2004

DISSECTING THE FUNCTIONS OF CARMOVIRUS AND TOMBUSVIRUS REPLICASE PROTEINS

Kottampatty Rajendran
University of Kentucky

ABSTRACT OF DISSERTATION

KS. Rajendran

The Graduate School

University of Kentucky

2004

DISSECTING THE FUNCTIONS OF CARMOVIRUS AND TOMBUSVIRUS REPLICASE PROTEINS

ABSTRACT OF DISSERTATION

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

By

KS. Rajendran

Lexington, Kentucky

Director: Dr. Peter D. Nagy, Associate Professor of Plant Pathology

Lexington, Kentucky

2004

ABSTRACT OF DISSERTATION

DISSECTING THE FUNCTIONS OF CARMOVIRUS AND TOMBUSVIRUS REPLICASE PROTEINS

Replication of genetic material is the most important and central process during the viral life cycle. Most RNA viruses assign one or more proteins translated from their own genome for replicating genomic RNAs. Understanding the various biochemical activities of these replication proteins is the aim of this dissertation research. The replicase proteins of *Turnip crinkle virus* (TCV) and *Tomato bushy stunt virus* (TBSV) were selected for this study. Both viruses have small, messenger-sense, single-stranded RNA genomes. Replicase proteins – p28/p88 of TCV and p33/p92 of TBSV- were expressed and purified from *E. coli* as N-terminal fusions to maltose binding protein. *In vitro* assays revealed that the recombinant p88 has RNA-dependent RNA polymerase (RdRp) and RNA-binding activities. Deletion of the N-terminal p28 domain in p88 resulted in a highly active RdRp, while further deletions at both N- and C-terminal ends abolished RdRp activity. Comparison of p88, the N-terminal p28-deletion mutant of p88 and a TCV RdRp preparation obtained from infected plants revealed remarkable similarities in RNA template recognition and plus and minus strands synthesis. Contrary to recombinant TCV

RdRp activities under similar experimental conditions. p33 preferentially binds to single-stranded (ss) RNA with positive cooperativity *in vitro*. The RNA binding activity was mapped to arginine/proline-rich motif (RPR-motif) at the C-terminus of p33 and the corresponding sequence in p92. The non-overlapping C-terminal domain of p92 also contained additional RNA-binding regions that flank the conserved RdRp motifs on both sides. Cooperative RNA binding by p33 suggested inter-molecular interactions between p33 monomers. Indeed the yeast two-hybrid and surface plasmon resonance assays revealed interactions between p33 and p33 and also between p33 and p92. The sequence involved in the protein-protein interactions was mapped to the C-terminal region in p33, proximal to RPR-motif. Within this region, mutations introduced at two short stretches of amino acid residues were found to affect p33:p33 and p33:p92 interactions *in vivo* and also decreased the replication of a TBSV-defective interfering RNA in yeast, a model system, supporting the significance of these protein interactions in tombusvirus replication.

KEYWORDS: *Tomato bushy stunt virus*, *Turnip crinkle virus*, RNA-dependent RNA polymerase, RNA-protein interactions, protein-protein interactions

Kottampatty S. Rajendran

Date: 04/21/04

DISSECTING THE FUNCTIONS OF CARMOVIRUS AND TOMBUSVIRUS REPLICASE PROTEINS

By

KS. Rajendran

Peter D. Nagy
Director of Dissertation

Lisa J. Vaillancourt
Director of Graduate Studies

Date: 04/21/04

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctors degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

DISSERTATION

KS. Rajendran

The Graduate School

University of Kentucky

2004

DISSECTING THE FUNCTIONS OF CARMOVIRUS AND TOMBUSVIRUS REPLICASE PROTEINS

DISSERTATION

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

By

KS. Rajendran

Lexington, Kentucky

Director: Dr. Peter D. Nagy, Associate Professor of Plant Pathology

Lexington, Kentucky

2004

ACKNOWLEDGEMENTS

First of all I would like to thank my grand parents, parents and my sister who helped me reach this stage through their enormous amount of hard work and prayers. I am indebted to them forever and dedicate this dissertation to them. Peter D. Nagy, my major advisor, is an avid researcher and was never short of pushing me to excel in science. His never-diminishing interest in science kept me busy with my work throughout my graduate studies. A big thanks to Peter. I also place my appreciation for my committee members: Dr. Christopher L. Schardl, Dr. Said A. Ghabrial and Dr. Arthur G. Hunt, for serving in my committee and helping me when approached. I thank Dr. Mark Young, my outside examiner for evaluating my dissertation and giving a seminar to the department. We students are very lucky to have a highly active and caring Director of Graduate Studies Dr. Lisa Vaillancourt and her initiatives like Graduate Students News Letter and Departmental Web pages are to be lauded. Thanks to Judit Pogany, who, insightful and experienced with laboratory methods, was very helpful in troubleshooting experiments. Hearty appreciation and thanks goes to Natasha, Chi-Ping and Gaby who kept the work environment lively and enjoyable. I also thank Murali, my long time friend from my under-graduate years, for providing me with good company and food for all these years. Murali, Natasha and Oles Vaskin made me feel “at home” and I am grateful to them for being good friends. Josef Stork tolerated all my friendly punching, shared Hungarian jokes and stories to keep me awake in the evening hours in the lab and arranged several eat-outs in the commonwealth village lawns. Thanks Stori. I also thank Jannine Baker for her Halloween party and for occasional lunch-outs in addition to her strides in keeping the lab worker-friendly. I also like to appreciate the material, technical and advisory help provided by Tadas, Zivile, Saulius, Elena, Josef Gal. Additionally, I wish to thank the professors, the friendly office staff, students, post-docs and technical staff of the department for their help in one way or the other in making this dissertation possible. Thank You Dave for your “how-to” lessons on Mac system.

In the end I reserve my special thanks and appreciation for my wife who always puts my interest first before hers. Every one of my friends says that I am lucky to have her. Indeed I am. Thank You Anu. I am grateful to Dr. YR. Sarma, my father-in law and my previous boss for his unflagging support and encouragement. I also thank my mother-in law for her kindness and blessings.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FILES

Rajendra.pdf …… 3.6 Mega Bytes

CHAPTER ONE

INTRODUCTION

REPLICATION OF SINGLE STRANDED POSITIVE SENSE RNA VIRUSES

RNA viruses with positive strand polarity are the largest group of plant viruses and also include many important human pathogens such as hepatitis C virus (HCV), poliovirus, common cold rhinoviruses and severe acute respiratory syndrome (SARS) coronovirus. This group includes viruses from over one third of all virus genera and causes severe diseases in plants, humans and animals. They are intracellular obligate parasites and utilize host cellular factors in concert with their own genome-coded proteins for expression and replication of their genomes. The replication process of these viruses is remarkably similar despite differences in their genome organization, gene expression, virion morphology, mode of spread and host range. Therefore, studies on RNA viruses are of great importance as it might lead one day to the development of a common strategy to counter infections by many different viruses.

Genomic RNA replication is the central process in the reproductive cycle of positive sense RNA viruses. Viral RNAs participate in many different processes in the viral infection cycle such as replication, translation, movement and encapsidation. Upon invading the host cells, the positive strand genomic RNA serves as mRNA for the expression of 5' proximal open reading frames (ORFs), which usually code for replicase proteins. Replicase proteins recruit viral RNAs and also presumably some host factors to assemble replication complexes that are usually associated with host membranous structures (4, 98, 132), One of the viral replicase proteins possess the RNA dependent RNA polymerase (RdRp) activity which recognizes the 3' end of the viral genomic RNA (promoters) and direct the synthesis of complementary RNA (negative strand) intermediates, which in turn are used as templates for genomic RNA (positive strand) synthesis. The ratio of positive to negative strand viral RNAs is higher in infected cells leading to asymmetrical accumulation of genomic RNAs. This ratio is likely regulated by interactions between various *cis*-elements located in both strands of RNA and replicase proteins (93). This suggests that RNA-RNA, RNA-protein and protein-protein interactions between viral RNA, viral replicase and host factors are important in regulating the viral replication process.

REPLICASE PROTEINS OF POSITIVE STRAND RNA VIRUSES

VIRAL REPLICASE COMPLEX: RNA viruses code for their own replicase proteins, which are necessary for replication of viral RNAs. These proteins are encoded either in the 5' end of the polycistronic viral RNAs in non-segmented genomes or in separate monocistronic RNAs in segmented genomes. Viral replicase, also often referred to as RNA dependent RNA polymerase (RdRp), is a multi-subunit enzyme consisting of virus and host encoded proteins (4). The virus-coded subunits often include RNA polymerase (RdRp), RNA helicase, protease and methyltransferase activities associated with replicase complex. Some viruses, for example potexviruses pack polymerase, helicase and methyltransferase domains in one large (~160 kDa) protein (75). Many other groups of viruses such as alphavirus, tobamovirus code for more than one protein subunits with these activities divided among them. A unique strategy used by several different plant RNA viruses such as *Tobamovirus, Tombusvirus, Carmovirus, Furovirus, Tobravirus* and *Necrovirus* is to encode their two replicase subunits in the 5' most ORF in the genomic RNA and express the larger subunit as a readthrough product (159).

The comparative sequence analysis of replicase proteins revealed the presence of highly conserved motifs indicative of the above-mentioned enzymatic activities (65). However, only the viral subunit with polymerase motifs is found in all positive strand RNA viruses (65). Helicase motifs are usually found in almost all positive strand RNA viruses that have genomes larger than 5.8 kb (56, 65). RNA helicase is required during viral RNA replication for unwinding the local secondary structures formed due to intramolecular base pairing and the duplex of template strand and the newly synthesized complementary strand. Tombus- and Carmoviruses – the subject of this dissertation research have a genome of less than 5 kb and comparative sequence analysis revealed the lack of helicase- like motifs in their replicase proteins. Only viruses with capped RNA genomes - all members of Sindbisvirus-like supergroup - code for putative methyl-transferase-like motifs typically near N-terminus of replicase proteins (127). In addition to the above-mentioned enzymatic activities, virus encoded replicase subunits also perform several other functions that include recruitment of viral RNAs, virus- and host-coded proteins to replication sites inside the host cells.

Qβ replicase of single strand RNA coliphage Qβ is the best-studied viral replicase. This holoenzyme complex consists of four subunits: virus-coded RdRp (β subunit), host encoded

ribosomal protein S1 (α subunit) and elongation factors EF-Tu (γ subunit) and EF-Ts (δ subunit) (15). Ribosomal protein S1 binds to RNA template specifically and initiates the RNA polymerization by RdRp. The role of elongation factors remains to be understood fully. One other host protein termed as host factor was also found to be part of this enzyme complex for the synthesis of minus strand replication intermediates and affects the plus- and minus-strand ratio (142). The membrane fractionated RdRp preparations purified from plants infected with tombus-, tobamo- and bromo-virus contained virus-encoded replicase proteins in addition to some host factors (4, 68, 119, 133, 151). However, the identity and bio-chemical functions of these host factors in viral RNA replication are not yet established well (4). An integral membrane protein TOM1 of *Arabidopsis* has been found to be associated with TMV replicase and RdRp activity (44). The purified BMV replicase complex contained a 45 kDa protein similar to a p41 subunit of eukaryotic initiation factor 3 (eIF3) or a closely related protein (119). Poliovirus RdRp ($3D^{pol}$ protein) complex includes a host factor subunit with terminal uridylyltransferase (TUTase) activity which is believed to synthesize primer for RNA replication (5). Cucumber mosaic virus and turnip yellow mosaic virus RdRp also copurified with host factors whose identity is not characterized yet (48, 82). There is also a very little information on how these three components–viral RNA, viral replicase and host factors-interact with each other.

RdRp - SEQUENCE MOTIFS AND STRUCTURE: Comparative sequence analysis among viral RdRps, coded by positive-, negative- and double-strand RNA viruses, showed the presence of several conserved regions (7, 58). One of the most conserved regions – YGDD was also found to be present in RNA-dependent DNA polymerases or reverse transcriptases (RT) and DNA-dependent DNA polymerases (DdDp). Subsequent studies established the existence of four highly conserved motifs (A, B, C and D) in all of the above polymerases and a fifth motif (E), unique to RdRp and RT (45, 114). The motifs A-D are arranged in the same linear order with consistent inter-motif distances in RdRps and RTs but in DdDp, the inter-motif distances are inconsistent (114). Alignment of representative RdRp aminoacid sequences from all groups of positive-strand RNA viruses revealed the presence of eight (I - VIII) distinctly conserved motifs (65). The amino acid composition of motifs A, B, C and E is identical to motifs IV, V, VI and VII described above respectively.

The first crystal structure of RdRp emerged only recently in 1997 with poliovirus $3D^{pol}$ RdRp (45). Currently, the crystal structures of phi6-dsRNA bacteriophage (22), rabbit hemorrhagic disease virus (92), HCV (1, 16, 71) and reovirus RdRps are also available (145). All these structural data revealed that the RdRps fold into a "right hand" like structure with palm, fingers and thumb subdomains. The overall structure of RdRp resembled that of other polymerases. The palm subdomain of poliovirus RdRp consists of five motifs (A-E), of which four (A-D) are also observed in the palm subdomains of all other polymerases (94). The functions of motif A (or motif IV) include magnesium coordination and sugar selection and the two universally conserved aspartate residues are thought to be involved in these activities. Motif B (or Motif V) may play a role in the selection of ribose versus deoxyribose sugars. The two adjacent aspartates in motif C are conserved in all RdRps and are postulated to be the catalytic center and involved in divalent cation (Mg^{2+}) coordination. The functions of motif D is unclear, however it completes the palm domain of viral RdRp. Positioned between palm and thumb domains, hydrophobic residues of motif E may mediate a hydrophobic interactions between palm and thumb, thus stabilizing the structure as observed in poliovirus (94). Though the palm domain of poliovirus is remarkably similar to palm domains of other polymerases, the fingers and thumb domains are structurally different from corresponding domains in other polymerases.

The structure of HCV RdRp shows the major canonical features of polymerases. However, the extensive interactions between fingers and thumb domains give a closed -structure appearance with enzyme active site covered inside (71). The structure of HCV RdRp also revealed a β-loop, conserved among other viral RdRps in amino acid sequence and structure, which is termed as motif F (71). These studies show that the viral RdRps share remarkable structural similarities to other polymerases even though there is very little sequence identity between them.

RdRp - POLYMERASE ACTIVITY: The core activity of RdRp coded by RNA viruses is to synthesize complementary RNA molecules using template RNAs. This activity is required for genome replication, genomic and sub-genomic messenger RNA synthesis, RNA recombination and other processes (3). Many of the RdRps have been identified based on the conserved sequence motifs (65, 94). The actual biochemical functions of RdRp have been characterized only for a few (14, 38, 51, 52, 57, 62, 69, 75-77, 112, 117, 150) using highly purified RdRps

from heterologous expression hosts. HCV and poliovirus RdRps were studied in detail for polymerase, RNA-binding, RNA-recombination and oligomerization activities (13, 14, 38, 51, 62, 76, 77, 112, 117, 121, 146, 150). Poliovirus RdRp has been shown to synthesize complementary RNA in a primer-dependent manner (112, 121). At high RdRp concentration, template binding and elongation rate of polymerase is highly cooperative (112). The poliovirus polymerases produced from infected cells and heterologous system were shown to be identical in template specificity, elongation rate (121).

Similarly, earlier studies with HCV RdRp expressed in insect cells found that the RdRp synthesized complementary RNA by copy-back priming mechanism (14, 38, 76). However, recent studies indicate that *E. coli* expressed recombinant HCV RdRp can initiate RNA synthesis *de novo* without the need of primer (77, 160) as expected to happen in infected cells. Interestingly, the bovine viral diarrhea virus (BVDV) RdRp could utilize circular single stranded DNA as a template for RNA synthesis in addition to single stranded viral RNA (69). In contrast, RdRp encoded by a potyvirus exhibited preference for poly(A) over poly(dA) (54). The difference in template utilization by BVDV and potyvirus RdRp is surprising given the fact that both viruses have single stranded RNA messenger-sense genome coding for a large polyprotein. The reasons for such differences in viral RdRps are not understood yet. The bamboo mosaic potexvirus RdRp was also shown to initiate de novo RNA synthesis on viral RNA template (75).

In the host cellular milieu, RdRps work in concert with other accessory viral and host proteins in a replication complex. Such RdRp complexes were isolated from plants infected with viruses including bromovirus (48, 61, 118), tobamovirus (98), tombusviruses (86) and turnip crinkle virus (TCV) (140). *In vitro* replication assays revealed that tobamovirus could carry out complete replication cycles (98), whereas other plant purified RdRps could complete usually the first stage - transcription of complementary RNA - of replication process (86, 118, 140). RdRp complexes purified from TCV infected plants have been shown to transcribe (+) and (-) strands of satellite and defective interfering (DI) RNAs (140). The role of TCV RdRp in recognizing the recombination hotspot and replication enhancer in TCV satellite RNA was also reported (91). Similarly, tombusvirus RdRps purified from cucumber necrosis virus (CNV) and tomato bushy stunt virus (TBSV) – infected plants have been shown to efficiently recognize the *cis*-elements on viral RNAs and catalyze complementary RNA synthesis from both (+) and (-) strand templates (86). Tombusvirus RdRps also initiated complementary RNA synthesis on TCV-

associated satellite RNAs. Comparison of TCV RdRp and tombusvirus RdRps suggested similar but not identical template usage (86). CNV RdRp also efficiently recognized the enhancer and silencer elements of replication found in tombusvirus defective interfering RNAs (DI-RNAs) (104, 115). Important *cis*- acting viral RNAs elements involved in replication and recombination were discovered using partially purified tombusvirus and TCV RdRps derived from virus infected plants (43, 86, 87, 91, 104, 107, 108, 115). These studies mainly addressed the bio-chemical features of RdRp complex and the contribution of individual components is not well understood yet. There is still a wide gap in knowledge on how accessory viral replicase proteins and/or host factors regulate or modify the inherent polymerase activity and template specificity.

INTERACTIONS BETWEEN REPLICASE PROTEINS AND VIRAL RNA

With the existing knowledge, it can be postulated that viral RNA in the cellular milieu exists primarily as ribonucleoprotein complex at different stages of infection process. The protein components in the ribonucleoprotein complex likely vary depending on the stages of infection cycle the viral RNA goes through. For instance, during replication, viral RNAs are associated with replicase proteins and with movement proteins during inter- and extra -cellular movement, and then with coat protein during virus assembly.

Replicase proteins mediate several processes during viral RNA replication such as recruitment of RNAs to site of replication, recognition of *cis*-acting elements on viral RNAs to initiate the replication, synthesis of nascent plus and minus RNA strand, unwinding the template and nascent RNA duplex to free RNA strands for further rounds of replication, capping the 5' end and polyadenylating the 3' end of genomic RNAs etc. All of these processes require interactions between replicase proteins and viral RNA. Despite its significance, there is very little information available for RNA-binding properties of replicase proteins, while majority of the RNA-binding studies focused on movement and coat proteins of plant RNA viruses.

RNA-binding proteins are major regulators of gene expression in all forms of life. Structural and sequence analyses have identified conserved structural and sequence motifs in these proteins (10). These RNA binding motifs (RNP) also referred to as RNA recognition motif (RRM), are usually composed of 90-100 amino acids with RNP consensus sequence (RNP-CS), which is divided into two short regions called RNP1 and RNP2 and the whole region forms an RNA binding domain (RBD). The structural features of the RBD in many of the proteins are also

conserved typically with four anti-parallel β-strands packed against two α-helices (βαββαβ) (19). Most of the cellular RNA binding proteins contains one or more copies of RNP motifs. Surprisingly, RNP motifs are not as widely reported in viral RNA binding proteins as was for eukaryotic RNA binding proteins.

Interestingly, viral, bacteriophage and also ribosomal RNA binding proteins often contain short arginine rich sequences, referred as arginine-rich motif (ARM). Unlike RNP motifs, positions of arginine in or the structure of the ARM motif are not conserved between proteins (19). HIV1 *Rev* and *Tat* proteins are the structurally and functionally best- characterized ARM proteins with totally different structures (34). ARMs were also found in coat proteins of brome mosaic virus (BMV), cucumber mosaic virus (CMV), satellite panicum mosaic virus (sPMV) and tobacco streak mosaic virus (TSV), plum pox potyvirus cylindrical inclusion (CI) protein, delta antigen of hepatitis D virus, satellite bamboo mosaic virus P20 and rous sarcoma virus nucleocapsid proteins (37, 70, 147). Prior to my studies, there were no reports of ARM in viral replicase proteins.

Unlike the replicase proteins of plant RNA viruses, the RNA binding activities of replicase proteins of human viruses such as poliovirus and HCV have been studied in great details at both functional and structural levels. The poliovirus RdRp shows lower affinity for RNA, however at higher concentration the recombinant poliovirus polymerase binds to RNA with positive cooperativity and the minimal binding site on RNA consists of 10nt (13, 112). The studies thus far indicate that poliovirus polymerase binds RNA non-specifically as was reported for HCV polymerase (76). The crystal structures of polio and HCV viruses RdRp revealed that basic amino acid residues come together and form a lattice of positive potential which then attracts RNA and the complex is stabilized by hydrogen bonding around the contacting residues and nucleotides (45, 71).

INTERACTIONS BETWEEN VIRAL REPLICASE PROTEINS

Most RNA viruses code more than one replicase proteins with different functional domains distributed in different proteins. These viral proteins then associate together along with some host proteins and membranes to form an active replication complex. The interactions between viral protein subunits of replication complex are very important for formation of the complex as well as the maintenance of its structure and as a consequence for viral RNA

replication. Bromovirus group heralded protein-protein interaction studies between replicase proteins and also between viral replicase and host proteins. It was shown that BMV replicase protein 1a interacted with itself at the N-terminal methyltransferase-like domain and its C-terminal helicase-like domain interacted with N-terminus of 2a (97). The interactions of 1a with itself, 2a and genomic RNA were necessary for the assembly of spherular replicase complex on endoplasmic reticulum (93). Similarly, TMV replicase proteins 128K/183K were shown to interact with each other within the helicase-like domain in 128K subunit (42, 151). Interestingly, there are no reports of oligomerization of replication proteins containing polymerase domains (RdRp) such as TMV 183K and BMV 2a. On the other hand, oligomerization of poliovirus and HCV RdRp has been shown to be essential for efficient catalytic activity (112, 117, 150). The oligomeric poliovirus RdRp exhibited cooperative RNA synthesis indicating that polymerase-polymerase interaction is an important regulator of viral replication (150).

TOMBUSVIRUS AND CARMOVIRUS AS MODEL SYSTEMS

Tombusviruses and Carmoviruses, the subject of this dissertation research, are single stranded positive sense monopartite RNA viruses that replicate their genomes through asymmetrical replication process described above. Both Tombus- and Carmoviruses belong to the same family *Tombusviridae.* Tomato bushy stunt virus (TBSV) is the type species of the genus Tombusvirus, whereas Turnip crinkle virus (TCV) is the prominent member of the genus Carmovirus. Both TBSV and TCV are among the most intensively studied RNA viruses. Their genomic RNA is encapsidated in spherical particles made of 180 subunits of capsid proteins.

TBSV is the first icosahedral virus to be crystallized and the atomic structure determined (47) and is also one of the best-studied viruses from a structural perspective (46). TBSV is also an excellent model system to study viral replication and recombination due to several reasons. The first and foremost is that they generate defective interfering RNAs (DI-RNAs) during serial passage of viral RNAs in plants and protoplasts under experimental conditions (50) and also satellite RNAs in the natural infections (63). Second, they support replication of DI-RNA (replicon) in yeast (*Saccharomyces cerevisiae*) transformed with plasmids coding for replicase proteins and DI-RNAs (105, 111). Yeast, being an extensively studied genetic system, offers a useful tool to study replication, recombination and virus-host interactions. Third, full-length

cDNA clones of TBSV and related viruses, powerful experimental methods such as in vitro replication assays with purified viral RdRp complexes are available (86, 106-108).

DI-RNAs are small (~400-600) sub-viral molecules composed of non-contiguous RNA segments derived from viral genomes (63, 154). These RNAs are defective because they lack the ability to replicate independently of the helper virus. They generally lack protein-coding regions; therefore depend on replicase proteins of parent virus to replicate their genomes (63). The non-coding nature of DI-RNAs offers a unique advantage because unlike viral RNAs the changes in DI-RNA sequence do not directly affect expression of the essential genes from the helper virus. Because DI-RNAs are short and likely possess the regions of parent virus with important *cis*-acting regulatory elements, they replicate very efficiently and they can compete and interfere with the parent virus replication (154). Therefore DI-RNAs, with their small and simple genomes, non-coding nature and robust replication, are ideal for studies on replication, recombination and virus- host interactions.

TCV is also a well-studied virus with many of the advantages listed above for TBSV such as smaller genomes, ability to generate satellite RNAs and availability of cDNA clones, optimal virological methods etc (90, 135).

TBSV AND TCV GENOME

TBSV genome comprises of a single positive strand RNA of ~ 4.8 kb in length. It is neither 5'-capped nor 3'-polyadenylated (49). It contains five open reading frames (ORFs) flanked by untranslated regions at 5'-(~170nt) and 3'-(~350nt) ends (Figure1.1). The 5'- proximal ORFs 1 and 2 code for 33 and 92 kDa proteins, which are translated directly from the messenger-sense genomic RNA. The 92kDa (p92) protein is a readthrough product of p33 termination codon and hence, shares the entire p33 sequence at its amino-terminal end. Viral coat protein (CP), movement protein (p22) and host gene silencing suppressor protein (p19) are translated from ORFs 3, 4 and 5 respectively. The expression of these three proteins requires transcription of subgenomic viral mRNAs (130) because 3'-proximal genes are not accessible to eukaryotic host translation machinery (66).

The genome of TBSV DI-RNAs typically contains four non-contiguous regions derived genomic RNAs (Figure 1.1). RI contains the entire 5'-UTR plus the start codon of p33/p92 genes whereas RIV is derived from the very end of 3' UTR. RII includes internal coding sequences of

p92 gene (downstream of p33 stop codon) and RIII comprises of sequences from the 3' end of p22 gene and the beginning of 3' UTR. The prototypical DI-RNA, termed DI-72 is about 600 nt long although there have been DI-RNAs with varying size between 400 to 700 nt observed during TBSV infections (154).

TCV has a ~ 4 kb long single stranded positive sense RNA genome which codes for five proteins. Like TBSV RNA, TCV RNA is also neither 5'-capped not 3'-polyadenylated. Replicase proteins p28 and its readthrough product p88, related to p33/p92 of TBSV, are expressed directly from genomic RNA. The downstream genes coding for coat protein (CP) and movement proteins (p8 and p9) are translated from the sub-genomic RNAs transcribed from the genomic RNA (24).

Sub-viral molecules such as satellite C (satC) and satellite D (satD) RNAs are naturally associated with TCV infections (74, 134). satC is a chimeric RNA generated through recombination between satD and TCV genomic RNAs. satD has no sequence similarity with genomic RNA except for the last 7 bases at its 3' end (135). Among these RNAs, satC is infectious and increases the severity of symptoms induced by helper TCV virus whereas satD has no apparent effect on viral infection.

TBSV REPLICASE PROTEINS

TBSV replicase proteins p33 and p92 have been shown to be essential and sufficient for viral RNA replication (103, 133). They are also required for transcription of subgenomic RNAs. Production of p92 is attenuated with respect to p33 via translational readthrough of amber stop codon of p33. In effect, the relative accumulation in the infected protoplasts of p33 and p92 has been estimated to be 20:1(133). This type of translation regulation of replication genes is common in Sindbis virus, tobamovirus, tombusvirus and carmovirus groups (93). Modification of the ratio of polymerase to other replicase proteins resulted in quick reversion to wild type level indicating a tight regulation of concentration of replicase proteins in infected cells (55, 72). Replication proteins of TCV p28/p88 are similar to p33/p92 in terms of their organization, expression and sequence relatedness that prompted me to study the functions of TCV and TBSV replicase proteins with major emphasis on the later.

Since p92 has the hallmark motifs of RNA dependent RNA polymerase (RdRp) at its carboxy terminal half (Figure 1.2), it's likely the catalytic subunit of the viral replicase complex

(65, 94). Biochemical studies with partially purified RdRp complex from TBSV infected plants have shown that the complex contained both p33 and p92 proteins, recognized the *cis* regulatory elements on viral RNAs and initiated *de novo* synthesis of progeny strands from positive and negative template viral RNAs (86, 104, 107, 115).

The actual function of the abundantly produced p33, which is essential for replication and also for necrosis phenotype in infected plants (20), is not known yet. The amino acid sequence analysis of p33 has not revealed any known functional motifs or domains identified in similar replicase proteins. However the sequence analysis displayed two integral transmembrane domains at its amino-terminal half (Table 1.1). Indeed, recent studies with other closely related tombusviruses - cymbidium ringspot virus (CymRSV) and carnation Italian ringspot virus (CIRV) have shown that p33-like proteins possess peroxisomal and mitochondrial membrane targeting sequences respectively (128, 153). These results suggest that one of the functions of p33 may be to target viral replicase complexes to host membrane structures, a function similar to that of BMV 1a and TMV 128K replication proteins. These include binding to host membranes, viral RNAs, viral RdRp and host factors. Understanding some of these functions is the major objective of my dissertation research.

RESEARCH OBJECTIVES AND SYNOPSIS

The major objective is to understand the functions of TBSV-encoded replication proteins and their impact on viral RNA replication with major emphasis on RNA-protein and protein-protein interactions between viral RNA and replication proteins. The information gained from this study will help understand the big picture of composition, assembly and function of viral replicase complex and the contribution of individual host- and viral-encoded subunits.

To investigate the role of replicase proteins, I started out with expressing the TBSV and TCV replicase genes in *E.coli* and purified recombinant fusion proteins in collaboration with Judit Pogany and studied the functions of the replicase proteins of both viruses. In the course of study, I found out that the TCV p88 replicase was enzymatically active in transcribing the viral RNAs, whereas TBSV p92 was not functional under the conditions used for the assay. The TCV replicase work is presented in Chapter II. Though the *E.coli* produced TBSV replicase proteins were defective for polymerase functions, they were active in binding to viral RNAs and the work on their RNA binding activities is reported in Chapter III. In addition to interactions between

replicase proteins and viral RNA, the interactions between TBSV replicase proteins were also studied and presented in Chapter IV. The results from these above studies are summarized in Chapter V to explain various functional domains on replicase proteins and their roles in the replication process.

Table 1.1: Predictions of transmembrane domains (TMD)[a] in TBSV proteins using computer softwares listed in EXPasy server at http://www.expasy.org.

TBSV proteins	TMHMM	SOSUI	TopPred2	TMPRED	HMTOP	DAS	Consensus[b]
p33	83-102 132-154	25-47 80-102 131-153	82-102 135-155	83-102 136-154	14-32 84-101 132-150	34-36 88-98 136-150 266-273	83-102 132-155
p92	82-102 131-153	25-47 80-102 131-153	82-102 135-155	83-99 131-151 594-617 714-733	84-101 132-149	88-98 136-150	83-102 132-155
p41	69-91	None	17-37 67-87 298-318 357-377	10-36 71-92 131-152 233-266 278-296 320-340 340-362	20-37 68-85	None	None
p22	None	None	None	None	38-57	None	None
p19	None	None	None	None	None	None	None

Note: [a] The numbers in the table indicate the amino acid positions in the primary sequence. [b] The consensus is arrived at based on the predictions of a TMD by all the six softwares listed in the table.

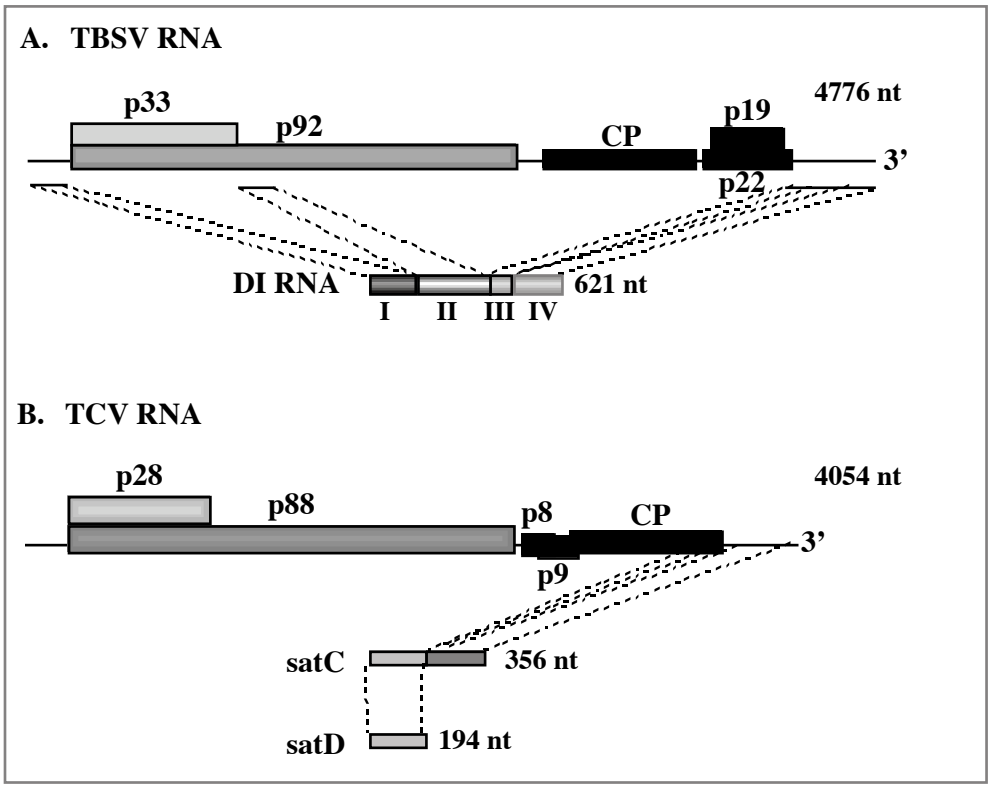


Figure 1.1: Schematic representation of TBSV and TCV genomes and the associated DI- and sat-RNAs. The genomic RNAs of TBSV and the related TCV have five open reading frames, of which two are expressed from the genomic RNAs (shown as open boxes) and three (shown as black boxes) are expressed from two subgenomic RNAs (not shown). The noncontiguous regions on TBSV genomic RNA, which make the DI-RNAs, are depicted with solid bars below the genomic RNA. The TCV-associated satC is a chimeric RNA containing sequences from satD at its 5' region and two 3' terminal regions derived from genomic TCV RNA. satD bears no significant similarity with the genomic TCV RNA (84).

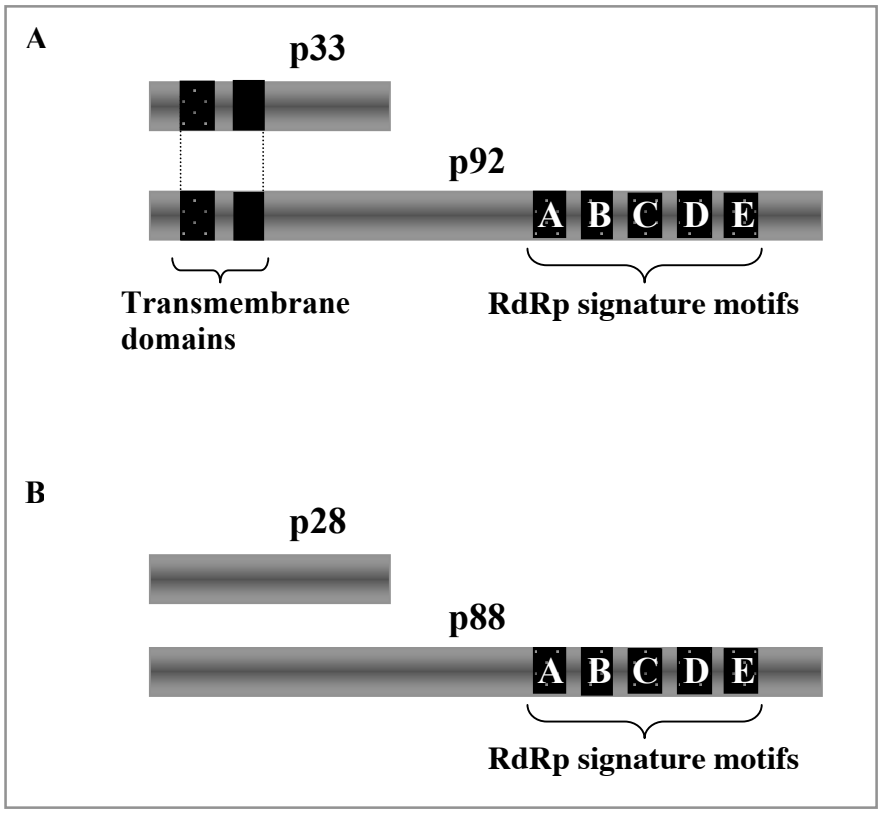


Figure 1.2: Replicase proteins of TBSV and TCV and their previously known functional domains. The inter-motifs distances (A-E) and the transmembrane domains are not represented to scale.

CHAPTER II

COMPARISON OF *TURNIP CRINKLE VIRUS* RNA-DEPENDENT RNA POLYMERASE PREPARATIONS EXPRESSED IN *E. COLI* OR DERIVED FROM INFECTED PLANTS

INTRODUCTION

Replication of RNA viruses is carried out by membrane-bound multi-subunit replicase complexes, which consist of viral- and host-coded proteins (17, 18, 68). The catalytic subunit of the viral replicase complexes is a viral-coded RNA-dependent RNA polymerase (RdRp). Biochemical features of RdRps for several positive-strand RNA viruses, including poliovirus (11), flaviviruses (38, 62, 76, 77, 160), plant poty-(52) and potexviruses (75) have been examined in some detail by using purified preparations obtained from heterologous expression systems. Another approach to obtain viral RdRps is to purify them from virus-infected cells. Indeed, this approach has been documented for several viruses, including bacteriophages (15), brome mosaic virus (BMV) (61, 118), cucumber mosaic virus (48), tobacco mosaic virus (98), turnip yellow mosaic virus (30, 136), alfalfa mosaic virus (120), potexviruses (113) and tombusviruses (86). The RdRp preparations were found to initiate complementary RNA synthesis either with short primers or in the absence of primers (*de novo* synthesis). Specificity in template selection and the stringency of promoter recognition varies among the purified RdRps (17, 18, 35). However, it has been shown that common sequence elements, such as CCA repeats, can be recognized by several RdRps (31, 158), and this suggests that there are significant similarities among some RdRps. This is further supported by the similar structures of three RdRps (1, 16, 22, 45, 71).

Turnip crinkle virus (TCV), a carmovirus, is a well-characterized model plus-strand RNA virus (17, 135). TCV has a small genome (4054 bases) with five genes of which two are required for replication, namely p28 and p88 (157). p88 overlaps p28 and it contains the signature RdRp motifs (94) in the unique C-terminal portion (Figure 2.1). The role of the p28 is not known. In addition, TCV infections are associated with several satellite (sat) RNAs (135), including satD

(194 nt) and satC (356 nt). With the help of an *in vitro* replicase assay based on an RdRp preparation obtained from TCV-infected plants these satellite RNAs were widely used to dissect *cis*-acting elements involved in RNA accumulation (43, 85, 87, 135, 139, 140).

In vitro analyses of sequences that affect minus- and plus-strand synthesis revealed the presence of promoter and enhancer elements that are required for or enhance accumulation of satC. The plus strand contains a 3'-terminal 29-base hairpin promoter that is required for complementary strand synthesis *in vitro* (139). The minus-strand satC contains two RNA elements, one called the 3'-proximal element, and the other, the 5'-proximal element, both of which can function as independent promoters *in vitro* (43). Another *cis*-acting element present on minus-strands of satC, which is important for plus-strand synthesis, is a 30 base hairpin (positions 180 to 209), denoted the motif1-hairpin. The motif1-hairpin was found to function as an RNA replication enhancer for primer independent complementary RNA synthesis (*de novo* initiation) (87). In addition, the motif1-hairpin also facilitates primer-dependent RNA synthesis by the TCV RdRp *in vitro* (88, 89, 91), which can lead to the generation of recombinant RNA molecules by a template switching mechanism (25, 90).

To dissect the functions of the TCV replicase proteins in more detail, I expressed and purified them from *E. coli*. I found that p88 alone possessed RdRp activity. Interestingly, the unique C-terminal fragment of p88, designated p88C also showed RdRp activity *in vitro*. Both p88 and p88C were capable of both *de novo* and primer-dependent RNA synthesis, similar to the plant TCV RdRp. In addition, all three RdRp preparations showed similar template selectivity as well as promoter recognition *in vitro*.

MATERIALS AND METHODS

Construction of expression plasmids. To express p88 gene, which has an UAG termination codon behind the overlapping p28 gene (see Figure 2.1), a mutation was introduced at position 817 to generate a tyrosine codon (157) using sequential PCR. First, the N- and C-terminal portions of p88 were generated separately using Pfu turbo polymerase (Stratagene) and either primer pair p88F (CCTCTTCTACACACACTC) and p28TYR (GTAGCGGACAAAAGAGATC) or primer pair p88TYR (GGGTGCTTGCGGGAGCTG) and p88-stop (CCGTAAGCTTGATT-

AGAGAGTTGTAGGGAATTCG). The N- and C-terminal PCR products were ligated together to generate the modified p88 gene. The ligated product was treated with *Hind*III, followed by cloning into pMAL-c2X (NEB) at the *Xmn*I-*Hind*III sites. Construct p88C was generated by ligating the PCR product (digested with *Hind*III and *Xba*I) obtained with primers 818F (GGAGTCTAGAGATACCATCAAGAGGATG) and p88-stop into pMAL-c2X between the *Hind*III and *Xba*I sites. The other constructs, which were used to express proteins shown in Figure 2.1B, were generated using a method similar to that described for p88C, except using *Eco*RI and *Xba*I sites. The primers used were the following: for construct p28: 88F and 28-stop (GGAGTCTAGACTAGCGGACAAAAGAGATC); for ΔN30/ C100: aa281F (GAGGAATTCA AGGTTCGACGCATCTTC) and aa674F (GAGTCTAGACTATGATTGTAACACTGGTAC); for ΔN100/C30: aa351F (GAGGAATTCGGA-AATCATACCCCTGTG) and aa744F (GAGTCT AGACTAGTCGTAGTACTCCTCAA- G); for ΔN100/C100: aa351F and aa674F. The resulting constructs expressed proteins ΔN100/C30 (24), ΔN30/C100 (amino acid positions 282-675) and ΔN100/C100 (amino acid positions 352-675).

Purification of p88 and its derivatives from *E. coli*. The generated expression constructs (see above), which were used to obtain proteins shown in Figure 2.1B, were introduced into Epicurian BL21-CodonPlus (DE3)-RIL (Stratagene). Protein expression was induced as recommended by the supplier using IPTG. After 8-10 hours induction at 14 ^{0}C, the cells were harvested, collected by centrifugation (5000 rpm for 5 min), resuspended and sonicated as recommended by the supplier, except reducing the amount of NaCl to 25 mM. The samples were then centrifuged again (15000 rpm for 5 min), followed by affinity-based chromatography (amylose column from NEB) following the supplied procedure. After thorough washing with the column buffer (except containing 25 mM NaCl), the proteins were eluted with maltose-containing column buffer (NEB). All steps were carried out on ice or in the cold room. The quality of the proteins obtained was checked by 10% SDS-PAGE analysis (131). Most of the RdRp studies (see below) were done with the fusion proteins (unless mentioned otherwise).

Plant TCV RdRp preparation. Turnip plants inoculated with TCV transcripts were used ten days after inoculation to isolate partially purified TCV RdRp preparations as described previously (86, 140). All steps were carried out on ice or in the cold room.

Preparation of RNA templates. RNA templates were prepared using T7 RNA polymerase (131), followed by removal of unincorporated nucleotides as described previously (85). Constructs containing full-length cDNAs of satC, satD and MDV [MDV-1 is a 221 nt satellite RNA associated with Qβ bacteriophage (8)] were obtained from Anne Simon, while DI-72 was from Andy White (156). The series of promoter-containing constructs of Figure 2.7 were obtained from Tadas Panavas (107).

RdRp assay. RdRp reactions were carried out as previously described for the plant TCV RdRp (86) for 1 hr at 25 ^{0}C (unless indicated otherwise). Briefly, the RdRp reactions were performed in the presence of 50 mM Tris-HCl, pH 8.2, 10 mM $MgCl_2$, 10 mM DTT, 100 mM potassium glutamate, 1.0 mM ATP, CTP and GTP, and 0.01 mM UTP (final concentration) and 0.5 μl ^{32}P UTP (ICN) in 50 μl total volume**.** In addition, RdRp reactions contained 165nM of template RNA. In some experiments, the amount of template RNA was reduced to 80nM in case of p88C due to its high RdRp activity (Figure 2.4 and Figure 2.5). The amounts of p88 and p88C (used as fusion proteins with MBP) were 3 μg/assay. After phenol/chloroform extraction and ammonium acetate/isopropanol precipitation, half the amount of the RdRp products was treated with S1 nuclease as described previously (88, 91). Subsequently, the RdRp products were analyzed on a 20 cm long denaturing 5% PAGE/8 M urea gels, followed by analysis with a PhosphorImager as described (85, 140). During the competition experiments, the same amount (50 nM) of template RNA, and increasing amounts (between 1 to 16-fold excess) of competitor RNAs were used in RdRp reactions similar to those described above.

RNA binding Studies. RNA probes labeled with ^{32}P UTP were prepared using T7 RNA polymerase as described above, except a 100-fold less unlabeled UTP was used. Approximately 20ng of labeled RNA was mixed with 2 μg of p88 or p88C preparations for 20 min at 25 ^{0}C in the presence of the RdRp buffer (see above) (86, 140). The volume of the plant TCV RdRp preparation used for RNA binding was 5 times less than that used during the RdRp assays. The samples were analyzed by electrophoresis on native 1% agarose gels performed at 200V for 1 hr at 4 ^{0}C in TAE buffer (131). Dried gels were analyzed using a PhosphorImager. During the

competition experiments, several fold excess amounts of cold competitor RNAs were added simultaneously with the ^{32}P-labeled probe to similar amounts of p88 (as shown in Figure 2.9).

RESULTS

Expression and purification of TCV p88 in *E. coli*. The 5' region of p88 gene of TCV overlaps with p28 gene (Figure 2.1), while the unique 3' portion codes for the putative RdRp, based on sequence comparison (17, 94). A mutated p88 gene of TCV, in which the readthrough stop codon was changed to a tyrosine codon (157), was cloned into an expression vector (pMAL c2X, NEB) and p88 was expressed as a C-terminal fusion protein with the maltose-binding protein (MBP) in *E. coli* as described in Materials and Methods. After affinity-based purification of the p88 fusion protein (Figure 2.2), I tested its RdRp activity first using minus-strand satC as template in an RdRp buffer described for TCV RdRp preparations purified from TCV-infected plants (85, 140). The RdRp products were analyzed by denaturing PAGE that revealed RdRp activity for the p88 preparation only when external template was added (not shown). The use of minus-strand satC templates in the p88-based RdRp assay resulted in three major bands in denaturing PAGE gels (Figure 2.3B). The major bands obtained with p88 were similar to the bands obtained with a control TCV RdRp preparation purified from plants (Figure 2.3B, note that I will use the term plant TCV RdRp for the control TCV RdRp preparation purified from plants). S1 nuclease digestion of the RdRp products revealed the presence of three major types of RdRp products: the first type constitutes nuclease-resistant template-sized complementary RNA; and the second type includes partially nuclease-sensitive, shorter than template-sized RNA products (Figure 2.3B). Earlier work with the plant TCV RdRp preparation has demonstrated that the template-sized RdRp product is generated by *de novo* initiation from the very 3' end of the minus-strand satC RNA [product D (de novo) in Figure 2.3A, see ref. (141)]. The shorter than template-sized product is generated by primer extension (self-priming from the 3' end of the template) initiating internally close to the motif1-hairpin replication enhancer [product M (supported by the motif-1 hairpin replication enhancer) in Figure 2.3A; also termed as large RNA in ref. (141)]. The third type, which migrates aberrantly on denaturing PAGE gels, is also partially S1 nuclease sensitive and it is likely generated by primer extension starting close to the 3' end of the template [product E (priming from close to the 3' end) in Figure 2.3A; also termed as small RNA in ref. (141)]. The S1 nuclease digestion profile of the *in vitro* products obtained

with the plant TCV RdRp preparation was similar to that obtained with p88 (Figure 2.3B, lane +). The above data, based on the similar-sized RdRp products and on the similar S1 nuclease digestion profile, suggest that RdRp products are likely generated by the same mechanism as described for the plant TCV RdRp products (140, 141).

Five different N- and C-terminally truncated versions of p88 were also generated, expressed, purified and tested as MBP fusion proteins *in vitro* for RdRp activity (Figure 2.1B). First, p28, which contains the N-terminal portion of p88, but lacks the putative polymerase domain of p88 and represents one of the two naturally expressed TCV replicase proteins (Figure 2.1) (157), had no RdRp activity *in vitro* (Figure 2.1B and data not shown). Second, p88C, which has the complete polymerase-like domain of p88 (Figure 2.1B), showed unusually high RdRp activity (Figure 2.3B). Removal of the MBP domain by cleavage with Factor Xa from the p88C fusion protein had no effect on RdRp activity of p88C (not shown). Third, three different truncations within the p88C domain (constructs ΔN100/C30, ΔN30/C100, ΔN100/C100, Figure 2.1B) made the preparations inactive in standard RdRp assays (Figure 2.1B and data not shown). Overall, these data confirm that the C-terminal portion of p88 contains the functional RdRp domain and they also suggest that the entire 88C domain of p88 is required for RdRp activity *in vitro*. In addition, the lack of TCV-specific RdRp activities of p28, the three truncated derivatives of p88C (Figure 2.1B) and the maltose-binding protein alone (data not shown) excludes the possibility that I purified a contaminating putative *E. coli* RdRp, which would be responsible for labeling the externally added templates *in vitro*. This is because the putative *E. coli*-derived RdRp, if any, should be present in all the above preparations, which were obtained from the same *E. coli* strain and purified using exactly the same procedure (Materials and Methods). Yet, only p88 and p88C preparations showed RdRp activities. In addition, the RdRp products obtained with p88 and p88C preparations were similar to those generated with the plant TCV RdRp (Figure 2.3B). Based on the above observations, I conclude that p88 and p88C preparations contain functional TCV RdRp. The RdRp activities of p88 and p88C were further characterized as shown below.

Characterization of RdRp activities of TCV p88 and p88C. To test the stability and activity of p88 and p88C, I performed the RdRp assays at various temperatures as shown in Figure 2.4 using minus-strand satC templates. Preparations of p88 and p88C showed RdRp activities from 4

to 37 ^{0}C temperatures (Figure 2.4A-B). The plant TCV RdRp was also active at these temperatures, although much longer exposure of the gels was needed to detect the small amount of products synthesized (Figure 2.4A-B). All three preparations produced mainly primer extension products (product M in Figure 2.4A) at low temperatures (4 and 12 ^{0}C), while *de novo* initiation was hardly detectable (product D in Figure 2.4A). Primer extension (product M in Figure 2.4A) was the highest for all three preparations at 30 ^{0}C. *De novo* initiation occurred with the highest efficiency at 25 ^{0}C and 30 ^{0}C for p88C and the plant TCV RdRp, while p88 generated the most template-sized products at 37 ^{0}C (product D, Figure 2.4A). S1 digestion of the RdRp products, however, revealed that a portion of the template-sized RdRp products (product D) for p88 obtained at 30 ^{0}C and 37 ^{0}C, but not at 25 ^{0}C, were S1 nuclease sensitive (data not shown), suggesting the presence of terminal transferase-like activity in the p88 preparation. Similar activity was not detected in the p88C or the plant TCV RdRp preparations. It is known that other viral RdRp preparations, such as the RdRp of Hepatitis C virus, obtained from *E. coli* may also contain terminal-transferase activity (76). Because the above terminal transferase-like activity in the p88 preparations was not detectable at 25 ^{0}C, I used this temperature in all the following studies.

To compare the kinetics of RNA synthesis by p88, p88C and the plant TCV RdRp, I conducted time course experiments at 25 ^{0}C. All three preparations produced primer extension products (product M) as quickly as 30 seconds (Figure 2.5). In contrast, detection of the template-sized *de novo*-initiated products (product D) required at least 5 minutes of incubation with all three preparations. The amounts of both *de novo*-initiated (product D) and primer extension products (product M) for each preparation increased over time, suggesting that each RdRp molecule can likely produce several nascent RNA products sequentially (RdRp is reused). The data in Figure 2.5 also demonstrate that p88C is different from p88 and the plant TCV RdRp by favoring primer extension from the internal replication enhancer (the motif1-hairpin, product M) over the production of template-sized *de novo* products (compare products M and D). For example, at the 60 minute time point, the amount of product M was ~70-fold more (normalized value) than that of product D in the p88C assays, while product M was only 13- and 21-fold more abundant for the plant TCV RdRp and the p88 RdRp than product D (Figure 2.5B, D, and F, also see Figure 2.4B). More diluted p88C preparations also resulted in ~70-fold differences between products M and D (not shown). The profile of RdRp products obtained with p88 and the

plant TCV RdRp was more similar to each other than to that obtained with p88C. The above observations indicate that the N-terminal portion of p88, which is not present in p88C, may somehow inhibit primer extension by more than two-fold. Further experiments will be needed to demonstrate a direct role for the N-terminal region of p88 in affecting selection between *de novo* synthesis and primer extension.

Comparison of template use by p88, p88C and the plant TCV RdRp. To compare the template use by p88, p88C and plant TCV RdRp, I performed RdRp assays using four pairs of plus- and minus-strand RNA templates. When compared to the above tested minus-strand satC, the plus-strand satC RNA was used inefficiently in all three assays (Figure 2.6A-C). All three preparations synthesized predominantly the full-length complementary products for the satC(+) template, based on the size and S1 nuclease-resistant nature of the products [Figure 2.6, lane satC(+)]. The second pair of templates was satD RNA, which is 194 nt and is associated with TCV infections (135). The third set of templates was DI-72 RNA, which is a 621 nt defective interfering RNA associated with *Tomato bushy stunt virus* (TBSV, a tombusvirus related to TCV) infections. The fourth pair of templates was the heterologous MDV RNA, a 221 nt satellite RNA associated with Qβ bacteriophage (8). The plus-strand satD, DI-72 and MDV were found to be poor templates in all three RdRp assays [Figure 2.6, lanes satD(+), DI-72(+), MDV(+)]. In contrast, the minus-strand satD as well as DI-72 were used efficiently in all three RdRp systems [Figure 2.6, lanes satD(-) and DI-72(-)]. In addition to the full-length complementary product (pointed by arrows in Figure 2.6), primer extension (i.e., extension from the 3' end of template, marked with asterisks in Figure 2.6) products were also obtained in all three RdRp assays with satD(-) and DI-72(-). Easily detectable levels of internal initiation (bracketed) at three positions were detected for DI-72(-) in all three RdRp assays. The nature of these products was characterized in more detail in a previous paper (86). Plus-strand MDV RNA was not used as a template for *de novo* initiation from the authentic 3' end in any of the three assays [there are no S1 nuclease resistant bands in Figure 2.6, lane MDV (+)]. Low levels of partially S1 nuclease sensitive products were observed in all three systems, although p88C gave the highest amount of these products. The above products obtained with MDV (+) are likely the result of inefficient primer extension based on their partially S1 nuclease-resistant nature. The lack of detectable amounts of *de novo*-initiated products for MDV (+) suggests that all three RdRp preparations

have comparable levels of selectivity in template usage. Surprisingly, however, the minus-strand MDV RNA was a rather efficient template in all three assays [Figure 2.6, lanes MDV (-)]. I also tested an additional heterologous template, namely tRNA from yeast, which was not used as a template in any of the three RdRp assays (data not shown). Overall, these experiments demonstrated that, in general, plus-strand RNAs were poorer templates than minus-strand RNAs. In addition, the three RdRp preparations were shown to have comparable levels of template specificity. This is somewhat surprising since p88C only includes the polymerase domain, while the plant TCV RdRp preparation contains small amount of p88 and excess amount of free p28 and some host components (140).

Comparison of promoter recognition by the *E. coli*-expressed and the plant TCV RdRps. The TCV RdRp has been shown to recognize short (11 nt) linear sequences and a hairpin sequence with short single-stranded tail for plus- and minus-strand RNA synthesis, respectively (43, 139). To test whether p88 and p88C are also capable of recognizing a minimal TCV promoter and other viral and artificial sequences, I tested the RdRp activity with seven different RNA sequences as shown in Figure 2.7. The selected sequences were separately fused to the 3' of a chimeric heterologous sequence (see construct Anc-MDV, Figure 2.7). The rationale for use of heterologous sequences is that they can facilitate correct RNA folding, which in turn reduces the extent of primer extension in the *in vitro* RdRp system (43, 107). The heterologous sequence used in this work consists of negative-strand MDV sequence and a 3' 'anchor' sequence derived from TBSV (+) (Figure 2.7). This anchor sequence was found previously to reduce the level of RNA synthesis from MDV sequences with the tombusvirus RdRps (107). Accordingly, the heterologous sequence alone supported RNA synthesis only at a very low level in all three *in vitro* TCV systems tested here (Figure 2.7, lane Anc-MDV). By adding the 11 nt minimal promoter for positive-strand synthesis (i.e., the 3' proximal sequence, ref. 43) derived from the minus-strand satC to the 3' end of Anc-MDV, the extent of *de novo* initiation was increased 12-, 143-, and 25-fold for p88, p88C and the plant TCV RdRp, respectively (see construct cTCV, Figure 2.7). The proper and efficient recognition of the 3' proximal sequence of satC(-) suggests that all three *in vitro* systems can recognize the TCV promoter sequence in a similar manner. The higher level of increase for the p88C preparation, which has ten-fold higher activity in RNA synthesis for satC (-) templates (Figure 2.4) than p88 or the plant TCV RdRp (not shown) may

be the result of the presence of a larger fraction of active RdRp molecules in the p88C preparation (see Figure 2.4).

To test whether the *in vitro* systems can recognize heterologous promoter sequences, I included cPR21, cPR11, and gPR sequences, which are derived from the related TBSV. cPR21 and cPR11 represent the minimal promoters for plus-strand (constructs cPR21 and cPR11, Figure 2.7), while gPR represents the "core" promoter sequence for minus-strand synthesis for TBSV (107). These experiments demonstrated efficient *de novo* initiation for the three constructs in all three *in vitro* systems (Figure 2.7, lanes cPR21, cPR11, and gPR). Addition of artificial AU-rich and GC-rich sequences (constructs A/U and G/C, Figure 2.7) to the 3' end of the Anc-MDV sequence resulted in reduction in RNA synthesis (Figure 2.7, lanes A/U and G/C), when compared to the level obtained with cTCV. The above experiments demonstrated that the promoter selectivity of all three preparations is comparable based on the tested promoter sequences.

Template competition experiments. To further study the role of sequences in template selection by p88 and p88C, I have used template competition experiments, which may be more similar to *in vivo* conditions, where the viral replicase can encounter different templates and promoters and must choose among them. The template RNA was cPR11 (shown in Figure 2.7), which consists of Anc-MDV at the 5' end and the TBSV-derived 11 nt core (+)-strand synthesis promoter at the 3' end. cPR11 is recognized efficiently by both p88 and p88C when present as a single template in the *in vitro* assay (Figure 2.7). Three different competitor RNAs were used in these experiments, including satC(-), satD(-) and satC(+), which are excellent, good and poor templates for the TCV RdRp, respectively (Figure 2.6). When compared to the template RNA, the above competitor RNAs give different sized RdRp products to allow for separation of template and competitor-derived products. By applying the same amount of template RNA, while increasing the amount of the competitor RNA in the RdRp assays, I have calculated the IC_{50} value for the above templates. The IC_{50} value determines the amount of competitor RNA needed to reduce *de novo* initiation from the given template RNA to 50% of the level obtained in the absence of the competitor (modified from (26, 101). The data obtained in the competition experiments revealed that satC(-) was the most competitive, and satD(-) showed moderate level of competitiveness, while satC(+) was the least competitive in both the p88 and p88C *in vitro*

assays (Figure 2.8). The IC_{50} values for the three competitors, however, were markedly different in the p88 and the p88C assay systems with lower levels of competitors competing more efficiently in the p88C assays (Figure 2.8). The reason for the observed differences in template competition between the p88 and p88C is not known. I speculate that p88 and p88C may differ in the kinetics of complementary RNA synthesis (i.e., how quickly the RdRp is released from the template after termination) or need different amount of time for RNA binding, both of which can affect the chance for the repeated use of the RdRps.

RNA binding by p88 and p88C. Gel mobility shift assays were used to characterize the ability of p88, p88C and the plant TCV RdRp to bind RNA. p88 and the plant TCV RdRp were found to bind efficiently to the minus-strand satD (Figure 2.9A). Binding of p88C to the minus-strand satD was also significant, although the pattern of the shift was different from that observed with p88 and the plant TCV RdRp (Figure 2.9A). For example, in addition to the fully shifted band (which is the major band obtained with p88 and the plant TCV RdRp), a fraction of RNAs was only partially shifted by binding to p88C. This suggests that not only p88, but the p88C also contains an RNA binding domain.

RNA binding by p88 was also tested in competition experiments, in which increasing amounts of "cold" competitor RNAs, either satC(+) or tRNA, were used in addition to a constant amount of labeled satC(+) template (Figure 2.9B). This experiment demonstrated that while satC(+) is a good competitor, tRNA was a poor competitor (Figure 2.9B-C).

DISCUSSION

Viral RdRps purified from infected plant cells are multisubunit enzymes (17, 18). Comparison of activities associated with the plant-derived viral RdRps and the single subunit RdRps obtained from heterologous systems should help define functions for the subunits. I undertook these studies using TCV for which the plant-derived RdRp is available (140). This preparation contains a small amount of p88 and excess amount of p28 and some host components (140). I have shown that the heterologously expressed TCV p88 protein, in the absence of free p28 and eukaryotic host factors, has RdRp activity. In addition, the N-terminally truncated p88, designated as p88C, which lacks the p28 overlapping domain (Figure 2.1), is not

only an active RdRp, but it has 10-fold higher activity than the full-length p88 when applied in the same amount (not shown). The approximately ~60 kDa p88C is one of the smallest RdRps with size similar to the $3D^{pol}$ of poliovirus (11). In contrast to $3D^{pol}$, p88C can initiate complementary RNA synthesis *de novo* (see below).

Comparison of template recognition and use by the *E. coli*-expressed p88 and p88C with that of the well-characterized plant TCV RdRp (140) revealed surprising similarities among the single-unit p88 and p88C and the multi-subunit plant TCV RdRp preparations. The similarities include the following features: (i) the ability to initiate *de novo* on both plus- and minus-strand templates (this was confirmed indirectly by showing the nuclease resistant nature of the template-sized products in Figure 2.3); (ii) the ability to extend on primers (self-priming); (iii) the ability to use the minus-strand templates more efficiently than the plus-strand templates; and (iv) the ability to use the TCV-associated satC and satD as well as the related TBSV-derived DI-72 templates more efficiently than the heterologous MDV RNA (Figure 2.6) or tRNA (not shown). It is possible that free p88, which dissociated from other subunits that are normally part of the replicase, is present in our plant TCV RdRp preparation. The presence of free and active p88 in our preparation would explain the similar results obtained with p88 obtained from E. coli and the plant TCV RdRp preparation. This explanation, however, is unlikely since the plant TCV RdRp was purified as a large complex (>500000 Da).

Comparison of promoter recognition by p88 and p88C with the plant TCV RdRp revealed that all three RdRps recognized efficiently and correctly the 3' proximal minimal positive-strand initiation promoter of satC (Figure 2.7) defined previously by Guan *et al*. (43). In addition, all three RdRps also recognized the minimal positive-strand initiation promoter for the related TBSV as efficiently as the corresponding TCV promoter. In contrast, the unrelated Anc-MDV template or a template carrying artificial GC-rich sequences at the 3' end was recognized poorly by all three RdRps. These data suggest that these RdRps have the ability to discriminate against some templates at comparable levels. The discrimination against some host RNAs may be manifested by their lack of binding to the RdRp, since tRNA was found to be a poor competitor in RdRp binding (Figure 2.9) and is not used as a template for complementary RNA synthesis by p88 and p88C (not shown). Overall, it is surprising that no significant difference in template recognition was found among these preparations (Figure 2.6). This suggests that the minimal p88C RdRp has many of the features characteristic of the multi-subunit plant TCV

RdRp preparation *in vitro*. This will make the small, single component p88C an attractive system for future *in vitro* studies.

The most significant difference between p88C and either p88 or the plant TCV RdRp was the enhanced ability of p88C to generate a primer extension (self-primed) product that is facilitated by the presence of the motif1-hairpin replication enhancer (87). It was found that ~70-fold more primer extension product than the *de novo* initiated products were obtained with p88C after 60 min incubation. The corresponding values for p88 and the plant TCV RdRp was only 21- and 13-fold. Since p88C lacks the p28 region, while that is present in both p88 and the plant TCV RdRp, it is possible that the role of the overlapping region (or in case of the plant TCV RdRp, free p28 may also play a role) is to inhibit the ability of the RdRp to use (self-) primers. Since self-priming is probably an *in vitro* artifact, it is possible that free p28 or the overlapping region of p88 may be involved in reduction of incorrectly initiated products or in stimulating correctly initiated products. Further experiment will be needed to dissect the role of the free p28 and the overlapping region of p88 in TCV replication.

Kinetic studies with the three RdRp preparations demonstrated that RNA synthesis is relatively rapid in these systems. It took 30 seconds to obtain detectable amounts of the self-primed product (product M in Figure 2.5) in all three systems. In contrast, 5 minutes were needed in all three systems to detect the template-sized *de novo* initiated products (product D in Figure 2.5). This observation suggests that initiation of RNA synthesis by self-priming at the motif1-hairpin replication enhancer is a faster process than *de novo* initiation at the 3' terminal promoter. Also, the increasing amounts of RdRp products during prolonged incubations suggest, although do not prove, that the RdRp molecules are re-used sequentially several times during the reaction. This also suggests that these RdRp preparations are stable during prolonged incubations.

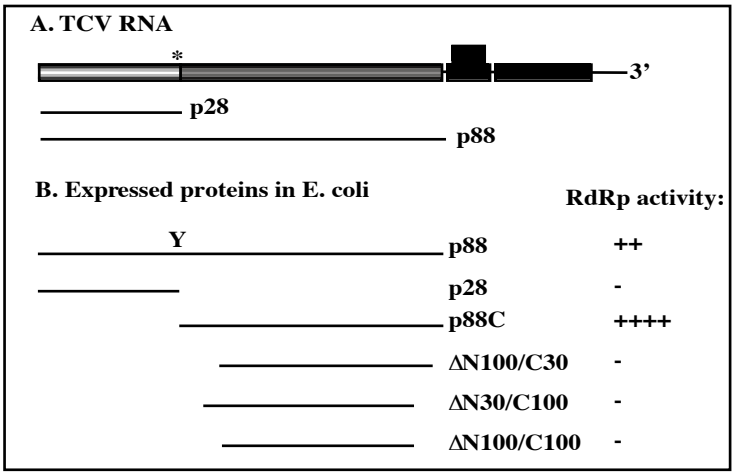


Figure 2.1. RdRp activity of the recombinant TCV p88 and its derivatives obtained from *E. coli*. **(A)** Schematic representation of the TCV genome. TCV RNA contains five open reading frames of which two are expressed from the genomic RNAs (shown by shaded boxes) and three (shown by black boxes) are expressed from two subgenomic RNAs. p88 is expressed by leaky termination of the stop codon (marked with an asterisk) of the p28 gene. **(B)** Schematic representation and RdRp activities of p88 and its derivatives expressed and purified from *E. coli*. The leaky termination codon in p88 was altered to a tyrosine (Y) codon (see Materials and Methods), which was shown to be functional *in vivo* (157). The RdRp activity of each of the six separately expressed and purified proteins (fused to the maltose-binding protein) was tested in a standard *in vitro* assay using ^{32}P-labeled UTP and PAGE analysis of the RdRp products (not shown). Only p88 and p88C showed detectable levels of RdRp activity.

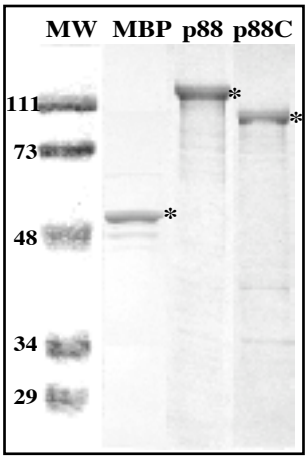


Figure 2.2. A representative SDS-PAGE analysis of the purified recombinant proteins expressed in *E. coli*. The gel was stained with Coomassie-blue. Asterisks mark the full-length proteins. Due to the expression strategy with pMAL c2X, the maltose binding protein (lane MBP) contains a C-terminal extension, resulting an ~50 kDa protein. Note that p88 and p88C are fused with the maltose-binding protein to aid affinity-based purification. The smaller products are likely generated by either premature termination of translation of p88 and p88C genes or protein degradation, since their amounts were not reduced by intensive washing of the columns during affinity-based purification (not shown).

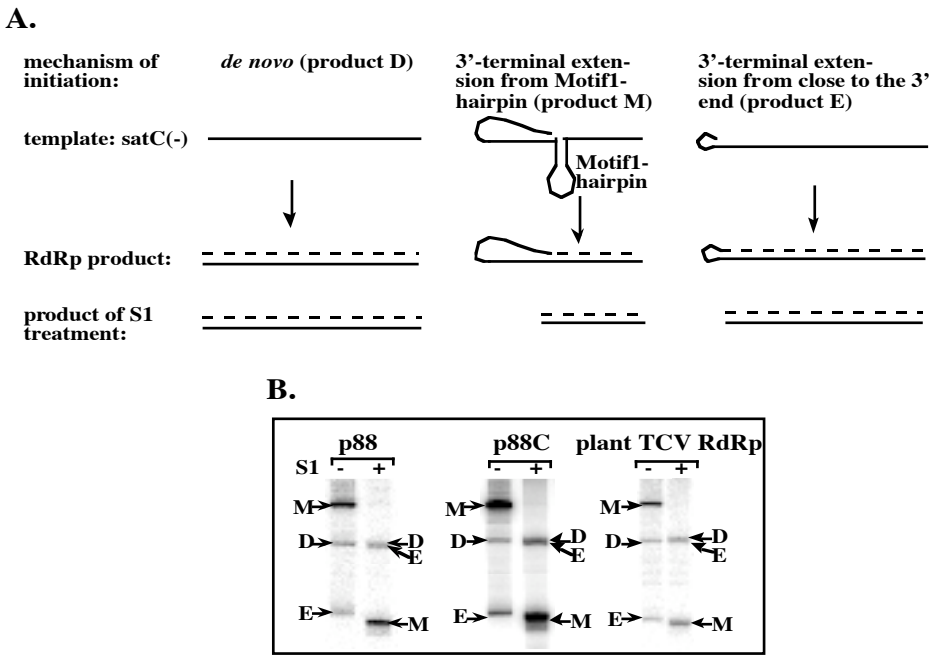


Figure 2.3. Comparison of the RdRp products obtained with the recombinant p88 and p88C with the TCV RdRp purified from plants using satC(-) as template. (A) Schematic representation of the three major products generated by the plant TCV RdRp *in vitro* using satC(-) templates. The mechanism of the formation of these products is explained schematically based on data from Song and Simon (140, 141). Solid lines represent the template, while the broken line depicts the complementary RNA product (labeled during the RdRp reaction). S1 nuclease treatment removes the single-stranded regions of the RdRp products under the conditions used (91, 141). The names of the products reflect their generation: product D: de novo initiation, product M- motif1-hairpin mediated primer extension, product E: 3' end mediated primer extension. (B) Representative denaturing gel analyses of radiolabeled RNA products synthesized by *in vitro* transcription with the recombinant p88 and p88C and the plant TCV RdRp. - and + signs above the lanes indicate untreated and S1 nuclease treated samples prior to PAGE analysis. Products M, D and E are marked with arrows. Note that product E runs aberrantly (i.e. much faster) under these conditions in the untreated samples due to the highly stable hairpin structure of this product (141). Also, S1-treated product E migrates close to product D under the conditions used. The exposure time was different for each set of RdRps due to the different activities of the RdRp preparations. The amount of p88C applied in the RdRp assay was 40% of p88, based on Coomassie-stained SDS-PAGE gels (not shown).

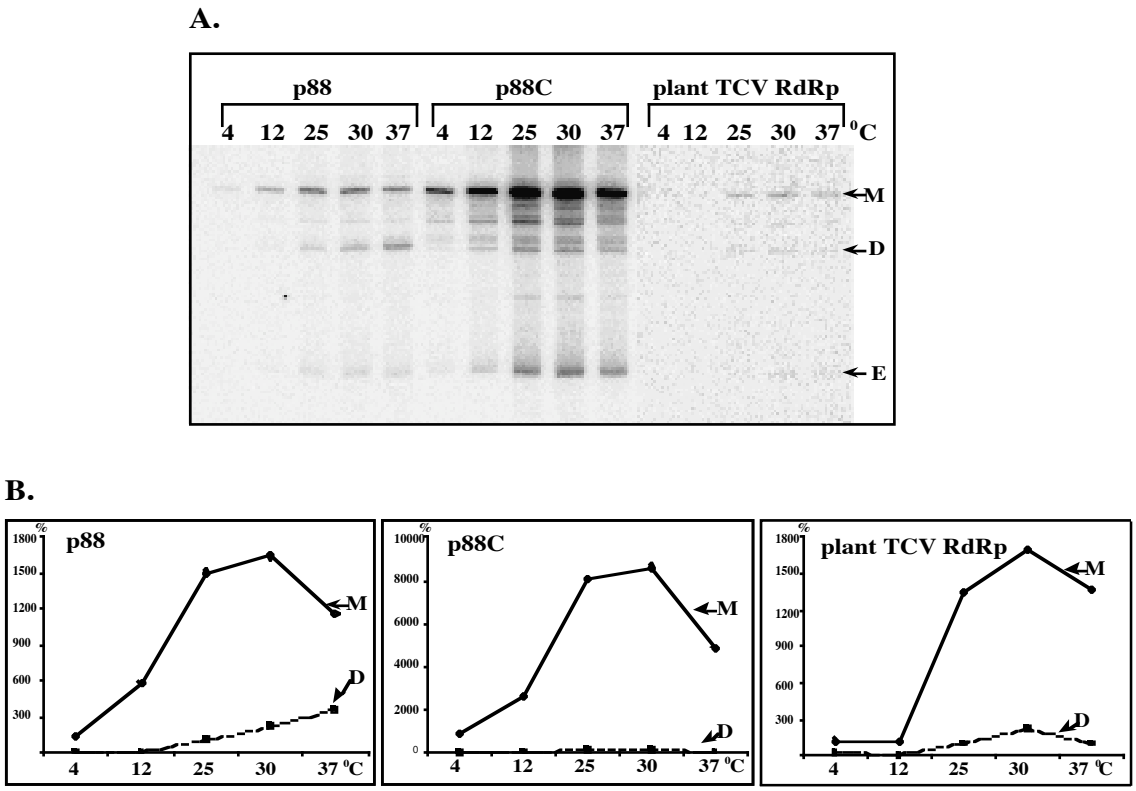


Figure 2.4. The effect of temperature on the RdRp activity of the recombinant p88 and p88C and the plant TCV RdRp. (A) Representative denaturing gel analyses of radiolabeled RNA products synthesized by *in vitro* transcription with p88, p88C and the plant TCV RdRp using satC(-) as template. Products are marked as shown in the legend to Figure 2.3. The exposure time was different for each RdRp assay due to the increased activity for p88C. (B) The relative amount of M and D RdRp products was measured at various temperatures as shown. The amount of product D obtained at 25 ^{0}C was selected as 100% separately for each RdRp preparation. Note that the increased level of product D at 30 ^{0}C and 37 ^{0}C is due to terminal transferase activity present only in the p88 preparation as confirmed by S1 nuclease digestion (not shown). The terminal transferase activity was not significant at 25 ^{0}C (not shown). Terminal transferase activity in the p88 preparation was confirmed by obtaining template-sized product (similar in size to product D) when the RdRp reaction contained ^{32}P UTP in the absence of the other three nucleotides. This terminally labeled product, unlike product D, was fully S1 nuclease sensitive (not shown). Product E was not included since its amount was variable in separate experiments.

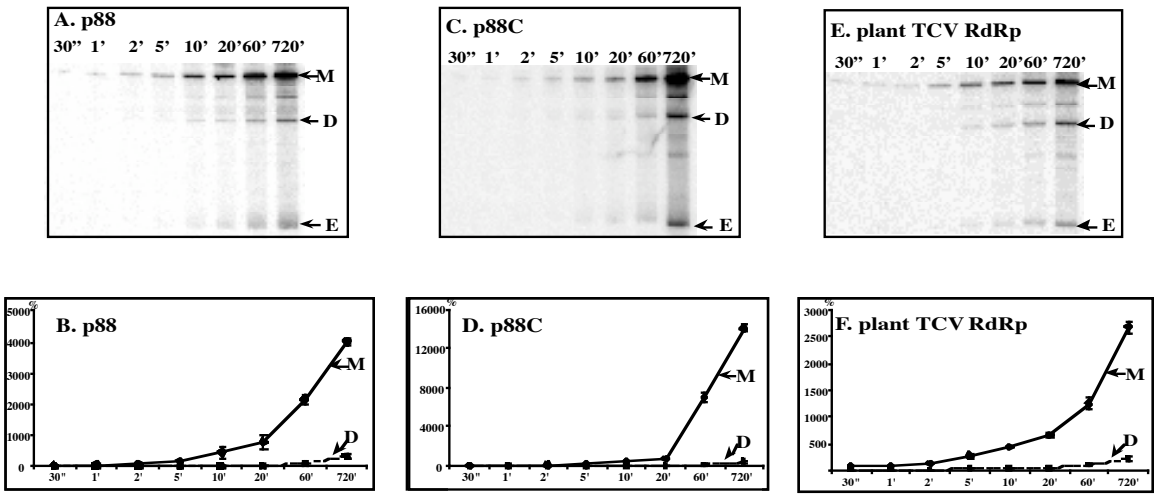


Figure 2.5. Comparison of the kinetics of RNA synthesis by the recombinant p88 and p88C with that of the plant TCV RdRp. Standard RdRp reactions were performed at 25^0C for a given amount of time using satC(-) as template. Products are marked as described in the legend to Figure 2.3. Panels (A), (C) and (E) show representative denaturing gel analyses of radiolabeled RNA products synthesized by *in vitro* transcription with p88, p88C and the plant TCV RdRp. Panels (B), (D) and (F) show the relative amounts of M and D RdRp products made during various lengths of incubation time as shown. The amount of product D obtained after 60 minutes of incubation was selected as 100% separately for each RdRp preparation.

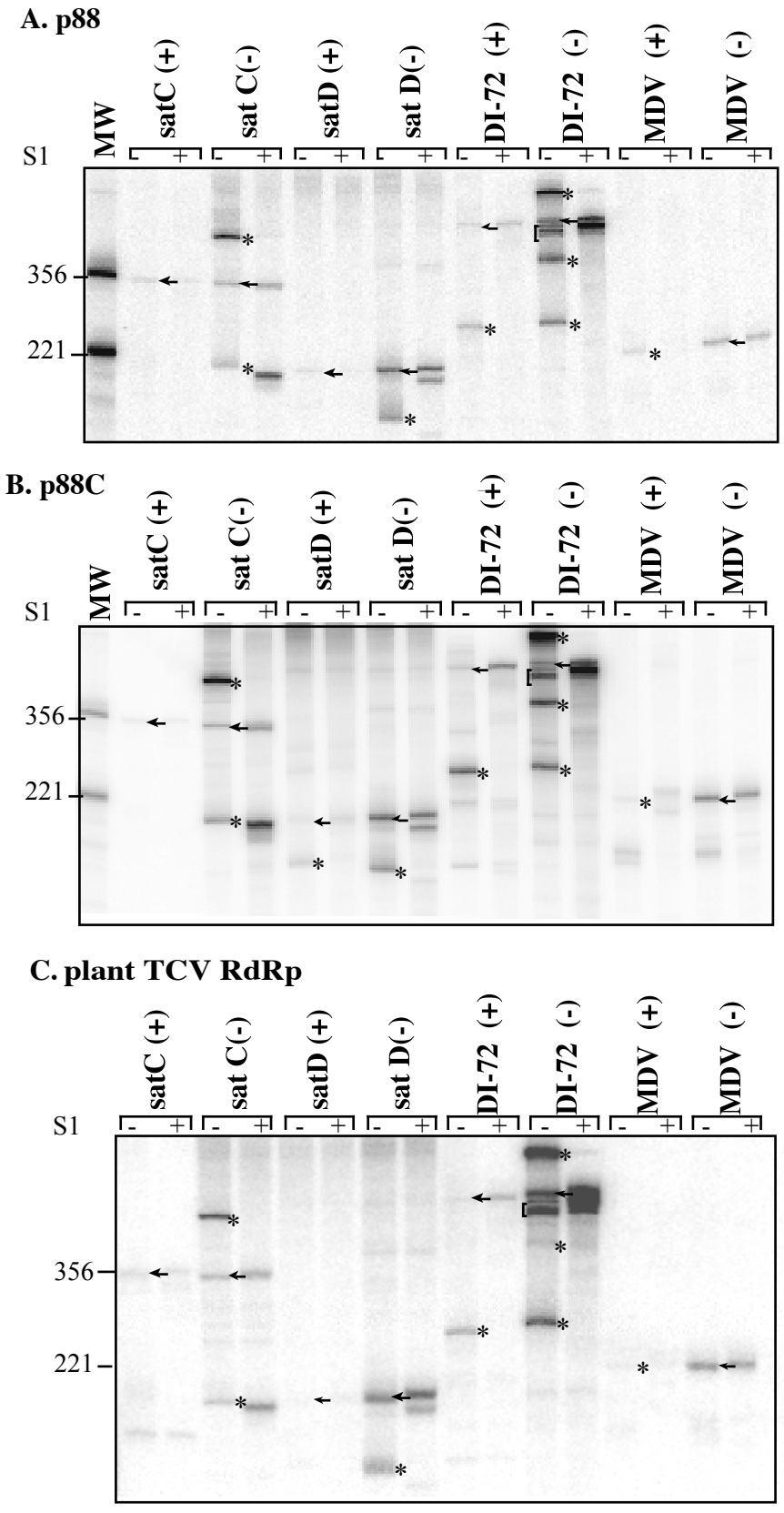

A. p88
MW
satC (+)
sat C(-)
satD (+)
sat D(-)
DI-72 (+)
DI-72 (-)
MDV (+)
MDV (-)
S1
356
221
B. p88C
C. plant TCV RdRp

Figure 2.6. Comparison of template use of recombinant p88 and p88C and the plant TCV RdRps. Representative denaturing gel analyses of radiolabeled RNA products synthesized by *in vitro* transcription with (A) p88; (B) p88C; and (C) the plant TCV RdRp are shown. Full-length plus or minus strands of the TCV-associated satC and satD RNAs, the TBSV-associated DI-72 RNA and the Qβ-associated MDV RNA were used in equal amount (1 μg) as templates. Half the amount of RdRp products was treated with single-strand specific S1 nuclease (lanes marked with +). Arrows mark template-sized RdRp products that were resistant to S1 nuclease treatment. Asterisks depict the primer extension products that are partially S1 nuclease sensitive. These products change their migration pattern after S1 nuclease treatment. MW, single-stranded RNA size-markers (in bases) obtained by T7 transcription using satC (356 nt) and MDV (221 nt) clones as templates.

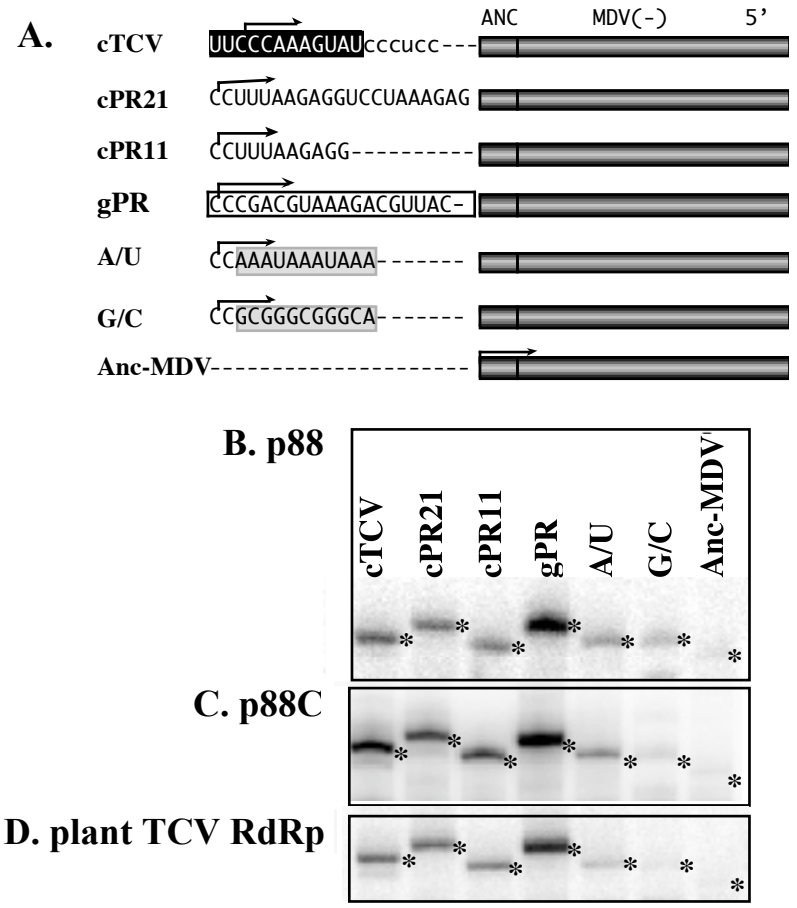


Figure 2.7. Comparison of promoter recognition by the recombinant p88 and p88C and the plant TCV RdRps. (A) Schematic representation of constructs tested in *in vitro* RdRp assays. The actual 3' end sequence of each construct is shown in 3' to 5' orientation. The core plus-strand initiation promoter for satC(-) (43) is shown in a black box (construct cTCV). Constructs cPR21 and cPR11 contain the extended and the core plus-strand initiation promoter for minus-strand DI-72 of TBSV (107). The core minus-strand synthesis promoter for DI-72 of TBSV is boxed (construct gPR). The artificial AU-rich and GC-rich promoter sequences are shown in gray boxes (constructs A/U and G/C). The sites of expected initiation products are shown with arrows. In addition to the shown sequences, each construct contains the same 5' sequences derived from minus-strand MDV (221 nt) and a 17 nt anchor (Anc) sequence from TBSV (positions 4666-4682). (B-D) Representative denaturing gel analyses of radiolabeled RNA products synthesized by *in vitro* transcription with p88, p88C and the TCV plant RdRp. Template-sized products are labeled with asterisks. The order of samples is the same in each panel.

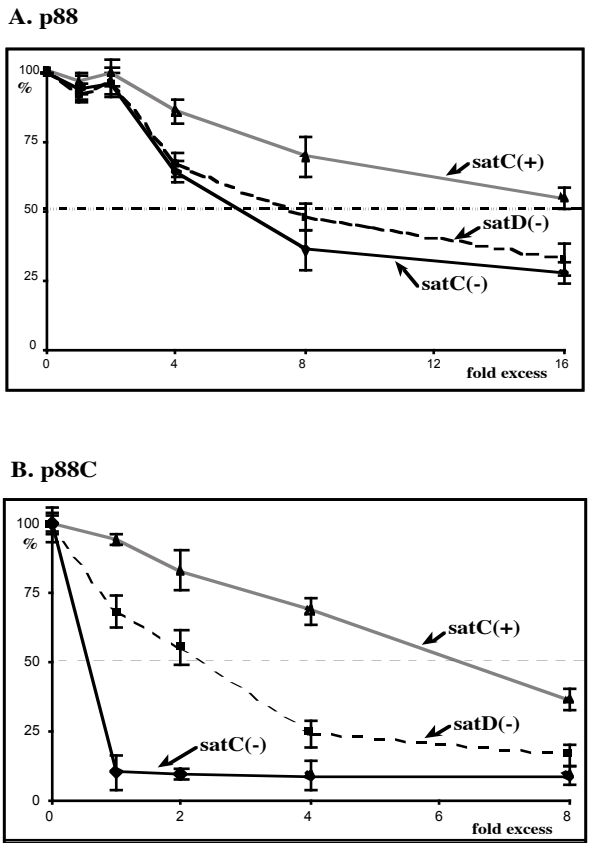


Figure 2.8. Determination of IC_{50} values for satC(+), satC(-) and satD(-) in RdRp reactions containing p88 (A) or p88C (B). The amount of radiolabeled, template-sized RdRp products synthesized on +PR11 templates (Figure 2.7) was measured in the absence (100%) or in the presence of a competitor RNA. The three competitor RNAs, satC(+), satC(-) and satD(-) were used separately. The relative amount of competitor (in comparison with the constant amount of template RNA) is shown below the graph. Panels (A) and (B) show the results obtained with the recombinant p88 and p88C, respectively.

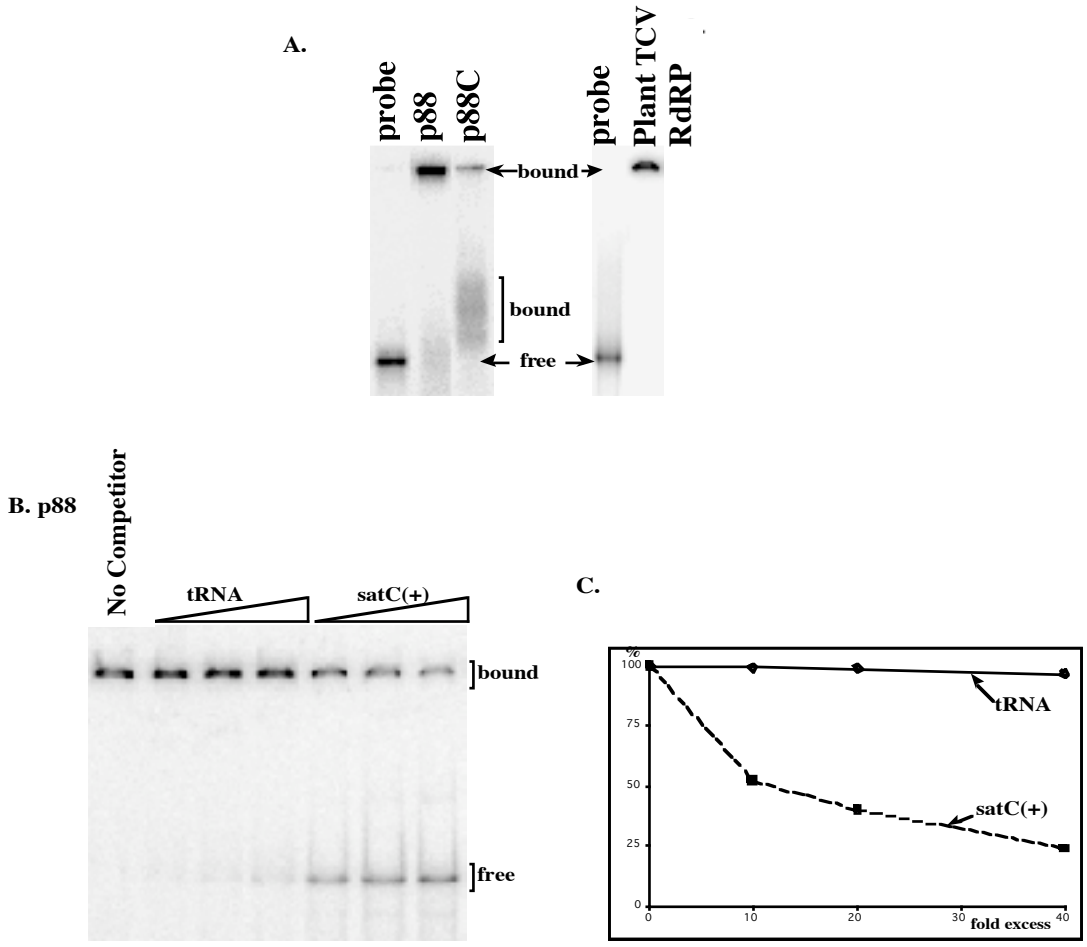


Figure 2.9. (A) A representative gel mobility shift analysis of radiolabeled RNA bound to p88, p88C or the plant TCV RdRp. The bound and the free RNAs are marked at the sides. Comparable amounts of p88 and p88C were used for RNA binding (see Materials and Methods for details). Note that the bound probe to p88C resulted in two types of products: a slow migrating product at the top of the gel and a diffuse fast migrating product above the free probe. The first lane on the left shows the ^{32}P-labeled satD(-) template without added proteins (lane probe only). (B) Template competition experiments for RdRp binding. A representative gel mobility shift analysis of radiolabeled satC(+) RNA bound to p88 in the absence or presence of cold competitor RNAs (the competitors shown on the top of the gels). The bound and the free RNAs are marked at the sides. The first lane on the left shows the ^{32}P-labeled template bound to p88 without added competitors (no competitor). (C) A graphical presentation of the competition experiments. The relative amount of competitor (in comparison with the constant amount of ^{32}P-labeled RNA probe) is shown below the graph.

CHAPTER III

CHARACTERIZATION OF THE RNA-BINDING DOMAINS IN THE REPLICASE PROTEINS OF *TOMATO BUSHY STUNT VIRUS*

INTRODUCTION

Viral-coded replicase proteins are essential for replication of tombusviruses, similar to other positive strand RNA viruses (64, 103, 133). *Tomato bushy stunt virus* (TBSV), the type species of the genus *Tombusvirus*, codes for p33 and p92 replicase proteins (130), which are essential for TBSV replication (103, 133). Based on the presence of the signature motifs for RNA-dependent RNA polymerase (RdRp) within the C-terminal portion of p92, the predicted function of the p92 is to synthesize viral RNA progenies (94, 133). The function(s) of p33 in TBSV replication is currently unknown.

An interesting feature of tombusviruses is that the larger replicase protein is expressed from the genomic RNA via a ribosomal readthrough mechanism of the p33 termination codon (130, 133). Thus, the N-terminal portion of p92 overlaps with p33. Both p33 and p92 have been proposed to be part of the TBSV replication complexes (133). Accordingly, Western-blot analysis of the partially purified TBSV RdRp preparation revealed the presence of both p33 and p92 in the transcriptionally active RdRp fractions (Pogany and Nagy, unpublished).

Both replicase proteins are localized to membranous structures in infected cells, the putative sites of tombusvirus replication (21, 133). The membrane localization domains in the replicase proteins are present within the N-terminal, overlapping domain (128, 129). Based on their essential roles in tombusvirus RNA replication, it is possible that both tombusvirus replicase proteins can bind to the viral RNA in infected cells. For example, the viral genomic RNA must be recruited to the viral replicase complex after translation (3, 17, 18). In addition, the viral replicase complex, which likely contains both viral- and host-coded proteins, must synthesize a complementary (minus-strand) RNA on the genomic (+) RNA. This is followed by robust plus-strand RNA synthesis utilizing the minus-strand intermediates (60). Tombusviruses also synthesize subgenomic RNAs for the expression of their 3' proximal genes (130). In

addition, tombusviruses are frequently associated with defective interfering (DI) RNAs that are generated from the genomic RNA by multiple deletions (155). Synthesis (replication) of all these RNAs requires both replicase proteins *in vivo* (21, 103, 133). Accordingly, a recently developed *in vitro* assay based on partially-purified RdRp from tombusvirus-infected cells was used to demonstrate that the TBSV RdRp could synthesize complementary RNA on both plus- and minus-stranded TBSV templates (86). Further studies confirmed that the tombusvirus RdRp (obtained for either TBSV or the closely related *Cucumber necrosis virus*, CNV) could bind to TBSV RNA *in vitro* and recognize *cis*-acting sequences, such as the genomic and complementary promoters (107, 108) and a replication enhancer (104), which leads to RNA transcription *in vitro*.

To gain insights into the functions of the tombusvirus replicase proteins, I tested their abilities to bind RNA *in vitro*. I demonstrated that both p33 and p92 of TBSV could bind to TBSV-derived RNA sequences *in vitro*. Gel mobility shift assay performed with a series of truncated recombinant p33 purified from *E. coli,* revealed that an arginine/proline-rich motif (termed RPR motif), which is conserved among tombusviruses, is critical for efficient RNA-binding. The corresponding region in the overlapping domain of p92 may also bind to RNA, although this has not been confirmed. This is because I also found that p92 had additional RNA-binding domains within its non-overlapping C-terminal region; therefore mutations within the RPR motif in p92 did not abolish RNA-binding. Overall, these experiments provide direct evidence that both TBSV replicase proteins bind to viral RNA and this feature might be important for the functions of these proteins during tombusvirus infection.

MATERIALS AND METHODS

Construction of expression plasmids: The full length TBSV cDNA clone [T-100, generous gift of Andy White, (49)] was used to amplify the p33 open reading frame (ORF) using primers 3 and 4 (Table 3.1). The PCR product was treated with Polynucleotide kinase to blunt the ends, digested with *Xmn*I and cloned into a protein expression vector pMAL-c2X (NEB) to generate expression construct p33 (Figure 3.1 and Table 3.2). Construct p33 contains the in-frame fusion of the MBP and the p33 ORF. To express the p92 ORF, I used a mutant TBSV clone [pHS-175, supplied by H. Scholthof, (133)], which had the amber (UAG) termination codon replaced by a

tyrosine codon at the end of the p33 ORF. The mutated p92 gene was amplified via PCR using primers 3 and 5 (Table 3.1) and template pHS-175. The obtained PCR product was cloned into pMAL-c2x to generate construct p92 (Figure 3.1 and Table 3.2), as described for construct p33 above. A similar strategy was used to generate construct p92C (Figure 3.1) by PCR using T-100 template (49) and primers 6 and 11 (Table 3.1), except for the *Xba*I digestion of the PCR product. All other deletion constructs of p33 and p92 (Figure 3.7 and 3.8) were generated by the method described for p33 with the exception that the obtained PCR products were digested with *EcoR*I and *Xba*I prior to cloning (Table 3.2). The primers used to generate the expression plasmids are listed in Table 3.1.

The expression construct N (Figure 3.7) was generated by digestion of construct p33 with *EcoN*I, followed by religation. Constructs R1, R2 and R4 (Figure 3.8) were generated by digesting the p92C clone with *Hind*III, *BstB*I/*Hind*III and *EcoN*I/*Hind*III respectively, followed by religation.

Constructs R15 to R22 (Figure 3.8B) were generated using alanine/serine scanning mutagenesis via two separate PCRs for each construct carried out with the R6 template. One PCR product represented sequences coding for the N-terminal portion of R6, while the other PCR product for the C-terminal portion of the R6 protein. The obtained PCR products (the primer pairs used for PCR are shown in Table 3.2 for R15 to R22) were digested with *Nhe*I, and the appropriate PCR products representing the modified N- and C-terminal portions of R6 were ligated together, and then reamplified by PCR. The PCR products were then digested with *EcoR*I and *Xba*I before cloning into pMAL-c2X.

Purification of p33 and p92 proteins and their derivatives from *E. coli*:

The expression and purification of the recombinant TBSV proteins were carried out as described earlier for the TCV p88 replicase protein (122). Briefly, individual expression plasmids (see above) were transformed into Epicurion BL21-CodonPlus (DE3)-RIL (Stratagene). The overnight cultures from the transformed bacterial cells were diluted to 1:100 in rich growth medium (10*g* tryptone, 5*g* yeast extract, 5*g* NaCl) containing 0.2% glucose and ampicillin 100 μg/ml and grown at 37^0C till the bacterial density reached OD_{600} 0.6 - 0.8. Protein expression was then induced at 14^0C with 0.3 mM IPTG (isopropylthiogalactopyranoside) for 8 to 10 h. The induced cells were harvested at 4000*g* at 4^0C for 10 min, resuspended in ice-cold column buffer

(10mM Tris-HCl pH 7.4, 1mM EDTA, 25mM NaCl, 10mM β-mercaptoethanol), sonicated on ice to disrupt the cells and centrifuged at 9000*g* for 30 min at 4^0C. The supernatant was added to equilibrated amylose resin column (NEB), washed thoroughly with 20 volume of column buffer and then eluted with 10 mM maltose in column buffer. All protein purification steps were carried out in a cold room. The purified recombinant proteins were analyzed in 10% SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) for their purity. The cleavage of MBP/p33 protein was carried out using *Factor Xa* protease (NEB) as recommended by the manufacturer (1 μg of *Factor Xa* was applied for 50 μg of MBP/p33). To measure the amount of purified recombinant proteins, I used the Bio-Rad protein assay, which is based on the Bradford method.

Preparation of RNA templates:

The DNA template representing the 82 bp region III of DI-72 RNA of TBSV was generated by PCR using DI-72XP as a template (156). The primers used were #253 (5'-TTGGAAATTCTCC-TTAGCGAGTAAGACAGACTC-3') and #23 (5' GTAATACGACTCACTATAGGGACCCAA CAAGAGTAACCTG-3'). The PCR template also included the T7 promoter to facilitate synthesis of RNA probes. Both labeled and unlabeled RNAs for gel mobility shift and competition experiments were prepared *in vitro* using T7 RNA polymerase, as described earlier (85, 91). The labeled RNA probes were obtained using [^{32}P] UTP in the T7 transcription reaction (86), followed by removal of free nucleotides using P-30 micro Bio-Spin columns (BioRad). Template DNA was removed by *DNase* I, followed by purification of the RNA transcript with phenol/chloroform extraction and 95% ethanol precipitation. The pellet was washed with 70% ethanol to remove residual salts. The RNA transcripts were quantified by UV spectrophotometer (Beckman), followed by either 1% agarose or 5% polyacrylamide gel electrophoresis (86).

For competition experiments (Figure 3.3), I used unlabeled minus-strand region III of DI-72 RNA. The dsRNA competitor was generated via annealing the positive and negative strands of region III RNA. The dsDNA competitor was the PCR amplified region III DNA fragment, while the ssDNA competitor was an artificially synthesized 51 bp oligo DNA (5'-CCCAGACC-CT CCAGCCAAAGGGTAAATGGGAAAGCCCCCCGTCCGAGGAGG-3'). RNA constructs AU and GC (Figure 3.4) were obtained from Chi-Ping Cheng (27).

Gel mobility shift assay:

Approximately 1 μM of protein was incubated with 5 ng of radioactively labeled minus-stranded region III RNA probe (see above) in a binding buffer [50 mM Tris-HCl pH 8.2, 10mM $MgCl_2$, 1mM EDTA, 10% Glycerol, 200 ng yeast tRNA (Sigma) and 2U of RNase inhibitor (Ambion)] at 25^0C for 30 min (122). After the binding reaction, the samples were analyzed by either 4% non-denaturing polyacrylamide gel electrophoresis performed at 200V or 1% agarose gel electrophoresis run at 100V in TBE buffer in a cold room (122). The gels were dried, exposed and analyzed in a PhosphorImager (Molecular Dynamics Inc.) and quantified using ImageQuant v.1.2 (Amersham). For competition experiments, unlabeled competitors (applied in molar excess as indicated in Figure 3.3-4) were added simultaneously with the labeled RNA probe to the binding reaction. The experimental binding curves were statistically fit to data using Excel spreadsheet.

Northwestern assay:

Approximately equal amount (~2 μg) of recombinant proteins were run in 10% SDS-PAGE and then transferred to PVDF membranes (138). The membranes were renatured at room temperature in a renaturation buffer [10mM-Tris-Hcl pH 7.5, 1mM EDTA, 50mM NaCl, 0.1% Triton X-100 and 1X Denhardt's reagent (44)] (131)with three changes of buffer for 20 min each. The membranes were probed with ^{32}P-labeled RNA (minus-stranded region III, see above) for 1 h, washed three times with the renaturation buffer, air-dried and analyzed using a PhosphorImager.

Biosensor analysis:

The surface plasmon resonance (SPR) experiments were carried out using BIACORE X (Biacore Inc, NJ) at 25^0C as recommended by the manufacturer. Briefly, the running buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3mM EDTA, 0.05% surfactant P-20) was filtered and degassed every time before use. An SA sensor chip (Biacore) was used to immobilize a 20-mer RNA from the 3' end of gTBSV RNA with a biotin label at the 5' end **(**5'-UGUAACGUCUUU-ACGUCGGG- 3', Dharmacon Inc.). The surface of the chip was first preconditioned with three 1-min pulses of 50 mM NaOH in 1M NaCl. Flow cell 1 (Fc1) was used to immobilize 440 resonance units (RU) of RNA, while flow cell 2 (Fc2) was kept as a control surface to account for non-specific binding and bulk refractive index changes upon injection of protein samples, as

suggested by the manufacturer. The recombinant protein samples were diluted with running buffer to 1 μM final concentration before injection. The interactions between RNA and the recombinant proteins were analyzed in real time through a sensogram (Figure 3.2), in which the resonance units (RU) were plotted as a function of time. One RU is equivalent to a change in adsorbed mass of 1pg/mm^2 of the sensor surface (BIA Applications Handbook, Biacore, NJ). All the data shown in Figure 3.2 were corrected based on data obtained from the control (Fc2).

RESULTS

Expression, purification and RNA-binding by recombinant p33 and p92 replicase proteins.

In order to obtain sufficient amounts of soluble TBSV p33 and p92 proteins (Figure 3.1A) for biochemical studies, I over-expressed them in *E. coli* as maltose binding (MBP) fusion proteins. This expression strategy allowed for affinity-based purification of the recombinant p33 and p92 proteins as shown in Figure 3.1B. Standard gel mobility shift experiments with a ^{32}P-labeled RNA probe, representing the 82 nt TBSV replication enhancer, termed region III(-) (35; this probe was used in all the experiments, unless stated otherwise) (104), demonstrated that both the recombinant p33 and p92 could bind to RNA efficiently (Figure 3.1C). In contrast, comparable amounts of bovine serum albumin (BSA) and the full length MBP alone (expressed and purified under the same conditions as the recombinant p33 and p92) did not bind to the RNA probe efficiently (Figure 3.1C), ruling out the possibility that MBP or any contaminating proteins from *E.coli* contribute to RNA-binding activity. Interestingly, the N-terminally truncated version of p92 that included the whole unique sequence of p92 (termed p92C, see Figure 3.1A) also bound to RNA efficiently (Figure 3.1C, lane p92C). This observation suggests that p92 contains a minimum of two RNA-binding regions, one in the overlapping region and another in the unique region (see below). Since p33 and p92 have a common RNA-binding region within the overlapping sequence, I used mostly the recombinant p33 in the experiments below (unless stated otherwise).

To test if the presence of the N-terminal MBP fusion might affect the ability of p33 to bind to RNA and/or affect the migration of the RNA-protein complex in the gel, I compared a purified recombinant p33 preparation, which had the MBP cleaved off by protease *Factor Xa* [sample p33 (cleaved), Figure 3.1D], with the uncleaved recombinant MBP/p33 fusion protein

preparation (sample MBP/p33, Figure 3.1D) in a gel mobility shift assay (Figure 3.1E-F). I found that the overall efficiencies of RNA-binding by the recombinant p33 and MBP/p33 fusion protein were similar (Figure 3.1E-F), suggesting that the MBP fusion does not alter the ability of p33, and likely p92, to bind to RNA. Also, the mobility of the RNA-protein complex was the same for the two preparations (Figure 3.1E-F). Therefore, I used the MBP fusion proteins in the following experiments.

To corroborate the results obtained from gel mobility shift analysis, I carried out surface plasmon resonance (SPR) measurements with a Biacore biosensor as described in the Materials and Methods section. Briefly, the SPR provides data about real-time protein-RNA interactions (29), by measuring the change in refractive index that takes place between the immobilized RNA and the protein that is being passed in an aqueous buffer over the surface of the chip. The refractive index of the medium changes near the chip surface due to change in mass resulting from RNA-protein interaction (29). For this study, I fixed a 20 nt long, 5'-biotinylated TBSV RNA, which includes the minimal genomic (i.e., minus-strand initiation) promoter (108) to the surface of a streptavidin coated chip. The recombinant p33, p92, p92C and MBP proteins diluted with running buffer were passed separately over the immobilized RNA (Figure 3.2). These experiments confirmed that p33 (Figure 3.2A), p92 (Figure 3.2B) and p92C (Figure 3.2C) could bind efficiently and stably to the RNA, while MBP (Figure 3.2D) could not. Detailed kinetic measurements on p33/p92 and RNA interactions will be published elsewhere.

Preferential binding of recombinant p33 to single-stranded RNA.

To test if the recombinant p33 could bind only to single-stranded RNA or to other nucleic acids as well, I used various nucleic acids in template competition experiments, which were evaluated by using gel mobility shift assay. Briefly, the same amounts of ^{32}P-labeled region III(-) RNA probe and purified recombinant p33 were used in the presence of increasing amounts of unlabeled competitors, such as ssRNA, dsRNA, ssDNA and dsDNA (Figure 3.3). All these sequences were derived from the same region of the TBSV genome (minus- or double-stranded region III). These experiments demonstrated that the ssRNA was the best competitor in binding to p33, while ssDNA competed moderately well (Figure 3.3B). In contrast, dsRNA and dsDNA templates were poor competitors under the experimental conditions used. Overall, the

data suggest that p33 is a single-stranded nucleic acid binding protein with the highest preference towards ssRNA.

I also tested if the recombinant p33 could preferentially bind to a TBSV-derived ssRNA sequence by using four different, but comparable-sized ssRNAs in competition experiments as shown in Figure 3.4. One competitor was region III(-) [named DI-RIII(-), Figure 3.4A], while the other three RNAs were nonviral. One of these competitor ssRNAs consisted of an AU-rich (named AU, Figure 3.4A) and another a GC-rich (named GC, Figure 3.4A) artificial sequences (27), while the fourth RNA was tRNA. Comparison of the abilities of these RNAs to compete with the ^{32}P-labeled ssRNA probe in binding to p33 based on the gel mobility shift experiments revealed that the region III(-) was far the best competitor among the ssRNAs (Figure 3.4B-C). The artificial AU-rich RNA was a moderately good competitor, while the artificial GC-rich RNA and the tRNA were poor competitors (Figure 3.4B-C). These experiments suggest that p33 preferably binds to the TBSV-derived sequence and at a lower efficiency to an AU-rich sequence. Note that the artificial AU and GC competitor RNAs have double- versus single-stranded regions comparable in length to that present in DI-RIII (-) (Figure 3.4A). Thus, increased binding by p33 to DI-RIII (-) is likely due to the favorable sequence/structure of region III (-).

Cooperative binding of p33 to ssRNA.

In order to characterize the binding behavior of p33 to TBSV RNA, I incubated progressively increasing amounts of recombinant p33 proteins in the presence of a ^{32}P-labeled probe followed by gel mobility shift assay. In the presence of small amount of p33 (samples on left side of Figure 3.5A), the RNA probe was found mostly in unbound form, while increasing the amount of p33 in the binding reaction resulted in rapid transition of the RNA probe to bound form (samples on right side in Figure 3.5A). The absence of intermediately shifted bands, resulting from limited binding of the probe by p33**,** between the retarded and the free probe in the gel (Figure 3.5A), suggests that most of the RNAs are either coated with p33 or not bound to p33 at all- depending on the amount of protein present in the RNA-binding reactions. The hill coefficient of 1.8 for RNA-binding by p33 also support that p33 binds to RNA in a co-operative manner (Figure 3.5A). Interestingly, when I used a truncated p33 containing only a 60 amino acid segment of p33 that includes the RNA-binding domain (construct C10, Figure 3.5B), then I

still observed cooperative RNA-binding by this truncated p33, as is clearly evident from the rapid transition from the totally free to completely bound state of probe RNA with a marginal rise in protein concentration (Figure 3.5B).

The ability of the truncated p33 to bind co-operatively to RNA was surprising (because of the small size of the protein); therefore I also used the three-membrane sandwich method as described by Pata *et al.* (36) to confirm the above finding. Briefly, the first polysulfone membrane, which has low affinity to both RNA and protein, was expected to retain only large RNA/protein complexes (36). The second nitrocellulose membrane can bind to small protein/RNA complexes (those, which have not been retarded on the first membrane), while the third, positively charged membrane retains all unbound RNA. Increasing amounts of C10 protein were added to the same amount of RNA probe, as described in Figure 3.6. Aliquots of the binding reactions were filtered through the three-membrane sandwich, followed by detection of the retarded ^{32}P-label on each membrane (Figure 3.6). These experiments demonstrated that the RNA probe was retarded as part of a large complex in the presence of elevated amounts of C10 proteins (see membrane #1, Figure 3.6), while the formation of small RNA/protein complex was relatively inefficient (see membrane #2, Figure 3.6). Overall, the shown data are most consistent with the model of co-operative RNA-binding, which predicts that most of the RNA is present in a coated form (seen as a large complex), when sufficient amount of the truncated p33 is present in the binding reaction.

Mapping the RNA-binding site in p33.

In order to identify the RNA-binding domain in p33, and in the corresponding region of p92, I made a series of 13 overlapping constructs with various truncations within the p33 gene (Figure 3.7A). The truncated p33 proteins were over-expressed in *E. coli*, followed by affinity purification of the proteins as MBP fusions (see above), and SDS-PAGE analysis **(**Figure 3.7B). The ability of the obtained truncated p33 proteins to bind to the ^{32}P-labeled probe was tested in gel mobility shift experiments (Figure 3.7C). These experiments revealed that the C-terminal segment of p33 contains an RNA-binding domain (lane C1, Figure 3.7C), while the N-terminal segment does not bind to RNA under the *in vitro* conditions (lane N, Figure 3.7C). Further testing of C-terminally nested segments of p33 indicated that the central portion of the C-terminal region harbors the RNA-binding domain (compare lanes C2-C4 versus C5-C6, Figure

3.7C). This was further supported by the observation that deletions of the C-terminal 20-56 amino acids in p33 did not affect RNA-binding (lanes C7-C8, Figure 3.7C). Deletion of 103 amino acids from the C-terminus, however, abolished RNA-binding (lane C9, Figure 3.7C). Deletions starting from both the N- and C-termini defined that the shortest p33 derivative that still bound to RNA was 30 amino acid long and covered the central portion of the C-terminal region (lane C11, Figure 3.7C). Deletion of an 8 amino acid portion of construct C11, which has the following sequence: TGRPRRRP, completely abolished RNA-binding (lane C12, Figure 3.7C). Indeed, all the p33 derivatives that carried the above arginine/proline-rich motif bound to the RNA probe efficiently, while those lacking this domain did not bind to the RNA probe (Figure 3.7C). Based on these data, I propose that the arginine/proline-rich motif, which I call RPR-motif, might be the primary RNA-binding site in p33 and p92.

To confirm that the RNA-binding by the truncated p33 derivatives is an inherent feature of these proteins, and not due to the presence of a contaminating protein from *E. coli*, I analyzed the RNA-binding ability of a selected group of short deletion mutants of p33 in Northwestern assay (Figure 3.7D). Briefly, the full-length p33 and four of the purified recombinant p33 derivatives, C9, C10, C11 and C12 (Figure 3.7A) were subjected to SDS-polyacrylamide gel electrophoresis, followed by blotting to a PVDF membrane. This was followed by probing the membrane with a ^{32}P-labeled RNA probe as described in the Materials and Methods section. Importantly, I could detect the protein/RNA complex by autoradiography in the portion of the membranes that contained proteins p33 (WT), C10 and C11, but I could not detect RNA/protein complex in case of C9 and C12 (Figure 3.7D, panel Northwestern). Therefore, the Northwestern analysis confirmed the results obtained in the above gel mobility shift experiments that the RPR motif is the core region in RNA-binding in p33.

Mapping additional RNA-binding sites in p92.

In order to identify the RNA-binding domains present in the unique segment of p92 (see p92C in Figure 3.1A), I used a series of deletion derivatives of p92C, which were over-expressed and purified from *E. coli*. Surprisingly, many of the expressed truncated p92C proteins were unstable in *E. coli*, preventing us to generate large enough numbers of p92C derivatives that could have been useful to pinpoint the RNA-binding sites precisely (not shown). To this end, I was able to obtain 14 truncated p92C proteins in large enough amounts suitable

for biochemical assays (Figure 3.8A). Nested truncations of 82-133 amino acids from the C-terminus of p92C resulted in reduced RNA-binding, suggesting that this portion of p92C contributes to RNA-binding, although the residual RNA-binding by these proteins was still significant (lanes R1-R2, Figure 3.8C). Deletion of 434 amino acids from the C-terminal end of p92C abolished RNA-binding (R4, Figure 3.8C). p92C derivatives carrying segments from the central portion of the protein were found to bind to RNA (lanes R5-R11, Figure 3.8C). In contrast, p92C derivatives, which lacked the above central segment and the C-terminal segment, did not bind efficiently to RNA (lanes R12-R13, Figure 3.8C). Interestingly, protein R14 that contained the 303 amino acid C-terminal segment of p92C bound to RNA, though with reduced efficiency when compared to p92C (lane R14, Figure 3.8C). Overall, analysis of RNA-binding by the above deletion series of p92C revealed that two segments of p92C are involved in RNA-binding: the central segment (represented by R6 and R7) and the very C-terminal segment (~131 amino acids). These two segments can bind to the RNA independently of each other, although each shows somewhat reduced efficiency when compared to p92C.

To further delineate the ~108 amino acid long central RNA-binding site in p92C, I introduced separately five clusters of 5-7 alanine/serine mutations into construct R6 (Figure 3.8A) as shown schematically in Figure 3.8B. I targeted short regions that contained clusters of positively charged amino acids, such as arginine, lysine and histidine (lanes R15-R19, Figure 3.8B). Surprisingly, all these mutated proteins retained their abilities to bind to RNA (lanes R15-R19, Figure 3.8C), suggesting that (i) either the selected amino acids are not involved in RNA-binding, or (ii) several amino acids, possibly located at different parts of the central segment of the p92C protein, are brought together by protein folding to form an RNA-binding groove as is the case for other viral RdRps (16, 45). Therefore, it is possible that the cluster mutagenesis approach was not effective since it did not modify all the important amino acids at once. Therefore, I decided to introduce 27-45 amino acids deletions into construct R6 as shown schematically for constructs R20-R22 (Figure 3.8B). These deletion derivatives bound to RNA poorly (especially protein R22, but see also R20 and R21, Figure 3.8C). Overall, this analysis suggests that an RNA-binding region might be present within the 45 amino acid long segment of the central region in p92. Since I had unexpected difficulties in obtaining many other mutants (due to protein stability problems during expression in *E. coli*, see above), I could not further

map the actual amino acids involved in RNA-binding for the central segment in p92C using this approach.

DISCUSSION

Binding of the replicase proteins to the viral RNA is predicted to be important during many steps of the viral infectious cycle, including the recruitment of the viral RNA to the site of replication, recognition of cis-acting elements during replication, complementary RNA synthesis, etc (3, 17, 18, 60). To gain insights into the replication process of tombusviruses, I have characterized the ability of p33 and p92 replicase proteins to bind to RNA. Although I find that binding of p33 to ssRNA is the strongest, ssDNA was also bound by p33 with moderate efficiency (Figure 3.3). In contrast, binding to dsRNA or dsDNA was poor, suggesting that p33 is a single-stranded nucleic acid binding protein. Interestingly, the ability of p33 replicase protein to bind single-stranded nucleic acids is similar to other plant viral proteins, such as movement proteins and coat proteins (6, 9, 28, 32, 33, 54, 73, 100, 102, 124, 126, 147). In spite of its ability to bind to viral as well as nonviral RNAs, p33 showed preference for the TBSV-derived sequence, which was the best substrate among four different, similar-sized ssRNA templates (Figure 3.4). An artificial AU-rich template was also bound by p33 moderately well. In contrast, a GC-rich RNA and tRNA bound poorly to p33 (Figure 3.4). Overall, the observed selectivity of p33 in RNA-binding may not be enough for p33 and possibly p92 to bind to only TBSV-related RNAs in infected cells. It is possible that other factors or a combination of factors are needed to achieve such level of selectivity in template use.

Based on p33-RNA binding experiments, I propose that that both p33 (Figure 3.5) and p92 (data not shown) bind RNA in a co-operative manner. Results (Figure 3.5A) show typical "all or none" behavior characteristic of co-operative binding (39, 80, 83). The RNA bound by the recombinant p33, which was either fused to MBP (Figure 3.5A) or was cleaved off the MBP domain (Figure 3.1F), stayed in the well, probably due to the large size of the complex. In contrast, the p33 or p33 fused to MBP did migrate into the gel in the absence of RNA, as shown in Figure 3.1D. In addition, a 60 amino acid long truncated p33 (protein C10, Figure 3.7A) also bound to RNA in a co-operative manner (Figure 3.5B). Indeed, detection of large complexes

between the truncated p33 and the RNA probe in the three-membrane sandwich assay (Figure 3.6) is also consistent with co-operative RNA-binding by this truncated p33.

The proposed ability of p33 to bind RNA in a co-operative manner suggests that, after the initial binding by the first p33 to the RNA, subsequent binding of additional p33 molecules to the same RNA is greatly facilitated not only by the stabilizing effect coming from binding to the RNA, but also by p33-p33 (protein-protein) interactions. This may lead to complete coating of the RNA with p33 molecules. In support of this model, I found p33-p33 interactions *in vitro* (Rajendran and Nagy, unpublished observation).

Binding to RNA in co-operative manner can be advantageous for both the viral RNA and the replicase proteins, since this may increase the stability of RNA-protein complexes inside the infected cells. Therefore, it is not surprising that many viral proteins, including 2D of poliovirus, NS5B RdRp of *Hepatitis C virus* (150), viral coat proteins and plant viral movement proteins (6, 9, 28, 32, 33, 54, 73, 100, 102, 124, 126, 147) were found to bind viral RNAs in a co-operative manner. The functional significance of co-operative binding by p33 and p92 is currently not known. It is possible that p33 can coat the viral ssRNAs in infected cells, which may be beneficial during template recruitment and/or replication. The p33 coated viral RNAs may be more resistant to nucleases and less accessible to host-mediated gene silencing than free viral RNAs (149, 152). Since co-operative binding depends on the amounts of replicase proteins and viral RNAs present in the cells, it is likely that this feature may be important for the function of p33, which is 20-fold more abundant than the p92 replicase protein in the infected cells (133). The ability to bind co-operatively to RNA, possibly in combination with p33, may also be important for the function of p92 during replication. For example, co-operative binding between p92 and p33 may facilitate recruitment of the less abundant p92 proteins to the viral RNA in infected cells. Alternatively, binding of p92 to the viral RNA templates in a co-operative manner may enhance its RdRp activity, as it has been shown for the Poliovirus 2D (RdRp) protein that binds RNA co-operatively (51, 112). In addition, it has been shown that the poliovirus 2D protein forms large complexes, which might be involved in virus replication (78).

Deletion analysis revealed that p33 has one high-affinity RNA-binding site in its C-terminal region. This RNA-binding region includes the arginine/proline-rich (RPR) motif, which has been shown to be essential for RNA-binding (compare proteins C11 and C12 in Figure 3.7C). Mutagenesis of the RPR motif in p33 has also revealed that this domain is essential for

the replication of a tombusvirus and an associated DI RNA *in vivo* (109). Mutations within the RPR motif have affected subgenomic RNA synthesis as well (109). These observations demonstrate that the RPR motif plays a central role in viral RNA synthesis/metabolism.

The RPR motif is highly conserved among the replicase proteins of tombusviruses and the related *Turnip crinkle virus* (TCV) (Figure 3.9B). Accordingly, a recombinant TCV p88 protein, which is the RdRp protein similar to p92 of TBSV, contains an RPR motif in its N-terminal portion and it has been shown to bind RNA efficiently *in vitro* (122). Moreover, the RPR motif in TBSV is similar to the previously proposed ARM-motif (arginine-rich motif), which is present in several viral- and host RNA-binding proteins (34), including the trans-activator protein (Tat) of *Human immunodeficiency virus-1* (Figure 3.9C) (12, 23). The structure of the ARM-motif, however, is rather different in several RNA-binding proteins. Therefore, it is not known if the ARM-motif functions similarly in different proteins (34).

Protein analysis software predicts that the RPR motif in TBSV p33 constitutes a hydrophilic pocket and it is exposed to solvent (Figure 3.9A). This structural prediction is compatible with the proposed function of the RPR motif in RNA-binding. The secondary structure analysis predicts that the RPR motif itself has mainly turns (Figure 3.9A), suggesting that this motif may take stable conformation upon binding to RNA. Indeed, the ARM-motif in the HIV-1 Tat protein (116, 137) constitutes flexible turns that can adopt to two different conformations upon binding to two different viral RNAs (137).

In addition to its role in p33 functions, the RPR motif is likely functional in p92 as well, since its N-terminal sequence (the so-called overlapping domain) is identical with p33. I could not test this *in vitro*, however, since p92 has additional RNA-binding sites, one in the central part (named RBR2 region, located in the vicinity of the RdRp signature motifs, Figure 3.10A) and another in the C-terminal segment (named RBR3, Figure 3.10A), which can facilitate binding to RNA in the absence of the RPR motif (see construct p92C, Figure 3.1B). Nevertheless, separate mutagenesis of the RPR motif in p33 and p92, followed by testing virus replication in protoplast, revealed that the RPR motif in p92 is essential for tombusvirus replication (109). Therefore, based on (i) sequence identity between p33 and the overlapping domain of p92, and (ii) the *in vivo* requirement of the wild type RPR motif in p92 for tombusvirus replication, I propose that the RPR motif in p92 is a functional RNA-binding site. However, the actual function(s) of the RPR motif in p92 will need further studies.

Sequence comparison of the central RBR2 RNA-binding domain within the unique p92C segment of TBSV with other tombus- and related viruses revealed that this 45 amino acid region contains several highly conserved amino acids, including motif-F (Figure 3.10B). Interestingly, an RNA-binding region, which includes the motif-F, has also been defined for the NS5B RdRp protein of HCV (16), as shown in Figure 3.10B. This region has been shown for HCV RdRp to include some of the basic amino acids that line up the RNA-binding groove and bind to RNA template during viral replication (16). In addition, the highly conserved arginine and isoleucine or leucine in motif-F has been found to bind to rNTPs (71). These observations suggest that this 45 amino acid long region in TBSV p92 may posses the amino acids responsible for binding to both rNTPs as well as to the template RNA. Based on its similarity in location to RNA-binding region in the HCV NS5B, the function of the RBR2 RNA-binding region in p92 of TBSV may be "to channel" the RNA template towards the active site in the RdRp. The significance and possible function(s) of the C-terminal RBR3 RNA-binding region in TBSV p92 is currently under further investigation.

In summary, our *in vitro* studies defined the RNA-binding regions in p33 and p92 replicase proteins of TBSV, which likely play major role(s) in tombusvirus RNA replication/metabolism. These results will open the way for future studies on the functions of these proteins in particular, and on tombusvirus replication in general.

Table 3.1. List of primers used for PCR to generate expression constructs

Primer	Nucleotide positions[a]	Sequence
3	169-186	GAGACCATCAAGAGAATG
4	1056-1039	CT(A/T)TTTGACACCCAGGGA
5	2622-2605	TC(A/C)AGCTACGGCGGAGTCGG
6	1057-1075	GGAGGCCTAGTACGTCTAC
10	1056-1039	GCAGTCTAGACTATTTGACACCCAGGGA
11	2622-2605	GCAGTCTAGATCAAGCTACGGCGGAGTC
48	616-632	GAGGAATTCTACGCTACCCTACCTAG
49	1477-1493	GAGGAATTCGTCTCCGAGAGGGATAG
50	2016-2002	GAGTCTAGATCAAATCCCCAGATGACG
75	706-723	GAGGAATTCTGTCTGGTGGTTGAGCCG
76	796-813	GAGGAATTCACTGGGCGCCCTCGTCGA
77	952-969	GAGGAATTCGATGTCATATTGCCTTTG
78	871-888	GAGTCTAGACTATCTATTCTCTGGACTGTT
79	1381-1399	GAGGAATTCGCCCACCGACTTGGGTATG
92	973-990	GAGTCTAGACTAGACAAAACAGCATCCAAT
93	1306-1323	GAGGAATTCACTCCACAACCTACCAAA
139	1624-1641	GAGTCTAGACTAAACAGCTTTCATCAGCTT
183	667-678	GAGGAATTCGCACGAGCACACATGGAG
184	727-738	GAGTCTAGACTATTTACCCTTAAGTTCCCT
383	820-837	GAGGAATTCTATGCGGCAAAGATCGCA
385	1537-1554	GAGTCTAGACTAAGGTGCTGGGTCACCCTT
664	2059-2076	GAGTCTAGACTAATTCCTGCGTTCGACAAT
665	2209-2226	GAGTCTAGACTAGTTTCGAACCATCTTCCA
777	1357-1371	GCGGGCTAGCGCTGCAGCGGTTTGTGAGAAGG
778	1319-1335	GCGGGCTAGCAGCCGCAGCTCCAAAGGCTCCTTTG
779	1450-1465	GCGGGCTAGCGCTGCAGCCGTGGAGAGTCTGC
780	1412-1428	GCGGGCTAGCAGCCGCAGCCGCACCGCTGTAGTATG
781	1498-1513	GCGGGCTAGCGCTGCCTTGACTACCTTCGTAA
782	1466-1482	GCGGGCTAGCCGCGGAGACAGGTGTGATAT
783	1558-1573	GCGGGCTAGCGCTGCGGTGATTCAGCCTCGAA
784	1520-1536	GCGGGCTAGCAGCCGCAGCAGACGTCGATATCTTCT
785	1582-1597	GCGGGCTAGCGCTGCGTACAATGTGGAACTTG
786	1550-1566	GCGGGCTAGCGGCCTGAATCACCCGAGGTG
787	1615-1630	GCGGGCTAGCATGGAATCCAAGCTGA

[a]Position on the TBSV genomic RNA.

Table 3.2. List of primers and templates used for PCR to generate expression constructs

Constructs	Primers[a]	Templates
p92	3/5	pHS175
p33	3/4	T-100
C1	48/10	p33
C2	183/10	p33
C3	75/10	p33
C4	76/10	p33
C5	383/10	p33
C6	77/10	p33
C7	183/92	p33
C8	48/78	p33
C9	48/184	p33
C10	75/78	p33
C11	76/78	p33
C12	383/78	p33
p92C	6/11	p92
R3	6/50	p92C
R5	93/385	p92C
R6	93/139	p92C
R7	79/139	p92C
R8	79/385	p92C
R9	79/50	p92C
R10	49/50	p92C
R11	49/11	p92C
R12	661/664	p92C
R13	661/665	p92C
R14	661/11	p92C
R15	93/778, 777/139	R6
R16	93/780, 779/139	R6
R17	93/782, 781/139	R6
R18	93/784, 783/139	R6
R19	93/786, 785/139	R6
R20	93/784, 787/139	R6
R21	93/782, 785/139	R6
R22	93/782, 787/139	R6

[a]Two separate PCR reactions, performed with two primer pairs, were used to generate several constructs, as described in the Materials and Methods section.

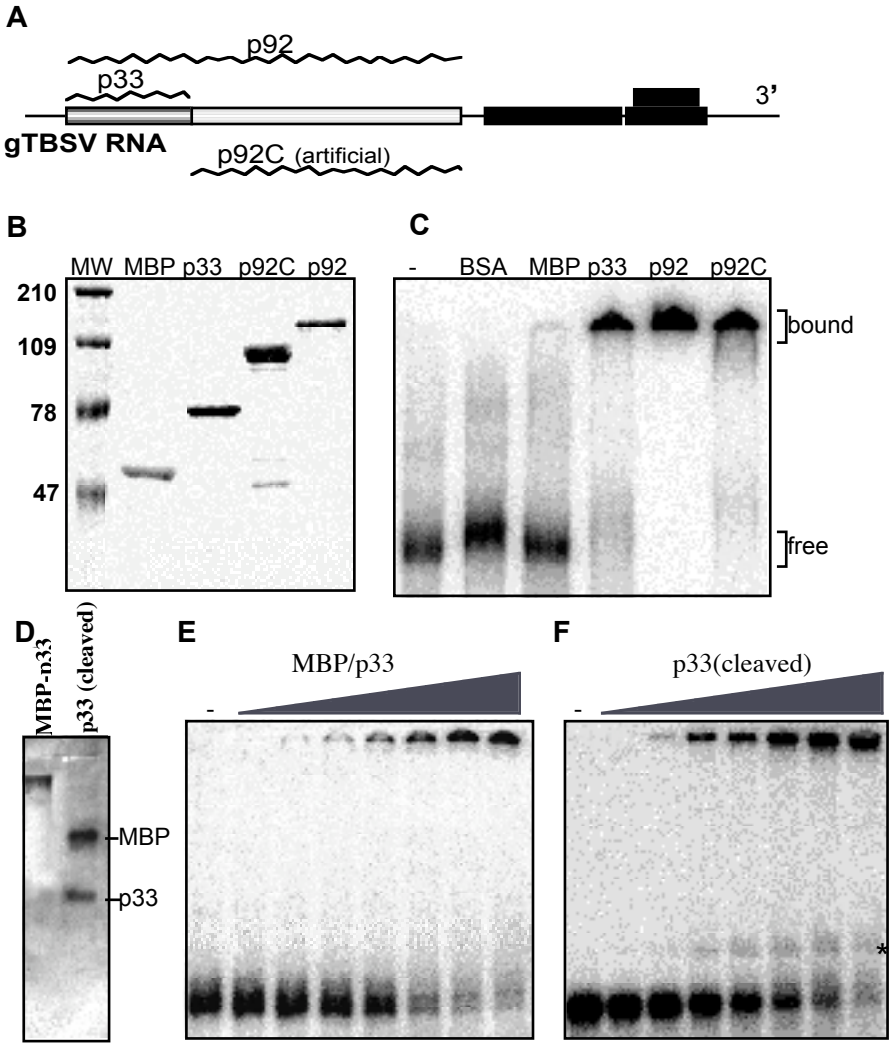


Figure 3.1. RNA-binding by the recombinant p33 and p92 replicase proteins of TBSV *in vitro*. (A) Schematic representation of the expression strategy of the replicase genes in TBSV. The plus- strand genomic RNA is used in the infected cells to produce the replicase proteins p33 and p92 (wavy lines above the boxes represent the individual replicase proteins expressed from the TBSV RNA). p92 is translated via ribosomal readthrough of the stop codon at the end of the p33 open reading frame. An artificial deletion derivative of p92, termed p92C, was also generated to produce the unique, non-overlapping portion of p92 protein, which contains the signature motifs of RNA-dependent RNA polymerases. The other genes shown in black boxes are expressed from subgenomic RNAs (130). (B) SDS-PAGE analysis of purified recombinant TBSV replicase proteins from *E. coli*. The p33, p92 (133) and the truncated p92C genes were cloned into pMAL-c2X and expressed as C-terminal fusion proteins with the maltose binding protein. The fusion proteins were purified using amylose resin affinity chromatography and analyzed in a 10% SDS-

PAGE gel. Lane MW shows molecular weight markers (in thousands) on the left; while the other lanes contain samples purified from *E. coli*: lane 2, MBP/lacZ fusion protein; lane 3, MBP/p33; lane 4, MBP/p92C; lane 5, MBP/p92. (C) A gel mobility shift assay showing interactions between the recombinant TBSV replicase proteins and TBSV RNA. The 82 nt ^{32}P-labeled minus-stranded region III RNA was separately incubated with bovine serum albumin (BSA) and one of the recombinant proteins (1 μM) as shown, in a binding buffer at 25^{0}C for 30 min and then analyzed in 4% non-denaturing polyacrylamide gel. The unbound, free RNA probe and the shifted (bound) RNA/protein complexes are marked on the right. (D) Mobility of the recombinant MBP/p33 and the p33 (after cleavage with *Factor Xa*) in the absence of RNA probe in a 1% agarose gel. The electrophoresis was performed under the same conditions as in panels E-F. The proteins were stained with Coomassie blue. (E-F) Comparison of the RNA-binding abilities of two recombinant p33 preparations either fused with MBP or cleaved off the MBP. The gel mobility shift assays were performed as in panel C, except increasing amounts of MBP/p33 (0.03, 0.06, 0.13, 0.27, 0.65, 1.3 and 2.6 μM protein/ per lane) or recombinant p33 (cleaved) (0.03, 0.06, 0.13, 0.25, 0.50, 1.0, and 2.0 μM total protein/per lane) were applied. The samples were analyzed using 1% agarose gel electrophoresis run at 100V in a cold room. Note that the faint band located between the fully shifted (at the top) and free (at the bottom) RNA bands (marked with *) in panel F was not consistently detectable when I repeated these experiments.

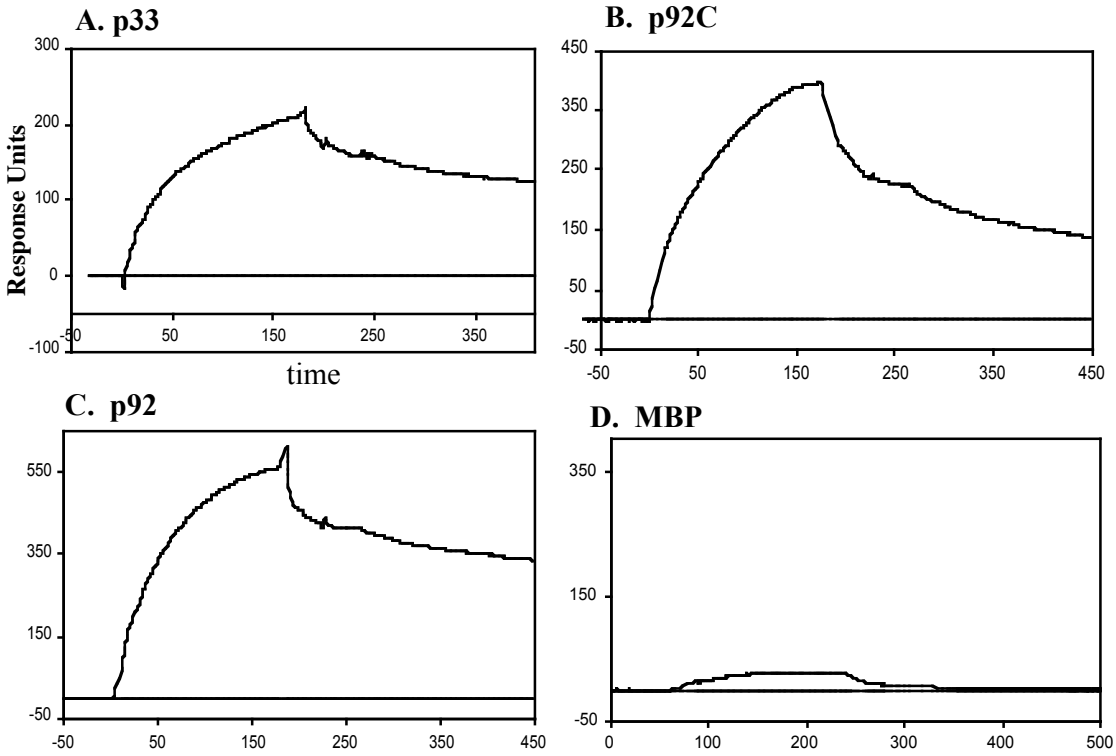


Figure 3.2. SPR analysis of interactions between the TBSV RNA and the recombinant replicase proteins. SPR analysis was carried out using BiacoreX, as described in the Materials and Methods. A 5'-biotinylated 20-nt RNA derived from the 3' end of TBSV RNA (107) was immobilized (440 RU) onto a streptavidin-coated sensor chip. The purified recombinant proteins (1 μM) in a binding buffer were passed over the RNA-coated surface of the chip, and the change in mass due to interaction between RNA and protein altered the refractive index of the medium, which was recorded in real time in a sensogram. The time allowed for association and dissociation phases was 180 sec and 200-260 sec, respectively, in each protein-RNA interaction assay. Interactions between the target RNA and p33, p92C, p92 (all three were tested as MBP fusion proteins) and MBP are shown in panels A, B, C and D, respectively. The sensogram data were corrected for non-specific binding based on data obtained using a control surface (shown as the baseline), which was free of RNA. Note that the RU units were not normalized; therefore they should not be compared directly for proteins of different molecular mass. The small positive response in the MBP test (in panel D) is likely due to loose, non-specific binding, which was subsequently washed out with the buffer as is evident from the curve reaching the baseline rapidly during the washing step.

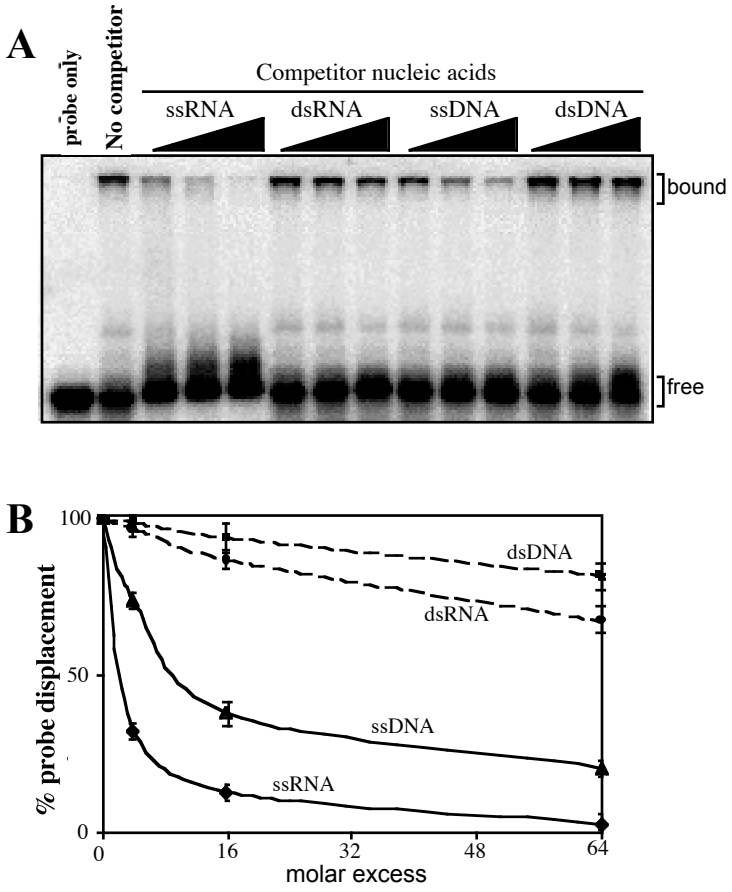


Figure 3.3. Preferential binding of the recombinant p33 to single-stranded (ss) RNA. (A) 1 μM of purified recombinant p33 was incubated with ^{32}P-labeled ssRNA probe, representing the 82 nt minus-stranded region III in the absence or presence of increasing amounts (in 4, 16 or 64-fold excess) of unlabeled competitors (as shown on the top of the figure). The experiments were repeated twice. (B) Graphical representation of data obtained in panel A. The extent of competition was quantified as the percentage of displaced labeled RNA probe from the bound complex (indicated on the right in panel A) using a PhosphorImager and Image Quant (v1.2).

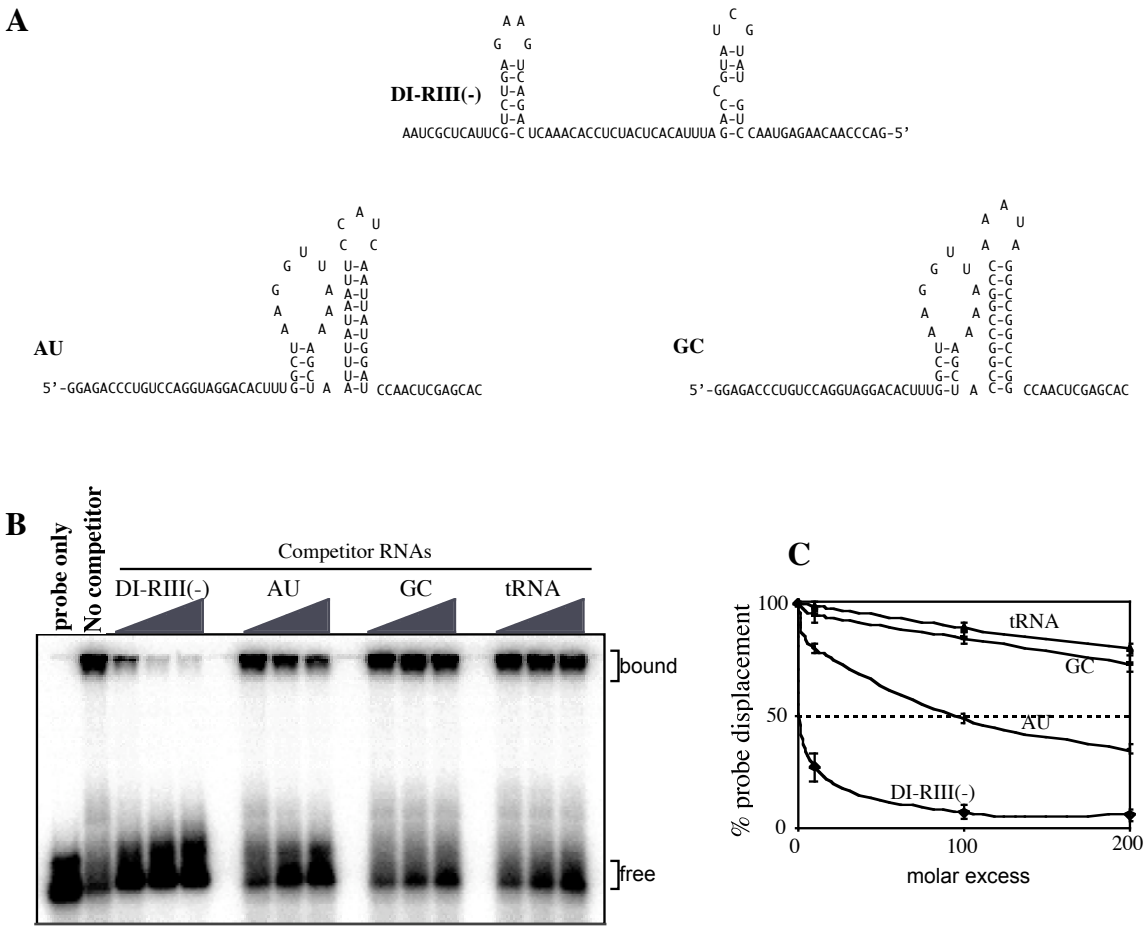
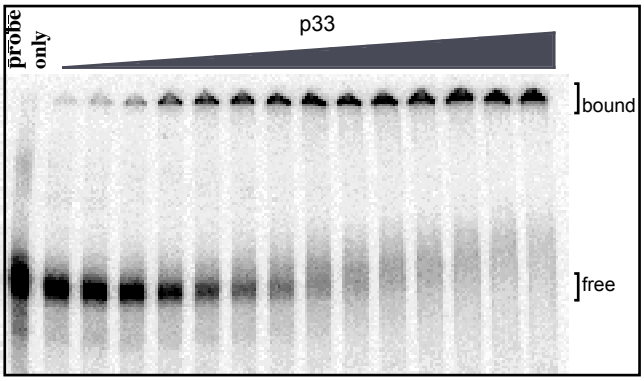
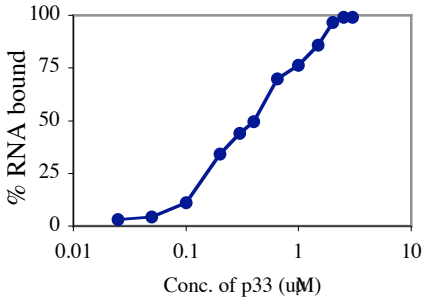
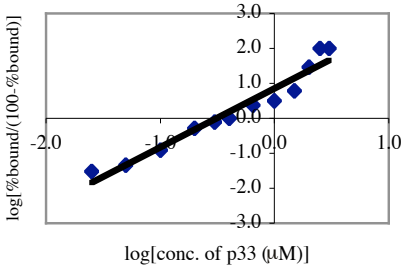


Figure 3.4. Testing binding preference of the recombinant p33 to RNA. (A) Predicted secondary structures of the competitor RNAs, including the 82 nt minus-stranded region III of TBSV [DI-RIII(-)], the similarly sized artificial AU and GC templates (27). Note that the double- versus single-stranded regions are comparable in length in these competitor RNAs. (B) Unlabeled competitor RNAs at increasing amounts (in 10, 100 or 200-fold excess) were added to the mixture containing the labeled probe (82 nt minus-stranded region III, see Figure 3.1B) and 1 μM purified recombinant p33 and the bound complexes were analyzed in gel mobility shift assay. The tRNA was from yeast. (C) Graphical representation of data obtained in panel B. The quantification of the experiment was done as described in the legend to Figure 3.3. The experiments were repeated twice.

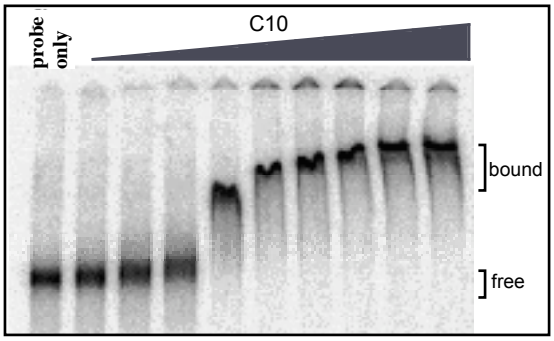
A

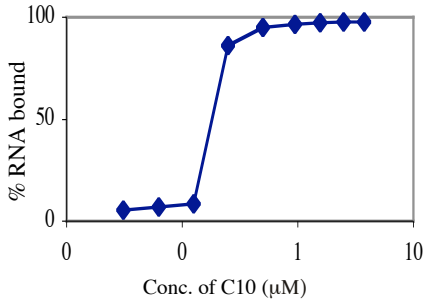

probe only
p33
bound
free


% RNA bound
100
75
50
25
0
0.01
0.1
1
10
Conc. of p33 (uM)


log[%bound/(100-%bound)]
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
-2.0
-1.0
0.0
1.0
log[conc. of p33 (μM)]

B


probe only
C10
bound
free


% RNA bound
100
50
0
0
0
1
10
Conc. of C10 (μM)

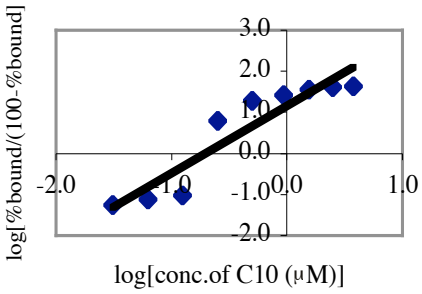

log[%bound/(100-%bound)]
3.0
2.0
1.0
0.0
-1.0
-2.0
-2.0
-1.0
0.0
1.0
log[conc.of C10 (μM)]

Figure 3.5. Cooperative RNA binding by the full length and truncated recombinant p33. (A) Increasing molar concentrations of p33 were incubated with ^{32}P-labeled probe [region III (-), Figure 3.1B] and the RNA-protein complex was resolved in a gel mobility shift assay. The middle panel shows the semi-log plot of the percentage of RNA bound versus molar concentration of p33 determined by using a PhosphorImager. The bottom panel shows a Hill plot of the experimental data obtained from the binding assay (top of panel A). To deduce the Hill coefficient, I used the points that correspond to the middle of the curve in the semi-log graph (as cooperativity is negligible at the extremes). (B) The gel mobility shift assay was performed with a truncated recombinant p33, termed C10 (Figure 3.7), which contained only a 60 amino acids long region including the RNA binding site in the recombinant p33. The middle and bottom graphs are prepared as described in panel (A) above.

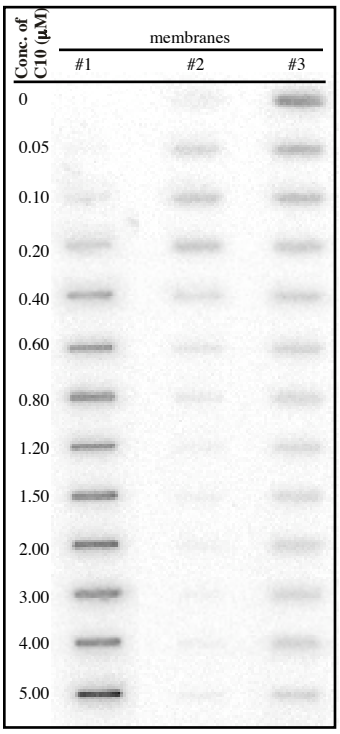


Figure 3.6. Three membrane-sandwich experiments to demonstrate co-operative RNA binding by a truncated recombinant p33. The purified recombinant C10 protein (Figure 3.7, the applied amount is shown on the left side) was incubated with ^{32}P-labeled region III(-) RNA probe for 30 minutes at 25^{0}C, followed by filtering the reaction products through a three-membrane sandwich (112) in a slot-blot apparatus. The membranes were then washed twice with (300 μl) binding buffer, air-dried and analyzed using a PhosphorImager. The membranes were used in the following order: polysulfone on the top (column #1), nitrocellulose in the middle (column #2), and Hybond N+ on the bottom (column #3). Note that the unbound label at the highest protein concentration likely represents labeled ribonucleotides (which, in spite of a purification step, are still present in the RNA probe as a carry over from the labeling reaction) that were unbound to protein C10.

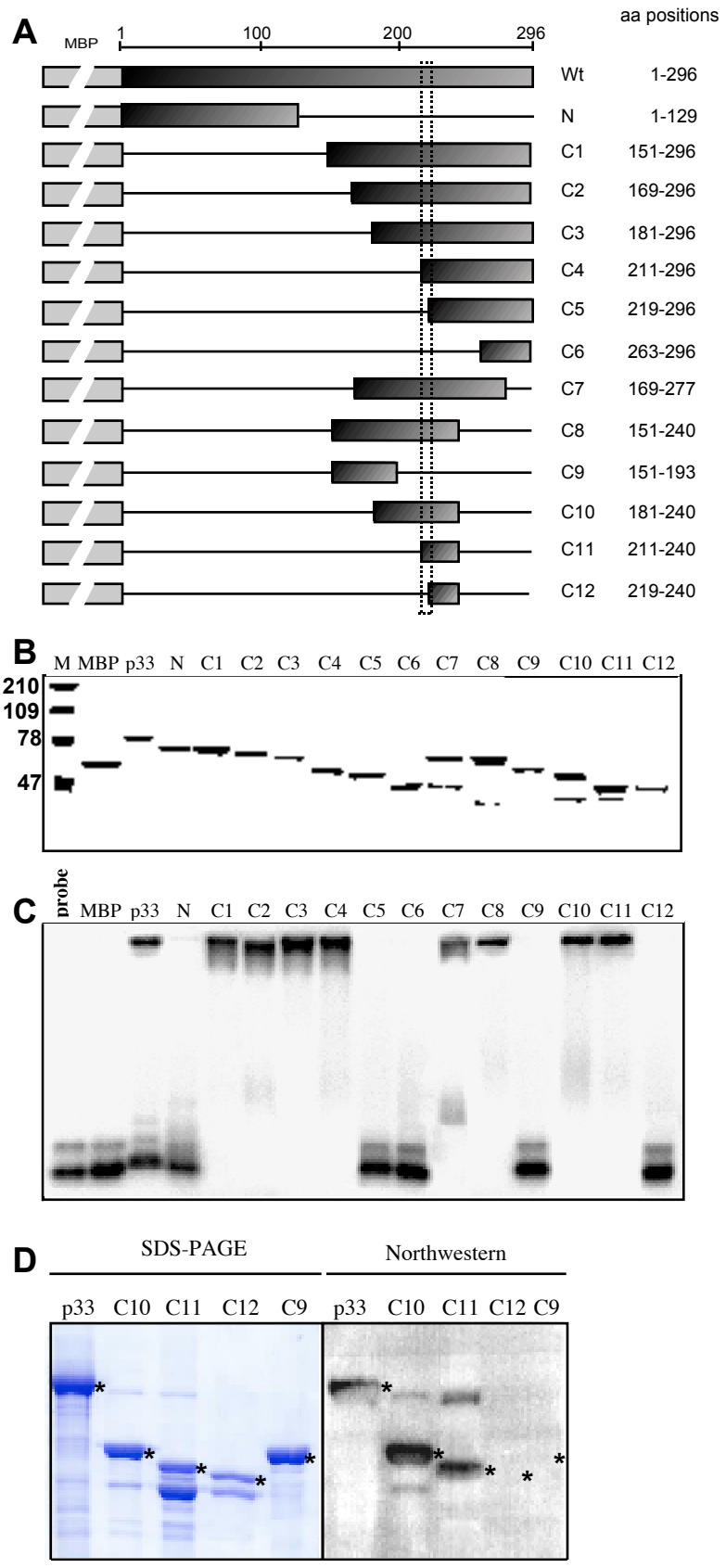

A
aa positions
MBP
1
100
200
296
Wt 1-296
N 1-129
C1 151-296
C2 169-296
C3 181-296
C4 211-296
C5 219-296
C6 263-296
C7 169-277
C8 151-240
C9 151-193
C10 181-240
C11 211-240
C12 219-240
B
M MBP p33 N C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12
210
109
78
47
C
probe
MBP p33 N C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12
D
SDS-PAGE
Northwestern
p33 C10 C11 C12 C9
p33 C10 C11 C12 C9

Figure 3.7. Mapping the RNA binding domain in the recombinant p33. (A) A schematic representation of the deletion derivatives of p33. The names of the constructs and the positions of the amino acid present in the truncated proteins are shown on the right. These truncated p33 proteins were expressed in *E. coli* as fusions to MBP (indicated schematically by a dotted box). The shaded boxes indicate the portions of p33 protein that were present in given expression constructs. The horizontal lines represent the deletions. (B) SDS-PAGE analysis of the purified recombinant proteins in a 10% polyacrylamide gel stained with Coomassie brilliant blue. The lane MW refers to molecular weight markers (in kDa). (C) RNA binding activities of the truncated p33 proteins. The labeled RNA probe and the gel mobility shift assay were as described in the legend to Figure 3.1B. Equimolar concentrations (2 μM) of proteins were used for gel shift assay. (D) Northwestern analysis of selected truncated p33 proteins. The purified recombinant proteins (~2 μg) were run in 10% SDS-PAGE as shown in the left panel, transferred to a PVDF membrane and then probed with a ^{32}P-labeled probe [region III(-), Figure 3.1B]. The positions in the Northwestern blot, which represent a particular recombinant protein, were marked with asterisks.

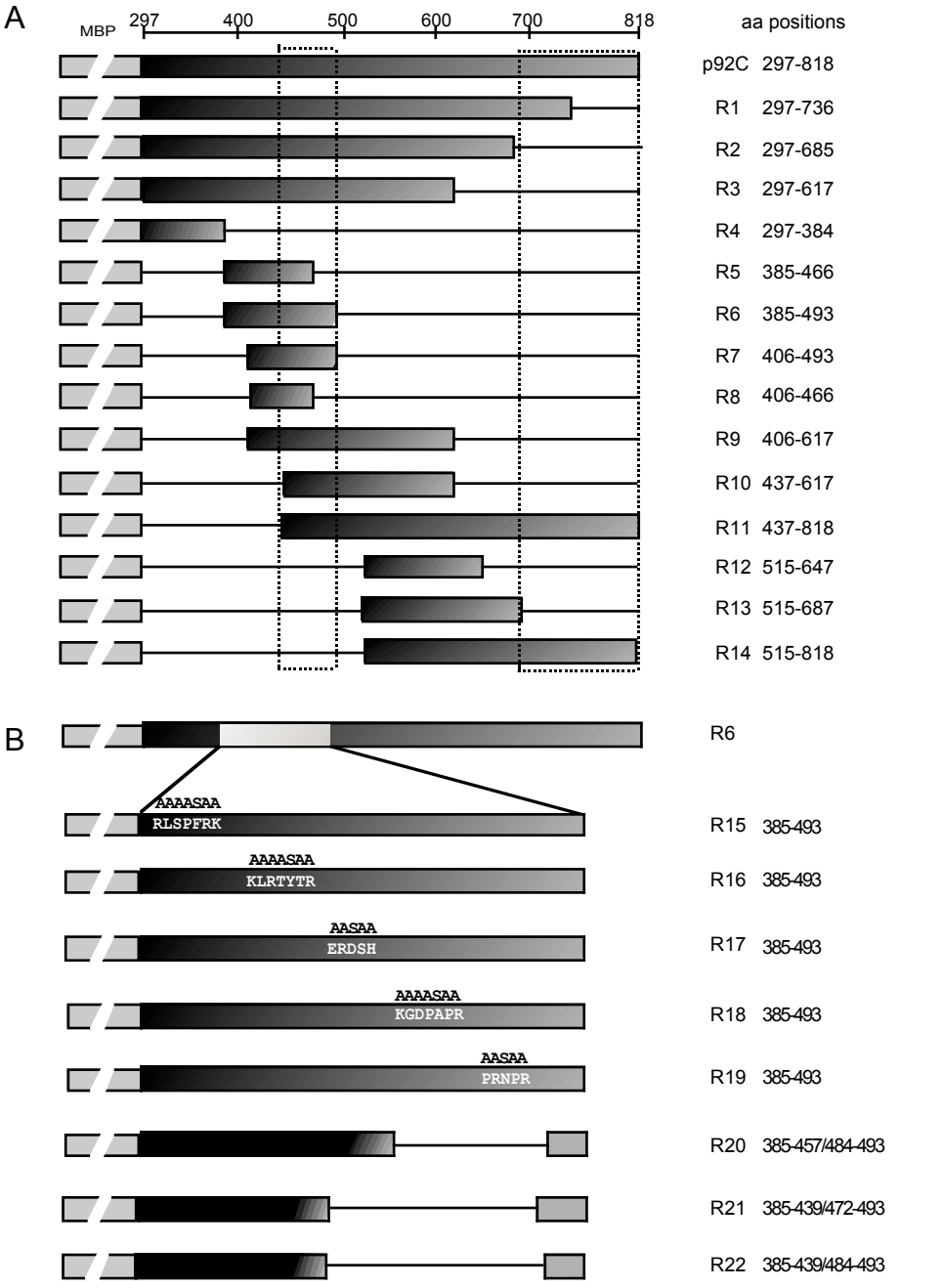
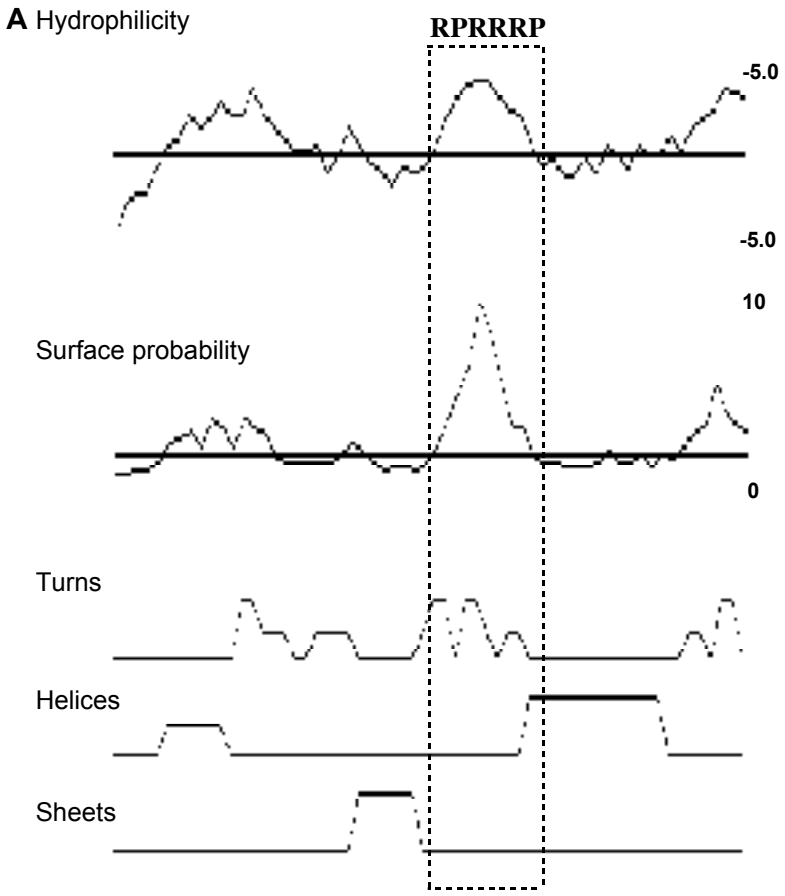

A
MBP
297
400
500
600
700
818
aa positions
p92C 297-818
R1 297-736
R2 297-685
R3 297-617
R4 297-384
R5 385-466
R6 385-493
R7 406-493
R8 406-466
R9 406-617
R10 437-617
R11 437-818
R12 515-647
R13 515-687
R14 515-818
B
R6
AAAASAA
RLSPFRK
R15 385-493
AAAASAA
KLRTYTR
R16 385-493
AASAA
ERDSH
R17 385-493
AAAASAA
KGDPAPR
R18 385-493
AASAA
PRNPR
R19 385-493
R20 385-457/484-493
R21 385-439/472-493
R22 385-439/484-493
C
Probe alone
p92
p92C
R-series
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Probe alone
R-series
15 16 17 18 19 20 21 22

Figure 3.8. Mapping the RNA binding domains within the unique portion of the p92 protein, termed p92C. (A) A schematic representation of the deletion derivatives of p92C. The names of the constructs and the positions of the amino acid present in the truncated proteins are shown on the right. These truncated p92C proteins were expressed in *E. coli* as fusions to MBP (indicated schematically by a dotted box). The shaded boxes indicate the portions of the p92C protein that were present in given expression constructs. The horizontal lines represent the deletions. (B) A schematic representation of clustered alanine/serine scanning and deletion mutagenesis of a segment in p92C to map the RNA binding site. The alanine/serine scanning mutations were targeted at five different groups of basic amino acid clusters as shown. Expression constructs R20 to R22 were made by deleting two or more of the basic amino acid clusters as indicated by the straight lines. (C) RNA binding activities of the truncated recombinant p92C proteins were analyzed in a standard gel mobility shift assay (see Figure 3.1B for details). Equimolar concentrations of proteins (~2 μM) were used in the binding assay in the presence of a ^{32}P-labeled RNA probe [region III (-), see Figure 3.1B]. The samples containing particular recombinant proteins are indicated above the lanes using the same numbering as in panels A and B above.



B

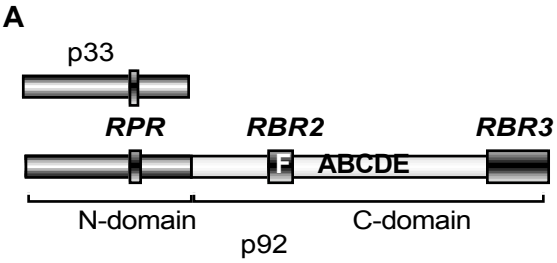
STG----RPRRRPYA	TBSV
STG----RPRRRPYA	CNV
STG----RPRRRPYA	AMCV
RAG-----ARRRPYA	CIRV
STG----RPRTRPYA	CymRSV
STG----RPRTRPYA	CRSV
STADV---PRRVRQK	TCV
KVK-----PRRNRYA	MNeSV

C

STG---RPRRRPYA	TBSV p33
SYGRKKRPRRRPPQ	HIV1 Tat

Figure 3.9. Primary and secondary structure analysis of the RNA binding region in p33. (A) SEQWeb v1.1 from GCG Wisconsin package was used to predict the secondary structure, surface probability and hydrophilicity of the RNA binding region in p33. The arginine/proline-rich motif involved in RNA binding is boxed. (B) Sequence alignment of the arginine/proline-rich motifs present in the p33-like replicase proteins of different tombusviruses and the related

proteins in carmoviruses. The conserved motif is boxed. The following abbreviations were used: CNV, *Cucumber necrosis virus*; AMCV, *Artichoke mottled crinkle virus*; CIRV, *Carnation Italian ringspot virus*; CymRSV, *Cymbidium ringspot virus*; CRSV, *Carnation ringspot virus* (a dianthovirus); TCV, *Turnip crinkle virus* (a carmovirus); MNeSV, *Maize necrotic streak virus*; and HIV-1, *Human immunodeficiency virus-1*. (C) Alignment of the arginine/proline-rich motif in TBSV p33 and arginine-rich motif in HIV-1 Tat, which is involved in RNA binding (12, 23). The residues important for RNA binding in the Tat protein are underlined, while the arginine-rich sequence is boxed.



B

```
TBSV    RD SHLTTFVKAE KIS-TSKGDP APRVIQPRNP RYNVELGRYL RHMES
CNV     RD SHLTTFVKAE KIS-TAKGDP APRVIQPRNP RYNVELGRYL RHMES
AMCV    RD SHLTTFVKAE KIS-TSKGDP APRVIQPRNP RYNVELGRYL RHMES
CymRSV  RD SHLTTFVKAE KIS-TSKGDP APRVIQPRNP RYNVELGRYL RHMES
CIRV    RD SHLTTFVKAE KIS-TSKGDP APRVIQPRNP RYNVELGRYL RHMES
TCV     KD AELKTFVKAE KINFTAKKDP APRVIQPRDP RYNIEVGKYL KPYEH
MNeSV   RD SHLTTFVKAE KIS-TSKGDP APRVIQPRSP RYNVELGRYL RHLES
PoLV    RD AYLKTFVKAE KINITLKPDP APRVIQPRDP RINVEVGRYL KPLEP
CRSV    RD SHLTTFVKAE KIS-TSKGDP APRVIQPRNP RYNVELGRYL RHMES
HCV     ID TTIMAKNEVF CVQPEKGGRK PARLIVFPDL GVRVCEKMAL YDVVS

Cons    .D ..l.tfvkae k!..t.kgdp apRvIqprdp r.nVe.gryL ...es
```

Figure 3.10. Locations of the three RNA-binding regions in TBSV p92. (A) A schematic representation of the three mapped RNA-binding regions, RPR, RBR2 and RBR3, in p92. The N-terminal portion of p92, which overlaps with p33, is indicated as a dark box, while the unique C-terminal segment is shown as a light box. The signature motifs of RdRps are marked with letters A-F (94). The names of the RNA-binding regions are shown above the filled boxes. (B) Sequence alignment of the RBR2 region in TBSV and in other replicase proteins of tombusviruses and related viruses as well as the RNA-binding region in NS5B of HCV. The RBR2 region in TBSV is indicated by a solid line above the sequence, while the RNA-binding sequence in NS5B of HCV (16) is marked with a dotted line underneath of the sequence. The conserved motif F is boxed. The consensus (cons) sequence is also shown at the bottom. The abbreviations used for viruses are the same as in the legend to Figure 3.9, except that I also included *Pothos latent virus* (PoLV, Tombusviridae) and HCV.

CHAPTER IV

INTERACTION BETWEEN THE REPLICASE PROTEINS OF *TOMATO BUSHY STUNT* VIRUS IN VITRO AND IN VIVO

INTRODUCTION

Viral RNA replication requires the assembly of replicase complexes in infected cells. These replicase complexes, which are found within characteristic membrane-containing structures such as multivesicular bodies, are known to contain the viral RNA template(s) and viral- and host-coded proteins (17, 93). These factors are likely held together by protein-protein and protein-RNA interactions within the replicase complex. For example, the 1a and 2a replicase proteins of *Brome mosaic virus* (BMV), which are present within the BMV replicase, have been demonstrated to interact with each other using co-immunoprecipitation and the yeast two-hybrid assays (59, 95, 96). In addition, 1a is also capable of intermolecular interaction with other 1a proteins, which might be important to bring two or more 1a proteins into complex with the 2a protein (97). The relevance of these intermolecular interactions was confirmed in replication studies in plant protoplasts using selected BMV mutants (97). Similarly, the 126K protein of *Tobacco mosaic virus* (TMV) was shown to form intermolecular interaction with other 126K proteins and with the 186K RNA-dependent RNA polymerase (RdRp) protein (42, 151). Using temperature sensitive 126K mutants in the yeast two-hybrid assay and for replication studies in protoplasts, Goregaoker *et al*. (42) verified that the 126K-186K interaction is essential for TMV replication. Interestingly, analysis of purified recombinant 126K-containing complexes by electron microscopy revealed the formation of ring-like structures (possibly hexamers), though the function of this structure is currently unknown (41). Interaction between various replicase proteins has also been demonstrated for several other plus-stranded RNA viruses, including poliovirus (2, 53, 78, 121), hepatitis C virus (146), cucumoviruses (144) and potyviruses (132).

Tombusviruses, including *Tomato bushy stunt virus* (TBSV) and *Cucumber necrosis virus* (CNV), are positive-stranded RNA viruses of plants. The ~4.8 kb genomic (g)RNA codes for 5 proteins, of which two, namely p33 and p92, are essential for replication (130). The

signature motifs of RdRp are present in p92(94), which is produced via ribosomal readthrough of the p33 stop codon. It has been estimated that the amount of p92 in infected cells is about 20-fold lower than p33 (133). Both p33 and p92 proteins are present in the highly active partially-purified RdRp preparation obtained from TBSV and CNV-infected plants (J. Pogany and P. D. Nagy, unpublished), suggesting that they might play direct roles in viral RNA synthesis. In vitro studies with the above CNV RdRp preparation demonstrated that the CNV RdRp could recognize essential promoter sequences (107), replication enhancers (104, 106) and a replication silencer (115) element during RNA synthesis. The significance of these cis-acting RNA elements in Tombusviruses has been confirmed using plant protoplasts (*Nicotiana benthamiana* and cucumber) (36, 106, 115, 125) and yeast, a model host (105).

The function of p33 is not yet known, although it is likely involved in RNA replication, because (i) p92 could not support viral replication in the absence of p33 in plant protoplasts (103, 109), and (ii) p33 has been demonstrated to bind to the viral RNA in vitro (109, 122). The RNA binding region in p33, and the corresponding overlapping (pre-readthrough) region in p92, has been mapped to an arginine-proline rich motif (RPR motif, Figure 4.1), which is similar to the RNA-binding sequence of Tat protein in the *Human immunodeficiency virus-1* (122). Mutations within the RPR motif in p33 affected gRNA and subgenomic (sg)RNA synthesis (109), and RNA recombination (110), suggesting that p33 is a multifunctional protein. The RPR motif is also important for the function of p92 (109). In addition, the TBSV and CNV replicase proteins likely bind to intracellular membranes, based on studies with the closely related replicase proteins of *Carnation Italian ringspot virus* and *Cymbidium ringspot virus*, which have been shown to localize to multivesicular bodies (membranous structures derived from peroxisomes or mitochondria) (21, 129).

Co-purification of p33 and p92 proteins in the RdRp preparation supported a model that direct interactions between these proteins might contribute to the stability of the RdRp complex. To test this hypothesis, I examined if p33 and p92 could interact directly with each other using surface plasmon resonance (SPR) analyses with purified recombinant proteins and the yeast two-hybrid system. I demonstrated that p33 could interact with p92, and with other p33 molecules (intermolecular interaction). I identified two short regions within the C-terminal portion of p33, which promoted the aforementioned interactions. The significance of p33:p33 and p33:p92 interactions was confirmed in a model tombusvirus replication system in yeast by

expressing p33 and p92 proteins carrying site-specific mutations within the region needed for protein interaction. Overall, these experiments supply direct evidence that the TBSV replicase proteins interact with each other, and that this interaction is important for the functions of the replicase proteins during tombusvirus infection.

Materials and Methods

Biosensor Assay

The recombinant proteins used for the biosensor assays were expressed as fusions to maltose binding protein (MBP) in *E. coli* and purified as described by Rajendran *et al.* (2003). The surface plasmon resonance (SPR)-based biosensor assays were carried out using BIACORE X (Biacore, Inc., Piscataway, N.J.) at 25 ^{0}C. The running buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% surfactant P-20) and the protein immobilization buffer (10 mM sodium acetate, pH 5.0) were filtered and degassed before use. The sensor chip CM-5 and the amine coupling kit were purchased from Biacore. The purified recombinant MBP-p33 (see above) was dialyzed in the immobilization buffer to achieve a net negative charge to facilitate the coupling reaction and then immobilized onto CM-5 sensor chip by using the amine-coupling chemistry as recommended by the manufacturer. Briefly, the surface of both flow cells on the sensor chip was activated with a solution containing N-hydroxysuccinimide and N-ethyl-N'-(dimethylaminopropyl)-carbodimide for 7 min at the rate of 5μl/min. The purified MBP-p33 (100 μg/ml) was injected over the chip on flow cell #1 (Fc1) to immobilize the protein up to 12500 RU, which corresponds to ~12.5 ng/mm^2, whereas flow cell #2 (Fc2) was left as the control surface to account for non-specific binding and bulk refractive index changes upon injection of protein samples, as suggested by the manufacturer. Both flow cells were then deactivated using 1.0 M ethanolamine-HCl, pH 8.5, for 7 min at 5 μl/min and washed by repeated injections with the running buffer.

The protein-protein binding reactions were performed by injecting the purified test proteins (1μM) that had been dialyzed in the running buffer, at the flow rate of 20 μl/min for 3 min followed by 3 min injection of the running buffer at the same flow rate. The interactions between the immobilized p33 and the tested recombinant proteins were analyzed in real time through a sensogram (Figure 4.2), in which the resonance units (RU) were plotted as a function

of time. One RU is equivalent to a change in adsorbed mass of 1pg/mm^2 of the sensor surface (BIA Applications Handbook, Biacore, NJ). All the data shown in Figure 4.2 were corrected based on data obtained from the control (Fc2).

Yeast two-hybrid assay

Yeast two-hybrid experiments were performed using the Matchmaker 3 two-hybrid system (BD Biosciences Clontech) based on the Yeast Protocol Handbook (BD Biosciences Clontech). To clone p33 and p92 genes into the yeast two-hybrid vectors pGADT7 and pGBKT7 (BD Biosciences-Clontech), I PCR-amplified them using the full-length T100 (49) and pHS175 clones (133), which had the stop codon at the end of the p33 gene mutated to tyrosine to facilitate the production of p92 protein. The PCR products were cloned into plasmid vectors using *Eco*RI and *Bam*HI restriction sites. Similar cloning strategy was used to make a series of deletion mutants in p33 and p92 with the help of primers listed in Table 4.1. The sequences of the primers are described in Table 4.2. PCR-based mutagenesis was used to create alanine substitution mutations as described by Rajendran and Nagy (2003). To make internal deletions, sequences on both sides of deletions were PCR amplified separately with primers containing unique restriction sites (I selected the sites in such a way to keep the original amino acid sequence), followed by treatment with the appropriate enzymes and ligation. The ligated products were then re-amplified by PCR using the end primers containing *Eco*RI and *Bam*HI restriction sites. Finally, I cloned the PCR products into the yeast two-hybrid vectors as described above.

Yeast strain AH109 (BD Biosciences Clontech) was co-transformed with 500 ng of individual plasmids using polyethylene glycol-lithium acetate (PEG/LiAC) method. The transformants were cultured on synthetic drop-out (SD) medium containing 2% (v/v) glucose and all amino acids except leucine and tryptophan (SD-Leu-Trp) at 30 ^{0}C for 3-4 days as described in the Matchmaker 3 manual. The positive and negative control transformations were performed with the plasmids supplied by the manufacturer.

Single colonies from each transformant were grown in 2 ml of SD medium containing 2% (v/v) glucose without leucine and tryptophan at 30 ^{0}C for 24 hrs (three repeats per each experiment). The culture densities were normalized with SD-Leu-Trp, followed by three times 10-fold serial dilution. 10 μl of the diluted cultures were grown on SD-Leu-Trp, SD-Leu-Trp-

His and SD-Leu-Trp-His-Ade (leucine, tryptophan, histidine and adenine) at 30 ^{0}C for 3-4 days.

Replication assay for TBSV DI RNA in Yeast

The yeast system for replication of DI RNA described by Panavas and Nagy (2003) (105) was adopted to study the role of interactions between replicase proteins (identified in the yeast two-hybrid assay) in DI RNA replication. Both p33 and p92 proteins were mutated separately at amino acid positions 244, 246, 274 and 276 by substituting alanine for tyrosine, arginine, phenylalanine and tyrosine respectively using the Quick Change mutagenesis kit (Stratagene). The yeast plasmids expressing wild type and/or mutant replicase proteins (p33 and p92) were co-transformed with pYC-DI72 which codes for DI-RNA into yeast strain SC1 using the PEG/LiAc method. The transformants were grown at 30 ^{0}C in SD medium containing 2%(v/v) galactose without leucine, tryptophan and uracil (SD-Leu-Trp-Ura). Individual colonies for each transformant were cultured in SD-Leu-Trp-Ura broth containing galactose at 24 ^{0}C, 30 ^{0}C and 34 ^{0}C for 36, 24, and 20 hrs, respectively. The cells were then collected by centrifugation at 500 g for 2 min and re-suspended in the RNA extraction buffer (50 mM sodium acetate, 10 mM EDTA and 1% SDS) and mixed thoroughly by vortexing in the presence of 50% water-saturated phenol. The mixture was then heated to 65 ^{0}C for 4 min, placed on ice for 2 min and centrifuged at 21,000 g for 4 min at room temperature. The aqueous phase was then mixed with 2 volumes of 95% ethanol and 0.1 volume of 3M sodium acetate pH 5.2. After 30 min at –20 ^{0}C, the RNA was pelleted by centrifugation at 21,000 g and washed with 70% ethanol for 5 min at 21,000 g at 4 ^{0}C. The RNA pellet was air-dried and resuspended in sterile water. The RNA was then electrophoresed in 1.5% agarose gel and transferred onto Hybond-XL membrane (Amersham) using a semi-dry transblot apparatus (BioRad) (109). The membrane was then treated with 2X SSC for 3 min, air-dried and UV-cross-linked using a crosslinker (Stratagene). The membrane was then probed with DI RNA-specific riboprobes, exposed to a phosphorscreen and analyzed using a PhosphorImager (Molecular Dynamics, Inc.) as described by Panaviene *et al*. (2003) (109).

Preparation of crude protein extracts from yeast, SDS-polyacrylamide gel electrophoresis, western blotting and immunodetection

Crude protein extracts from yeast were obtained as described by Kushnirov (2000) (67).

Briefly, transformed SC1 yeast strains were grown in 2 ml of SD-Leu-Trp-Ura broth containing 2% galactose at 30 ^{0}C for 24 hrs. The cells were then harvested and resuspended in 100 µl sterile water. The cells were treated with 100 µl 0.2M NaOH at room temperature for 5 min, followed by a quick spin to remove the aqueous phase containing the alkali. The pellet was then resuspended in 50 µl 1X SDS-PAGE loading buffer and the samples were boiled for 5 min. The cell debris was cleared by a quick spin and 6 µl of the supernatant was run in 8% SDS-PAGE gel. The protein samples were then transferred from the gel onto PVDF membrane (Biorad). Immunodetection of p33 protein in the above samples was performed by using Anti-His monoclonal antibodies (Amersham) against the 6xHis-tag present at the N-terminus of the p33 protein (Panaviene *et al*., submitted).

RESULTS AND DISCUSSION

Interaction between the recombinant p33 and p92 replicase proteins in vitro

The TBSV and CNV p33 and p92 replicase proteins are present in the partially-purified RdRp complex, which is capable of complementary RNA synthesis on added plus- or minus-stranded TBSV RNA templates (86). The RdRp complex might be held together by protein-RNA interactions and/or protein-protein interactions between p33 and p92. To test if interaction between p33 and p92 (p33:p92) or between p33 and p33 (p33:p33) occurs, I used an in vitro protein interaction assay based on surface plasmon resonance (SPR) measurements. For this assay, I expressed recombinant p33 and p92 proteins in *E. coli* as fusion proteins to the maltose binding protein (MBP) as described by Rajendran and Nagy (2003). The p92 carried a tyrosine codon in place of the p33 stop codon to ensure the expression of the full-length p92 protein in *E. coli* (Figure 4.1A). This tyrosine mutation did not interfere with the wild-type (wt)-like function of p92 in a two-component complementation-based system, in which p33 and p92 are expressed from separate TBSV RNAs (103, 109). After affinity purification of the recombinant p33 and p92 (Figure 4.1B), (122), I carried out SPR measurements with a Biacore biosensor as described in the Materials and Methods section. Briefly, the SPR assay provides real-time protein-protein interaction data (29), by measuring the change in refractive index at the surface of the sensor chip due to change in mass resulting from protein-protein interaction between the immobilized protein and the protein that is being passed over the sensor chip. For this study, I fixed the

purified full-length MBP-p33 fusion protein onto the surface of a CM5 chip using amine-coupling reaction (Biacore). Then, comparable amounts of purified recombinant p33, p92, and MBP proteins diluted with the running buffer were passed separately over the immobilized MBP-p33 (Figure 4.2). As predicted, I obtained strong signals (between 2000 – 8000 RU, resonance unit) when p33 and p92 were tested (Figure 4.2), demonstrating that both p33 and p92 interacted strongly with the immobilized MBP-p33. In contrast, the "MBP only" sample resulted in a signal that was only slightly above the background (Figure 4.2), indicating that MBP interacted weakly with the immobilized MBP-p33 protein. It is also possible that the weak signal is due to non-specific binding of MBP to the dextran layer present on the surface of the sensor chip.

It is important to emphasize that the SPR signal (RU value) depends not only on the amount (the absolute number) of interacting molecules, but also on the molecular weight (MW) of the particular protein. Thus, the higher RU value for p33:p92 interaction than for p33:p33 interaction does not necessarily mean that the interaction between p33 and p92 is "stronger" than between p33 and p33, because p92 has significantly higher MW than p33. In addition to the above observations, the SPR measurements also indicate that the binding between p33:p92 or between p33:p33 is stable (i.e., the dissociation was slow). Overall, based on the SPR measurements, I propose that the recombinant p33 can interact stably with other p33 molecules as well as with p92 in vitro.

To test what region in p33 is responsible for protein binding, I tested the interaction between p33 and two truncated p33 proteins, termed p33N and p33C (Figure 4.1A). While p33N contains the N-terminal half of p33 with the putative membrane-spanning domain plus additional N-terminal sequences, p33C includes the C-terminal half of p33 with the RNA-binding domain (Figure 4.1A). Note that p33N and p33C do not overlap with each other and they have similar MW values. The SPR analysis with purified recombinant proteins demonstrated that p33C could efficiently interact with p33, while the interaction of p33N with p33 was weak (see comparable RU values for p33N and MBP-p33) (Figure 4.2). Interestingly, the interaction between p33 and p33C was rather strong (e.g., higher RU value in spite of the lower MW for p33C than for the full-length p33) when compared to p33:p33 interaction in spite of the comparable molar amounts of proteins used in these assays. This suggests that the C-terminal region in p33 promotes interaction with p33, while the N-terminal hydrophobic domain

in the full-length p33 may inhibit p33:p33 interaction. The opposing effects of the N- and C-terminal domains on p33:p33 interaction might serve as a regulatory function in vivo, which might affect the strength (stability) of the p33:p33 interaction. For example, the inhibitory effect of the N-terminal region of p33 might be negated when it is buried in membranous structures (due to the presence of the membrane-spanning domain, Figure 4.1A). Thus, the strength of p33:p33 interaction in cells might depend on the proximity of suitable membranes interacting with the N-terminal region of p33. Alternatively, it is also possible that the inhibitory effect of the N-terminal domain in p33 is due to the conditions in the in vitro assay, which allow the exposure of the hydrophobic sequences within the N-terminal domain due to the lack of membranes. Overall, these in vitro SPR experiments confirmed that p33 could interact with the C-terminal half of p33 and possibly with the corresponding p33C sequence in p92.

The C-terminal sequence in p33 is essential for p33:p33 and p33:p92 interactions in the yeast two-hybrid assay

To validate the above in vitro SPR-based observations on the interactions between p33:p33 and p33:p92, I also tested the interaction between the TBSV replicase proteins with the yeast two-hybrid assay (Matchmaker 3, BD Biosciences Clontech). This assay is based on the modular nature of the GAL4 transcription factor, where the DNA-binding domain (DB) and the transcription activation domain (AD) are fused separately to the test proteins. If there is interaction between the test proteins, then the BD and AD domains will be able to activate transcription allowing the yeast strain to grow on selective media. Using the above yeast two-hybrid assay, I found that yeast expressing p33 or p33N in any pair-wise combinations such as p33:p33, p33:p33N, p33:p33C, p33N:p33N, did not grow on medium or high stringency media (Figure 4.3). In contrast, I did observe strong positive interaction when p33C was used in combination with p33C (Figure 4.3). Thus, there is some discrepancy between the SPR measurements and the yeast two-hybrid data regarding the full-length p33. This might be due to the presence of the N-proximal hydrophobic membrane-spanning domains in p33 (Figure 4.1), which might interfere with the nuclear localization of the fusion proteins that is required for detection of positive interactions in the yeast two-hybrid assay. Accordingly, protein interaction studies with full-length membrane-associated proteins have often led to negative results in the yeast two-hybrid assay (143, 148). On the other hand, p33C lacking the membrane-spanning

domain interacted with p33C in the yeast two-hybrid assay (Figure 4.3), as predicted from the in vitro SPR analysis. Overall, the above experiments with the yeast two-hybrid assay confirmed that p33C could stably interact with p33C in vivo.

To test if p33C can also interact with the p92 protein in vivo, I used two truncated p92-derivatives in the yeast two-hybrid assay. The longer of the two proteins, termed p33C/92C (Figure 4.4), contained the p92 sequence except the N-terminal membrane-spanning domain (see explanation above), whereas the shorter protein, termed p92C, contained only the unique readthrough portion of p92 (Figure 4.4). The yeast two-hybrid test revealed that p33C interacted with p33C/92C, while it did not interact with p92C (Figure 4.4). Thus, the obtained data suggests that the interaction between p33 and p92 includes similar sequences with those responsible for intermolecular interaction between p33:p33. In contrast, the unique readthrough portion of the p92 (i.e., p92C, Figure 4.1), which contains the RdRp motifs, is unlikely to interact with p33 in vivo.

In general, the p33:p33 and p33:p92 interaction domains in TBSV p33 and p92 show similar arrangement to those found for the distantly-related 126K/183K of TMV replicase proteins, which are also expressed via ribosomal readthrough mechanism (18). For example, the C-terminal helicase-like domain of 126K protein interacted with the corresponding regions in 126K (intermolecular interaction) and in 183K, while the readthrough portion of the 183K protein carrying the RdRp motifs did not participate in this interaction (42). Additional similarity between the TBSV and TMV replicase proteins is the presence of an RNA-binding sequence in proximity to the protein interaction domain (99). In spite of the similar arrangement of the above domains, the replicase proteins of TMV and TBSV likely perform different functions as well. For example, 126K might function as an RNA helicase and a capping enzyme (41, 81), whereas p33 lacks these motifs.

Taken together these results suggest that the C-terminal region in p33 and the corresponding overlapping (pre-readthrough) sequence in p92 carry domains, which are likely responsible for assembly and/or stability of the replicase complex, while the major function of the readthrough portion of p92 might be complementary RNA synthesis.

Defining short amino acid stretches essential for p33:p33 interaction in yeast

To further delineate the sequences essential for the observed p33:p33 and p33:p92 interactions, I made a deletion library derived from p33C and tested it for interaction with p33C in the yeast two-hybrid assay (Figure 4.5). These experiments revealed that deletion of 26 or 46 amino acids (aa) from the C-terminus of p33C (in combination with deletion of 13 aa from the N-terminus) did not interfere with interaction with p33C (see constructs $p33_{181\text{-}270}$ and $p33_{181\text{-}250}$, Figure 4.5). Additional deletions of 10 or 36 aa from the C-terminus (constructs $p33_{181\text{-}240}$ and $p33_{181\text{-}214}$ Figure 4.5) resulted in weak interaction with p33C on the medium stringency media, while no interaction was seen on the high stringency media. Thus, these experiments indicated that there is a 10 aa stretch in p33, termed site 1 (located between position 240 and 250 in p33) that is essential for p33:p33 interaction.

To test if there is additional sequence in p33, which might affect p33:p33 interaction, I made a further series of combined deletions from both the N- and the C-terminus of p33C, as shown in Figure 4.5. When tested in combination with p33C, I found that constructs that included a 13 amino acid stretch, termed site 2, between positions 271 and 283 (for example, constructs $p33_{271\text{-}290}$ and $p33_{263\text{-}283}$ Figure 4.5) interacted with p33C, based on yeast growth on both medium and high stringency media. Overall, these experiments defined site I and II that were essential for interaction with p33C.

To confirm that the above two regions (i.e., site 1 and site 2) can independently interact with p33, I expressed $p33_{217\text{-}250}$ (which includes site 1) and $p33_{251\text{-}296}$ (which includes site 2) in *E. coli* as MBP fusion proteins, followed by affinity purification as described in Materials and Methods (122, 123). Comparable amounts of the purified $p33_{217\text{-}250}$ and $p33_{251\text{-}296}$ proteins were then used in an SPR analysis when p33 was fixed on the surface of the chip (see above). These experiments demonstrated that both truncated p33 proteins interacted efficiently with p33 in vitro (Figure 4.5B). Therefore, I conclude that site 1 and site 2 within p33 can facilitate the interaction with p33 and possibly with p92 in vitro.

Albeit site 1 and site 2 in p33 were able to promote interaction with the C-terminal region of p33 independently of each other in vitro and in the yeast two-hybrid assay, it is possible that they are part of a larger p33:p33 and p33:p92 interaction domain. A more detailed structure/function analysis of p33 will be necessary to identify the entire interaction domain and the critical amino acids within.

An interesting observation from this and earlier works (122) is that the protein interaction and the RNA-binding regions are both located in close vicinity within the C-terminal half of p33, and the corresponding sequence in p92. Thus, it is possible that the RNA-binding and p33:p33 interaction processes might affect each other. Accordingly, I observed cooperative RNA-binding by p33 in vitro (122), which suggests that p33:p33 interaction might stabilize binding of p33 to the viral RNA. The functional significance of cooperative binding by p33 is currently not known.

Mutations within interaction site 1 and site 2 of p33 and p92 inhibit replication of the TBSV replicon in yeast

To test if the above defined interaction site 1 and site 2 are important for replication, I took advantage of the recently developed TBSV DI RNA replicon system in yeast (105). The major advantage of this heterologous replication system for this work is that it allows the separate expression of the CNV (very closely related to TBSV, Figure 4.6) p33 and p92 proteins from plasmids together with a TBSV-derived DI RNA replicon (termed yDI-72 RNA), which is expressed from the galactose-inducible GAL1 promoter (105). The wt CNV p33 and p92 can efficiently support replication of DI-72 RNA even after the suppression of yDI-72 RNA transcription by glucose (105). Using the above yeast-based tombusvirus replicon system, I separately mutagenized p33 and p92 proteins within site 1 and site 2 and tested the abilities of the resulting mutated proteins to support DI-72 RNA replication in vivo. Tyrosine (Y_{244}) and arginine (R_{246}) within site 1, and phenylalanine (F_{274}) and tyrosine (Y_{276}) within site 2 (Figure 4.6) were mutated to alanines separately in both p33 and p92. These residues were mutated because they are highly conserved among Tombusviruses (Figure 4.6), and that protein-protein interfaces often comprise of tyrosine and arginine residues surrounded by hydrophobic residues (40, 79). When tested in the two-hybrid assay, Y_{244}-A and R_{246}-A mutations within site 1 decreased the number of yeast colonies on high stringency media, suggesting that these mutations affected p33:p33 interaction (Figure 4.7B). Importantly, Y_{244}-A and R_{246}-A mutations present in either p33 or p92 supported very low level of DI-72 RNA replication in yeast at 30 ^{0}C (Figure 4.7C) or 34 ^{0}C (not shown). I conclude that these amino acids are important for the function of both p33 and p92 proteins. Because the mutated p33 proteins were present in comparable amount with the wt p33 in yeast (Figure 4.7, Western), I suggest that the Y_{244}-A and

R_{246}-A mutations decreased the strength (stability) of interactions between the replicase proteins, which might have led to the assembly of less stable or less active replicase complexes in yeast. Similar to the site 1 mutants, the site 2 mutants, F_{274}-A and Y_{276}-A, also affected p33:p33 interaction and replication of DI-72 RNA in yeast. However, the observed effects were less pronounced for site 2 mutants than for the site 1 mutants on p33:p33 interaction in the yeast two-hybrid assay (Figure 4.7) and on DI-72 RNA replication in yeast expressing either mutated p33 or p92 (Figure 4.7). Overall, these results support the model that p33:p33 and/or p33:p92 interactions are essential for Tombusvirus replication.

In summary, for the first time I demonstrated direct interaction between the replicase proteins of a Tombusvirus using an in vitro SPR method with purified recombinant proteins and in vivo using the yeast two-hybrid assay. The identified replicase protein interaction domain seems to be important for replication of a model Tombusvirus template, suggesting that this interaction is essential for the assembly and/or function of the Tombusvirus replicase.

Table 4.1 List of primer pairs used to make constructs described in text

Constructs[a]	Primers
Surface Plasmon Resonance Assay	
p33	3/4
p92	3/5
p92C	6/11
p33C	48/10
$p33_{219-250}$	383/91
$p33_{251-296}$	1188/10
Yeast Two Hybrid Assay	
p92	473/475
p92C	187/475
p33C/92C	48/475
p33	473/474
$p33_{1-160}$ (p33N)	473/620
$p33_{168-296}$ (p33C)	183/474
$p33_{181-214}$	75/1131
$p33_{181-240}$	75/1132
$p33_{181-250}$	75/1187
$p33_{181-270}$	75/1143
$p33_{181-296}$	75/474
$p33_{211-240}$	76/1131
$p33_{211-250}$	76/1187
$p33_{211-296}$	76/474
$p33_{219-296}$	383/474
$p33_{241-257}$	1141/1142
$p33_{241-270}$	1141/1143
$p33_{241-296}$	1141/474
$p33_{251-296}$	1188/474
$p33_{258-296}$	1189/474
$p33_{263-283}$	77/1202
$p33_{271-290}$	1200/1201
$p33_{271-296}$	1200/474
$p33_{241-270Y244A}$	1206/1143
$p33_{241-270R246A}$	1207/1143
$p33_{271-290F274A}$	1203/1201
$p33_{271-290Y276A}$	1204/1201
DI-RNA replication Assay in Yeast	
p33Y244A	1212/1213
p33R246A	1214/1215
p33F274A	1218/1219
p33Y276A	1220/1221
p92Y244A	1212/1213
p92R246A	1214/1215
p92F274A	1218/1219
p92Y276A	1220/1221

[a] Subscript numbers refer to the amino acids and/or substitutions the constructs carry.

Table 4.2. List of primers and their sequences used for PCR to generate constructs described in the text

Primer #	Nucleotide positions[a]	Sequence[c]
3	169–186	GAGACCATCAAGAGAATG
4	1056–1039	CT(A/T)TTTGACACCCAGGGA
5	2622–2605	TC(A/C)AGCTACGGCGGAGTCGG
6	1057–1075	GGAGGCCTAGTACGTCTAC
10	1056–1039	GCAGTCTAGACTATTTGACACCCAGGGA
11	2622–2605	GCAGTCTAGATCAAGCTACGGCGGAGTC
48	616-632	GAGGAATTCTACGCTACCCTACCTAG
75	706-723	GAGGAATTCTGTCTGGTGGTTGAGCCG
76	796-813	GAGGAATTCACTGGGCGCCCTCGTCGA
77	952-969	GAGGAATTCGATGTCATATTGCCTTTG
183	667-678	GAGGAATTCGCACGAGCACACATGGAG
187	1057-1075	GAGGAATTCGGAGGCCTAGTACGTCTAC
383	820-837	GAGGAATTCTATGCGGCAAAGATCGCA
473	169-186	GAGGAATTCGAGACCATCAAGAGAATG
474	1036-1053	GAGGGATCCCTATTTGACACCCAGGGACTC
475	2602-2619	GAGGGATCCTCAAGCTACGGCGGAGTCGAG
620	628-645	GAGGGATCCCTACGACAGTTTTTCCCTAGG
1131	868-885	GAGGGATCCCTAATTCTCTGGACTGTTCTT
1132	790-807	GAGGGATCCCTAAGGGCGCCCAGTGGACGC
1141	886-906	GCGGAATTCAGACTAATCTACCAGAGGGTG
1142	916-936	GAGGGATCCCTAGACGCAGTCTTTGTCCATGAT
1143	955-975	GAGGGATCCCTAAATAGCCAAAGGCAATATGAC
1187	898-915	GAGGGATCCCTACTCGATCATCACCCTCTG
1188	916-933	GCGGAATTCATCATGGACAAAGACTGC
1189	934-951	GCGGAATTCGTCAGGTATGTTGACAGG
1200	976-993	GCGGAATTCGGATGCTGTTTTGTCTAT
1201	1018-1035	GAGGGATCCCTACTGTGAGCCCCATAGTGC
1202	991-1018	GAGGGATCCCTACTCCACTCCATCCGGATA
1203	976-1002	GCGGAATTCGGATGCTGTGCTGTCTATCCGGATGGA
1204	976-1008	GCGGAATTCGGATGCTGTTTTGTCGCTCCGGATGGAGTGGAG
1206	886-912	GCGGAATTCAGACTAATCGCCCAGAGGGTGATGATC
1207	886-918	GCGGAATTCAGACTAATCTACCAGGCGGTGATGATCGAGATC
1212	866-898[b]	GAGAATAGATTGATAGCCCAGAGAGTGATGATC
1213	866-898[b]	GATCATCACTCTCTGGGCTATCAATCTATTCTC
1214	872-904[b]	AGATTGATATACCAGGCAGTGATGATCGAGATC
1215	872-904[b]	GATCTCGATCATCACTGCCTGGTATATCAATCT
1218	956-988[b]	GCTATTGGATGCTGCGCTGTCTACCCGGATGGA
1219	956-988[b]	TCCATCCGGGTAGACAGCGCAGCATCCAATAGC
1220	962-994[b]	GGATGCTGCTTTGTCGCCCCGGATGGAGTGGAG
1221	962-994[b]	CTCCACTCCATCCGGGGCGACAAAGCAGCATCC

a Nucleotide positions relative to complete sequence of TBSV genome or b CNV genome
c The underlined nucleotide sequences are mutated to generate alanine substitutions.

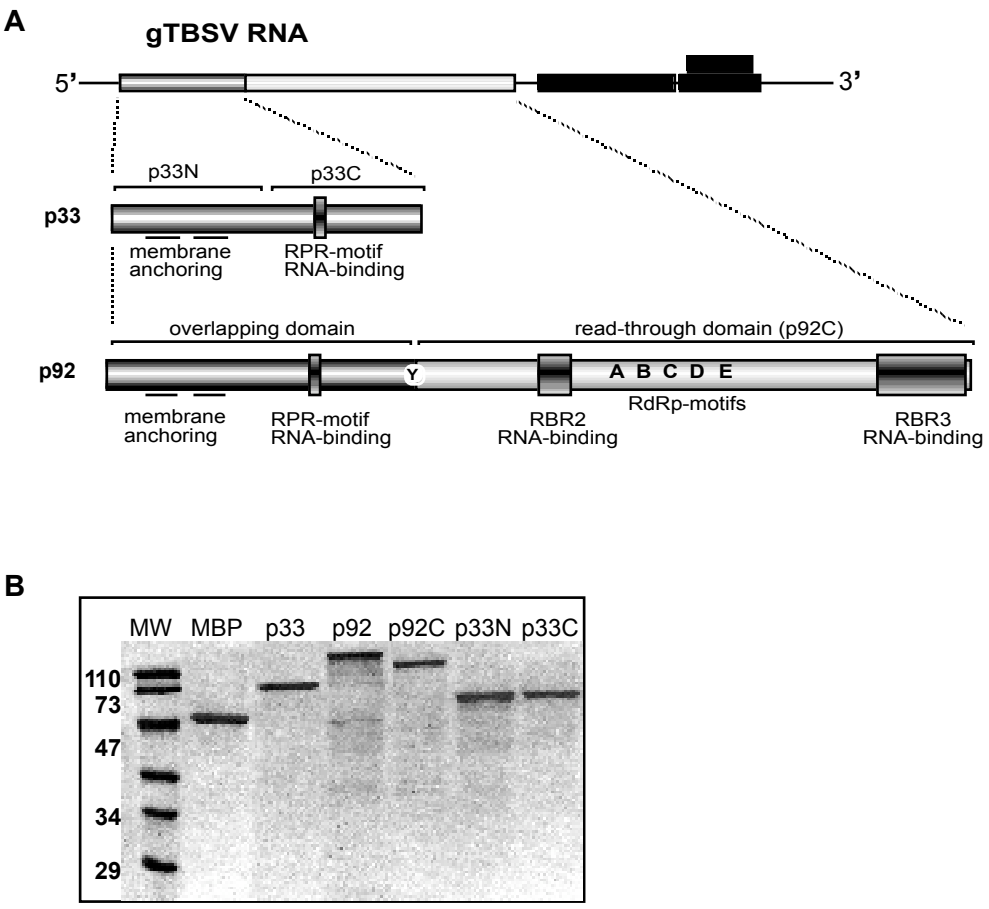


Figure 4.1. Schematic representation of the TBSV replicase proteins. (A) The five open reading frames coded by the TBSV genomic RNA (~4,800 nt in length) are indicated with boxes, and the noncoding sequences with solid lines. The known and predicted functional domains in the two overlapping replicase proteins, p33 and p92, are shown schematically. Y represents an amber stop codon-to-tyrosine codon mutation in p92 to facilitate its expression in *E. coli* and yeast. (B) SDS-PAGE analysis of purified recombinant p33 and p92 proteins and their derivatives, expressed as MBP fusion proteins in *E. coli*.

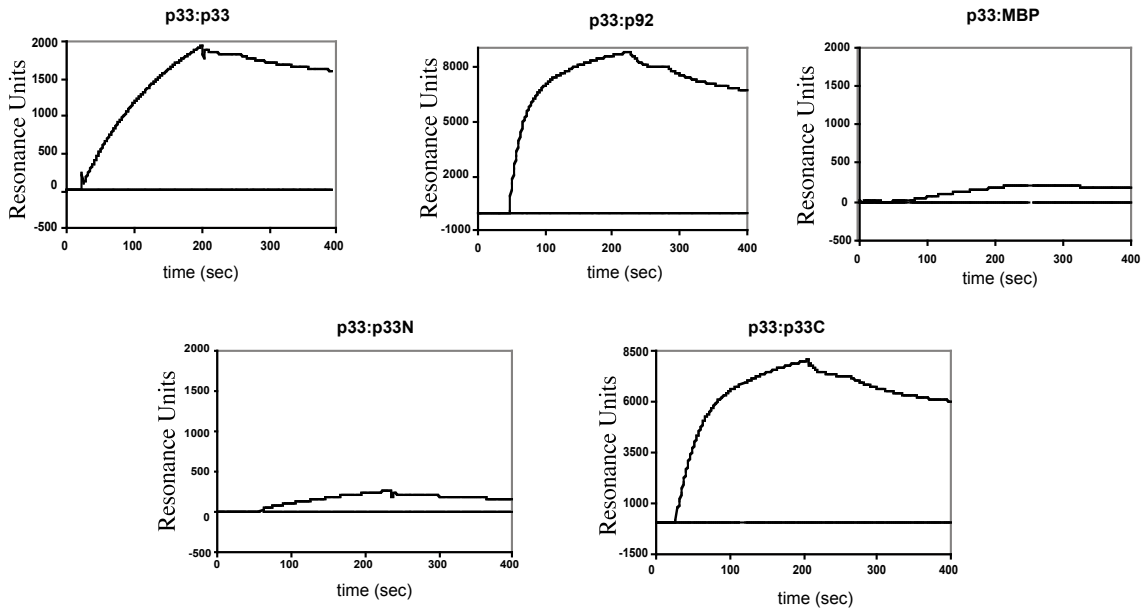


Figure 4.2. Surface plasmon resonance analysis of interactions between TBSV replicase proteins. The p33 and p92 replicase proteins and their derivatives, expressed in *E. coli*, were affinity-purified as C-terminal fusions to MBP. The purified MBP-p33 was immobilized on CM-5 sensor chip, whereas p33, p92, MBP, p33N and p33C (1 μM each) were separately injected in the running buffer over the chip. The interaction data shown were subtracted from the data from control surface (baseline) to account for bulk effects of running buffer and/or non-specific binding.

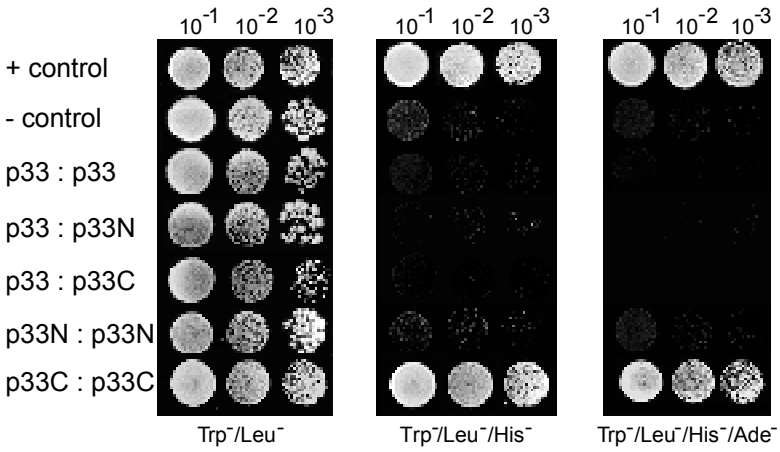


Figure 4.3. Defining interactions between p33:p33 proteins of TBSV in the yeast two-hybrid assay. The yeast two-hybrid plasmids harboring GAL4 activation (AD) and binding domains (BD) fused to the p33 gene or its derivatives were co-transformed into AH109 cells. The transformed cells were serially diluted 10-fold and grown on three different selection media as shown. Growth on Trp^-/Leu^- indicates successful transformation of AH109 cells with plasmids. Growth on $Trp^-/Leu^-/His^-$ and $Trp^-/Leu^-/His^-/Ade^-$ indicates transcriptional activation of HIS and ADE marker genes through interactions between AD and BD fusion proteins. The positive and negative controls included in the Matchmaker 3 system were used here. Proteins p33N and p33C represent N- and C-terminal halves of p33 protein (Figure 4.1A).

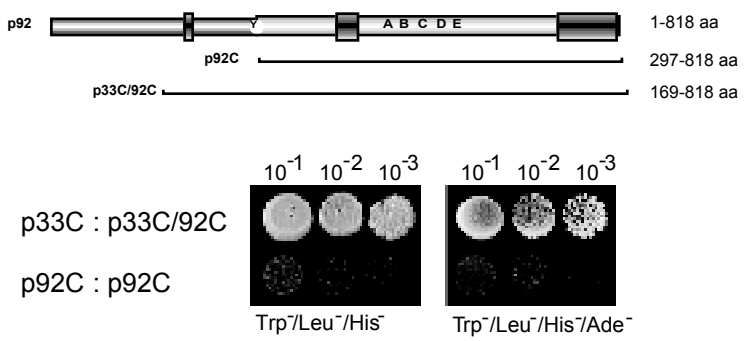


Figure 4.4. Interaction between the TBSV p33 and p92 derivatives in yeast. **A**. Schematic representation of the full-length p92 ($p92_{1-818}$) and its deletion derivatives tested in the yeast two-hybrid assay. The positions of the amino acids (aa) present in particular p92 derivatives are shown on the right. **B**. Data obtained in the yeast two-hybrid assay is shown with the specified constructs. See details in the legend to Figure 4.3.

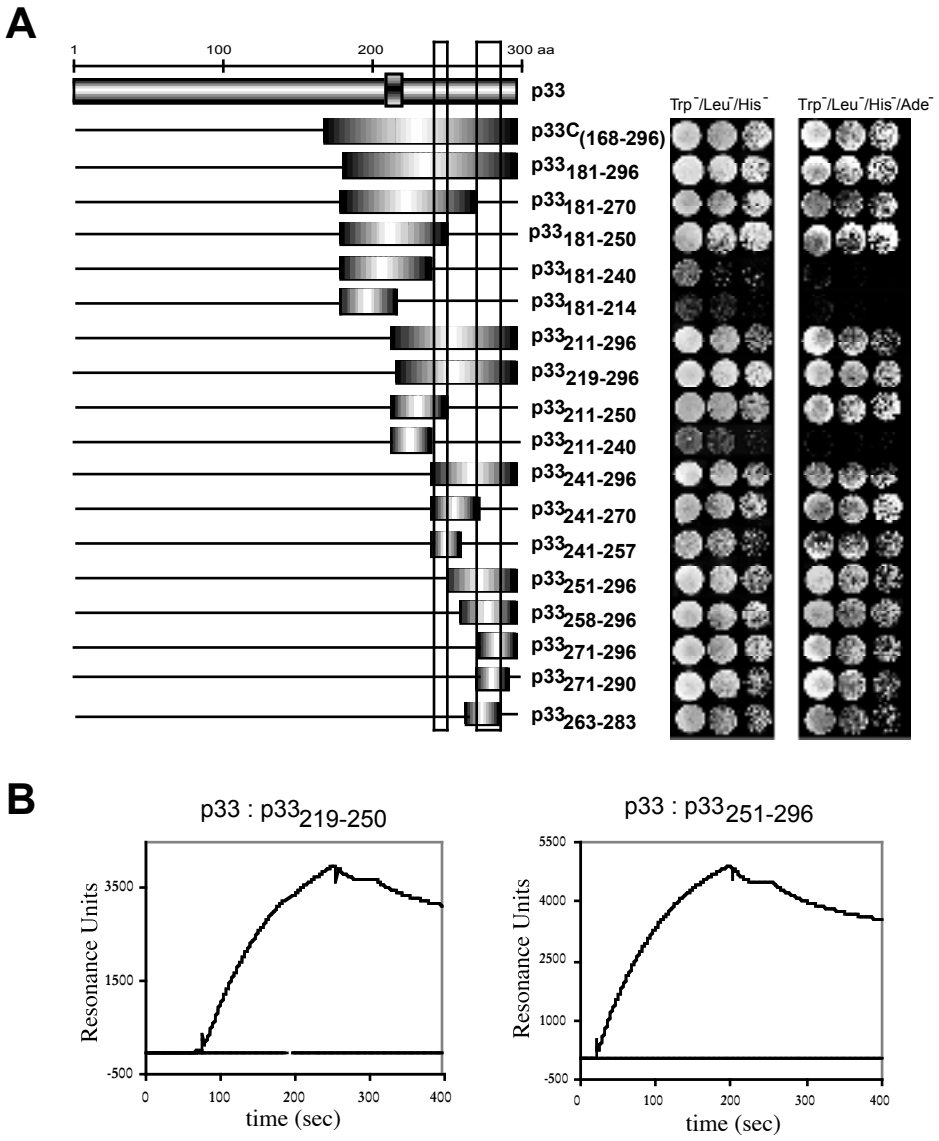


Figure 4.5. Defining short regions in p33 that promote p33:p33 interaction. (A) Schematic representation of p33 and its deletion derivatives tested in the yeast two-hybrid assay is shown on the left. The deleted sequences are indicated with solid lines. The two sites (site 1 and site 2) involved in p33:p33 interaction are boxed. The amino acid positions in particular p33 derivatives are shown in the center, whereas data obtained in the yeast two-hybrid assay with p33C ($p33_{169-296}$) as bait are shown on the right. See details in the legend to Figure 4.3. (B) Biosensor analysis of interactions between p33 and its two deletion derivatives, $p33_{219-250}$ containing interaction site 1 and $p33_{251-296}$ containing interaction site 2. The biosensor experiments were performed as described in the legend to Figure 4.2.

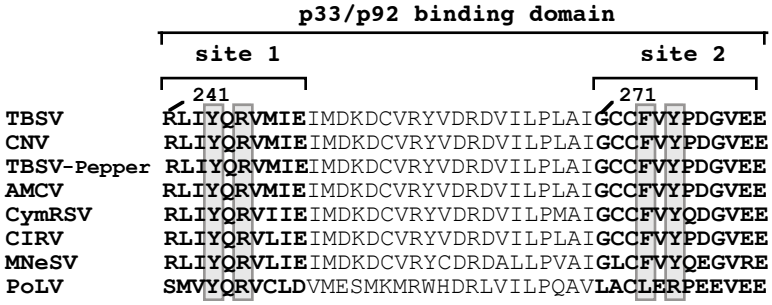


Figure 4.6. Sequence alignment of p33:p33/p92 interaction sites in Tombus- and related viruses. Interaction sites 1 and 2 are shown with bold-faced letters. The conserved amino acids selected for mutagenesis in the CNV sequence are boxed. The following abbreviations were used: TBSV, *Tomato bushy stunt virus* (cherry strain, top, and pepper strain, third row); CNV, *Cucumber necrosis virus*; AMCV, *Artichoke mottled crinkle virus*; CymRSV, *Cymbidium ringspot virus*; CIRV, *Carnation Italian ringspot virus*; MNeSV, *Maize necrotic streak virus*; and *Pothos latent virus* (PoLV, Tombusviridae).

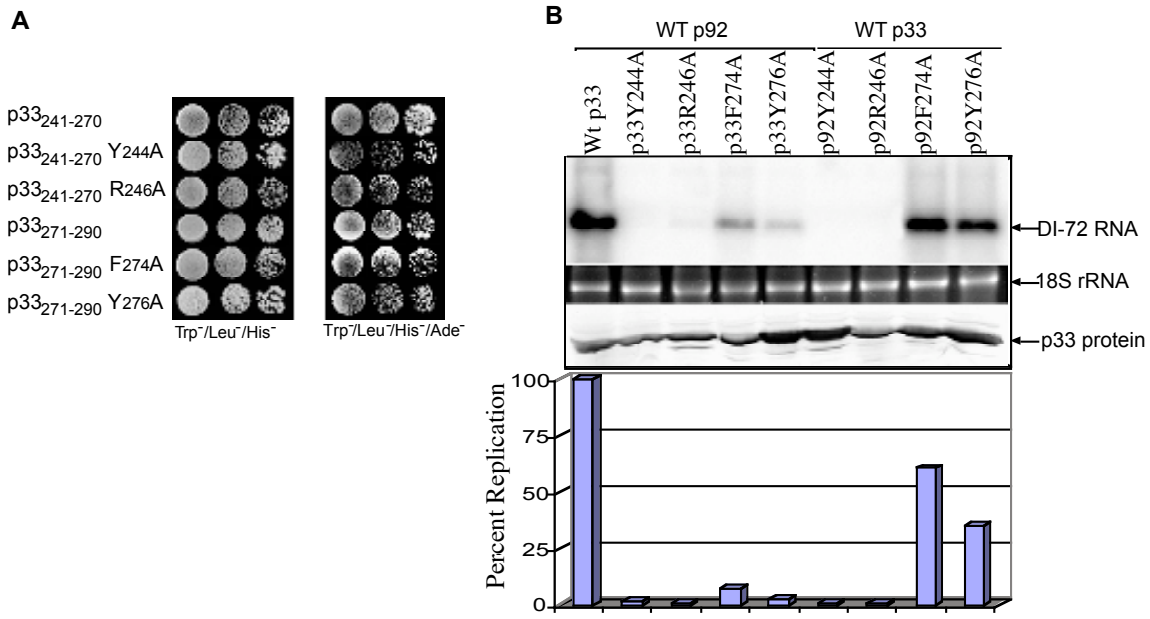


Figure 4.7. Effect of mutations within p33:p33/p92 interaction sites 1 and 2 on DI RNA replication in yeast. (A) Two-hybrid interactions between TBSV $p33C_{169-296}$ and the mutants (Figure 4.6) generated by substituting selected amino acids with alanine in the protein interaction sites. Note that the mutations were introduced into truncated p33 proteins, either $p33_{241-270}$ or $p33_{271-290}$ as shown. (B) Northern blot analysis of the effects of the above mutations on replication of DI-72 RNA of TBSV. Yeast strain Sc1 was co-transformed with three plasmids coding for full-length wild type and/or full-length mutants of CNV replicase proteins (p33 and p92) as shown and a plasmid coding for the TBSV DI-72 RNA replicon. The position of the DI-72 RNA in the total RNA isolated from transformed yeast cells is shown on the right. The ethidium bromide-stained agarose gel with the 18S ribosomal RNA is shown in the middle as loading control. Western blot analysis of the amount of p33 present in yeast is shown below the agarose gel. The relative level of DI-72 RNA replication in yeast, based on quantification of the Northern blot using PhosphorImager and ImageQuant (v1.2) software from three separate experiments, was plotted in the graph.

CHAPTER FIVE

DISCUSSION AND SUMMARY

For viruses, replication of their genetic material is the major step in their life cycle. They carry out this important event by taking advantage of host cellular machinery and their own proteins marked for this purpose. These replication proteins are the subjects of the dissertation research and I chose to study their role in replication using tombusvirus as a model system. The major emphasis was given to RNA dependent RNA polymerase (RdRp), RNA-binding and protein-binding activities of replicase proteins.

For various biochemical analyses, I required pure proteins in large quantities. For this purpose, I collaborated with Judit Pogany, a post-doctoral associate in the laboratory of Dr. Peter Nagy to produce large quantities of proteins from *E. coli* expression system. Replicase proteins of TBSV (p33 and p92) and its close relative TCV (p28 and p88) were expressed in parallel in *E. coli* as a N-terminal fusion to maltose binding protein (MBP) and then purified using amylose resin-based affinity chromatography. The main goal here is to understand the functions of p33 and p92 proteins of TBSV in virus replication. TCV replicase proteins were chosen to compare and contrast with TBSV proteins for their biochemical functions such as RdRp and RNA-binding activities.

When tested *in vitro*, surprisingly, only p88 exhibited RdRp activity by synthesizing complementary RNA on the input template RNA, while p92 did not under identical assay conditions. The result is surprising because both p88 and p92 proteins share significant sequence similarity at the aminoacid level. Several modifications in experimental procedures that were carried out for p92 did not yield an active RdRp. Nevertheless, I proceeded with TCV p88 to study the characteristics of RdRp *in vitro*.

TCV p88 is a readthrough protein of the p28, therefore shares the entire p28 sequence at its N-terminus. A deletion derivative of p88, referred as p88C, which lacked the entire p28 sequence, was found to be 10-fold more active than the full-length p88. There are only a few plant viruses - tobacco vein mottling potyvirus (52), bamboo mosaic potexvirus (75) - for which the RdRp activity of purified recombinant replicase proteins was demonstrated. For other plant RNA viruses – brome mosaic virus (118), tobacco mosaic virus (98)**,** turnip crinkle virus (140),

tomato bushy stunt virus and cucumber necrosis virus (86, 105) - the RdRp function of replicase proteins were demonstrated with partially-purified RdRp complexes from plant or yeast cells. However, for most of other RNA viruses the RdRp function was deduced solely from sequence motifs (65, 94). p88C is one of the highly active RdRps and similar to HCV RdRp and poliovirus RdRp ($3D^{pol}$) in molecular size. Highly efficient p88C raises an obvious question on the effect of p28 sequences at the N-terminus of p88 in RNA synthesis. Does the N-terminus play any regulatory role in RNA synthesis *in vivo*? Or is it an artifact of deletions and heterologous expression system? Is p28 domain a structural constraint for the RdRp function p88 protein, thus reducing the efficiency of RNA polymerization? Answers to these questions will elevate our understanding of biochemistry of polymerase functions.

Further deletions in p88C in both N- and C-terminal directions (ΔN30/C100 and ΔN100/C30) resulted in loss of RdRp activity. None of these deletions affected the conserved RdRp-specific motifs in p88C, which suggests that the regions flanking these motifs are important either for maintaining proper folding of RdRp or involved in other functions such as binding to RNA template during RNA synthesis. Studies with recombinant TBSV replicase proteins showed that the regions flanking the polymerase motifs were involved in RNA-binding. Indeed, p88C was found to bind to template RNA in gel-shift assay, however the regions involved in RNA-binding are yet to be mapped. Therefore, all these results, taken together, suggest that the loss of RdRp activity in p88C deletion mutants may be due to loss of or inefficient template binding during RNA synthesis. It's also possible that the deletions in p88C affected the proper maintenance of polymerase structures leading to loss of RdRp activity.

The *E. coli* expressed recombinant TCV RdRp's (p88 and p88C) behavior in *in vitro* RdRp assay is remarkably similar to TCV RdRp purified from virus-infected plants. Plant purified TCV RdRp is presumably a multi-subunit enzyme complex containing both p28 and p88 along with some host components as opposed to single unit p88 or p88C. The replication of viral RNAs is a two step process where the genomic plus strand is used as template to synthesize complementary minus stand in the first step and then in the second step the duplex is unwound so that the newly synthesized minus strands are used as template to make genomic RNAs. In *in vitro* assay RdRp mainly performs transcription of complementary RNA on input RNA templates. Therefore, comparison between single subunit recombinant p88 or p88C and multi-subunit plant purified TCV RdRp effectively lies at this one step and not the entire replication

cycle of viral RNAs. For full cycle of replication, more factors than just polymerase domain are likely required as suggested for several viruses like TMV, BMV and CNV (86, 98, 118, 119). However, for dissecting the role of various *cis*-acting elements in viral RNA, interactions between RNA and polymerase and polymerase functional studies, this single-subunit RdRp is an ideal system. All these results clearly demonstrate for the first time that p88 protein of TCV is the RdRp, the polymerase activity lies within the readthrough portion (p88C) and the RdRp activity of p88 or p88C is similar to that of plant purified TCV RdRp.

Having characterized the RdRp and RNA binding activities of TCV p88 and p88C, I shifted my focus to understanding the role of p33 and p92 of TBSV. p92 has been assigned as RdRp based on sequence analysis, but p33 does not contain any known functional motifs. However, protein topology prediction programs predicted the presence of two integral transmembrane helices at the N-terminal half of p33. Experimental evidence for p33-like proteins of related tombusviruses also concur with the notion that p33 and p92 are membrane proteins (128, 153). Being a membrane protein with no known catalytic motifs and its presence in large quantities relative to polymerase protein - p92 (20:1), p33 is hypothesized to recruit itself, viral RNA and p92 RdRp molecules to replication complex and target the complex to intracellular host membranes, which is usually the sites of replication for most RNA viruses. In order to validate the hypothesis, first the macromolecular interactions between replicase proteins and viral RNA have to be understood.

Accordingly, I demonstrated that both p33 and p92 interacted with viral RNA. The RNA binding region was mapped to an arginine-proline rich motif (RPR-motif) in the C-terminal half of p33 sequence. Bioinformatic analysis showed that the RPR-motif region is surface-exposed hydrophilic pocket. It's also interesting that RPR-motif is highly conserved in tombus- and also in carmo-viruses. This RPR-motif also has a perfect sequence identity with the RNA-binding basic region of Tat proteins of some HIV1 strains. The sequence identity between p33 and Tat is unlikely to have any evolutionary significance considering the fact that these two proteins have no sequence or structural similarity other than preponderance of arginine and proline residues.

The RNA-binding of p33 is characterized by its preference to single stranded RNA and cooperativity, a process in which binding of one molecule of p33 to RNA favors the binding of additional p33 molecules. This process ensures that the viral RNA is partly or fully coated with p33 proteins, which may be beneficial to keep (+) and (-) strands separated from forming duplex

during replication and also to protect the viral RNAs from host nucleases including ribonuclease dicer complex involved in gene silencing mechanism. It may also help to recruit viral RNAs for replication from translation by preventing host ribosomal access to RNA by tightly binding to it. Since p92 also has p33 sequence, cooperative binding of both p33 and p92 to RNA could be a major factor that brings all three viral components of replication together. As demonstrated for poliovirus and HCV RdRp (112, 150), cooperative binding of RdRp to viral RNA template could enhance its processivity leading to efficient transcription of viral RNAs.

p92 also has two additional RNA binding regions (RBR) in its readthrough portion, referred as p92C. These regions are located on both sides of RdRp catalytic motifs at the primary sequence level. RBR1 includes the motif F, also found in other RdRps which functions as NTP binding site (71). With all the polymerase motifs, RBRs and similar molecular size (~ 53 to 63 kDa), p92C portion of p92 may be functionally analogous to HCV and poliovirus RdRps. However, both poliovirus and HCV RdRps have been shown to oligomerize and oligomerization was found to be critical for their activity. For p92 RdRp, the oligomerization status is not known and also the evidence for p92 RdRp activity is lacking. It will be interesting to find out the strength of p92C binding to RNA and compare it with the strength of RNA binding by p33. p92C binding to RNA has to be transient in order to allow the RdRp or RNA to move along during polimerization. It was shown using UV-cross linking experiments that TMV 126 kDa was able to specifically bind to the 3' end of genomic RNA, whereas polymerase domain couldn't (99). The authors reasoned that non-binding of polymerase domain to RNA might be due to weaker binding. Generally the regulatory proteins like *Tat* and *Rev* of HIV 1 bind to RNA target with much higher affinity than the polymerases like poliovirus RdRp. The binding strength will shed some light on the role of p33 domain on p92 protein.

The observation of cooperative binding by p33 brings to the light the protein – protein interactions between p33 molecules and also between p33 and p92. The yeast two-hybrid, SPR and pull down (not shown) assays clearly shows that p33 interacts with itself and also with p92. There are two interaction sites in p33 and mutations in these two sites affected the replication of DI-RNA in yeast. Careful analysis of data indicates that mutations in site I affect the yeast growth in yeast two-hybrid assay and DI-RNA replication in yeast more dramatically than do mutations in site II. However, further analysis is required to confirm the contributions of site I and II independently to the overall binding strength of p33. These two sites are located close to

each other and also proximal to RNA binding-motif at the C-terminus of p33. The proximity of both RNA- and protein-binding regions within a span of less than 90 amino acids suggests that they all fall into one large functional domain at the C-terminus of p33. Confirming to this point, the sequence alignment of several members of tombusvirus genus reveals that the C-terminal segment of p33-like proteins is highly conserved where as N-terminal segment is less conserved. This suggests the importance of the C-terminal amino acid residues for the function of p33.

Protein-protein interactions between p33 and p92 could be a force in addition to RNA-binding for recruiting the less abundant p92 polymerase to the site of replication. It has been demonstrated for TMV that immuno-affinity purified RdRp complex contained 126 kDa (analogous to p33) and 183 kDa (analogous to p92) proteins at 1:1 molar ratio and the RdRp could synthesize minus strand RNAs on plus strand templates (151). Also large excess of free 126 kDa proteins were found in infected plants. Similarly, p33 could associate with p92 to form polymerase complex whereas the free excess p33 could be involved in recruiting viral RNA and also possibly some host factors to the replication sites.

Overall, it's reasonable to assume that p33 is a multifunctional protein: the N-terminus of p33 anchors to the membranes and the C-terminus engages in binding to other p33 molecules, viral RNA and p92 and possibly some host factors. All these functions are likely required for assembling all the viral replication components together (Figure 5.1). It would be interesting to find out if p33 is also involved in recruiting host factors to this complex.

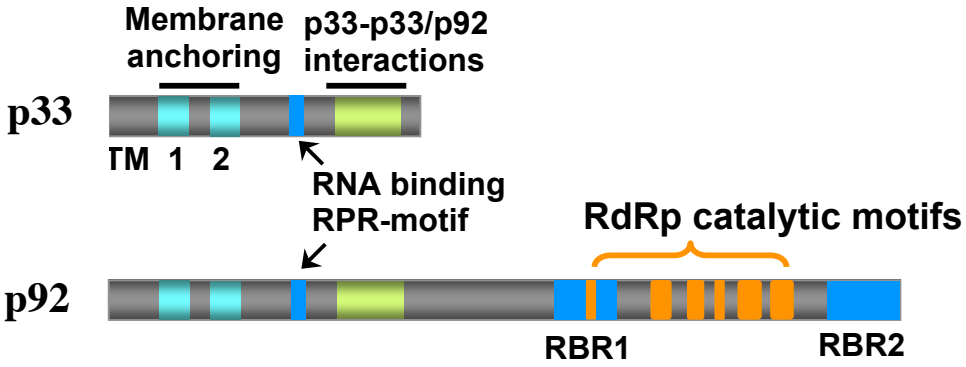


Figure 5.1 Schematic representation of functional domains in replicase proteins of tomato bushy stunt virus. TM refers to transmembrane domains. RNA binding regions are abbreviated as RBR (shown in blue).

REFERENCES

1. **Ago, H., T. Adachi, A. Yoshida, M. Yamamoto, N. Habuka, K. Yatsunami, and M. Miyano.** 1999. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Structure Fold Des **7:**1417-26.
2. **Agol, V. I., A. V. Paul, and E. Wimmer.** 1999. Paradoxes of the replication of picornaviral genomes. Virus Res **62:**129-47.
3. **Ahlquist, P.** 2002. RNA-dependent RNA polymerases, viruses, and RNA silencing. Science **296:**1270-3.
4. **Ahlquist, P., A. O. Noueiry, W. M. Lee, D. B. Kushner, and B. T. Dye.** 2003. Host factors in positive-strand RNA virus genome replication. J Virol **77:**8181-6.
5. **Andrews, N. C., and D. Baltimore.** 1986. Purification of a terminal uridylyltransferase that acts as host factor in the in vitro poliovirus replicase reaction. Proc Natl Acad Sci U S A **83:**221-5.
6. **Ansel-McKinney, P., and L. Gehrke.** 1998. RNA determinants of a specific RNA-coat protein peptide interaction in alfalfa mosaic virus: Conservation of homologous features in ilarvirus RNAs. J Mol Biol **278:**767-85.
7. **Argos, P.** 1988. A sequence motif in many polymerases. Nucleic Acids Res **16:** 9909-16.
8. **Axelrod, V. D., E. Brown, C. Priano, and D. R. Mills.** 1991. Coliphage q beta RNA replication: RNA catalytic for single-strand release. Virology **184:**595-608.
9. **Baer, M. L., F. Houser, L. S. Loesch-Fries, and L. Gehrke.** 1994. Specific RNA binding by amino-terminal peptides of alfalfa mosaic virus coat protein. Embo J **13:**727-35.
10. **Bandziulis, R. J., M. S. Swanson, and G. Dreyfuss.** 1989. RNA-binding proteins as developmental regulators. Genes Dev **3:**431-7.
11. **Barton, D. J., B. J. Morasco, and J. B. Flanegan.** 1996. Assays for poliovirus polymerase, 3D(pol), and authentic RNA replication in HeLa s10 extracts. Methods Enzymol **275:**35-57.
12. **Bayer, P., M. Kraft, A. Ejchart, M. Westendorp, R. Frank, and P. Rosch.** 1995. Structural studies of HIV-1 tat protein. J Mol Biol **247:**529-35.

13. **Beckman, M. T., and K. Kirkegaard.** 1998. Site size of cooperative single-stranded RNA binding by poliovirus RNA-dependent RNA polymerase. J Biol Chem **273:**6724-30.

14. **Behrens, S. E., L. Tomei, and R. De Francesco.** 1996. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. Embo J **15:**12-22.

15. **Blumenthal, T., and G. G. Carmichael.** 1979. RNA replication: Function and structure of qbeta-replicase. Annu Rev Biochem **48:**525-48.

16. **Bressanelli, S., L. Tomei, A. Roussel, I. Incitti, R. L. Vitale, M. Mathieu, R. De Francesco, and F. A. Rey.** 1999. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Proc Natl Acad Sci U S A **96:**13034-9.

17. **Buck, K. W.** 1996. Comparison of the replication of positive-stranded RNA viruses of plants and animals. Adv Virus Res **47:**159-251.

18. **Buck, K. W.** 1999. Replication of tobacco mosaic virus RNA. Philos. Trans. R. Soc. Lond. Biol. Sci. **354:**613-27.

19. **Burd, C. G., and G. Dreyfuss.** 1994. Conserved structures and diversity of functions of RNA-binding proteins. Science **265:**615-21.

20. **Burgyan, J., C. Hornyik, G. Szittya, D. Silhavy, and G. Bisztray.** 2000. The ORF1 products of tombusviruses play a crucial role in lethal necrosis of virus-infected plants. J Virol **74:**10873-81.

21. **Burgyan, J., L. Rubino, and M. Russo.** 1996. The 5'-terminal region of a tombusvirus genome determines the origin of multivesicular bodies. J Gen Virol **77 (Pt 8):**1967-74.

22. **Butcher, S. J., J. M. Grimes, E. V. Makeyev, D. H. Bamford, and D. I. Stuart.** 2001. A mechanism for initiating RNA-dependent RNA polymerization. Nature **410:**235-40.

23. **Calnan, B. J., B. Tidor, S. Biancalana, D. Hudson, and A. D. Frankel.** 1991. Arginine-mediated RNA recognition: The arginine fork. Science **252:**1167-71.

24. **Carrington, J. C., L. A. Heaton, D. Zuidema, B. I. Hillman, and T. J. Morris.** 1989. The genome structure of turnip crinkle virus. Virology **170:**219-26.

25. **Cascone, P. J., T. F. Haydar, and A. E. Simon.** 1993. Sequences and structures

required for recombination between virus-associated RNAs. Science **260:**801-5.

26. **Chapman, M. R., and C. C. Kao.** 1999. A minimal RNA promoter for minus-strand RNA synthesis by the brome mosaic virus polymerase complex. J Mol Biol **286:**709-20.
27. **Cheng, C. P., J. Pogany, and P. D. Nagy.** 2002. Mechanism of DI-RNA formation in tombusviruses: Dissecting the requirement for primer extension by the tombusvirus RNA dependent RNA polymerase in vitro. Virology **304:**460-73.
28. **Citovsky, V., D. Knorr, G. Schuster, and P. Zambryski.** 1990. The p30 movement protein of tobacco mosaic virus is a single-strand nucleic acid binding protein. Cell **60:**637-47.
29. **Crouch, R. J., M. Wakasa, and M. Haruki.** 1999. Detection of nucleic acid interactions using surface plasmon resonance. Methods Mol Biol **118:**143-60.
30. **Deiman, B. A., K. Seron, E. M. Jaspars, and C. W. Pleij.** 1997. Efficient transcription of the tRNA-like structure of turnip yellow mosaic virus by a template-dependent and specific viral RNA polymerase obtained by a new procedure. J Virol Methods **64:**181-95.
31. **Deiman, B. A., P. W. Verlaan, and C. W. Pleij.** 2000. In vitro transcription by the turnip yellow mosaic virus RNA polymerase: A comparison with the alfalfa mosaic virus and brome mosaic virus replicases. J Virol **74:**264-71.
32. **Donald, R. G., and A. O. Jackson.** 1996. RNA-binding activities of barley stripe mosaic virus gamma b fusion proteins. J Gen Virol **77 (Pt 5):**879-88.
33. **Donald, R. G., D. M. Lawrence, and A. O. Jackson.** 1997. The barley stripe mosaic virus 58-kilodalton beta protein is a multifunctional RNA binding protein. J Virol **71:**1538-46.
34. **Draper, D. E.** 1999. Themes in RNA-protein recognition. J Mol Biol **293:**255-70.
35. **Dreher, T. W.** 1999. Functions of the 3'-untranslated regions of positive strand RNA viral genomes. Annu Rev Phytopathol **37:**151-174.
36. **Fabian, M. R., H. Na, D. Ray, and K. A. White.** 2003. 3'-terminal RNA secondary structures are important for accumulation of tomato bushy stunt virus DI-RNA s. Virology **313:**567-80.
37. **Fernandez, A., S. Lain, and J. A. Garcia.** 1995. RNA helicase activity of the

plum pox potyvirus CI protein expressed in *Escherichia coli*. Mapping of an RNA binding domain. Nucleic Acids Res **23:**1327-32.

38. **Ferrari, E., J. Wright-Minogue, J. W. Fang, B. M. Baroudy, J. Y. Lau, and Z. Hong.** 1999. Characterization of soluble hepatitis C virus RNA-dependent RNA polymerase expressed in *Escherichia coli*. J Virol **73:**1649-54.
39. **Ferrari, M. E., W. Bujalowski, and T. M. Lohman.** 1994. Co-operative binding of *Escherichia coli* SSB tetramers to single-stranded DNA in the (SSB)35 binding mode. J Mol Biol **236:**106-23.
40. **Gallet, X., B. Charloteaux, A. Thomas, and R. Brasseur.** 2000. A fast method to predict protein interaction sites from sequences. J Mol Biol **302:**917-26.
41. **Goregaoker, S. P., and J. N. Culver.** 2003. Oligomerization and activity of the helicase domain of the tobacco mosaic virus 126- and 183-kilodalton replicase proteins. J Virol **77:**3549-56.
42. **Goregaoker, S. P., D. J. Lewandowski, and J. N. Culver.** 2001. Identification and functional analysis of an interaction between domains of the 126/183-kda replicase-associated proteins of tobacco mosaic virus. Virology **282:**320-8.
43. **Guan, H., C. Song, and A. E. Simon.** 1997. RNA promoters located on (-)-strands of a subviral RNA associated with turnip crinkle virus. RNA **3:**1401-12.
44. **Hagiwara, Y., K. Komoda, T. Yamanaka, A. Tamai, T. Meshi, R. Funada, T. Tsuchiya, S. Naito, and M. Ishikawa.** 2003. Subcellular localization of host and viral proteins associated with tobamovirus RNA replication. Embo J **22:**344-53.
45. **Hansen, J. L., A. M. Long, and S. C. Schultz.** 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. Structure **5:**1109-22.
46. **Harrison, S. C., A. J. Olson, C. E. Schutt, F. K. Winkler, and G. Bricogne.** 1978. Tomato bushy stunt virus at 2.9 a' resolution. Nature **276:**368-373.
47. **Harrison, S. C.** 1980. Virus crystallography comes of age. Nature **286:**558-9.
48. **Hayes, R. J., and K. W. Buck.** 1990. Complete replication of a eukaryotic virus RNA in vitro by a purified RNA-dependent RNA polymerase. Cell **63:**363-8.
49. **Hearne, P. Q., D. A. Knorr, B. I. Hillman, and T. J. Morris.** 1990. The complete genome structure and synthesis of infectious RNA from clones of tomato bushy stunt virus. Virology **177:**141-51.

50. **Hillman, B. I., J. C. Carrington, and T. J. Morris.** 1987. A defective interfering RNA that contains a mosaic of a plant virus genome. Cell **51:**427-33.

51. **Hobson, S. D., E. S. Rosenblum, O. C. Richards, K. Richmond, K. Kirkegaard, and S. C. Schultz.** 2001. Oligomeric structures of poliovirus polymerase are important for function. Embo J **20:**1153-63.

52. **Hong, Y., and A. G. Hunt.** 1996. RNA polymerase activity catalyzed by a potyvirus-encoded RNA-dependent RNA polymerase. Virology **226:**146-51.

53. **Hope, D. A., S. E. Diamond, and K. Kirkegaard.** 1997. Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3ab. J Virol **71:**9490-8.

54. **Huang, C. Y., Y. L. Huang, M. Meng, Y. H. Hsu, and C. H. Tsai.** 2001. Sequences at the 3' untranslated region of bamboo mosaic potexvirus RNA interact with the viral RNA-dependent RNA polymerase. J Virol **75:**2818-24.

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

56. **Kadare, G., and A. L. Haenni.** 1997. Virus-encoded RNA helicases. J Virol **71:**2583-90.

57. **Kamen, R.** 1970. Characterization of the subunits of q-beta replicase. Nature **228 :**527-33.

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

59. **Kao, C. C., R. Quadt, R. P. Hershberger, and P. Ahlquist.** 1992. Brome mosaic virus RNA replication proteins 1a and 2a from a complex in vitro. J Virol **66:**6322-9.

60. **Kao, C. C., P. Singh, and D. J. Ecker.** 2001. De novo initiation of viral RNA-dependent RNA synthesis. Virology **287:**251-60.

61. **Kao, C. C., and J. H. Sun.** 1996. Initiation of minus-strand RNA synthesis by the brome mosaicvirus RNA-dependent RNA polymerase: Use of oligoribo-nucleotide primers. J Virol **70:**6826-30.

62. **Kao, C. C., X. Yang, A. Kline, Q. M. Wang, D. Barket, and B. A. Heinz.** 2000. Template requirements for RNA synthesis by a recombinant hepatitis C virus RNA-dependent RNA polymerase. J Virol **74:**11121-8.

63. **Knorr, D. A., R. H. Mullin, P. Q. Hearne, and T. J. Morris.** 1991. De novo generation of defective interfering RNAs of tomato bushy stunt virus by high multiplicity passage. Virology **181:**193-202.

64. **Kollar, A, and J. Burgyan.** 1994. Evidence that ORF 1 and 2 are the only virus-encoded replicase genes of cymbidium ringspot tombusvirus.Virology **201:**169-72.

65. **Koonin, E. V.** 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J Gen Virol **72 (Pt 9):**2197-206.

66. **Kozak, M.** 1989. The scanning model for translation: An update. J Cell Biol **108:** 229-41.

67. **Kushnirov, V. V.** 2000. Rapid and reliable protein extraction from yeast. Yeast **16:**857-60.

68. **Lai, M. M.** 1998.Cellular factors in the transcription and replication of viral RNA genomes: A parallel to DNA-dependent RNA transcription. Virology **244:**1-12.

69. **Lai, V. C., C. C. Kao, E. Ferrari, J. Park, A. S. Uss, J. Wright-Minogue, Z. Hong, and J. Y. Lau.** 1999. Mutational analysis of bovine viral diarrhea virus RNA-dependent RNA polymerase. J Virol **73:**10129-36.

70. **Lee, E. G., A. Alidina, C. May, and M. L. Linial.** 2003. Importance of basic residues in binding of rous sarcoma virus nucleocapsid to the RNA packaging signal. J Virol **77:**2010-20.

71. **Lesburg, C. A., M. B. Cable, E. Ferrari, Z. Hong, A. F. Mannarino, and P. C. Weber.** 1999. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. Nat Struct Biol **6:**937-43.

72. **Li, G. P., and C. M. Rice.** 1989. Mutagenesis of the in-frame opal termination codon preceding nsp4 of Sindbis virus: Studies of translational readthrough and its effect on virus replication. J Virol **63:**1326-37.

73. **Li, Q., and P. Palukaitis.** 1996. Comparison of the nucleic acid- and NTP-binding properties of the movement protein of cucumber mosaic cucumovirus and

tobacco mosaic tobamovirus. Virology **216:**71-9.

74. **Li, X., Heaton LA, Morris TJ, Simon AE.** 1989. Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated denovo. Proc Natl Acad Sci U S A **86:**9173-9177.

75. **Li, Y. I., Y. M. Cheng, Y. L. Huang, C. H. Tsai, Y. H. Hsu, and M. Meng.** 1998. Identification and characterization of the *Escherichia coli*-expressed RNA-dependent RNA polymerase of bamboo mosaic virus. J Virol **72:**10093-9.

76. **Lohmann, V., F. Korner, U. Herian, and R. Bartenschlager.** 1997. Biochemical properties of hepatitis C virus NS5b RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. J Virol **71:**8416-28.

77. **Luo, G., R. K. Hamatake, D. M. Mathis, J. Racela, K. L. Rigat, J. Lemm, and R. J. Colonno.** 2000. De novo initiation of RNA synthesis by the RNA-dependent RNA polymerase (ns5b) of hepatitis C virus. J Virol **74:**851-63.

78. **Lyle, J. M., E. Bullitt, K. Bienz, and K. Kirkegaard.** 2002. Visualization and functional analysis of RNA-dependent RNA polymerase lattices. Science **296:** 2218-22.

79. **Ma, B., T. Elkayam, H. Wolfson, and R. Nussinov.** 2003. Protein-protein interactions: Structurally conserved residues distinguish between binding sites and exposed protein surfaces. Proc Natl Acad Sci USA **100:**5772-7.

80. **McGhee, J. D., and P. H. von Hippel.** 1974. Theoretical aspects of DNA-protein interactions: Co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. J Mol Biol **86:**469-89.

81. **Merits, A., R. Kettunen, K. Makinen, A. Lampio, P. Auvinen, L. Kaariainen, and T. Ahola.** 1999. Virus-specific capping of tobacco mosaic virus RNA: Methylation of GTP prior to formation of covalent complex p126-m7gmp. FEBS Lett **455:**45-8.

82. **Mouches, C., C. Bove, and J. M. Bove.** 1974. Turnip yellow mosaic virus-RNA replicase: Partial purification of the enzyme from the solubilized enzyme-template complex. Virology **58:**409-23.

83. **Murphy, K. P., V. Bhakuni, D. Xie, and E. Freire.** 1992. Molecular basis of co-

operativity in protein folding. III. Structural identification of cooperative folding units and folding intermediates. J Mol Biol **227:**293-306.

84. **Nagy, P. D., and J. Pogany.** 2003. Role of enhancer and silencer elements in replication of defective interfering RNAs and satellite RNAs associated with tombusvirus and carmovirus infections, p. 171-187. *In* X. Zhang (ed.), RNA viruses: Mechanisms of replication and transcription, vol. 37/661. Research Signpost, Trivandrum.

85. **Nagy, P. D., C. D. Carpenter, and A. E. Simon.** 1997. A novel 3'-end repair mechanism in an RNA virus. Proc Natl Acad Sci U S A **94:**1113-8.

86. **Nagy, P. D., and J. Pogany.** 2000. Partial purification and characterization of cucumber necrosis virus and tomato bushy stunt virus RNA-dependent RNA polymerases: Similarities and differences in template usage between tombusvirus and carmovirus RNA-dependent RNA polymerases. Virology **276:**279-88.

87. **Nagy, P. D., J. Pogany, and A. E. Simon.** 1999. RNA elements required for RNA recombination function as replication enhancers in vitro and in vivo in a plus-strand RNA virus. Embo J **18:**5653-65.

88. **Nagy, P. D., and A. E. Simon.** 1997. New insights into the mechanisms of RNA recombination. Virology **235:**1-9.

89. **Nagy, P. D., and A. E. Simon.** 1998. In vitro characterization of late steps of RNA recombination in turnip crinkle virus. I. Role of motif1-hairpin structure. Virology **249:**379-92.

90. **Nagy, P. D., and A. E. Simon.** 1998. In vitro characterization of late steps of RNA recombination in turnip crinkle virus. II. The role of the priming stem and flanking sequences. Virology **249:**393-405.

91. **Nagy, P. D., C. Zhang, and A. E. Simon.** 1998. Dissecting RNA recombination in vitro: Role of RNA sequences and the viral replicase. Embo J **17:**2392-403.

92. **Ng, K. K., M. M. Cherney, A. L. Vazquez, A. Machin, J. M. Alonso, F. Parra, and M. N. James.** 2002. Crystal structures of active and inactive conformations of a caliciviral RNA-dependent RNA polymerase. J Biol Chem **277:**1381-7.

93. **Noueiry, A. O., and P. Ahlquist.** 2003. Brome mosaic virus RNA replication: Revealing the role of the host in RNA virus replication. Annu Rev Phytopathol

41:77-98.

94. **O'Reilly, E. K., and C. C. Kao.** 1998. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. Virology **252:**287-303.

95. **O'Reilly, E. K., J. D. Paul, and C. C. Kao.** 1997. Analysis of the interaction of viral RNA replication proteins by using the yeast two-hybrid assay. J Virol **71:**7526-32.

96. **O'Reilly, E. K., N. Tang, P. Ahlquist, and C. C. Kao.** 1995. Biochemical and genetic analyses of the interaction between the helicase-like and polymerase-like proteins of the brome mosaic virus. Virology **214:**59-71.

97. **O'Reilly, E. K., Z. Wang, R. French, and C. C. Kao.** 1998. Interactions between the structural domains of the RNA replication proteins of plant-infecting RNA viruses. J Virol **72:**7160-9.

98. **Osman, T. A., and K. W. Buck.** 1996. Complete replication in vitro of tobacco mosaic virus RNA by a template-dependent, membrane-bound RNA polymerase. J Virol **70:**6227-34.

99. **Osman, T. A., and K. W. Buck.** 2003. Identification of a region of the tobacco mosaic virus 126- and 183-kilodalton replication proteins which binds specifically to the viral 3'-terminal t RNA-like structure. J Virol **77:**8669-75.

100. **Osman, T. A., R. J. Hayes, and K. W. Buck.** 1992. Cooperative binding of the red clover necrotic mosaic virus movement protein to single-stranded nucleic acids. J Gen Virol **73 (Pt 2):**223-7.

101. **Osman, T. A., C. L. Hemenway, and K. W. Buck.** 2000. Role of the 3' t RNA-like structure in tobacco mosaic virus minus-strand RNA synthesis by the viral RNA-dependent RNA polymerase in vitro. J Virol **74:**11671-80.

102. **Osman, T. A., P. Thommes, and K. W. Buck.** 1993. Localization of a single-stranded RNA-binding domain in the movement protein of red clover necrotic mosaic dianthovirus. J Gen Virol **74 (Pt 11):**2453-7.

103. **Oster, S. K., B. Wu, and K. A. White.** 1998. Uncoupled expression of p33 and p92 permits amplification of tomato bushy stunt virus rnas. J Virol **72:**5845-51.

104. **Panavas, T., and P. D. Nagy.** 2003. Yeast as a model host to study replication

and recombination of defective interfering RNA of tomato bushy stunt virus. Virology **314:**315-25.

105. **Panavas, T., and P. D. Nagy.** 2003. The RNA replication enhancer element of tombusviruses contains two interchangeable hairpins that are functional during plus-strand synthesis. J Virol **77:**258-69.

106. **Panavas, T., Z. Panaviene, J. Pogany, and P. D. Nagy.** 2003. Enhancement of RNA synthesis by promoter duplication in tombusviruses. Virology **310:**118-29.

107. **Panavas, T., J. Pogany, and P. D. Nagy.** 2002. Internal initiation by the cucumber necrosis virus RNA-dependent RNA polymerase is facilitated by promoter-like sequences. Virology **296:**275-87.

108. **Panavas, T., J. Pogany, and P. D. Nagy.** 2002. Analysis of minimal promoter sequences for plus-strand synthesis by the cucumber necrosis virus RNA-dependent RNA polymerase. Virology **296:**263-74.

109. **Panaviene, Z., J. M. Baker, and P. D. Nagy.** 2003. The overlapping RNA-binding domains of p33 and p92 replicase proteins are essential for tombusvirus replication. Virology **308:**191-205.

110. **Panaviene, Z., and P. D. Nagy.** 2003. Mutations in the RNA-binding domains of tombusvirus replicase proteins affect RNA recombination in vivo. Virology **317:** 359-72.

111. **Pantaleo, V., L. Rubino, and M. Russo.** 2003. Replication of carnation Italian ringspot virus defective interfering rna in saccharomyces cerevisiae. J Virol **77:** 2116-23.

112. **Pata, J. D., S. C. Schultz, and K. Kirkegaard.** 1995. Functional oligomerization of poliovirus RNA-dependent RNA polymerase. RNA **1:**466-77.

113. **Plante, C. A., K. H. Kim, N. Pillai-Nair, T. A. Osman, K. W. Buck, and C. L. Hemenway.** 2000. Soluble, template-dependent extracts from Nicotiana benthamiana plants infected with potato virus X transcribe both plus- and minus-strand rna templates. Virology **275:**444-51.

114. **Poch, O., I. Sauvaget, M. Delarue, and N. Tordo.** 1989. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. Embo J **8:**3867-74.

115. **Pogany, J., M. R. Fabian, K. A. White, and P. D. Nagy.** 2003. A replication silencer element in a plus-strand RNA virus. Embo J **22:**5602-11.

116. **Puglisi, J. D., R. Tan, B. J. Calnan, A. D. Frankel, and J. R. Williamson.** 1992. Conformation of the tar RNA-arginine complex by NMR spectroscopy. Science **257:**76-80.

117. **Qin, W., H. Luo, T. Nomura, N. Hayashi, T. Yamashita, and S. Murakami.** 2002. Oligomeric interaction of hepatitis C virus NS5b is critical for catalytic activity of RNA-dependent RNA polymerase. J Biol Chem **277:**2132-7.

118. **Quadt, R., and E. M. Jaspars.** 1990. Purification and characterization of brome mosaic virus RNA-dependent RNA polymerase. Virology **178:**189-94.

119. **Quadt, R., C. C. Kao, K. S. Browning, R. P. Hershberger, and P. Ahlquist.** 1993. Characterization of a host protein associated with brome mosaic virus RNA-dependent RNA polymerase. Proc Natl Acad Sci U S A **90:**1498-502.

120. **Quadt, R., H. J. Rosdorff, T. W. Hunt, and E. M. Jaspars.** 1991. Analysis of the protein composition of alfalfa mosaic virus RNA-dependent RNA polymerase. Virology **182:**309-15.

121. **Racaniello, V. R., and R. Ren.** 1996. Poliovirus biology and pathogenesis. Curr Top Microbiol Immunol **206:**305-25.

122. **Rajendran, K. S., and P. D. Nagy.** 2003. Characterization of the RNA-binding domains in the replicase proteins of tomato bushy stunt virus. J Virol **77:**9244-58.

122a. **Rajendran, K. S., and P. D. Nagy.** 2004. Interactions between the replicase proteins of tomato bushy stunt virus in vitro and in vivo. Virology. In Press.

123. **Rajendran, K. S., J. Pogany, and P. D. Nagy.** 2002. Comparison of turnip crinkle virus RNA-dependent RNA polymerase preparations expressed in *Escherichia coli* or derived from infected plants. J Virol **76:**1707-17.

124. **Rao, A. L., and G. L. Grantham.** 1996. Molecular studies on bromovirus capsid protein. II. Functional analysis of the amino-terminal arginine-rich motif and its role in encapsidation, movement, and pathology. Virology **226:**294-305.

125. **Ray, D., and K. A. White.** 2003. An internally located RNA hairpin enhances replication of tomato bushy stunt virus rnas. J Virol **77:**245-57.

126. **Richmond, K. E., K. Chenault, J. L. Sherwood, and T. L. German.** 1998.

Characterization of the nucleic acid binding properties of tomato spotted wilt virus nucleocapsid protein. Virology **248:**6-11.

127. **Rozanov, M. N., E. V. Koonin, and A. E. Gorbalenya.** 1992. Conservation of the putative methyltransferase domain: A hallmark of the 'Sindbis-like' super-group of positive-strand RNA viruses. J Gen Virol **73 (Pt 8):**2129-34.

128. **Rubino, L., and M. Russo.** 1998. Membrane targeting sequences in tombusvirus infections. Virology **252:**431-7.

129. **Rubino, L., F. Weber-Lotfi, A. Dietrich, C. Stussi-Garaud, and M. Russo.** 2001. The open reading frame 1-encoded ('36k') protein of carnation Italian ringspot virus localizes to mitochondria. J Gen Virol **82:**29-34.

130. **Russo, M., J. Burgyan, and G. P. Martelli.** 1994. Molecular biology of tombusviridae. Adv Virus Res **44:**381-428.

131. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor, New York.

132. **Schaad, M. C., P. E. Jensen, and J. C. Carrington.** 1997. Formation of plant RNA virus replication complexes on membranes: Role of an endoplasmic reticulum-targeted viral protein. Embo J **16:**4049-59.

133. **Scholthof, K. B., H. B. Scholthof, and A. O. Jackson.** 1995. The tomato bushy stunt virus replicase proteins are coordinately expressed and membrane associated. Virology **208:**365-9.

134. **Simon, A., Engel H, Johnson RP, Howell SH.** 1988. Identification of regions affecting virulence, RNA processing and infectivity in the virulent satellite of turnip crinkle virus. Embo J **7:**2645-2651.

135. **Simon, A. E.** 1999. Replication, recombination, and symptom-modulation properties of the satellite RNAs of turnip crinkle virus. Current Topics in Microbiology and Immunology **239:**19-36.

136. **Singh, R. N., and T. W. Dreher.** 1997. Turnip yellow mosaic virus RNA-dependent RNA polymerase: Initiation of minus strand synthesis in vitro. Virology **233:**430-9.

137. **Smith, C. A., V. Calabro, and A. D. Frankel.** 2000. An RNA-binding chameleon. Mol Cell **6:**1067-76.

138. **Soldevila, A. I., W. M. Havens, and S. A. Ghabrial.** 2000. A cellular protein with an RNA-binding activity co-purifies with viral ds RNA from mycovirus-infected helminthosporium victoriae. Virology **272:**183-90.

139. **Song, C., and A. E. Simon.** 1994. RNA-dependent RNA polymerase from plants infected with turnip crinkle virus can transcribe (+)- and (-)-strands of virus-associated RNAs. Proc Natl Acad Sci U S A **91:**8792-6.

140. **Song, C., and A. E. Simon.** 1995. Synthesis of novel products in vitro by an RNA-dependent RNA polymerase. J Virol **69:**4020-8.

141. **Song, C., and A. E. Simon.** 1995. Requirement of a 3'-terminal stem-loop in in vitro transcription by an RNA-dependent RNA polymerase. J Mol Biol **254:**6-14.

142. **Spiegelman, S., N. R. Pace, D. R. Mills, R. Levisohn, T. S. Eikhom, M. M. Taylor, R. L. Peterson, and D. H. Bishop.** 1968. The mechanism of RNA replication. Cold Spring Harb Symp Quant Biol **33:**101-24.

143. **Stagljar, I., and S. Fields.** 2002. Analysis of membrane protein interactions using yeast-based technologies. Trends Biochem Sci **27:**559-63.

144. **Suzuki, M., M. Yoshida, T. Yoshinuma, and T. Hibi.** 2003. Interaction of replicase components between cucumber mosaic virus and peanut stunt virus. J Gen Virol **84:**1931-9.

145. **Tao, Y., D. L. Farsetta, M. L. Nibert, and S. C. Harrison.**2002. RNA synthesis in a cage--structural studies of reovirus polymerase lambda3. Cell **111:**733-45.

146. **Tellinghuisen, T. L., and C. M. Rice.** 2002. Interaction between hepatitis C virus proteins and host cell factors. Curr Opin Microbiol **5:**419-27.

147. **Tsai, M. S., Y. H. Hsu, and N. S. Lin.** 1999. Bamboo mosaic potexvirus satellite RNA (satBamV RNA)-encoded p20 protein preferentially binds to satBamV RNA. J Virol **73:**3032-9.

148. **Van Aelst, L., M. Barr, S. Marcus, A. Polverino, and M. Wigler.** 1993. Complex formation between Ras and Raf and other protein kinases. Proc Natl Acad Sci USA **90:**6213-7.

149. **Vance, V., and H. Vaucheret.** 2001. RNA silencing in plants--defense and counterdefense. Science **292:**2277-80.

150. **Wang, Q. M., M. A. Hockman, K. Staschke, R. B. Johnson, K. A. Case, J. Lu,**

S. Parsons, F. Zhang, R. Rathnachalam, K. Kirkegaard, and J. M. Colacino. 2002. Oligomerization and cooperative RNA synthesis activity of hepatitis C virus RNA-dependent RNA polymerase. J Virol **76:**3865-72.

151. **Watanabe, T., A. Honda, A. Iwata, S. Ueda, T. Hibi, and A. Ishihama.** 1999. Isolation from tobacco mosaic virus-infected tobacco of a solubilized template-specific RNA-dependent RNA polymerase containing a 126k/183k protein heterodimer. J Virol **73:**2633-40.
152. **Waterhouse, P. M., M. B. Wang, and T. Lough.** 2001. Gene silencing as an adaptive defence against viruses. Nature **411:**834-42.
153. **Weber-Lotfi, F., A. Dietrich, M. Russo, and L. Rubino.** 2002. Mitochondrial targeting and membrane anchoring of a viral replicase in plant and yeast cells. J Virol **76:**10485-96.
154. **White, K. A.** 1996. Formation and evolution of *tombusvirus* defective interfering RNAs. Seminars in Virology **7:**409-416.
155. **White, K. A., and T. J. Morris.** 1994. Recombination between defective tombusvirus RNAs generates functional hybrid genomes. Proc Natl Acad Sci USA **91:**3642-6.
156. **White, K. A., and T. J. Morris.** 1999. Defective and defective interfering RNAs of monopartite plus-strand RNA plant viruses. Curr Top Microbiol Immunol **239:** 1-17.
157. **White, K. A., J. M. Skuzeski, W. Li, N. Wei, and T. J. Morris.** 1995. Immunodetection, expression strategy and complementation of turnip crinkle virus p28 and p88 replication components. Virology **211:**525-34.
158. **Yoshinari, S., P. D. Nagy, A. E. Simon, and T. W. Dreher.** 2000. CCA initiation boxes without unique promoter elements support in vitro transcription by three viral RNA-dependent RNA polymerases. RNA **6:**698-707.
159. **Zaccomer, B., A. L. Haenni, and G. Macaya.** 1995. The remarkable variety of plant RNA virus genomes. J Gen Virol **76 (Pt 2):**231-47.
160. **Zhong, W., A. S. Uss, E. Ferrari, J. Y. Lau, and Z. Hong.** 2000. De novo initiation of RNA synthesis by hepatitis C virus nonstructural protein 5b polymerase. J Virol **74:**2017-22.

Vita

KS. Rajendran was born in Kottampatty, a small hamlet in South Indian state of TamilNadu on January 8th, 1973. He graduated with B. Sc from TamilNadu Agricultural University, Madurai, India in 1993 and moved to New Delhi for his Masters Degree in the department of Plant Pathology at Indian Agricultural Research Institute from 1994-97. For M. Sc dissertation, he studied the viroid diseases of *Coleus blumei,* a popular foliage ornamental plant. He indexed the existing Coleus cultivars available in Delhi and the surrounding regions, studied their transmission and host range properties using nucleic acid spot hybridization, reverse-polyacrylamide gel electrophoresis and northern blot techniques. At the completion of M. Sc, he went on to work as a Scientist for two years (1997-99 at Indian Institute of Spices Research, Kozhikode, India. He investigated the etiology of a black pepper disease that showed symptoms similar to viral infections. He then accepted the research assistantship from the Fall, 1999 at the Dept. of Plant Pathology, University of Kentucky for his doctoral program.