*I site (underlined)]. Plasmids pCITE-CP-Rasp2 and pCITE-X4-Rasp2 include a stop codon at the end of the viral sequence preventing fusion to the His tail present in the plasmid. However, the viral sequences would be fused in frame.
with an N-terminal S-tag present in the vector. Plasmid pCITE-X4-Rasp2-HA and plasmid pCITE-X4-Rasp1-HA were constructed as follows: to create plasmid pCITE-X4-Rasp2-HA, a cDNA fragment containing the entire X4 coding region from the Rasp2 isolate fused to the HA tag was amplified using primers F-X4-Rasp2 [5’-ATATCCATGGGGGCTTCTCCCTCCCT-3’] and R6-X4-Rasp2 [5’-TAGAAAGATCTTACCGTAATCAGGAAACGTCGTACGGATACTGGCAGTTAGAGCG AGTAAC-3’], complementary to nts 2859-2879 of ToRSV-Rasp2 and containing the HA tag and a BglII site (underlined). Similarly, a cDNA fragment containing the entire X4 coding region from the Rasp1 isolate fused to the HA tag was amplified using primers F1-X4-Rasp1 [5’-ATATCCATGGGGTTTTTTCCTGGC-3’, corresponding to the 5’ end of the X4-Rasp1 coding region and containing an NcoI site (underlined)] and R6-X4-Rasp1 [5’-TAGAAAGATCTTACCGTAATCAGGAAACGTCGTACGGATACTGGCAGTTAGAAGC AGTGAC-3’] complementary to the 3’ end of the X4-Rasp1 coding region and containing the HA tag and a BglII site (underlined) to create plasmid pCITE-X4-Rasp1-HA. The fragments were digested by NcoI and BglII and inserted in the NcoI-BamHI sites of plasmid pCITE-4a(+). 

To construct plasmid pCITE-MP, a cDNA fragment containing the entire coding region for the full length movement protein (MP) was amplified by PCR from the cDNA clone of ToRSV RNA2 (Rott et al., 1991) using primers FMP [5’-GATTCCATGGTCTCTTCTAGGAAC-3’, corresponding to nts 2880-2893 of ToRSV-Rasp2 (RNA2)] and RMP [5’-GATTCTCGAGCTGAACAGAAGAATTC-3’, complementary to nts 4021-4037 of ToRSV-Rasp2 RNA2]. The amplified fragment was cleaved by NcoI and XhoI and inserted into the corresponding site of plasmid pCITE-4a(+) (Novagen) vectors and the pET21d (Escherichia coli expression vectors).

The plasmids were used as a template for coupled in vitro transcription/translation reactions using the TNT® Rabbit Reticulocyte lysate system following the manufacturer's protocol (Promega). The labelled [35S] methionine translation products were separated by SDS-PAGE. Immunoprecipitation of the in vitro translation products were conducted as previously described (Hans and Sanfacon, 1995).

To express proteins in E. coli, the following plasmids were constructed. Plasmid pET-CP-His and pET-X4-TM-His (a truncated form of X4 lacking the N-terminal region including a putative transmembrane domain) were constructed as follows. For pET-X4-TM-His plasmid: a PCR fragment was amplified from the cDNA clone of ToRSV RNA2 (Rott et al., 1991)
using primers FX4-TM1502 [5´- CATTCCATGGAAATGCCGTCTTGCGCAACC -3´, corresponding to nts 1502-1523 of ToRSV-Rasp2 RNA2] and RX4-TM [5´- CGGCTCGAGCTGGCAGTTAGAGCGAGTAAC -3´, complementary to nts 2858-2879 of ToRSV-Rasp2 RNA2]. For pET-CP-His, a PCR fragment was amplified from the cDNA clone of ToRSV RNA2 (Rott et al., 1991) using primers FCP-pET4038 [5´- CATGCCATGGGCGGTCTGGCAAGAAGGT -3´, corresponding to nt 4038-4058 of ToRSV-Rasp2 RNA2] and RCP-pET [5´- GCCCTCGAGCCACGCCCGAAAGGATTT -3´, complementary to nts 5705-5723 of ToRSV-Rasp2 RNA2]. Plasmid pET-CP-150-His (predicted C-terminal extension in the CP, nts 5571-5723 of ToRSV RNA2) and pET-CP-360-His (predicted exposed loop in the CP, nts 5198-5558 of ToRSV RNA2) were constructed using the following primers: for pET-CP-360-His: primers F145 [5´- CGAACCATGGCTAGGCAACATCCTTTATCTCTG -3´, corresponding to nts 5198-5219 of ToRSV-Rasp2 RNA2] and RCP [5´- GAACTCGAGACTAGGTCCATAGAAAGAAAATC -3´, complementary to nts 5558-5535 of ToRSV-Rasp2 RNA2] and for pET-CP-150-His: primers F147 [5´- CGAACCATGGCTAAGAAAGAAGTCGGCACGCCTAG -3´, corresponding to nts 5571-5593 of ToRSV-Rasp2 RNA2] and R148 [5´- CGAACCTCGAGCCACGCCCGAAAGGATTTTTCC -3´, complementary to nts 5723-5699 of ToRSV-Rasp2 RNA2]. The amplified fragments were cleaved by NcoI and XhoI enzymes and inserted into the corresponding site of the pET24d (Novagen) expression vector. In all these plasmids, the coding regions for the ToRSV proteins were fused in frame with a poly-histidine tail. We also produced a version of the CP in which a stop codon was introduced at the end of the CP coding region. As a result, the CP was not fused to a poly-histidine tail. This plasmid, termed pET-CP, was constructed by transferring the Nco-XhoI fragment containing the CP coding region with a stop codon from pCITE-CP into the corresponding site of the pET24d vector. Plasmids were transformed into E. coli BL21 (DE3). Bacteria were grown until the OD600 of the culture reached approximately 0.6. Protein expression was induced by adding 1 mM IPTG for 2 h at 37°C. Purification of the recombinant X4-TM His-tagged ~53 kDa protein (Rasp2) was carried out under denaturing conditions as recommended by the manufacturer (Qiagen).
3.2.3 Preparation of plant extracts and subcellular fractionations

ToRSV-Rasp1 was propagated routinely in *Cucumis sativus* var. Straight Eight. Plants were inoculated at leaf stage 5 to 7 using 40 μl of fresh leaf extracts of ToRSV-infected cucumber plants diluted 1:10 (w:v) in phosphate buffered saline (PBS). To produce post-nuclear plant extracts, one g of leaf tissue from healthy and ToRSV-infected *C. sativus* was collected 7 days post-inoculation (dpi) and briefly ground in 4 ml of homogenization buffer (Schaad *et al.*, 1997). After centrifugation at 3,700xg for 10 min, the supernatant was collected (S3 fraction). For subsequent subcellular fractionation, the S3 fraction was overlaid on a 20 to 45 % sucrose gradient and centrifuged at 34,000xg for 4 h as described previously (Han and Sanfacon, 2003). Protein content in the S3 fraction or in the sucrose gradient fractions were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and analyzed by immunoblotting.

3.2.4 Purification of ToRSV particles

Virus particles were purified from ToRSV-Rasp1-infected tissues essentially as described (Stace-Smith, 1966). Briefly, harvested symptomatic leaves from approximately 200 ToRSV-infected cucumber plants were ground in 0.5 M borate buffer (pH 6.8). After freezing overnight at -20 C, the extract was centrifuged at 10,000 x g for 30 min to eliminate large debris. The supernatant was further clarified using an ammonium sulphate precipitation step (15 g ammonium sulphate per 100 ml of extract). After centrifugation at 10,000 x g for 30 min, the supernatant was collected. Virus particles were pelleted from this clarified extract by centrifugation at 140,000 x g for 2 h. The pellets were resuspended in 9 mls of 50 mM borate buffer (pH 6.8) and CsCl was added to a density of 0.68 g/ml. The CsCl gradients were formed by an overnight centrifugation at 190,500 x g. Virus particles were extracted from the gradient and dialysed overnight against 20 mM borate buffer (pH 6.8). The purified virus preparation was diluted (1:100 phosphate buffer saline pH 7.4) and the protein content was analysed by SDS-PAGE. Proteins were visualized by silver staining (Invitrogen) or immunoblotting.

3.2.5 Immunoblotting

Proteins were separated by SDS-PAGE and transferred to polyvinylidene fluoride (PVDF) membranes (Bio-Rad). Membranes were incubated with the rabbit polyclonal antibodies
raised against the C-terminal region of X4 (X4(64-65), used at 1:5000 dilution, this study),
denatured purified viral particles (α-virion, used at 1:5000 dilution) (Sanfacon et al., 1995),
the central region of the movement protein (α-MP antibodies, used at 1:5000 dilution)
(Wieczorek and Sanfacon, 1993) or the proteinase (α-pro antibodies, used at 1:5000 dilution)
(Chisholm et al., 2007). Rabbit polyclonal antibodies raised against the Bip protein (donated
by M. Chrispeels) were also used. The secondary antibody was goat anti-rabbit conjugated
with horseradish peroxidase (1:15000 dilution, Bio/Can). Membranes were developed using
an ECL-development kit (Amersham Bioscience).

3.2.6 Immuno-electron microscopy

Highly purified ToRSV-Rasp1 particles (2 μg/ml) were diluted 1:10 in phosphate buffer
saline (buffer supplemented with 0.1% Tween-40, PBST buffer). Formvar/carbon coated
grids were allowed to float on 10 μl of purified virions for 2 min. Grids were incubated with
blocking buffer (PBS, 1% BSA, 0.1% Tween-20) for 10 min and transferred to 10 μl of
X4(64-65) antiserum or pre-immune antiserum (diluted 1:100 in PBST) for 1 hour. Grids
were rinsed once in PBST buffer, and 5 times in distilled water and blocked again for 30 min
in blocking buffer. The samples were washed (as described above) and incubated with anti-
rabbit secondary antibody conjugated with 10 nm gold particles (1:30 in blocking buffer,
Sigma) for 1h. Grids were incubated in blocking buffer for 30 min, rinsed and stained with 2
% uranyl acetate. Samples were viewed at 27,000 to 80,000 X magnification using a JEOL
100CX transmission electron microscope (TEM) operated at 80 kV.

For immunocapture experiment, formvar/carbon coated grids were allowed to float on α-
virion antibody (diluted 1:30 in PBST) for 1 hour. After the incubation, grids were rinsed
once in PBST buffer and 5 times in distilled water. Grids were incubated for 30 min in
blocking buffer. The samples were washed (as described above) and incubated with anti-
rabbit secondary antibody conjugated with 10 nm gold particles (1:30 in blocking buffer,
Sigma) for 1h. Grids were incubated in blocking buffer for 30 min, rinsed and stained with 2
% uranyl acetate. Samples were viewed at 27,000 to 80,000 X magnification using a JEOL
100CX transmission electron microscope (TEM) operated at 80 kV.

3.2.7 Trypsin digestion of ToRSV particles

TPCK-Trypsin (Sigma) stock solution was prepared at a concentration of 2 mg/ml in 0.001 N
HCl, and serial dilutions were made in phosphate buffer 67 mM (NaH₂PO₄, pH 7.6).
Purified ToRSV particles (800 ng/μl) were incubated with various dilutions of Trypsin (0, 0.6, 3, 16, 80, 400 ng/μl) in a final volume of 50 μl reaction for two hours at 37°C. The digestion reactions were stopped by adding 2X SDS-PAGE protein loading buffer. The protein content of each sample was examined by immunoblotting after SDS-PAGE.

3.2.8 Liquid chromatography-mass spectrometry with peptide mass fingerprinting (LC-MS/MS)

ToRSV was purified as discussed above (Stace-Smith, 1966; Stace-Smith, 1984) and the proteins were separated by 12% sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) followed by staining with mass spectrometry-compatible silver staining polyacrylamide gel (SilverQuest, Invitrogen). Two fragments were carefully cut: fragment 1, which contained doublet protein bands of 53, 55 kDa and fragment 2 which contained a 60 kDa protein band. The in-gel samples were sent for mass spectrometry. In a separate experiment a few μg of purified virus prep were sent for mass spectrometry. The samples were digested with trypsin. The extracted peptides were analyzed by LC-MS/MS using an Applied Biosystems QStar1 mass spectrometer, which is a quadrupole time-of-flight hybrid instrument equipped with nano-electrospray ionization source. The sample was first separated by reverse phase chromatography over a two hour gradient before spraying into the mass spectrometer. The MS/MS data were analyzed using a protein identification search engine algorithm called: MASCOT (http://www.matrixscience.com). The data was searched against the heterologous non-identical MSDB database, which is a protein sequence database designed specifically for mass spectrometry applications.

3.3 Results

3.3.1 Preparation of antibodies against the X4 protein

To characterize the ToRSV-Rasp1 X4 protein in infected plants, we initially raised rabbit polyclonal antibodies against a His-tagged recombinant protein that correspond to the C-terminal 15 kDa X4 protein of the ToRSV-Rasp2 (Fig. 3.1A,α-X4(64-65)). This region of the protein was later found to be relatively well conserved between the Rasp1 and Rasp2 isolates (Fig.3.11). Immunoblotting experiments confirmed that the X4(64-65) antibodies recognized the X4-C recombinant protein (Fig. 3.1B).
**Fig. 3.1 Genome organisation of ToRSV and GFLV and production of the X4(64-65) antibodies.** (A) Genome organization of ToRSV protein domains (subgroup C) and GFLV (subgroup A) nepovirus. GFLV lacks the X4 protein. The lower portion of the figure represents the X4 protein domain of Rasp2 isolate. The blue box at the C-terminal end of the protein represents the region of the ToRSV-Rasp2 X4 protein, which was used to produce polyclonal antibodies. The yellow box at the C termini X4, represents the protein that was expressed for X4 (64-65) antibody production. (B) Specificity of the X4(64-65) antibody tested by immunobloting. Immunoblot analysis was conducted using the purified recombinant X4-C protein and the X4 (64-65) antibodies. Molecular marker (Mr).

As I discussed in Chapter 2, the predicted size for the X4 protein differs among ToRSV isolates *in vivo*. The X4 proteins from the Rasp1 and Rasp2 isolates were expressed *in vitro* using coupled *in vitro* transcription/translation system. The apparent molecular mass of the proteins on SDS-PAGE was consistent with their calculated sizes. As expected the X4-Rasp1-HA protein (calculated molecular mass of 87.4 kDa, taking into account a 4 kDa N-terminal S-tag and a 1 kDa C-terminal HA-tag) migrated significantly slower than the X4-Rasp2 protein (calculated molecular mass of 76 kDa, with the N-terminal S-tag and C-terminal HA-tag) (data not shown). The untagged X4-Rasp2 (calculated molecular mass 75
kDa, including the N-terminal S-tag) and coat protein (calculated molecular mass of 66 kDa for the CP including the N-terminal S-tag) were also synthesized in vitro.

To test the specificity of the antibodies, immunoprecipitation was conducted using in vitro translated X4-Rasp2 and CP with different sets of antibodies specific for X4, CP and MP. The X4(64-65) antibody immunoprecipitated the X4 protein but did not immunoprecipitate the CP protein (Fig. 3.2A lanes 3 and 5), showing its specificity for the X4 protein in the immunoprecipitation assay. In addition X4(64-65) antibodies immunoprecipitated the X4-Rasp1–HA and the X4-Rasp2-HA translated proteins.

I also tested the previously described α-virion antibodies, which were raised against a denatured preparation of purified virus particles (Sanfacon et al., 1995). As expected, the α-virion antibodies recognized the CP (Fig. 3.2A lane 4). The α-virion antibodies also recognized the X4 protein (Fig. 3.2A lane 6), suggesting that the X4 protein may have been present in the purified virus preparation used as an antigen or that the X4 protein and the CP may share a common epitope. Two other α-virion antibodies which were used in enzyme-linked immunosorbent assay (polyclonal antibody for ELISA test, Agdia) were tested against the CP. These antibodies were produced against native virus particles. The antibodies did not immunoprecipitated in vitro synthesized CP (Fig 3.2B, lane 3 and 4), suggesting that they may recognize conformational epitope presented in intact virions but not in the CP subunit.
**Fig. 3.2 Specificity of the X4(64-65) and α-virion antibodies tested by immunoprecipitation assays.** (A) Immunoprecipitation analyses were conducted using in vitro translated CP and X4 with the X4(64-65) and the α-virion antibodies as indicated above each lane. The starting material (X4-Rasp2 and CP produced by in vitro translation) is shown in lanes 1-2. Immunoprecipitation of the X4-Rasp2 and CP-Rasp2 proteins with each antibody is shown in lanes 3-6 as indicated above each lane. (B) Specificity of various antibodies that was raised against ToRSV virion (α-virion) or the X4-C protein (the X4(64-65) antibodies), in immunoprecipitation assay using in vitro translated CP. CP was translated in vitro and immunoprecipitated with the antibodies. The antibodies used in each lane are as follows: X4(64-65) (lane1), α-virion (was made against denatured virus prep, lane 2), α-virion (ELISA antibodies, lane3), α-virion (monoclonal antibody, lane 4), and no antibodies (lane 5).

### 3.3.2 Detection of a 60 kDa protein by the X4(64-65) antibodies in ToRSV-infected plant extracts

Post-nuclear extracts derived from ToRSV-Rasp1 infected or mock-inoculated cucumber and *N. benthamiana* plants were tested by immunoblotting using the X4(64-65) antibodies. A
predominant protein with an apparent molecular mass of approximately 60 kDa was detected by the X4(64-65) antibodies in extracts derived from ToRSV-Rasp1 infected plants but not in healthy plant extracts (Fig. 3.3A). The protein migrated faster on SDS-PAGE than expected for the full-length X4 protein (calculated molecular mass of 82.6 kDa for X4-Rasp1). I will refer to this protein as the 60 kDa protein detected by the X4(64-65) antibodies for the remainder of this chapter. The X4(64-65) antibodies also detected a faint protein band of ~58-60 kDa in the PYB isolate (the band is not visible in the scanned Fig. 3.3B). Antibodies against the N-terminal region of X4 did not detect any specific band corresponding to the full length X4-Rasp2 (Carrier et al., 2001) or the 60 kDa protein detected by X4(64-65) antibodies (data not shown). All subsequent experiments with ToRSV infected plants were conducted using the Rasp1 isolate and the X4(64-65) antibodies.

Fig. 3.3 Detection of a 60 kDa protein by the X4(64-65) antibodies in infected plant extracts. (A) Immunoblot analysis of healthy (H) or ToRSV-Rasp1 infected (Inf) C. sativus and N. benthamiana plant extracts. Proteins were separated by SDS-PAGE (12%) and detected by immunoblotting using X4(64-65) antibodies. (B) Immunoblot analysis of plant extracts derived from plants infected with various isolates. The blots were probed using α-virion antibodies (denatured) (lanes 1-6) or X4(64-65) antibodies (lanes 7-12). Plants were infected with the following ToRSV isolates: lane 1 Rasp1 (Cucumber),
3.3.3 Co-fractionation of the 60 kDa protein detected by the X4(64-65) antibodies with empty and full virus particles in sucrose gradient fractionation assay

To study the sub-cellular distribution of the 60 kDa protein detected by the X4(64-65) antibodies, post-nuclear extracts derived from T oRSV-infected leaves were separated by sucrose gradient fractionation. Fractions were collected from the bottom of the gradient and were analyzed by immunoblotting using the X4(64-65) antibodies. I also tested for the presence of the movement protein, coat protein and Bip (a protein marker of the endoplasmic reticulum, ER) (Fig. 3.4A). The MP and the ER marker were detected in fractions 1-4 at the bottom of the gradient (Fig. 3.4A), a result consistent with previous observations (Han and Sanfacon, 2003). The 60 kDa protein detected by the X4(64-65) antibodies co-fractionated with the proteins recognized by the α-virion antibody in fractions 5-8 and 11-13 (Fig. 3.4A panel 1 and 2). ToRSV produces empty and full virus particles, which have been referred to as top and bottom components to reflect their sedimentation pattern in sucrose or CsCl gradients (Stace-Smith, 1966). The two sucrose gradient peaks, which contained proteins recognized by the α-virion and X4(64-65) antibodies, may correspond to the fractionation of full and empty virus particles. To verify this, I used the α-virion antibodies in immunocapture experiments. Electron microscopy (EM) analysis of the immunocaptured virus samples revealed that fraction 6 contained predominantly full virus particles, while fraction 12 contained a majority of empty virus particles (Fig. 3.4B). The co-fractionation of the 60 kDa protein detected by the X4(64-65) antibodies with the virus particle suggests that this protein is associated with the virion. Larger protein bands were also detected by the X4(64-65) antibodies in sucrose gradient fractionations (Fig.3.4C).
Fig. 3.4 Immunoblot analysis and electron microscopy of subcellular fractionation of proteins by sucrose gradient. (A) Immunoblot analysis of proteins in the sucrose gradient fractionation. Infected plant extracts were prepared and fractionated on a 20-45% sucrose density gradient. The proteins from each fraction were separated by SDS-PAGE (12%) and detected by immunoblotting using the X4(64-65), α-virion, α-MP and α-Bip antibodies. Bip is a resident protein of the ER and is labelled as ER marker in the figure. (B) Electron micrographs of empty and full virus particles from fraction 6 and 12 obtained after sucrose gradient fractionation. (C) Immunoblot analysis of proteins in the sucrose gradient fractionation. In this experiment the proteins were fractionated in 15 fractions, the X4(64-65) antibodies also detected higher protein
mass (80-120 kDa) In the viral fractions and lower protein mass on top of the gradient, the nature of these proteins needs further investigation. The 60 kDa detected by X4(64-65) antibodies in sucrose gradient is shown by the arrows. Molecular marker (Mr).

3.3.4 Co-purification of the 60 kDa protein detected by the X4(64-65) antibodies with the virus particles

To further investigate the possibility that the 60 kDa protein detected by the X4(64-65) antibodies is associated with the virions, I purified virus particles from ToRSV-infected cucumber using differential centrifugation, an ammonium sulphate clarification step and CsCl gradient centrifugation. The protein content of this preparation was separated by SDS-PAGE and analyzed by silver staining and immunoblot analysis (Fig. 3.5A). Silver staining detected a predominant band corresponding to a doublet protein with an estimated molecular mass of 53-55 kDa, as well as minor bands of apparent molecular masses of 60 kDa and 120 kDa (Fig. 3.5 panel 1). The α-virion antibodies recognized all proteins detected by silver staining (Fig. 3.5A panel 2 and Fig. 3.5B). This was expected since a similar preparation was originally used to produce this antibody. The X4(64-65) antibodies detected a 60 kDa protein, which co-migrated with the 60 kDa band detected by silver staining and with the α-virion antibodies (Fig. 3.5A panel 3). These antibodies also detected a protein of approximately ~120 kDa, which may correspond to dimeric forms of the 60 kDa protein. Another possibility is that this protein corresponds to a putative X3-X4 intermediate precursor of the mature X4 protein. The cleavage site between X3 and X4 (Rasp2) is a suboptimal cleavage site as described before (Carrier et al., 2001). The presence of the 60 kDa protein in highly purified virus preparation suggests that it is associated with virions. Other viral proteins (e.g. proteinase, movement protein) were not detected in association with the purified virus particles (Fig. 3.5A, panel 4 and data not shown).
Fig. 3.5 Co-purification of the 60 kDa protein detected by the X4(64-65) antibodies with purified virus particles. (A) Protein content of a highly purified virus preparation was separated by SDS-PAGE (12%) and examined by silver staining (lane 1) or by immunoblotting using X4(64-65), α-virion and α-MP antibodies (lanes 2-4). (B) Examination of the relative proportion of the 60 kDa protein detected by the X4(64-65) antibodies in a sucrose gradient fraction (lanes 1-2), two separate highly purified virus preparation (lanes 3-6) and after extensive storage of a highly purified virus preparation at 4 °C (lanes 7-8). Proteins were separated by SDS-PAGE as above and detected by immunoblotting using α-virion and X4(64-65) antibodies.

The 53 and 55 kDa protein doublet, which was recognized by α-virion antibodies but not by the X4(64-65) antibodies, likely corresponds to the CP. However, the apparent molecular masses of these proteins were smaller than those expected for the full-length CP (calculated molecular mass of 62 kDa), raising the possibility that they may correspond to truncated forms of the CP lacking the C-terminal portion of the protein as previously described for other nepoviruses (Seitsonen et al., 2008 and references therein). It is possible that the 60 kDa protein band detected by both α-virion antibodies and the X4(64-65) antibodies correspond to a mixture of the two proteins each recognized by a single antibody (Fig. 3.5)
A). Alternatively, the 60 kDa protein band could contain a single protein that is recognized by both antibodies.

The relative intensity of the 53, 55 kDa doublet bands detected by the α-virion antibodies and of the 60 kDa band detected by the α-virion and X4(64-65) antibodies varied from one virus preparation to another (Fig. 3.5B). In crude extracts and partially purified virus (sucrose gradient fraction), the 60 kDa band appeared as predominant (Fig. 3.5B, lanes 1-2, and Fig. 3.3B). In highly purified virus preparation, the relative intensity of the 60 kDa band was much lower compared to the 53, 55 kDa bands and varied from one preparation to another (Fig. 3.5B, lanes 3-6). After prolonged storage of purified virus preparation at 4°C, the 60 kDa band was not detectable by either antibody (Fig. 3.5B, lanes 7-8). It should be noted that many intact virus particles were detected in older virus preparations by EM, in spite of the apparent absence of the 60 kDa protein band (data not shown). The possible nature of this protein band was investigated further below.

### 3.3.5 Immunogold labelling of purified virus particles with X4(64-65) antibodies

Immunogold labelling experiments were conducted using the X4(64-65) antibodies and purified virus particles. The X4(64-65) antibodies decorated both empty and full virus particles (Fig. 3.6A panels 1-3). A pre-immune rabbit antibody did not bind to the virions (panel 4). Antibodies specific for the ToRSV movement protein (MP) or proteinase did not label the virions (data not shown) confirming that the MP does not co-purify with the virus particle. This result suggests that the 60 kDa protein detected by the X4(64-65) antibodies is associated with the virion and indicates that the epitope(s) recognized by the X4(64-65) antibodies is exposed on the surface of the virions.
Fig. 3.6 Immunogold labelling of purified virus particles using the X4(64-65) antibodies.
Immunogold labelling of ToRSV particles using X4(64-65) antibodies (panel 1-3). The preimmune serum showed no specific labelling (panel 4).

3.3.6 Susceptibility of the 60 kDa protein recognized by the X4(64-65) antibodies to trypsin digestion

To further test the topology of the 60 kDa protein which is associated with the virions, I conducted limited trypsin digestion experiments using the purified virus preparation. This preparation was digested with various concentrations of trypsin. The partially digested proteins were separated by SDS-PAGE and detected by immunoblotting using α-virion antibodies (Fig. 3.7). Several fragments were apparently protected from the digestion when lower concentrations of trypsin were used (Fig. 3.7, lanes 2-4). These fragments likely correspond to portions of the proteins which are buried within the structure of the virion after digestion. Alternatively, they could also correspond to highly structured regions in the protein. The 60 kDa protein detected by the X4(64-65) antibodies was resistant to trypsin digestion at lower concentration of trypsin (lane 7-8). At higher concentration of trypsin, smaller fragments were not detected with these antibodies (lane 9-12). This result suggests that the portion of the 60 kDa protein, which is recognized by the X4(64-65) antibodies, is exposed at the surface of the virion.
3.3.7 The X4 protein is not detected in association with virions using liquid chromatography-mass spectrometry with peptide mass fingerprinting (LC-MS/MS)

To obtain additional information on the nature of the 60 kDa protein recognized by the X4 antibodies, LC-MS/MS experiments were conducted. In a first experiment, the 60 kDa protein band was cut out from the SDS-PAGE gel and sent for analysis. In a second experiment, a preparation of purified virus was sent for analysis. In both cases, the protein fragments detected by the LC-MS/MS corresponded predominately to the ToRSV coat protein. The Rubisco protein, a high abundance protein in plants, was also detected. The X4 protein was not detected in the samples, suggesting that it was absent from the virus preparation or that its concentration was below the detection threshold of this method. It is possible that the high concentration of the two dominant proteins (CP and Rubisco) competed for the ionization process with low abundance proteins.

3.3.8 Re-examination of the specificity of the X4(64-65) antibodies and the α-virion antibodies

Since X4 was not detected in LC-MS/MS and since the 60 kDa protein found in association with purified virus particles was recognized by both α-virion and α-X4 antibodies in immunoblotting, I decided to re-examine the specificity of the X4(64-65) and α-virion
antibodies. Immunoblotting was conducted using full-length or truncated versions of X4 and CP expressed in E. coli using the pET expression vector. All proteins were fused to a His tail, except for the full-length CP (pET-CP) protein which was not fused to any tags. This is important since the X4-C fusion protein originally used as an antigen for the production of the X4(64-65) antibodies also included a His tail. As expected, the anti-virion antibodies detected the full-length CP. Surprisingly, they did not detect the predicted C-terminal extension in the CP (CP-150) or the predicted surface exposed loop (CP-360) (Fig. 3.8A. panel 1 and 4 also Fig. 3.8 B panel 1, lane 2). The anti-virion antibodies did not detect two truncated forms of the X4 protein (X4-C fusion protein and X4-TM fusion protein, which lacks the N-terminal region of the protein). I was unable to express the full-length X4 protein in E. coli and could not test whether the α-virion antibodies would detect this protein in the immunodetection assay. As expected, the X4(64-65) antibodies detected the X4-C and X4-TM fusion proteins (Fig. 3.1B and Fig. 3.8A panel 5, Fig. 3.8B panel 2 lane 3 and 4). However, the X4(64-65) antibodies also cross-reacted with other E. coli expressed proteins including the untagged full-length CP in immunoblotting assays (Fig. 3.8A, panel 2 and 5 also Fig. 3.8B, panel 2). This result was unexpected and was in contrast with the specificity of the X4(64-65) antibody observed using the immunoprecipitation assay. It is noteworthy that the full-length CP-His expressed in E. coli comigrated with the 60 kDa protein recognized by the α-virion antibodies detected in ToRSV-infected plant extracts (Fig. 3.8B, compare lanes 2 and 5). This result suggests that this 60 kDa protein may correspond to the full-length CP.
Fig. 3.8 Immunoblot analyses of CP, CP-150, CP-360, X4(64-65), X4-TM (Rasp2) proteins.

(A) Panel 1 and 2 (from left to right): Immunoblotting using the α-virion antibodies (lanes 1 and 2) or X4(64-65) antibodies (lanes 3 and 4) and total E. coli extracts in which expression of the full length ToRSV CP was induced (lane 2 and 4) or not (lane 1 and 3). Panel 3, 4 and 5: Immunoblotting using the α-His antibody (panel 1), α-virion antibodies (panel 2) and X4 (64-65) antibodies (panel 3) and total extracts from E. coli induced for the expression of CP-360-His, CP-150-His and X4-C-His (as indicated above each lanes).  

(B) Panel 1: Immunoblotting using the α-virion antibodies and E. coli induced for the expression of proteinase protein (lane 1), CP (lane 2), none induced X4-TM protein (lane3), purified X4-TM protein (lane 4), purified virus particles (lane 5), induced X4-TM protein (lane 6). Panel 2: Immunoblotting using the X4 (64-65) antibodies and E. coli induced for the expression of proteinase protein (lane 1), CP-His (lane 2), X4-TM protein (lane 3), X4-TM purified protein (lane 4), purified virus particles (lane 5). Molecular marker (Mr).
3.3.9 *In vitro* co-immunoprecipitation of X4 and CP proteins with the X4(64-65) antibodies

Because the X4(64-65) antibodies recognized X4 but not CP in the immunoprecipitation assay, we used this assay to investigate possible protein-protein interactions between X4 and CP. Co-immunoprecipitation was conducted using the *in-vitro* translated CP (Rasp2) and X4 (Rasp2) proteins and also the X4(64-65) and α-virion antibodies. As shown before, the X4(64-65) antibody recognized X4 but not CP (Fig. 3.9A, lanes 3 and 5) while the α-virion antibody recognized both proteins (Fig. 3.9A, lane 7). The CP was co-immunoprecipitated with X4 by the X4(64-65) antibodies, suggesting an interaction (direct or indirect) between the CP and X4 proteins (Fig. 3.9A, lanes 8). I also used HA tags to test for possible interactions between X4 and CP. Full-length or truncated version of X4 were tagged with the HA epitope and mixed with an untagged version of the coat protein. Protein complexes were immunoprecipitated using anti-HA conjugated beads. The HA antibodies pulled down the X4-Rasp1-HA and X4-Rasp2-HA proteins (Fig. 3.9B, lanes 8 and 9). However, the amount of CP was similar in the presence or absence of X4-HA (compare lanes 8-9 to lane 7). Therefore there was no clear evidence of a protein–protein interaction between the X4 and the CP in this assay.
B

Fig. 3.9 Immunoprecipitation of the CP and the X4 protein. (A) Panel 1: in vitro translation of X4 full length Rasp2 (X4) and CP Rasp2 (CP). Panel 2: Immunoprecipitation assays. In order to verify the specificity of the antibodies, single in vitro translated (CP and X4) proteins were incubated with the X4 (64-65) or α-virion antibodies as indicated above each lane. X4 (64-65) antibodies immunoprecipitated the X4 protein (lane 3) but not the CP (lane 5). The α-virion antibodies immunoprecipitated both the CP (lane 4) and the X4 proteins (lane 6). In co-immunoprecipitation assays, proteins were mixed together and incubated with a single antibody as indicated above each lane (lanes 7-8). The α-virion antibodies immunoprecipitated both the CP and the X4 proteins (lane 7). X4 (64-65) antibodies immunoprecipitated the CP and the X4 proteins (lane 8). (B) In vitro translated full length CP (lane 1), mixture of full length Rasp1-X4 protein and CP (lane 2), mixture of Rasp2-X4 and CP (lane 3). Rasp1 and Rasp2 plasmids were fused to an HA tag at their C-terminal end. The proteins were incubated with the anti-HA conjugated beads and after centrifugation the unbound top fraction was loaded on the gel. The pellet fraction containing the immunoprecipitated proteins by the HA- beads were shown in lanes 7-9.

3.3.9.1 The MP did not co-immunoprecipitate with the CP in vitro

In a separate experiment I tested the possible interaction of the CP subunits with the MP by co-immunoprecipitation assay. In Cowpea mosaic virus, the MP was shown to bind specifically to intact CPMV virions and to the large CP, but not to the small CP (Carvalho et al., 2003). The entire coding region for the MP (Rasp2, ToRSV) was cloned into an expression vector to allow its expression by coupled in vitro transcription/translation system. The MP was fused to a poly-histidine tail and an S-tag. The CP was also expressed in vitro (Fig. 3.10, lane 1 and 2). The α-MP antibodies immunoprecipitated the in vitro translated MP (Fig. 3.10, lane3), but not the in vitro translated CP (lane 5). The α-virion antibodies
immunoprecipitated the *in vitro* translated CP (lane 6), but not the *in vitro* translated MP (lane 4). No co-immunoprecipitation was observed when the two proteins were mixed together, suggesting that the MP did not interact with the CP subunits in this assay (lane 7).

**Fig. 3.10 In vitro protein-protein interaction assays.** *In vitro* translation of CP Rasp2 (CP) and MP-Rasp2 (MP) (lanes 1 and 2), and immunoprecipitation assays (lanes 3-8). *In vitro* translated proteins were mixed and immunoprecipitated with the α-virion or α-MP antibodies as indicated above each lane. In a co-immunoprecipitation assay, the *in vitro* translated MP and the CP were mixed together and incubated with the α-MP antibodies, co-immunoprecipitation of the CP and the MP was not observed (lane 7). As a negative control, the *in vitro* translated MP was incubated in the absence of antibodies (lane 8).

### 3.3.10. Production of antibodies from different regions of the X4 and the CP

To further understand the nature of the X4, new antibodies from different regions of X4 and the CP were made (Pacific immunology company). Antigenicity computer prediction programs are based on predicting the hydrophobicity of different regions of the protein. Although they are not always reliable, they can be useful to design peptides for antibody prediction. The alignment of the sequences of X4 and CP were used to predict antigenic peptides that are common to all isolates. The suggested peptide sequences for the X4 and CP proteins that recognized all three isolates are as follows (Fig.3.11 and 3.12): for X4 : LQVLPQRVVLPECMDL (central repeat region), and LNLEEMDLSLHEYPEI (C-term region) and for CP, FEWPARLPDILDDKSEV (central
region) and NNPVGRPPENV (C-terminal extension). Immunoblotting were conducted using the X4 (3741-42, C-terminal region) and X4 (3743-44, repeat region) antibodies (Fig.3.11).

**Abs 2: central repeat region**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasp1</td>
<td>GGFSLPARKGVYVAPTIQGVVRAGLRAQKGFLNAVSAGIVAGARILKSKSQNWFRKSMIAHEYEVEGSALSITILECAGPVVQ</td>
</tr>
<tr>
<td>Rasp2</td>
<td>GGFSLPARKGVYVAPTIQGVVRAGLRAQKGFLNAVSAGIVAGARILKSKSQNWFRKSMIAHEYEVEGSALSITILECAGPVVQ</td>
</tr>
<tr>
<td>PYB</td>
<td>GGFSLPARKGVYVAPTIQGVVRAGLRAQKGFLNAVSAGIVAGARILKSKSQNWFRKSMIAHEYEVEGSALSITILECAGPVVQ</td>
</tr>
</tbody>
</table>

**Abs3: C-terminal region**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasp1</td>
<td>QEEESMKVPAVLEAASNIDDVTEAFFDDCLECESFYEDSYSEEDEAAEMAEVRFCRTMSELVAYSLTFVDAEGLRFKSHMYFRLVA</td>
</tr>
<tr>
<td>Rasp2</td>
<td>QEEEMFVPAVLEAASNIDDVTEAFFDDCLECESFYEDSYSEEDEAAEMAEVRFCRTMSELVAYSLTFVDAEGLRFKSHMYFRLVA</td>
</tr>
<tr>
<td>PYB</td>
<td>QEEEMFVPAVLEAASNIDDVTEAFFDDCLECESFYEDSYSEEDEAAEMAEVRFCRTMSELVAYSLTFVDAEGLRFKSHMYFRLVA</td>
</tr>
</tbody>
</table>

**X4-C protein**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasp1</td>
<td>YFRSFEEPLYSSRAFYSVKVKPVYRPKKFEGHIDCTCLODNMGEMERKSGVDAMRWCPRGLNLARKENTTRQEDKREWYQLR12LN</td>
</tr>
<tr>
<td>Rasp2</td>
<td>YFRSFEEPLYSSRAFYSVKVKPVYRPKKFEGHIDCTCLODNMGEMERKSGVDAMRWCPRGLNLARKENTTRQEDKREWYQLR12LN</td>
</tr>
<tr>
<td>PYB</td>
<td>YFRSFEEPLYSSRAFYSVKVKPVYRPKKFEGHIDCTCLODNMGEMERKSGVDAMRWCPRGLNLARKENTTRQEDKREWYQLR12LN</td>
</tr>
</tbody>
</table>

**Fig. 3.11 X4 protein: Alignment of 3 different isolates of ToRSV (Rasp1, Rasp2 and PYB).** The three peptides designed for antibody production are from the central repeat region and C-terminal region of the X4 protein (in colored boxes). Also the region of the X4 protein which was used to produce the X4(64-65) antibodies (X4-C peptide) at the C-terminal region of the protein (last 200 amino acids) is shown by the arrow.
Fig. 3.12 CP protein: Alignment of 3 different isolates of ToRSV (Rasp1, Rasp2 and PYB). The two peptides designed for antibody production are from the central region, CP (3739-40) and the C-terminal extension of the coat protein, CP (3737-38) (in colored boxes).
To test the specificity of the CP (3737-38) antibodies, immunoblotting were conducted using the antibodies against the C-terminal extension of the CP (polypeptide, TLETNNPVGRRPPENVD). This antibody reacted specifically with the *E. coli* expressed CP but not with the X4-Rasp1 protein (data not shown). The CP-2 antibody was made against a different region of CP (polypeptide, PARLPDILDDKSEV). This antibody also detected the *E. coli* expressed CP and did not recognise the X4-Rasp1 protein (data not shown). Both antibodies detected the full-length CP in infected plant extracts (60 kDa protein). The truncated CP (53-55 kDa) was detected in infected plant extracts by the CP (3739-40) antibodies but not by the CP (3737-38) antibodies, confirming that it is missing the C-terminal extension of the CP (Fig. 3.13).

The X4(3743-44) antibody (in the B3 repeat TPLVHLQESRMV) reacted specifically with the *E. coli* expressed X4-Rasp1 protein but not with the *E. coli* expressed CP (data not shown). The X4 (3741-42) antibody (a polypeptide against the C-terminal region of X4 PGRLLNAKRTYTRDD) also recognized the *E. coli* expressed X4-Rasp1 but not CP (data not shown). This is in contrast with the previous X4(64-65) antibodies that cross-reacted with the CP. The X4(3743-44) and X4(3741-42) antibodies did not detect the full length X4 or the 60 kDa protein that was detected by the X4(64-65) antibodies in infected extracts. Therefore the 60 kDa protein recognized by the X4(64-65) antibodies is not derived from the X4 coding region and probably corresponds to the full length CP. As shown above, the anti CP (3737-38) antibodies against the C-terminal extension recognized the 60 kDa protein confirming that this protein is the full length CP. Both new antibodies did detect a 45 kDa protein (Fig. 3.14B). A 45 kDa protein was also detected by the N-X4 antibodies. The nature of this protein needs further investigation. Also post-nuclear extracts derived from ToRSV-infected leaves were separated by sucrose gradient fractionation (as described in section 3.2.3). Fractions were collected from the bottom of the gradient and were analyzed by immunoblotting using the X4(3741-42) and X4(3743-44) antibodies (Fig.3.14B). In the sucrose gradient fractionation a 45 kDa protein was detected on top of the gradient, suggesting that it may be a soluble protein (Fig 3.14 B).
Fig. 3.13 Immunoblot analysis of the infected (I) and healthy (H) N. benthamiana extracts using the anti-virion, anti-CP (3739-40) and anti-CP (3737-38) antibodies. Sucrose gradient fractions from healthy and infected plant extracts were prepared and fraction number 5 and 6 containing virions were used for immunoblotting and tested against various CP antibodies. The full length CP (60 kDa) was detected by all CP antibodies, the CP (3737-38), against the C-terminal extension) only detected the upper band corresponding to the full length CP.
Fig. 3.14 Immunoblot analysis of the infected (Inf) and healthy (H) *N. benthamiana* extracts using the X4-N, X4(3743-44) and X4 (3741-42) antibodies. (A) Post-nuclear fractions (S3 fractions) were tested against various X4 antibodies. The full length X4 (expected molecular mass of 82 kDa) or the 60 kDa protein previously detected by the X4(64-65) antibodies were not detected in infected plant extracts by the X4(3743-44) or X4(3741-42) antibodies. A faint 45 kDa protein (shown by arrow) was detected that may correspond to a truncated form of the X4 protein. (B) Immunoblot detection of the X4 protein in sucrose gradient fractions in healthy and infected plant extracts with two new X4 antibodies, X4(3741-42) and X4(3743-44). The antibodies detected a 45 kDa protein and also degradation products at the top of the gradient. Molecular marker (Mr).
3.4 Discussion

The objective of this section of my research was to characterize the ToRSV X4 protein in planta. Towards this objective, I have used the X4(64-65) antibodies (a polyclonal antibody). The X4(64-65) antibodies were raised against a recombinant protein corresponding to the C-terminal 15 kDa of X4 (from the Rasp2 isolate) fused to a His tag. The X4(64-65) antibody recognized a 60 kDa protein in ToRSV infected plant extracts but not in healthy plant extracts. The 60 kDa protein co-fractionated with empty and full virus particles after sucrose gradient fractionation of infected plant extracts and co-purified with the virus particles. The results of trypsin digestion experiments and immunogold labelling suggest that the epitope of the 60 kDa protein recognized by the X4(64-65) antibody is exposed on the surface of the virus particle. The 60 kDa protein was in high concentration in plant virus extracts but in lower concentration after virus purification. After long storage of purified virus the 60 kDa protein was not detected by the X4(64-65) antibodies.

The α-virion antibody detected a 53-55 kDa doublet as well as a 60 kDa protein, which co-migrated with the 60 kDa protein recognized by the X4(64-65) antibody. These proteins were detected in plant extract as well as in purified virus preparations. As I explained in the introduction two forms of the CP are often observed in nepoviruses: the full-length form and a smaller form which has a C-terminal truncation. While the full-length form is predominant early in infection, the truncated form becomes predominant later in infection or during purification process as shown for BRV (Latvalla et al., 1998) and for TBRV (Demangeat et al., 1992). In the case of ToRSV, the C-terminal extension is predicted to be 54 amino acids (approximately 6 kDa). The difference in the apparent molecular mass between the 60 kDa and 53-55 kDa proteins detected by the α-virion antibodies would be consistent with the 60 kDa protein being the full-length CP and the 53-55 kDa protein corresponding to a truncated form of the protein lacking the C-terminal extension. Consistent with this suggestion, a full-length recombinant CP-His expressed in E. coli comigrated with the 60 kDa protein recognized by the α-virion antibodies (Fig. 3.8B). Also, the 60 kDa protein was recognized by all ToRSV CP antibodies available in the lab. In contrast, the antibodies raised against the C-terminal extension did not recognize the 53-55 kDa proteins confirming that they correspond to truncated forms of CP (Fig.3.13). The relative ratio of the 53-55 kDa and 60 kDa proteins recognized by the α-virion antibodies varied. The 60 kDa protein was predominant in crude infected plant extracts. In purified virus preparation, the ratio of the 60
kDa protein to the 53-55 kDa doublets was lower. After extended storage of purified virus preparation, the 60 kDa protein was not detectable by immunoblotting. The pattern of detection of the 60 kDa recognized by the α-virion antibody was similar to that observed for the 60 kDa recognized by the X4(64-65) antibody. This raised the possibility that the two antibodies recognized the same protein. Alternatively, two distinct 60 kDa proteins present in infected plant tissues are recognized by the two different antibodies.

The results from the mass spectrometry suggested that purified virus preparations contained predominately the coat protein and the Rubisco. Sequences corresponding to the X4 protein were not detected, although these results do not exclude the possibility that small amounts of the X4 protein were present in the preparations. These results raise the question of the nature of the 60 kDa protein detected by X4(64-65) antibodies.

There are two possible interpretations of the results. The first possible interpretation is that the 60 kDa protein recognized by the X4(64-65) antibody is a truncated form of the X4 protein. In this interpretation of the result, the 60 kDa protein recognized by the X4(64-65) antibody is a distinct protein that co-migrates with the 60 kDa full-length CP. The specificity of the X4(64-65) antibodies for the X4 protein but not for the CP protein in the immunoprecipitation assays is consistent with this interpretation of the results (Fig. 3.2). This interpretation of the results would suggest that the 60 kDa X4 protein is a virion-associated protein that becomes disassociated from the virions during the course of infection or during the course of virus purification. In addition to this protein there was also a ~120 kDa protein band detected by the X4(64-65) antibodies. The nature of this protein needs further investigation. If we assume that the 60 kDa protein is a truncated form of X4, one possibility is that the 120 kDa protein is an aggregated form (dimer) of the 60 kDa protein detected by the X4(64-65) antibodies. However, based on computer-based prediction of hydrophobicity (Chapter 1, Fig.1.11), X4 is not very hydrophobic. Another possibility is that the 120 kDa protein corresponds to a possible X3-X4 protein, a predicted intermediate precursor of the mature X4 protein that may be associated with the virion. This hypothesis would also suggest that the 60 kDa (or the 120 kDa protein) detected by the X4(64-65) antibodies are exposed at the surface of the virions. The fact that the α-virion antibodies recognized the in vitro synthesized X4 protein in the immunoprecipitation assay could suggest that the 60 kDa or the 120 kDa protein detected by the X4(64-65) antibodies was
present in the original virus preparation used for the production of these antibodies (Fig 3.5 A).

The second possible interpretation of the results is that the X4(64-65) antibody cross-react with the full-length 60 kDa CP but not with the truncated 53-55 kDa form of the CP. The cross-reactivity between the X4(64-65) antibodies and the E. coli expressed full-length CP in immunoblotting experiments is consistent with this suggestion. The lack of detection of X4 sequences in mass spectrometry using purified virus preparations is also consistent with this hypothesis. This interpretation of the results implies that the X4(64-65) antibody may recognizes an epitope present on the C-terminal extension of the CP and that this epitope is exposed at the surface of the virions. To reach a clear conclusion about the nature of the X4 protein and further understand the role of this protein in virus infection more antibodies were made from different regions of the CP and the X4. The anti-CP (3737-38) and CP (3739-40) antibodies detected the full length CP (60 kDa protein) in plant extracts. However, the anti-X4(3741-42) and (3743-44) antibodies did not detect either the 60 kDa protein or the predicted 82 kDa full length X4. The results from these antibodies were consistent with the results obtained by the N-terminus X4 antibodies. Also, preliminary results of the expression of X4-GFP protein in plants by biolistic delivery of the plasmid into N. benthamiana followed by immunodetection by the GFP antibodies did not allow the detection of the predicted full-length X4-GFP (data not shown). The observations that the X4 protein was not detected by four different antibodies in extracts derived from ToRSV-infected or transiently transfected plants (GFP antibodies, two new X4(3741-42) and X4(3743-44) antibodies and the N-terminal X4 antibodies) support the second interpretation in which the 60 kDa protein recognized by the X4(64-65) antibodies probably is the full length CP and this antibodies had cross reacted with the CP.

As mentioned above, the most likely interpretation of my results is that the X4(64-65) antibodies recognize an epitope within the C-terminal extension of the full-length CP. This interpretation of the results would imply that this C-terminal extension is exposed at the surface of the virion. This is consistent with the presence of an exposed C-terminal extension in other nepoviruses and comoviruses as I discussed in the introduction. The C-terminal extension can dissociate from the virus particle during purification or prolonged storage of the virus, therefore the X4(64-65) would not be able to detect the C-terminal extension in those preparations.
As I discussed in Chapter 1, nematodes acquire and retain virus particles but release the particle during their feeding process to allow plant-to-plant transmission. The CP C-terminal extension can be a potential site for interaction with the nematode receptor, since it can be released from the virus particle. In tomato black ring virus (TBRV) the C-terminal extension of the coat protein protrude from the virus particle surface and can be removed without disrupting the virion. This exposed detachable fragment may play a significant role in TBRV biology as does the protruding N-terminal fragment of potyvirus coat proteins in their transmission by aphids (Atreya et al., 1990).

The C-terminal extension in ToRSV is longer in size than the BRV C-terminal extension. This part of CP has no sequence homology with the CP of other nepoviruses or comoviruses sequenced so far (Seitsonen et al., 2008). Since the sequence is specific, the C-terminal extension is valuable for making valid antigenic targets for diagnostic and quick indication of ToRSV disease. The development of sensitive and reliable antibodies against ToRSV surface exposed CP sequences combined with regular tests would allow the recognition of the disease long before the symptoms become evident. Early detection might help in controlling the disease which may alleviate the economic impact of ToRSV on fruit crops all over the world.

In immunoprecipitation assays, the α-virion antibodies detect the full length X4 from Rasp2 or Rasp1 (Fig. 3.2A and data not shown). Also the X4(64-65) antibodies cross-reacted with the full length E. coli expressed CP in immunoblotting experiments. One possible explanation for these results is that the CP and X4 proteins share a common epitope. Alignment of the CP and the X4 did not show significant sequence homology (Fig. 3.12) although the possibility of a structural epitope can not be excluded.
Fig.3.15. Sequence alignment between the X4 protein and the CP of ToRSV. There is no significant homology suitable for a common epitope between the two sequences.
The X4 coding region is present in the RNA2 from all ToRSV infected isolates that were sequenced so far. However the X4 full length protein (e.g. 82 kDa in Rasp1) is not detectable in plant extracts. Other ToRSV viral proteins studied so far are readily detected by antibodies in infected plant extracts.

In ToRSV, each viral RNA codes for a large polyprotein. The polyprotein is cleaved by the viral protease (Pro) to mature and intermediate proteins. Therefore theoretically the cleavage products should have a ratio of 1:1. The absence of detection of the full-length X4 protein in plant extracts could be due to a shorter half-life of the X4 protein compared to other viral proteins like the MP or the CP, which are readily detected in infected plant extracts. This suggests that X4 may be subject to degradation.

Cauliflower mosaic virus (CaMV) particles are stable although its capsid protein (CP) precursor is very unstable in plant protoplasts when it is expressed outside the viral context (Leclerc et al., 1999). The capsid is assembled from three related proteins p44, p39 and p37. These proteins are all derived from bi-terminal processing of the 57 kDa CP precursor. The N-terminus of p44 contains a PEST motif (degradation motif, Chapter 1). It is suggested that the PEST signal within the N-terminus of the precursor targets the protein to the proteasome in plants. It is suggested that PEST related signals may act as degradation signals in plants (Baumberger et al., 2007), however the algorithm was developed for mammalian proteins and needs to be adapted for the plant system (Chapter 1 section 1.10.1) (Karsies et al., 2001).

Little is known about the degradation pathways in either the nucleus or the cytoplasm of plant cells (Karsies et al., 2001). Viruses can use the degradation pathways of their host for their own benefit. Poleroviruses (family Luteoviridae) suppress RNA silencing activity in N. benthamiana plants using the viral P0 protein which targets the Argonaute protein for degradation (Baumberger et al., 2007). Phosphorylation often plays a role in regulation of PEST signal target degradation. There are a few predicted phosphorylation sites in the X4 region of ToRSV isolates and there are a few predicted PEST motifs (Chapter 1). The number of predicted PEST motifs was greater in the X4 protein from the Rasp1 isolate. The PEST motifs were present within the protein repeats, therefore the greater number of the PEST motifs in Rasp1 may correlate with protein instability.
As discussed above, our results suggest that the full length X4 protein is an unstable protein. Another possibility is that since the cleavage between X3 and X4 is a suboptimal cleavage site, the protein may not have been cleaved from its precursor as efficiently as other viral proteins like the MP or the CP. Larger proteins are occasionally detected in ToRSV-plant extracts by all of the X4 antibodies but the nature of these proteins requires further investigation. It would be interesting to investigate the possibility of X3 as a stabilizing sequence for the X4 protein.

X4 is also a highly diverse protein among ToRSV isolates that contains tandem repeats and an extensive number of point mutations (Chapter 2). What is the possible function(s) of X4? In particular, why does the virus code for a protein that is absent from the genome of many other nepoviruses? Indeed, the coding region for X4 is retained in the virus genome even after extensive passages of the virus on herbaceous hosts. As mentioned in Chapter 2, a possible role in symptomatology or protein–protein interaction could be considered. The titre of ToRSV viral RNA remains at high concentration in recovered leaves in contrast with other nepoviruses. In Chapter 4, I will discuss preliminary results suggesting that X4 may be a silencing suppressor.

3.4.1. Summary and conclusion

Using X4(64-65) antibodies raised against the C-terminal region of X4, I have detected a 60 kDa protein in infected plants. A series of experiments aimed at analyzing the nature of this 60 kDa protein and its possible interaction with the virus particles allowed me to formulate two possible interpretations for the results. (1) The 60 kDa protein (or the possible 120 kDa precursor protein) recognized by the X4(64-65) antibody is a truncated form of the X4 protein, which is associated with virus particles and (2) the X4 antibodies recognize the C-terminal extension of the full-length CP. Recent new antibodies raised against different regions of X4 and CP, supported the second hypothesis. Based on the results from the X4 (3741-42), X4(3743-44) and the N-X4 antibodies, the predicted full length X4 (82 kDa) protein was not detected in ToRSV-infected plant extracts or the protein concentration was lower than the detection level by immunoblotting in infected plant extracts. This suggests that the full length X4 protein may be an unstable protein although further experiments are required to address this question.
3.5 Bibliography


Chapter 4

Preliminary evidence that the X4 protein may act as a silencing suppressor in plants

A version of this chapter will be submitted for publication. Jafarpour, B., Chisholm, J and Sanfacon, H. (2010) Preliminary evidence that the X4 protein may act as a silencing suppressor in plants.
4.1. Introduction

As I discussed in Chapter 1, RNA silencing is one of the natural plant defence mechanisms against virus infection. RNA silencing generally relies on a set of core reactions that are triggered by dsRNA, which is processed by the RNase III enzyme Dicer into small RNA duplexes that are 21–24 bp in length (Ding, 2000; Voinnet, 2005). The helicase selects one strand of the viral siRNA duplex to be the guide RNA in the mature RNA–induced silencing complex (RISC) (Schwarz et al., 2003). This strand of the siRNA duplex is then incorporated into the RISC complex. A core component of the RISC complex is an Argonaute protein (Ago), which is an RNase H-like enzyme (Baumberger and Baulcombe, 2005; Zhang et al., 2006) that guides sequence-specific degradation of the complementary RNA (Ding and Voinnet, 2007; Dunoyer and Voinnet, 2008; Lu, 2003).

The sequence-specificity of this process, together with the fact that dsRNA is a common product of virus replication, prompted the idea that virus-induced gene silencing is an antiviral defence response. The dsRNA in virus-infected cells is thought to be the replication intermediate that are produced during virus replication. Secondary structures within a specific strand of viral genomic RNA could also be recognized as dsRNA by the silencing machinery. As I explained above eventually this causes the siRNA/RNase complex to target the viral single-stranded RNA, and virus accumulation would slow down. Many plant viruses encode proteins that are suppressors of this RNA silencing process (Brigneti et al., 1998; Voinnet et al., 1999). These proteins influence the final steady-state level of virus accumulation (Lu, 2003). This, together with the discovery of siRNAs in virus-infected plant cells (Hamilton and Baulcombe, 1999) provided further evidence that RNA silencing is an antiviral defence system of plants. Viral suppressor proteins that counteract RNA silencing were first identified in plant cells (Anandalakshmi et al., 1998). It was later found that viruses of insect and mammalian cells, such as influenza virus also encode silencing suppressor proteins (Li, 2002; Li et al., 2004). This suggests that RNA silencing could have been an ancient antiviral defence mechanism in primitive eukaryotes and that its function has been conserved during the evolution of plants and animals (Baulcombe and Molnar, 2004).

As discussed in Chapter 1, one of the well studied plant virus suppressors of silencing is the 19 kDa protein (p19) from tomato bushy stunt virus (genus tombusvirus) (Silhavy et al., 2002). HC-Pro (potyviral helper component protease) was also one of the first identified
silencing suppressors (reviewed in Maia et al., 1996; Maia and Haenni, 1994; Roth et al., 2004; Urcuqui-Inchima et al., 2001). Viral silencing suppressor proteins are not highly conserved and the mode of action of these proteins against RNA silencing is diverse. Theoretically, viruses can combat RNA-silencing mediated defence in at least three ways: preventing the generation of siRNAs, inhibiting the incorporation of siRNAs into effector complexes and interfering with one of the effector complexes.

The recovery of plants from virus infection was first described for a nepovirus in 1928 (Wingard, 1928). Initial symptomatic infection of herbaceous plants of the Solanaceae family by tobacco ringspot virus was shown to be followed by attenuation or elimination of the symptoms in newly emerging leaves. Characterization of the recovery phenotypes observed in some natural virus infections provided additional experimental evidence for a link between RNA silencing and an antiviral defence mechanism (Covey et al., 1997; Ratcliff et al., 1997; Ratcliff et al., 1999). Plants infected with nepoviruses and caulimoviruses exhibit a response very similar to the virus-induced recovery. In plants infected with tomato black ring virus (TBRV, a nepovirus of subgroup B), the recovered leaves exhibit homology-dependent resistance to secondary infections and the virus titre was dramatically reduced in the recovered leaves (Covey et al., 1997; Ratcliff et al., 1997; Ratcliff et al., 1999). It has been demonstrated that recovery of Nicotiana benthamiana plants from the necrotic symptom induced by ToRSV is associated with RNA silencing. Unlike TBRV-infected plants, recovered leaves from ToRSV-infected plants did not show reduced virus titre (Jovel et al., 2007). ToRSV is different from TBRV in that it has a larger coding region in the two RNAs. The X4 domain in the N-terminal region of the ToRSV RNA-2 encoded polyprotein is absent in TBRV (Carrier et al., 2001).

4.1.1. Research hypothesis and objectives

The X4 is a variable protein among ToRSV isolates (Chapter 2). The high concentration of virus in ToRSV infected plants (Jovel et al., 2007) differs from what was previously observed with other nepoviruses and suggest that ToRSV has the ability to counteract or evade RNA silencing. Since the X4 protein is absent from the genome of other nepoviruses, we hypothesized that X4 may suppress silencing in plants. In this study, we tested the recovery phenotype and symptomatology of different ToRSV isolates. We used a transient expression assay and the green fluorescent reporter protein (GFP) to determine whether co-
expression of X4 and GFP could prevent the induction of RNA silencing directed at GFP. We show that when X4 is co-expressed with GFP in plants, the concentration of the GFP protein is higher than when the GFP is expressed alone.

4.2 Materials and methods

4.2.1 PCR amplification of a region of the RNA2 open reading frame coding for a portion of X3, the entire X4 and a portion of MP for the Chickadee isolates.

Chickadee isolate was obtained from Dr. Zongrang Liu (West Virginia). The exact origin of this specific Chickadee isolate is not known. However, a previously described Chickadee isolate was reported to have been introduced to New York State in rooted MM.106 apple layers (Bitterlin and Gonsalves, 1988). The Chickadee isolate was passaged in cucumber extensively before it was used in our laboratory. Therefore, it may have adapted to the cucumber host.

As I discussed in Chapter 2, leaves from *N. benthamiana* plants infected with Chickadee isolate were ground in liquid nitrogen and the total RNA was extracted using TRIzol reagent (Invitrogen). First–strand cDNA was synthesised using 1 μg of total RNA as a template and primer CP-R (5'-GTCAAGCTTGCCACGCCCGAAAGGAT-3', complementary to nts 5723-5707 of the RNA2 of ToRSV isolate Rasp2 and Superscript II Reverse Transcriptase (Invitrogen). For the PCR reaction, I used primers F208 (5'-GAGGCCGAATTGGCCTCAAAG-3', corresponding to nts 801-822 of Rasp2 RNA2, and R207 (5'-GCACCCGCATCAGAGGATC-3', complementary to nts 2998-2980 of Rasp2 RNA2.

4.2.2 Symptom evaluation in plants infected with Rasp1 and PYB ToRSV isolates

ToRSV-Rasp1 and ToRSV-PYB were inoculated on *Nicotiana benthamiana* plants. An abrasive (carborundum powder) was dusted on the leaves prior to inoculation. Plants were inoculated at leaf stage 5 to 7 by inoculating two leaves with approximately 40 μl of fresh leaf extracts of ToRSV-infected plants diluted 1:4 (w:v) in inoculation buffer (0.1 M phosphate buffer pH 7.4). The plants were grown in conviron climate chamber at 21°C and 27°C for symptom development and were assessed at regular interval for 22° days post inoculation.
4.2.3. Plasmid construction

Several agroinfiltration vectors were constructed. The pCITE-X4-Rasp1-HA and pCITE-X4-Rasp2-HA (containing the coding region for the X4 protein from Rasp1 and Rasp2 fused to an HA tag at the C-terminal end of the encoded protein) were described in Chapter 3. These plasmids were digested with *NcoI* and *BglII*. The resulting X4 (Rasp1 and Rasp2) fragments were inserted into the corresponding sites of an intermediate vector pBBI525, which contains a duplicated 35S promoter, the alfalfa mosaic virus translation enhancer and the nos polyadenylation signal (Sun et al., 2001). The plasmids were then digested with *HindIII* and *EcoRI* for the X4-Rasp2-HA plasmid and with *KpnI* and *EcoRI* for the X4-Rasp1-HA plasmid. The resulting fragments were inserted into the corresponding sites of a binary vector pBINplus (clontech Laboratory Inc.) resulting in pBIN-X4-Rasp1 and pBIN-X4-Rasp2. Plasmids pBIN-GFP allowing the expression of GFP and pBIN-p19 containing the coding region for the tomato bushy stunt virus (TBSV) suppressor of gene silencing were described previously (Zhang et al., 2005).

4.2.4 Agroinfiltration of *N. benthamiana* plants and immunoblotting

The binary vectors containing X4-Rasp1-HA, X4-Rasp2-HA, GFP and p19 were transferred into *Agrobacterium tumefaciens* LBA4044 (Invitrogen) by electroporation. Colonies confirmed to contain the X4 (Rasp1 and Rasp2) binary vector were used for agroinfiltration assays as described (Voinnet et al., 2003). The expression of GFP was tested in the presence or absence of the X4 protein or p19. TBSV p19 was used as a positive control for silencing suppression activity (Voinnet et al., 2003). After agroinfiltration, the plants were grown in a conviron climate chamber at 27°C for 6 days. The infiltrated area was collected for immunoblot analysis at 1-6 days post infiltration (dpi). One hundred mg of each leaf tissue was stored at -80°C until further use. Protein extracts were prepared by grinding the leaf under liquid nitrogen in the presence of extraction buffer (10mM KCl, 5 mM MgCl₂, 400 mM sucrose, 100mM Tris-HCl, pH 8.0, 10% [vol/vol] glycerol). Cell debris was removed by centrifugation at 3,700 X g at 4°C for 10 min. The supernatant was resolved on sodium dodecyl sulfate-polyacrylamide gels (12% acrylamide) and electroblotted to a nitrocellulose membrane. Immunodetection was carried out using a mouse monoclonal anti-GFP antibody (BD Bioscience). The secondary antibody was a goat anti mouse immunoglobulin G conjugated to horseradish peroxidase (Bio/Can). The immunostained protein was visualized by enhanced chemiluminescence detection with ECL kit (Amersham) according to the
manufacturer’s instructions. Coomassie staining of rubisco protein was shown as a loading control. To detect the X4-HA protein, immunoblotting was carried out using the anti-HA antibodies (Rat monoclonal IgG, Roche) followed by a secondary antibody, a goat anti mouse immunoglobulin G conjugated to horseradish peroxidase (Bio/Can) and visualized by enhanced chemiluminescence detection as described above.

4.3 Results

4.3.1 RT-PCR amplification of ToRSV Chickadee isolate and symptom development in ToRSV–infected Rasp1 and PYB ToRSV isolates

As I discussed earlier sequence analysis of the X4 region from (Rasp1 and PYB) isolates showed a surprising level of sequence variability. In Chapter 2, I have shown that the X4 protein varies in size among ToRSV isolates (Rasp1, Rasp2 and PYB) due to the insertion of a variable number of amino acid repeats. Similarly, after RT-PCR amplification of an additional isolate (Chickadee) a fragment that was larger in size than PYB or Rasp1 isolates was observed (Fig 4.1). The sequence of the X4 coding region in the Chickadee isolate has not yet been elucidated. The focus of this chapter was Rasp1, Rasp2 and PYB isolates of ToRSV. The X4 region of Chickadee isolate was not sequenced or inserted in the binary vector and was not agroinfiltrated or expressed in plants.

Here I present a detailed comparison of the symptoms induced by the two isolates that the X4 sequence is available: Rasp1 and PYB. The Rasp1 and PYB isolates of ToRSV induced symptoms on inoculated and systemic leaves in N. benthamiana. In plants inoculated with the Rasp1 isolate, recovery from symptomatic infection occurred at 27°C after 14 days post-infection. However, at lower temperatures (21°C), recovery did not occur and the plants died after 22 days post inoculation (dpi) as a result of necrosis induced upon ToRSV infection (Fig. 4.2C). In N. benthamiana plants infected with PYB1, symptoms were milder than those observed in Rasp1-infected plants and PYB1-infected plants recovered from infection even when cultivated at 21°C (Fig. 4.2F).
Fig. 4.1  PCR amplification of a region of the RNA2 open reading frame coding for a portion of X3, the entire X4 and a portion of MP in Rasp1, PYB and Chickadee isolates.

PCR fragments were amplified by RT-PCR using total RNA purified from *N. benthamiana* plants infected with ToRSV Rasp1, PYB1 and Chickadee isolates. Primer pair R207/F208 was used for the amplification (the RT-PCR product concentration of the lanes are not equal). RT-PCR reactions using total RNA extracted from Chickadee-infected plants resulted in the amplification of an even larger fragment (the RT-PCR product concentration of the lanes are not equal).

Fig. 4.2  Symptom development in ToRSV–infected Rasp1 and PYB ToRSV isolates.

ToRSV induced symptoms in inoculated and systemic leaves. The viral symptoms in *Nicotiana benthamiana* plants infected with Rasp1 (panels A-C) and PYB isolates (panels D-F) are shown after 6 (panels A and D), 13 (panels B and E) and 22 days post inoculation (panels C and F) at 21°C (panels 1, 3, 6, 8, 10) and 27°C (panels 2, 4, 7, 9, 11).
2, 4, 5, 7, 9, 11). Individual leaves shown in panels A and D after 6 days post-inoculation are inoculated leaves. Other panels show the entire plant. A picture is not shown for the 22 dpi time-point at 21° C for Rasp1, because at this time points all plants were dead and had been discarded.

4.3.2 Preliminary evidence that the X4 protein may act as a silencing suppressor in plants

To investigate the function of X4 as a silencing suppressor, we agro-infiltrated leaves of *N. benthamiana* with various combinations of agrobacteria carrying binary vectors allowing the expression of GFP, p19, X4-Rasp1-HA and X4-Rasp2-HA. The expression level of GFP alone was used as a control. The expression of GFP was decreased after 6 days as evidenced by the low level of GFP protein detected by immunoblotting. It was shown previously that transient expression of GFP in *N. benthamiana* plants after agroinfiltration, usually peaks at 60–72 hours post-infiltration and declines rapidly afterwards because of the induction of GFP-specific RNA silencing (Voinnet *et al.*, 2003). However, when plants are co-infiltrated with two strains of Agrobacterium allowing co-expression of GFP with a viral suppressor of silencing, silencing of GFP is inhibited and the expression level of GFP is increased. Co-expression of GFP with a virus suppressor of silencing (e.g., the p19 protein) allows sustained expression of GFP up to 10-14 days post agroinfiltration (Voinnet *et al.*, 2003). This assay is commonly used to assess silencing suppression activity of viral proteins.

As expected, co-expression of p19 and GFP resulted in enhanced and sustained expression of GFP compared to expression of GFP alone (Fig. 4.3 A and B). This was observed in two separate experiments: after 5dpi or in a time-course experiment (2-6) dpi. When pBIN-X4-Rasp2 was co-agro-infiltrated with pBIN-GFP, expression of GFP was enhanced after 5dpi. However, the protein expression level of GFP was not enhanced after 5dpi when pBIN-GFP was co-agro-infiltrated with pBIN-X4-Rasp1 (Fig. 4.3A). In a time course experiment, the GFP protein expressed individually was detectable at 2 and 4 dpi but was not detectable at 6 dpi, suggesting efficient induction of RNA silencing directed at GFP (Fig. 4.3B). When leaves were agro-infiltrated with both the pBIN-GFP and pBIN-p19 vectors, the expression of GFP was sustained and the concentration of GFP was higher at 4 to 6 days post infiltration (dpi) than at 2 dpi. In leaves infiltrated with a mixture of agrobacteria carrying pBIN-GFP and pBIN-X4 (Rasp1 and Rasp2), the concentration of GFP at 2 and 4 dpi was much stronger than in leaves infiltrated with GFP alone (Fig. 4.3A and B). However, this effect was
transient and the concentration of GFP dropped after 6 dpi. One possible interpretation for this result is that X4 transiently interferes with the induction of silencing directed at GFP.

![Image](image)

**Fig. 4.3 Expression levels of GFP in a transient expression assay conducted in the presence or absence of X4.** (A) The expression level of GFP detected 5 days post-agroinfiltration. GFP was expressed alone (-) or in combination with p19 or X4 (Rasp1 and Rasp2) in wild-type *Nicotiana benthamiana* plants as indicated above each lane. Plants agroinfiltrated with the X4-Rasp2 isolate show high level of GFP expression, however this was not observed in plants infiltrated with X4-Rasp1. (B) X4 transiently enhance the expression of GFP in *Nicotiana benthamiana* plants. GFP was expressed alone (-) or with p19 or X4 (Rasp2 and Rasp1) in wild-type *Nicotiana benthamiana* and samples were taken 2-6 dpi (top panel). Both, X4-Rasp1 and X4-Rasp2 enhance the GFP expression at 2 dpi. The lower panel in both A and B show the Coomassie blue staining of Rubisco as a loading control.

To determine whether the X4-Rasp1-HA and X4–Rasp2-HA proteins were stably expressed in the agroinfiltrated leaves, immunoblotting analysis was carried out using HA and X4(64-65) antibodies. The expected size for the full-length X4 is 82 kDa for Rasp1 and 71 kDa for Rasp2. The X4(64-65) antibodies did not detect any protein of the expected size for X4 in the agro-infiltrated areas (data not shown). However the HA antibodies did detect a smaller
protein band ~20 kDa one day post infiltration (dpi) in leaves expressing the X4 protein from the Rasp1 isolate (Fig. 4.4).

Fig. 4.4 Analysing the expression level of the X4 protein in infiltrated patches by immunoblotting. Immunoblotting was carried out using plant extracts of leaf patches infiltrated with agrobacteria carrying pBIN (p19), pBIN(X4-Rasp1-HA) and pBIN(X4-Rasp2-HA) as indicated above each lane. (-) the first two lanes in each panel, represent leaves that were infiltrated with agrobacteria carrying pBIN(X4-Rasp1-HA) and pBIN(X4-Rasp2-HA). (P19) the third and the fourth lanes in each panel represents leaves that were infiltrated with agrobacteria carrying pBIN(X4-Rasp1-HA) and pBIN(X4-Rasp2-HA) plus agrobacteria carrying pBIN (p19). Plant extracts were collected at 1, 2, 3, and 4 dpi and the proteins were detected using the HA antibodies. A protein band of ~20 kDa was present in Rasp1 isolate at 1dpi. pBIN(X4-Rasp2-HA) was shown by Rasp2 and pBIN(X4-Rasp1-HA) was shown by Rasp1 above each lane.

4.4 Discussion

We have previously shown that N. benthamiana plants inoculated with Rasp1 recover from infection at temperatures equal to or above 27°C and young leaves become symptom-free after an initial symptomatic systemic infection (Jovel et al., 2007). It was shown previously that at low temperature both virus and transgene triggered RNA silencing are inhibited. Therefore, in cold temperatures, plants become more susceptible to viruses. At low temperature, the level of virus or transgene-derived siRNAs is dramatically reduced. In contrast, RNA silencing is activated and the amount of siRNAs gradually increases with temperature rising (Szittya et al., 2003). Recovery from tomato black ring virus (TBRV, a subgroup B nepovirus) has been associated with the induction of RNA silencing and a concurrent dramatic reduction in virus titre in the recovered leaves (Ratcliff et al., 1997). ToRSV (Rasp1 isolate) accumulates to high titre in recovered leaves, in spite of active RNA
silencing directed at ToRSV sequences (Jovel et al., 2007). Although ToRSV-specific RNA silencing was detected in recovered leaves, the concentration of ToRSV-derived siRNAs was low and degradation of a sensor construct containing a ToRSV fragment was incomplete (Jovel et al., 2007). One possible interpretation for these results is that ToRSV encodes a suppressor of silencing, thereby allowing the virus to accumulate in recovered leaves. At lower temperatures (21°C) recovery does not occur in Rasp1 infected plants and the plants eventually die as a result of the necrosis induced by ToRSV infection (Fig. 4.2). PYB1 is closely related to Rasp1 and Rasp2 (Wang and Sanfacon, 2000). However, in N. benthamiana plants infected with PYB1, symptoms are milder than those observed in Rasp1-infected plants and PYB1-infected plants recover from infection even when cultivated at 21°C (Fig. 4.2). The X4 protein, which is absent from the TBRV genome is an attractive candidate for the silencing suppression function. Also the silencing suppressor activity of X4 may differ among ToRSV isolates. In cucumber mosaic virus (CMV) the severe subgroup (SD-CMV) has a stronger 2b suppressor activity than the mild subgroup (Q-CMV). Also it has been suggested that the variable domain in SD2b, which is absent in Q2b, is the major contributor to the stronger suppressor activity of SD2b (Ye et al., 2009). The availability of two closely related ToRSV isolates with diverging X4 sequences and differences in symptomatology provides us with an invaluable tool to address this question.

Based on the agroinfiltration experiments one possible interpretation of our result is that X4 transiently suppress silencing of the GFP reporter gene (Fig. 4.3). In leaves co-infiltrated with GFP and X4-Rasp1 at 5 dpi, the expression of GFP was not significantly enhanced compared to the control in which GFP was expressed alone. However, the GFP expression level was enhanced at 5 dpi when it was co-infiltrated with X4-Rasp2. In the time-course experiment, the co-expression of X4 (Rasp1 and Rasp2) with GFP allowed enhanced expression of GFP at early time-points. Altogether, these experiments suggest that both the X4-Rasp1 and X4-Rasp2 protein may be active as a suppressor of silencing in Nicotiana bentamiana plants. However, the possible silencing suppression activity of X4 will need to be further investigated by analyzing the levels of GFP RNA and siRNA in agro-infiltrated leaves.

The full length X4 or the 60 kDa protein was not detected in infiltrated plants with the HA antibodies (Fig. 4.4). The HA antibodies detected a 20 kDa protein in the Rasp1 isolate which does not correspond to the expected size for the full length X4-Rasp1 (82 kDa). The
nature of this 20 kDa protein needs to be further investigated. The HA tag was fused to the C-terminus of the X4 protein. Also the X4(64-65) antibodies were raised against the C-terminal end of the X4-Rasp2 protein did not detect the X4 protein in infiltrated plants (data not shown). The fact that the 20 kDa protein is detected by the HA antibodies at 1dpi but not by the X4(64-65) antibodies raised more questions regarding the nature of the 20 kDa protein. As I discussed in Chapter 3, X4 was not detected in infected plant extracts based on the X4-N, X4(3741-42) and X4(3743-44) antibodies. The results suggested that the X4 protein may be degraded or that the expression level of X4 protein in plants is below the detection level of the antibodies. It is also possible that a small fragment of X4 rather than the full length protein is active in silencing suppression. It would be interesting to test whether the X4 protein from PYB and Rasp1 isolates have different silencing suppression activities.

How does a protein that is not detectable and apparently is unstable in plants (Chapter 3), acts as an active suppressor of silencing? The silencing pathways are very diverse in plants (Brodersen and Voinnet, 2006), therefore the suppressor activity would be diverse as well. Many viral suppressor of silencing such as the TEV HC-Pro and the tombusvirus p19 and closterovirus p21 proteins act by binding to the ds-siRNA and prevent them from loading to the RISC complex. Most viral suppressors bind to long dsRNA or siRNA and prevent the production of siRNA or binding of siRNA to the AGO protein. However these are not the only suppressor activity pathways in plants. There are other pathways that prevent the silencing activity. One example is the 2b protein of Cucumoviruses that binds to the AGO protein and prevents silencing activity. The P0 of Polerovirus suppresses silencing through the proteasome mediated degradation pathway (Baumberger et al., 2007). This is a novel pathway to suppress silencing in plants. We have suggested that the 82 kDa X4 protein is an unstable protein (Chapter 3) and that it may act as a suppressor of silencing (this Chapter), the possibility that the X4 protein may suppress the silencing activity trough the proteasome pathway like the P0 protein of Polerovirus can not be excluded although further experiments need to be done (Chapter 5).

4.4.1 Summary and conclusion

The variability of the X4 protein among ToRSV isolates combined with the differences in symptomatology induced by these isolates and the preliminary evidence that X4 may act as a suppressor of silencing in plants, suggest that X4 could be involved in host specificity,
symptomatology and/or interaction with host defence responses. These results added a new level of understanding to the possible function of the X4 protein in ToRSV infection.
4.5 Bibliography


Chapter 5

General discussion and future projects
5.1 General discussion

In this thesis, I have studied the tomato ringspot virus X4 protein. The X4 protein does not have significant sequence identity with any other proteins available in the database and does not have a functional equivalent in the genome of nepoviruses of subgroup A and B. In Chapter 2, I have discussed the sequence diversity of the X4 protein among ToRSV isolates and have shown that insertion of large amino acid repeats and point mutations contribute to a high degree of sequence variability in this protein. In Chapter 3, I have studied a 60 kDa protein detected by the antibodies raised against the C-terminal region of X4. This antibody cross reacted with the full length CP. Based on this and on further experiment with the recent X4(3741-42) and X4(3743-44) antibodies, this 60 kDa protein corresponds to the full length CP. The predicted full-length X4 (82 kDa) protein was not detected in ToRSV-infected plant extracts. This suggests that the full length X4 protein may have a short half life. In Chapter 4, I presented preliminary results suggesting that X4 may suppress RNA silencing. Altogether, the results discussed in this thesis bring novel ideas about the nature of the X4 protein and raise new questions for future studies.

Based on the sequencing data, two series of tandem repeats were identified in the deduced amino acid sequence from the X4-Rasp1 protein. The first type of repeat is a 53 a. a. motif that was previously identified in the deduced a.a. sequence of the ToRSV-Rasp2 RNA2 encoded polyprotein (Rott et al., 1991). However, the number of copies of this repeat varied with the particular ToRSV isolate. The second type of repeat is 22 a. a. long and was not identified previously. The presence of two types of tandem repeats in the X4 protein among ToRSV Rasp1 and PYB isolates was reported for the first time. These repeats may represent novel classes of amino acid repeats as they did not share obvious sequence homology with known classes of repeats. Repeats can be found in proteins with diverse functions. It is believed that tandem repeats evolve from a common ancestor that contained a single repeat (Andrade, 2001). Evolution of repeat-containing proteins is relatively easy and “inexpensive”. In fact, large stable protein repeats may arise by simple gene duplication rather than the more complex process of α-helix and β-sheet creation. However, this may not be the only answer to the evolutionary success of protein repeats. It has been suggested that duplication of repeats allow proteins to acquire different molecular functions (e.g. protein–protein interaction or association with different ligands). The different number of repeats in ToRSV isolates may influence the interaction of X4 with other proteins. Further experiment such as co-immunoprecipitation assays of the X4 proteins (truncated forms and full length)
from different isolates or the yeast two hybrid system can be used to investigate the interaction of X4 with other viral proteins or with host proteins. It has also been suggested that an increase in the number of repeated domains might not alter the protein structure significantly and can in fact promote protein stability (Kloss et al., 2008; Kohl et al., 2003; Tripp and Barrick, 2004). However, there are also reports that the presence of protein repeats can result in the generation of misfolded and/or malfunctioning proteins, which are unstable or hypervariable (Usdin, 2008).

The full-length X4 is not detected with any of the antibody tested either in the context of virus infection or after transient expression by agroinfiltration or biolistic delivery (Chapter 3). The X4, CP and the MP are released from the RNA2 encoded polyprotein and theoretically the released proteins should be in a 1:1 ratio. The CP and the MP are readily detected by the CP and the MP specific antibodies in infected plant extract, however the full length X4 was not detected. Altogether this suggests that the X4 may be an unstable protein when it is expressed in plants. The cleavage site between X3-X4 is also a suboptimal cleavage site, therefore it is possible that X4 is not being cleaved from its precursor form properly (Carrier et al., 2001). Larger proteins of ~120 kDa are also detected in plant extracts by the X4 antibodies that may correspond to the X3-X4 intermediate precursor polyprotein, however further experiments are required to detect the nature of this proteins.

Based on the high degree of sequence variability of the X4 protein among ToRSV isolates, especially in the Rasp1 isolate that was extensively passaged in herbaceous hosts, I have suggested that the X4 protein may play a role in vector transmission, host specificity, symptomatology or interaction with host defence response such as RNA silencing (see Chapter 2 and 4). It was previously shown that in ToRSV systemic recovered leaves, the virus is in high concentration. Also, it was shown that RNA silencing is active during recovery and siRNA are detected in infected leaves although at low concentration (Jovel et al., 2007). This suggests that the virus may counteract RNA silencing by expressing a silencing suppressor. The high degree of sequence variability in the X4 coding region among ToRSV isolates may correlate with the different symptomatology and recovery phenotype in these isolates (see Chapter 4). Therefore, I investigated the possible function of the X4 protein as a suppressor of silencing in N. benthamiana plants.

In Chapter 4, I showed that the X4 protein transiently enhanced the expression of GFP in N. benthamiana plants. One possible interpretation for this result is that X4 suppresses RNA
silencing directed at GFP. However, we were not able to detect the expected full-length X4 protein in agroinfiltrated leaves by immunoblotting. Instead, a truncated 20 kDa protein was detected by the HA antibodies although the nature of this protein needs further investigation. This provides further support for the suggestion that the X4 protein is not a stable protein and it may be targeted for degradation \textit{in planta}. The possibility that X4 is an unstable protein could also explain why it only transiently enhanced GFP expression.

Because ToRSV-specific siRNAs are detected in infected plants, the possible silencing suppressor activity of ToRSV is not likely to function by eliminating the processing of dsRNA by the Dicer. One possible explanation is that ToRSV does not counteract RNA silencing completely. Another possibility is that ToRSV suppresses silencing downstream of the formation of the double stranded siRNA. As mentioned above, two different isolates of ToRSV, Rasp1 and PYB show differences in the strength of symptoms and recovery phenotype. The X4 protein in these two isolates also have different number of protein repeats. In tobacco etch virus (TEV), variants with hyposuppressor mutants were far less virulent than wild-type TEV (Torres-Barcelo \textit{et al.}, 2008). In cucumber mosaic virus (Ye \textit{et al.}, 2009)(CMV) the severe subgroup (SD-CMV) has a stronger 2b suppressor activity than the mild subgroup (Q-CMV). It has been suggested that the variable domain in SD2b, which is absent in Q2b, is the major contributor to the stronger suppressor activity of SD2b. Therefore a possible correlation between relative strength of silencing suppression activity of the X4 protein and the strength of the symptoms developed in ToRSV can not be excluded.

Another possible pathway for suppressing the silencing activity in plant will be to induce degradation of a key protein of the silencing pathway. As mentioned in Chapter 1, the 2b protein of Cucumoviruses binds to and inactivates an AGO1 effector protein of silencing. The P0 also targets the AGO to degradation since X4 is likely to be targeted for degradation, this raise the interesting possibility that it may also enhance degradation of other protein interaction partners, such as the key host proteins in the silencing pathway. However future experiment to test the decrease in the stability of these proteins in interaction with the X4 protein is required.

Do virus derived siRNAs down regulate the host genes implicated in defence response? It is possible that the virus derived siRNAs share homologies with plant defence genes and therefore target these genes to degradation. In this case it would be beneficial for the virus to only partially inhibit the silencing activity.
When I started this project, we only had the sequence of the Rasp2 isolate, the X4(64-65) antibodies and preliminary evidence that these antibodies recognize a 60 kDa protein in plants. The research presented in this thesis has increased our understanding about the variability of the X4 protein among ToRSV isolates and has provided preliminary evidence that X4 may act as a suppressor of silencing in plants. However, several questions regarding the nature of this protein and its function remain unanswered. To address these questions and to continue the research on the X4 protein, I discuss below some possible projects and ideas.

5.2 Future projects

As discussed above several important questions were raised in this thesis. These questions need to be addressed to provide a better understanding of the X4 protein. The first project would be to confirm the possible function of the X4 protein as a silencing suppressor in ToRSV infected plants. Using the GFP transient expression assay, it would be necessary to examine the concentration of GFP mRNA and siRNA extracted from patches co-infiltrated with agrobacteria carrying the pBIN-GFP plasmid and pBIN-X4 using Northern blot analysis. Since the sequence of the X4 protein (especially the region containing the tandem repeats) is different among ToRSV isolates, it would be interesting to compare the silencing suppressor activity of each of these isolates in *N. benthamiana* plants. I have produced a collection of deletion mutants of X4-Rasp1 and X4-Rasp2 (data not shown), which could be used to identify the protein domains involved in the silencing suppression. The next step would be to look at the ability of X4 to suppress local silencing or systemic (cell-to-cell or long-distance) spread of the RNA silencing signal in GFP-transformed *N. benthamiana* line 16c (transgenic line 16c, constitutively expressing green fluorescent protein) and analyse the effects of the X4 protein on GFP siRNA accumulation in wild-type *N. benthamiana* and transgenic *N. benthamiana* 16c plants. The result may provide further information on the mechanism of the X4 silencing suppressor activity in plants. It has been suggested that tobacco etch virus HC-Pro, as well as tombusvirus p19 and closterovirus p21, act by binding to the double-stranded siRNAs and prevent its loading into the RNA-induced silencing complex (Lakatos et al., 2006). It would be interesting to test the binding affinity of the X4 protein with the siRNA duplexes. Duplex siRNA-binding silencing suppressors inhibit the target RNA cleavage in planta. The X4 silencing suppressor activity possibly detected in *N. benthamiana* plants was transient, and may be correlated with the low level of expression of the protein (Chapter 4) or its possible instability (discussed in Chapter 3, 4 and above in this
chapter). It may be useful to add proteasome inhibitors in combination with agroinfiltration and minimize the possible degradation of this protein.

The second project would be to test whether the PEST signal identified in the X4 sequence (Chapter 1) is a possible degradation motif by point mutation analysis of the predicted PEST sequence. The X4 protein (or mutated versions of this protein) could be expressed in plant or protoplasts and their stability could be examined (Karsies et al., 2001). It would be interesting to look at predicted phosphorylation sites that are localized to the putative PEST region and look at the effect of possible post translational modifications in the regulation of the protein metabolic stability.

The third project would be to look at the suppressor activity of X4 in the degradation pathway in plants and its activity to bind to host proteins that are responsible in the silencing activity like the AGO1 effector protein. Also it would be interesting to look at the expression level of AGO1 when X4 is expressed in plant (Baumberger et al., 2007; Pazhouhandeh et al., 2006).

The fourth project would be to investigate the efficiency of nematode transmission in different ToRSV isolates. Based on selection pressure, nematode transmissibility may be abolished in the viral genome of ToRSV isolates after extensive passages of the virus in herbaceous hosts. It would be interesting to compare ToRSV isolates in the lab with different number of protein repeats (Chapter 2) with the nematode transmissible ToRSV isolate from the field. Serial mechanical passages of the virus would eliminate the selection pressure to keep sequences that are important for nematode transmissibility of the virus. Isolates of tobraviruses that have lost their nematode transmissibility after serial passages of the virus on herbaceous hosts contain RNA2 molecules of widely different lengths. Internal deletions in RNA2 removed a large portion of the CP gene which abolished the nematode transmissibility of the PPK20 isolate (Hernandez et al., 1996; Hernandez et al., 1997). Similarly, it would be interesting to use ToRSV nematode transmissible isolates from the field, sequence them and then do serial passaging by mechanical inoculation of infected extracts and look at the evolution of the X4 protein and the possible loss of its nematode transmissibility.

The fifth project is related and would involve sequencing a larger number of ToRSV isolates, especially the X4 region to further examine the diversity of X4 among ToRSV isolates. I
have already sequenced the X4 protein of the Rasp1 and PYB isolates (Chapter 2). Based on the X4 sequence, Rasp2 and PYB had a similar protein size and Rasp1 had an extra number of tandem repeats. Therefore, the X4-Rasp1 protein was larger in size than the other two. The results from reverse transcription-polymerase chain reaction (RT-PCR) showed that another isolate, Chickadee, was even larger in size compared to Rasp1 (Chapter 4). Therefore, it would be fascinating to gather more ToRSV isolates particularly the ones from North America. This study would provide valuable information on the evolution of ToRSV.

The sixth project would be to look at the stability of the X4 protein and the possible function of its protein tandem repeats in the process of protein degradation. I have discussed the possible role of tandem repeats in protein–protein or protein-ligand interaction, it would be useful to also look at the stability of X4 protein and study the possible role of the tandem repeats in protein degradation. The X4 protein sequence does have few predicted PEST sequences (data not shown) (a peptide sequence which is rich in proline (P), glutamic acid (E), serine (S), and threonine (T). This sequence is associated with proteins that have a short intracellular half-life. It is hypothesized that the PEST sequence acts as a signal peptide for protein degradation (Rogers et al., 1986). Studying the stability of the X4 protein can be performed by pulse-labelling and immunoprecipitation of the X4 protein, as previously conducted for the coat protein and the movement protein (Sanfacon et al., 1995).

The seventh project would be to look for protein-protein interaction of the X4 protein and host factors that play a critical role in virus-host interaction using N. benthamiana yeast two hybrid cDNA library system (Goodin et al., 2009). Since different ToRSV isolates have different symptoms in N. benthamiana and since X4 is a variable protein among these isolates, it would be interesting to compare the host factors interacting with X4 from various isolates and also compare the binding efficiency of the X4 protein to its interacting partner among the isolates.

The eighth project would be to investigate the nature of the larger proteins that were detected by the X4(64-65) antibodies in purified virus preparations or sucrose gradient (Chapter 3) by liquid chromatography-mass spectrometry (LC-MS-MS).

The final project would be to study the role of X4 or the X3-X 4 precursors in the RNA2 replication of ToRSV. The X4 protein may be representing a multifunctional protein in
ToRSV replication cycle. RNA1 replicates independently of RNA2 and RNA2 needs the RNA1 replication machinery for its own replication in GFLV. In order to perform similar experiment in ToRSV this requires infectious transcripts of ToRSV. The infectious transcripts are not available for ToRSV in our laboratory. However there are infectious transcripts for other nepoviruses. It would be interesting to replace the X3-X4 or the X4 in their coding region and make deletion mutation of the RNA2 and look for any possible accessory role of these proteins in the replication cycle. The 2a protein in RNA2 has an accessory role in RNA2 replication. In GFLV the 2a protein was located into juxtanuclear area in C. quinoa protoplasm (Gaire et al., 1999). It would be also interesting to look at the sub cellular localization of the X4 protein in plant cells. For this purpose, fusion protein of X4 with fluorescent protein such as GFP will be useful. I have constructed plasmids allowing the expression of the X4-Rasp2-GFP and GFP-X4-Rasp2 fusion proteins (data not shown). Similar expression vectors allowing the expression of X4 from different ToRSV isolates fused to GFP could be constructed and tested using biolistic delivery or agroinfiltration.

5.2.1 Potential impact of this research on the future on molecular virology research

It was previously believed that nepoviruses were especially susceptible to plant RNA silencing and that they did not encode a suppressor of silencing, explaining the symptom recovery phenomenon observed in plants infected with nepoviruses (Chapter 4). In this thesis, I provide preliminary evidence that suggests ToRSV X4 protein may play a role as a silencing suppressor in N. benthamiana plants. The suppressor activity may represent a novel class of suppressors. The presence of two types of protein repeats in the X4 protein that may represent novel classes of a. a. repeats is also an interesting new feature of the X4 protein. The results in this thesis will enrich our understanding of the possible function(s) of the ToRSV X4 protein and its possible protein–protein interaction with their hosts.
5.3 Bibliography


