IDENTIFICATION OF VIRAL AND HOST FACTORS INVOLVED IN TOMBUSVIRUS REPLICATION AND RECOMBINATION

Natalia Shapka
University of Kentucky
IDENTIFICATION OF VIRAL AND HOST FACTORS INVOLVED IN TOMBUSVIRUS REPLICATION AND RECOMBINATION

___________________________________________

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
Natalia Shapka
Lexington, Kentucky

Director: Dr. Peter D. Nagy, Associate Professor of Plant Pathology
Lexington, Kentucky
2006
ABSTRACT OF DISSERTATION
IDENTIFICATION OF VIRAL AND HOST FACTORS INVOLVED IN TOMBUSVIRUS REPLICATION AND RECOMBINATION

Rapid evolution of RNA viruses with mRNA-sense genomes is a major concern to health and economic welfare due to the devastating diseases these viruses inflict on humans, animals and plants. Rapid viral RNA evolution is frequently due to RNA recombination, which can be facilitated by recombination signals present in viral RNAs. Among such signals are short sequences with high AU contents that constitute recombination hot spots in Brome mosaic virus (BMV) and retroviruses. We have demonstrated that a defective interfering (DI) RNA, a model template associated with Tomato bushy stunt virus (TBSV), a tombusvirus, undergoes frequent recombination in plants and protoplast cells when it carries the AU-rich hot spot sequence from BMV. Similar to the situation with BMV, most of the recombination junction sites in the DI RNA recombinants were found within the AU-rich region. Our results support the idea that common AU-rich recombination signals might promote interviral recombination between unrelated viruses.

To test if host genes can affect the evolution of RNA viruses, we used a Saccharomyces cerevisiae single-gene deletion library, which includes ~80% of yeast genes, in RNA recombination studies based on a small viral replicon RNA derived from TBSV. The genome-wide screen led to the identification of five host genes, whose absence resulted in rapid generation of novel viral RNA recombinants. Thus, these genes normally suppress viral RNA recombination, but in their absence hosts become viral
recombination “hotbeds”. Four of the five recombination suppressor genes are likely involved in RNA degradation, suggesting that RNA degradation could play a role in viral RNA recombination. Overall, our results demonstrate for the first time that a set of host genes have major effect on RNA virus recombination and evolution.

Replication of the non-segmented, plus-stranded RNA genome of *Cucumber necrosis tombusvirus* (CNV) requires two essential overlapping viral-coded replication proteins, the p33 replication co-factor and the p92 RNA-dependent RNA polymerase. We have demonstrated that p33 is phosphorylated in vivo and in vitro by a membrane-bound plant kinase. Based on in vitro studies with purified recombinant p33, we show evidence for phosphorylation of threonine and serine residues adjacent to the essential RNA-binding site in p33. Our findings suggest that phosphorylation of threonine/serine residues adjacent to the essential RNA-binding site in the auxiliary p33 protein likely plays a role in viral RNA replication and subgenomic RNA synthesis during tombusvirus infections.

Key words: *Tomato bushy stunt virus*, RNA-dependent RNA polymerase, Defective interfering RNA, Yeast, Recombination
IDENTIFICATION OF VIRAL AND HOST FACTORS INVOLVED IN TOMBUSVIRUS REPLICATION AND RECOMBINATION

By

Natalia Shapka

__________________________
Director of Dissertation

__________________________
Director of Graduate Studies

Copyright © Natalia Shapka 2006.
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.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name

Date
IDENTIFICATION OF VIRAL AND HOST FACTORS INVOLVED IN TOMBUSVIRUS REPLICATION AND RECOMBINATION

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

Natalia Shapka

Lexington, Kentucky

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

Lexington, Kentucky

2006

Copyright © Natalia Shapka 2006.
# TABLE OF CONTENTS

List of Tables ............................................................................................................................ii
List of Figures ............................................................................................................................iii
List of Files ...............................................................................................................................v

Chapter I: Introduction ...........................................................................................................1
- RNA recombination in model virus systems .................................................................2
- Role of host factors in RNA recombination .................................................................2
- Tombusviruses are ideal model plus-stranded RNA viruses .................................3
- What the thesis will show ...............................................................................................5

Chapter II: The AU-rich RNA recombination hot spot sequence of Brome mosaic virus is functional in tombusviruses: Implications for the mechanism of RNA recombination
- Introduction .......................................................................................................................7
- Materials and Methods ..................................................................................................10
- Results ...............................................................................................................................14
- Discussion ........................................................................................................................25

Chapter III: Genome-wide screen identifies host genes affecting viral RNA recombination
- Introduction .......................................................................................................................49
- Materials and Methods ..................................................................................................51
- Results ...............................................................................................................................54
- Discussion ........................................................................................................................60

Chapter IV: Phosphorylation of the p33 replication protein of Cucumber necrosis tombusvirus adjacent to the RNA binding site affects viral RNA replication
- Introduction .......................................................................................................................77
- Materials and Methods ..................................................................................................80
- Results ...............................................................................................................................85
- Discussion ........................................................................................................................95

Chapter V: Summary .............................................................................................................113
- References .........................................................................................................................116
- Vita ..................................................................................................................................132
LIST OF TABLES

TABLE 2.1. Primers used for PCR.........................................................47
TABLE 3.1. Names and functions of the identified host genes.............75
LIST OF FIGURES

Fig. 2.1. The AU-rich hot spot sequence of BMV promotes recombination in DI RNA of TBSV .................................................................33
Fig. 2.2. Sequences of DI RNA recombinants around the junction sites.......................36
Fig. 2.3. The role of the RII sequence in DI RNA recombination.............................37
Fig. 2.4. Sequences of DI RNA recombinants obtained with RII deletion mutants.........38
Fig. 2.5. The effect of short AU-rich and GC-rich sequences on DI RNA recombination...40
Figure 2.6. Efficient binding of RII(-) to the recombinant p33 and p92 replicase proteins of TBSV in vitro .........................................................................................42
Fig. 2.7. Comparison of the level of primer extension obtained with RII(-) and GFP-derived sequences ......................................................................................44
Fig. 2.8. Models of AU-rich sequence-promoted RNA recombination in tombusviruses ....46
Fig. 3.1. Absence of CTL1, MET22/HAL2, HUR1, XRN1 and UBP3 host genes leads to enhanced recombination of TBSV DI-72 RNA replicon in yeast .........................65
Fig. 3.2. Schematic presentation of the DI-72 replicon with four regions and the recombinants with duplicated 3’ sequences (3’ part of RII, RIII and RIV) and 5’ deletions (RI and 5’ part of RII) .................................................................................66
Fig. 3.3. (A) Time-course experiment with hur1Δ and xrn1Δ co-expressing p33 and p92 reveals rapid generation of recombinants (probed with RIII(-)) after induction of DI-72 RNA transcription from plasmid pYC/DI-72 .........................................................67
Fig. 3.4. Deletion of Xrn1p increases the stability of recombinant RNAs and DI-72 replicon RNA .................................................................................................................68
Fig. 3.5. (A) Absence of PEP7, IPK1, CHO2 and DCI1 host genes leads to low frequency of recombination of TBSV DI-72 RNA replicon in yeast co-expressing p33 and p92 replicase proteins ........................................................................................................69
Fig. 3.6. Absence of SPE3 and SPT3 genes results in altered recombination profile with (A) DI-72 RNA replicon and (B) DI-AU-FP replicon..............................................................................................................71
Fig. 3.7. Comparison of recombination activity of DI-AU-FP replicon in xrn1Δ, pep7Δ, and parental yeast strains ........................................................................................................................72
Fig. 3.8. Schematic presentation of the DI-72 replicon-derived recombinants generated in hur1Δ strain expressing the DI-72 RNA replicon .................................................................................73
Fig. 3.9. Absence of viral recombinants in yeast DNA and RNA transcripts .................74
Fig. 4.1. Location of the predicted phosphorylation sites in the CNV replication proteins..100
Fig. 4.2. Detection of in vivo phosphorylated replication proteins of CNV and TCV .........101
Fig. 4.3. In vitro phosphorylation of p33 by a plant kinase ........................................102
Fig. 4.4. In vitro phosphorylation at S210 and T211 residues of p33 by PKC ..................103
Fig. 4.5. Comparison of accumulation of CNV gRNA carrying phosphorylation-mimicking mutations adjacent to the RPR-motif in p33/p92 gene .........................................................104
Fig. 4.6. Effect of phosphorylation-mimicking mutations on accumulation of (A) CNV (+) RNAs and (B) p33 replication protein ..............................................................................................................106
Fig. 4.7. Effect of phosphorylation-mimicking mutations on accumulation of DI-72 RNA in trans .................................................................................................................................107
Fig. 4.8. Phosphorylation-mimicking mutations affect the function of p33, and a lesser extent, p92 ......................................................................................................................109
Fig. 4.9. Effect of phosphorylation-mimicking mutations on DI RNA replication in yeast 110
Fig. 4.10. Nonphosphorylation-mimicking mutant of CNV shows delay in gRNA accumulation and symptom formation in *N. benthamiana* plants……………………111
LIST OF FILES

Shapka-Ph.D-dissertation.pdf
Chapter I

INTRODUCTION

Due to high frequency mutations and RNA recombination, RNA viruses can change their genomes frequently to develop new strains or viruses (Aaziz and Tepfer, 1999; Lai, 1992; Nagy and Simon, 1997; Worobey and Holmes, 1999). Therefore, it is not surprising that RNA viruses are widespread in nature and that they cause many diseases of humans, animals and plants. RNA recombination is especially powerful tool for viruses, because it can rapidly lead to dramatic changes in virus genomes by recombining or rearranging “battle-tested” (i.e., evolutionarily successful) sequences. Accordingly, the significant role of RNA recombination in emergence of new viruses or virus strains is well documented for numerous human, plant, animal, bacterial, insect and fungal viruses (Aaziz and Tepfer, 1999; Lai, 1992; Nagy and Simon, 1997; Worobey and Holmes, 1999). A recombinant virus may “jump species”. RNA recombination can also occur between viral and host sequences, thus leading to the emergence of recombinant viruses carrying novel host genes or gaining new functions. The best example is the incorporation of ubiquitin gene into the pestivirus genome that intensifies the disease symptoms and often leads to the death of the animal (Becher, Orlich, and Thiel, 2001). The increased pathogenicity of an influenza A virus hybrid was possibly due to recombination with a cellular RNA (Khatchikian, Orlich, and Rott, 1989). Understanding of the mechanism of RNA recombination is expected to help the development of safer vaccine strains of human viruses and more effective viral-based gene-delivery vectors in plants and animals.
RNA recombination in model virus systems

The major challenge in studying RNA recombination is that recombination is a chance event. In contrast to replication, RNA recombination probably does not need to occur in each virus-infected cell. Due to the complex nature of RNA recombination, experiments require careful design to bring together the necessary components of RNA recombination. Therefore, progress in this research area greatly benefits from studies with model viruses. Accordingly, studies on poliovirus, Brome mosaic virus (BMV), Carmoviruses, and Tombusviruses (Cascone, Haydar, and Simon, 1993; Jarvis and Kirkegaard, 1991; Kim and Kao, 2001; Lai, 1992; Nagy and Bujarski, 1993; Nagy and Simon, 1997; Pilipenko, Gmyl, and Agol, 1995; White and Morris, 1994c; Worobey and Holmes, 1999) have contributed greatly to our understanding of RNA recombination in general. Based on current models, the most frequent RNA recombination events are driven by the viral replicase, which is proposed to “jump” from one site or from one RNA molecule to another during RNA synthesis (Jarvis and Kirkegaard, 1991; Nagy and Simon, 1997). Accordingly, the viral replicase-driven template-switching mechanism has been demonstrated for Tombus- and Carmoviruses, and for BMV in vitro (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002; Kim and Kao, 2001; Nagy, Zhang, and Simon, 1998). In spite of the above studies, our understanding of RNA recombination is incomplete and it is not yet known whether host factors are involved in viral RNA recombination.

Role of host factors in RNA recombination
Unfortunately, there is no published information on direct roles of host genes in RNA recombination. Yet, indirect observations, such as the variable frequency of viral RNA recombination in different hosts, suggest that the host influences viral recombination. Thus, it is feasible to assume that host genes have significant roles in RNA recombination. Indeed, viruses rely very much on their hosts for enzymes, metabolites, energy sources and membrane surfaces for replication and other processes. For example, replication of RNA viruses, which is performed by the viral replicase, requires not only viral-coded replicase proteins, but host proteins, too. Accordingly, the role of host proteins in RNA replication has been demonstrated for an increasing number of RNA viruses (Ahlquist et al., 2003; Nagy and Pogany, 2006). In spite of the intensive studies on viral replication, most of the host factors are still unidentified and uncharacterized, whereas the role of host factors in viral RNA recombination is completely unknown.

**Tombusviruses are ideal model plus-stranded RNA viruses**

Tombusviruses, such as *Tomato bushy stunt virus (TBSV)* and *Cucumber necrosis virus (CNV)*, are important and emerging plant pathogens that are among the best-characterized viruses (Nagy and Pogany, 2006; Panavas et al., 2005d; Rajendran and Nagy, 2006; Serviene et al., 2006; Serviene et al., 2005). They are spherical viruses with monopartite (+)RNA genomes of ~4.8 kb (Russo, Burgyan, and Martelli, 1994; White and Nagy, 2004). They belong to supergroup 2 viruses that include important human and animal pathogens (e.g., HCV, flaviviruses, pestiviruses), and plant pathogens (luteoviruses, carmoviruses, and others). Tombusviruses code for five proteins including two replication proteins, termed p33
and p92 (Fig. 1.1), which are essential for replication (Oster, Wu, and White, 1998a; Panaviene, Baker, and Nagy, 2003; Scholthof, Scholthof, and Jackson, 1995b) and recombination (Panaviene and Nagy, 2003). Both p33 and p92 are translated from the genomic RNA and p92 is the result of translational readthrough of the p33 stop codon (Fig. 1.1) (Scholthof, Scholthof, and Jackson, 1995b; White and Nagy, 2004). Therefore, the N-terminal portion of p92 overlaps with p33. p92 is the RdRp (Panaviene, Panavas, and Nagy, 2005; Panaviene et al., 2004), while p33 plays a role in RNA template selection/recruitment and in the assembly of the viral RC (Monkewish et al., 2005; Panavas et al., 2005a; Pogany, White, and Nagy, 2005).

Due to recent major advances, tombusviruses are among the most advanced model viruses (Nagy and Pogany, 2006; Panavas et al., 2005d; Rajendran and Nagy, 2006; Serviene et al., 2006; Serviene et al., 2005; White and Nagy, 2004). Powerful in vitro assays based on highly-purified CNV (Panaviene, Panavas, and Nagy, 2005; Panaviene et al., 2004) and TBSV replicases are available (Nagy and Pogany, 2000a). In addition, tombusvirus replication can be studied in single plant cells (protoplasts), in whole plants and in yeast, an excellent model host (Panavas and Nagy, 2003b; Pantaleo, Rubino, and Russo, 2003).

Tombusviruses are frequently associated with defective interfering (DI) RNAs that are derived entirely from the genomic (g)RNA (Hillman, Carrington, and Morris, 1987). The most frequently occurring DI RNAs (~400-800 nt) contain four short noncontiguous segments of the gRNA without coding for functional genes (Fig. 1.1) (Hillman, Carrington, and Morris, 1987; Law and Morris, 1994; White and Morris, 1999). DI RNAs are excellent model templates for replication and they are used frequently in in vivo (Park, Desvoyes, and Scholthof, 2002; Qiu et al., 2001; Ray and White, 1999; Ray and White, 2003; Ray, Wu, and
White, 2003; White and Morris, 1999) with helper viruses and in in vitro studies (Panavas and Nagy, 2003a; Panavas et al., 2003; Pogany et al., 2003).

Altogether, the available in vitro replication assay based on purified tombusvirus RdRp (Nagy and Pogany, 2000a) and development of yeast for efficient replication of a tombusvirus replicon (Panavas and Nagy, 2003b; Pantaleo, Rubino, and Russo, 2003) makes tombusviruses ideal to study the RNA and protein factors involved in viral RNA replication.

**What the thesis will show**

The objectives of this work were to study RNA and protein, both viral and host origin, factors that affect viral RNA recombination and replication using tombusviruses as model viruses.

Chapter 2 will provide information on the roles of RNA sequences, namely AU-rich sequences, in promoting RNA recombination. Chapter 3 describes a high throughput screen performed in yeast to identify host factors affecting virus recombination. This work has led to the identification of host proteins that could suppress viral RNA recombination for the first time. Chapter 4 focuses on the posttranslational modification of the viral replicase proteins and the role of this modification in virus replication. At the end of the thesis, I will discuss the impact of these studies on revealing the roles of RNA elements, viral replication proteins and host factors in tombusvirus replication and recombination.
**Fig. 1.1. Genome organization of TBSV gRNA and a prototypical DI RNA.** p92 is the RdRp protein and it is translated from the genomic RNA via readthrough of the translational stop codon in p33. The second replicase protein, p33 is also required for replication. The prototypical DI RNA contains four non-contiguous segments (called RI to RIV) from the genomic RNA, which are represented with dotted lines. The closely related CNV has the same genome organization.
Chapter II

The AU-rich RNA recombination hot spot sequence of *Brome mosaic virus* is functional in tombusviruses: Implications for the mechanism of RNA recombination

INTRODUCTION

RNA recombination plays a major role during the evolution of plus-strand RNA viruses (Aranda et al., 1997; Fernandez-Cuartero et al., 1994; Strauss and Strauss, 1988; Worobey and Holmes, 1999). New viruses or strains may emerge via recombination between different viral RNAs or between viral and host RNAs (Strauss and Strauss, 1988; Wierzchoslawski et al., 2003; Worobey and Holmes, 1999). In addition to contributing to the genetic variability, RNA recombination is also proposed to function as a repair mechanism, which can salvage damaged or mutated viral RNAs and use them to generate infectious viral RNAs (Nagy and Simon, 1997). RNA recombination also plays a role in the formation of defective interfering (DI) RNAs associated with many animal and plant viruses (White and Morris, 1999). The first hallmark feature of DI RNAs is that they are derived from the parent (helper) virus via sequence deletion(s). The second is that DI RNAs are dependent on the helper virus for their replication, survival and/or spread. The best-known DI RNAs among plant viruses are those associated with tombusvirus infections (White and Morris, 1999).
The most popular model of RNA recombination is the template switching (copy choice) mechanism, which suggests that the viral RNA-dependent RNA polymerase (RdRp) switches templates during complementary RNA synthesis (Jarvis and Kirkegaard, 1991; Lai, 1992; Nagy and Simon, 1997). After the jump from the donor to the acceptor RNA, the RdRp resumes RNA synthesis using the nascent RNA (which has been made on the donor RNA) as a primer. Based on the role of base-pairing between the acceptor RNA and the primer (nascent RNA), RNA recombination events are divided into three categories: base-pairing dependent (similarity-essential), base-pairing assisted or base-pairing independent (similarity-nonessential) (Jarvis and Kirkegaard, 1991; Lai, 1992; Nagy and Simon, 1997).

RNA recombination is probably a chance event, thus each nucleotide in an RNA molecule may serve as a target for recombination. Most experimental data demonstrate that not all regions within an RNA are equally recombinogenic. The sequences that participate in RNA recombination at higher and lower frequencies are called hot and cold spots. Various models have been proposed to explain the occurrence of the observed hot and cold spots for different viruses (Jarvis and Kirkegaard, 1991; Lai, 1992; Nagy and Simon, 1997).

Tombusviruses are single-component plus-stranded RNA viruses of plants and they are known to support RNA recombination at high frequency (Borja et al., 1999; White and Morris, 1994b; White and Morris, 1994d; White and Morris, 1995). Although the involvement of the genomic RNAs in recombination is well documented for tombusviruses, DI RNAs associated with these viruses are the most popular templates in studies on RNA recombination. This is because (i) they are involved in RNA recombination with high frequencies (Borja et al., 1999; White and Morris, 1994b; White and Morris, 1994d; White and Morris, 1995); (ii) they do not contribute essential protein factors to replication; and thus
(iii) have higher genetic plasticity than the viral genomic RNA. Importantly, recombination in DI RNAs is thought to occur by using the same mechanism as recombination involving the viral genomic RNA (White and Morris, 1999). The replication process of tombusviruses and the associated DI RNAs is carried out by the replicase complex, which includes two viral proteins and unknown host factors (Nagy, Pogany, and Simon, 2001; Oster, Wu, and White, 1998b; Scholthof, Scholthof, and Jackson, 1995a). The tombusvirus DI RNAs, such as the Tomato bushy stunt virus (TBSV)-associated DI-73 (Fig. 2.1A), contain three or four noncontiguous genomic segments (White and Morris, 1999). The two or three sequence deletions leading to DI RNA formation are thought to be the consequence of viral replicase jumping on the genomic template and the deletions may occur in a step-wise manner (White and Morris, 1999).

Since it is possible that various RNA viruses may utilize similar RNA recombination pathways or mechanism, we wished to compare how different viruses recognize and use similar recombination promoting signals. This is not only important from mechanistic point of view, but it can also give practical observations about the possibility of interviral recombination, which could be facilitated by recognition of the same recombination promoting signals by different viruses. To initiate these studies, in this paper, we analyzed the effect of a recombination-promoting signal (a short AU-rich sequence), which is well defined for Brome mosaic virus [BMV, ref. (Nagy and Bujarski, 1997)], on tombusvirus recombination. The model recombination template was based on a DI RNA associated with TBSV. We observed that the AU-rich sequence of BMV could indeed promote RNA recombination in the model tombusvirus. The obtained data supports that viruses belonging to different supergroups, such as BMV and TBSV, can recognize the same recombination
promoting signals. However, the distribution of recombination sites was different between BMV and TBSV recombinants. We propose that, in addition to the role of the AU-rich sequence, a putative cis-acting replication element might also affect the selection of the recombination sites in the tombusvirus DI RNA.

MATERIALS AND METHODS

Construction of cDNA clones for TBSV DI RNAs. DI-FP recombination vector was constructed by deleting the N-terminal segment from dEGFP (BDBiosciences) open reading frame (ORF) using NcoI and construct DI-73dEGFP (J. Pogany and Nagy, unpublished), leaving only a 117 bp segment (termed FP, Fig. 2.1A) from dEGFP ORF (Fig. 2.1A). The resulting DI-FP contains, in addition to the truncated FP sequence between RI and RII in DI-73 (White and Morris, 1994b), the following unique restriction sites at the 5’ side of the FP sequence: XbaI, XhoI and NcoI; and at the 3’ side: MluI and BamHI.

All constructs that contained AU, GC2, R’ or AUs sequences (Fig. 2.1A) were obtained by inserting the PCR-amplified (using the primers listed in Table 2.1) AU sequence [including a segment from the 3’ end of BMV RNA3, between positions 197-242, (Nagy and Bujarski, 1997)], GC2 (Nagy and Bujarski, 1998), R’ (Nagy and Bujarski, 1996) or the artificial AUs (Nagy and Bujarski, 1997) segments into DI-FP either at the 5’ side of the FP sequence (at the unique XbaI and NcoI restriction sites) or at the 3’ side (at the unique MluI and BamHI restriction sites). The RII deletion constructs (derivatives of DI-AU1-FP-AU2) were obtained by amplifying of the RII and RIII/IV sequences of DI-FP with the primer sets shown in Table 2.1, followed by treatment with BamHI and Sall, and cloning to the similarly
treated DI-AU\(^1\)-FP-AU\(^2\) or DI-FP (Fig. 2.1A) constructs. All the clones were sequenced to confirm the desired changes.

Construct DI-152Bar* was made by replacing the FP region in DI-FP (Fig. 2.1A) with the PCR-amplified 152 bp C-terminal segment of the barstar ORF (Hartley, 2001) using the unique XbaI and BamHI sites. The PCR primers used are presented in Table 2.1.

**Preparation of CNV gRNA and DI RNA transcripts in vitro.** To generate the genomic (g)RNA of Cucumber necrosis virus (CNV) and the DI RNA transcripts, we linearized pK2/M5p20STOP and all the DI RNA clones, respectively, with Smal, followed by in vitro transcription with T7 RNA polymerase (Nagy, Pogany, and Simon, 1999; Nagy, Pogany, and Simon, 2001). In vitro generated transcripts of DI RNAs were purified from 1% agarose gel, followed by phenol/chloroform extraction, precipitation in 95% ethanol and washed three times with 70% ethanol to remove residual salts. The in vitro RNA transcripts were analyzed in 1% agarose gels and quantified by a UV spectrophotometer (Beckman).

For the in vitro RdRp experiments, RNA templates were obtained by in vitro transcription reaction with T7 RNA polymerase [see above and ref. (Cheng, Pogany, and Nagy, 2002; Nagy and Pogany, 2000b)]. The templates and primers used for PCR are listed in Table 2.1. The unincorporated nucleotides were removed by phenol/chloroform extraction and repeated ammonium-acetate/isopropanol precipitation (Cheng, Pogany, and Nagy, 2002; Nagy and Pogany, 2000b). The T7 transcription products were analyzed by 5% denaturing PAGE and the amounts of RNA were measured by UV spectrophotometer.
**Preparation and electroporation of protoplasts.** Nicotiana benthamiana protoplasts were prepared as described before (Panaviene, Baker, and Nagy, 2003). Briefly, N. benthamiana callus was treated with 0.5 g cellulysin and 0.1g macerase (Calbiochem) for 4.5h in protoplast incubation medium (Nagy, Pogany, and Simon, 2001) at 25 °C, and then washed twice with 0.5M mannitol and once with the electroporation buffer (10 mM HEPES, 10 mM NaCl, 120 mM KCl, 4 mM CaCl₂, 200 mM mannitol). For electroporation, we used 5×10⁵ protoplasts and 5 µg of CNV gRNA and 1 µg of DI-RNA. Electroporation was performed with Gene pulser II (Biorad), under the following settings: 0.2kV voltage and 0.5µF capacity. After electroporation, the samples were left on ice for 30min, followed by adding 1.8ml of protoplast culture medium (Nagy, Pogany, and Simon, 2001). Protoplasts were incubated in 35×10 mm petri dishes in the dark for 24-48h at 22 °C.

N. benthamiana plants were inoculated with 3 µg of CNV gRNA and 1 µg of DI-RNA (prepared as described above), using rub-inoculation as described earlier (Nagy, Pogany, and Simon, 2001). Plants were incubated for 10 days in a temperature-controlled room (~22 °C).

**Total RNA extraction from protoplasts and plants and RNA analysis.** Total RNA was extracted from protoplast using phenol/chloroform method (Nagy, Pogany, and Simon, 2001). Aliquots of total RNA were analyzed on 1.2% agarose gels. RNA samples were treated with formamide at 85 °C before loading on the gel. For Northern blot analysis, RNA was transferred to Hybond XL membrane (Amersham-Pharmacia) by electro-transfer and hybridized with DI-72-specific probes [RI(-), ref. (Panaviene, Baker, and Nagy, 2003)]. Hybridization with ³²P-labeled RNA probes was performed in ULTRAhyb hybridization
buffer at 68°C using the recommended conditions (Ambion). The $^{32}$P-labeled RNA probes were made in an in vitro transcription reaction with T7 RNA polymerase in the presence of a $[\alpha-^{32}\text{P}]\text{UTP}$ and a PCR template obtained with primers #15 (5’-GTAATACGACTCAGCTATAGGGCATGTGCTTGTGTTG-3’) and #20 (5’-GGAAATTCTCCAGGATTCTC-3’) (Panaviene, Baker, and Nagy, 2003).

**RT-PCR analysis, cloning and sequencing of recombinant DI RNAs.** To obtain recombinant DI RNA clones, reverse transcriptase (RT) reaction was done using primer #106 (5’-ACCTGGAAGCTTATGCCAGATTTACACTCATC-3’) and 2µl of total RNA (Panaviene, Baker, and Nagy, 2003). This was followed by amplification using PCR with primers #106 and #380 (5’-GGACGAATTCCATAATTATTATCTTTAFTTG-3’).

The RT-PCR products were digested with EcoRI and HindIII followed by gel isolation and ligation into similarly treated pUC19 vector. The clones for sequencing were selected after restriction digestion with EcoRI and HindIII. The sequencing was done using CEQ Cycle Sequencing Kit (Beckman) and primer #553 (5’-GTAAAACGACGGCCAGTG-3’).

**In vitro RNA binding studies.** The RNA probe in the gel shift experiments was RIII(-)/cPR11 (Panavas and Nagy, 2003a) containing the 82 nt minus-stranded RIII replication enhancer and the 11 nt minimal promoter for plus-strand synthesis (23). The labeling was done with $^{32}$P UTP using T7 RNA polymerase (Panavas and Nagy, 2003a). Competitor RNAs (Fig. 2.6) were prepared with T7 polymerase on PCR-amplified templates using primers described in Table 2.1. The gel shift experiments were performed according to Rajendran and Nagy (Rajendran and Nagy, 2003). Briefly, various amounts of competitor
RNAs (used in 5x, 15x and 45x excess over the constant amount of $^{32}$P UTP-labeled RNA probe) were mixed with 1 µM of recombinant p33 and p92 preparations in the presence of 50 mM Tris-HCl, pH 8.2, 10 mM MgCl$_2$, and 10 mM DTT, 10% glycerol, 2.4U RNase inhibitor and 100 ng tRNA. After 10 min pre-incubation at 25$^\circ$C, ~2 ng of $^{32}$P UTP-labeled RNA probe was added to each RNA-binding reactions, followed by further incubation for 25 min. Then, the samples were analyzed by electrophoresis on native 4% polyacrylamide gels run at 200V for 60 min at 4$^\circ$C in Tris-Glycine buffer (25mM Tris, 190 mM Glycine, 5 mM EDTA, pH 8.5). Dried gels were analyzed using a phosphorimager. The recombinant proteins were purified from *E. coli* as described earlier (Rajendran and Nagy, 2003; Rajendran, Pogany, and Nagy, 2002).

**RdRp assay.** The CNV RdRp preparations were obtained from systemically infected *N. benthamiana* leaves as described by Nagy and Pogany (Nagy and Pogany, 2000b). The RdRp reactions contained ~1 µg of RNA transcripts (the amounts of RNA templates were adjusted based on their sizes to have similar molar amounts of templates in each reaction). RdRp reactions were carried out as previously described (Nagy and Pogany, 2000b). The reaction was terminated by adding 5 µl 10% SDS followed by phenol/chloroform extraction and ammonium-acetate/isopropanol precipitation (Nagy and Pogany, 2000b). The RdRp products were analyzed by 5% denaturing PAGE in the presence of 8M urea, followed by Phosphorimager analysis. The data for each sample were normalized based on the number of templated UTP incorporation (Nagy and Pogany, 2000b).

**RESULTS**
**Rationale:** To test if different viruses could recognize the same recombination signal, we chose an AU-rich sequence, which is a well-defined recombination-promoting signal in BMV (Nagy and Bujarski, 1997). This AU-rich sequence was tested in a recombination system based on DI RNA associated with TBSV, a tombusvirus. BMV and TBSV belong to different supergroups of viruses and their RdRps are only distantly related. In addition, tombusviruses do not code for an RNA helicase-like protein, which was shown to participate in RNA recombination in BMV (Nagy et al., 1995).

Previous works defined BMV-derived and artificial AU-rich sequences that served as recombination-promoting signals (Fig. 2.1) during BMV replication in vivo (Nagy and Bujarski, 1996; Nagy and Bujarski, 1997; Nagy and Bujarski, 1998; Nagy, Ogiela, and Bujarski, 1999). Frequent generation of recombinants was only observed when the two recombining RNA templates carried the same or similar AU-rich sequences. Because recombination sites in BMV were located within the AU-rich regions, it was proposed that the BMV replicase recognized the AU-rich signals during the recombination events (Nagy and Bujarski, 1996; Nagy and Bujarski, 1997; Nagy and Bujarski, 1998; Nagy, Ogiela, and Bujarski, 1999). We wanted to test if these AU-rich sequences could also promote RNA recombination in tombusviruses.

**The control TBSV DI RNA-based recombination vector does not support RNA recombination.** To study RNA recombination in tombusviruses, first we developed a recombination vector, which facilitated the construction and testing of additional parental DI RNAs (Fig. 2.1A). The recombination vector, named DI-FP, is based on TBSV DI-73 RNA.
carrying a 117 nt insert of nonviral origin (derived from dEGFP, BDBioscience) between regions I and II (RI and RII; Fig. 2.1A). To avoid possible artifactual RNAs (i.e., unwanted recombinant-like RNAs) made during plasmid propagation in *E. coli* and/or during RNA transcription with T7 polymerase, we gel-isolated DI-FP RNA transcripts prior to their use. In addition, to exclude those DI RNAs that might be generated spontaneously from the helper genomic RNA (gRNA) during the infection, a process known as *de novo* DI RNA formation (Rochon, 1991; White and Morris, 1999), we chose the heterologous *Cucumber necrosis virus* (CNV, closely related to TBSV) as the helper virus. CNV can support efficiently the replication of TBSV DI-73 RNA and DI-FP vector (ref. (Panaviene, Baker, and Nagy, 2003) and Fig. 2.1C).

The gel-isolated DI-FP transcripts were co-electroporated with CNV gRNA transcripts into *N. benthamiana* protoplasts as described earlier (Panaviene, Baker, and Nagy, 2003). We followed the accumulation of DI-FP RNA in protoplast (termed “zero passage”, Fig. 2.1C) by using Northern blotting with a DI RNA-specific probe [RI(-), Fig. 2.1A] and RT-PCR (primers of #160 and #380, Fig. 2.1A). These experiments detected only the parental DI-FP RNAs in the zero passage protoplast samples, suggesting the lack of recombinant accumulation (Fig. 2.1C). The total RNA obtained after 48 hours of incubation in zero passage protoplasts was used for electroporation into a new batch of protoplasts (termed “first passage”). Northern blot and RT-PCR analyses revealed that the DI-FP RNA recombination vector replicated efficiently and stably in the first passage protoplasts (Fig. 2.1D, E and F, lanes 1-4). Note that we made sequential passages of the progeny DI RNAs from one batch of protoplasts to another, since there is no cell-to-cell spread of
tombusviruses in protoplasts. Altogether, we performed three sequential passages, yet the DI-FP vector did not appear to support recombinant DI RNA accumulation (not shown).

An AU-rich sequence supports RNA recombination efficiently in TBSV-associated DI RNA. To test the effect of an AU-rich sequence on tombusvirus recombination, we inserted the 69 nt long AU sequence (Fig. 2.1B; this sequence is identical with AU1), which had been tested previously for recombination in BMV, into DI-FP in such a way that an identical copy of the AU sequence flanked the dEGFP-derived FP region at the 5’ and 3’ sides, respectively (termed AU1 and AU2 in construct DI-AU1-FP-AU2, Fig. 2.1A), resulting in a repeated sequence. It was predicted, based on the BMV results (Nagy and Bujarski, 1997) that recombination might take place between the repeated copies of the AU sequence, thus resulting in deletion of one of the AU copies plus the FP sequence (Fig. 2.1A) located in between the repeated AU sequences.

Incubation of the protoplast cells after co-electroporation of the gel-purified DI-AU1-FP-AU2 RNA with the CNV gRNA resulted in efficient amplification of the parental-sized DI RNA, suggesting that the insertions did not debilitate DI RNA replication (Fig. 2.1C). No recombinants were detected in these protoplasts by Northern blotting, while RT-PCR analysis did show the occurrence of novel, ~100-250 bp shorter than input RNAs (not shown). This indicated that the putative recombinant DI RNAs were present in the zero passage protoplasts at low levels that could only be detected by the more sensitive RT-PCR. However, a passage of the total RNA to a new batch of protoplasts followed by incubation resulted in recombinant-like DI RNAs in 100% (15 out of 15) of the experiments based on total RNA (Fig. 2.1D, lanes 6-9), Northern blot (Fig. 2.1E) and RT-PCR analyses (Fig. 2.1F). Interestingly, the sizes of the novel DI RNA recombinants were variable, suggesting that the
recombination events were “imprecise” in nature. Indeed, cloning and sequencing of these putative recombinants confirmed that (i) they derived from DI-AU₁-FP-AU₂ RNA via deletions; and (ii) the recombination junction sites were different in many of the recombinants (Fig. 2.2A). All of the 5’ deletion sites in the 15 recombinants sequenced were located within the AU sequence (i.e., AU¹), while the 3’ deletion sites were clustered mostly within RII, which flanks the inserted sequences in DI-AU₁-FP-AU² (Fig. 2.2A). Overall, we did not find the generation of precise (homologous) recombinants between the duplicated AU sequences, which were the most common recombinants in the BMV system (Nagy and Bujarski, 1997).

Based on the distribution of the recombination junctions, the AU² sequence did not seem to influence the selection of recombination sites in DI-AU₁-FP-AU². Therefore, we deleted AU² to generate construct DI-AU₁-FP (Fig. 2.1A). Testing the recombination activity of DI-AU₁-FP, in a way similar to that described above for DI-AU₁-FP-AU², revealed that DI-AU₁-FP supported RNA recombination as efficiently as DI-AU₁-FP-AU² did (Fig. 2.1D-F, lanes 11-14). Even more importantly, the distribution of recombination sites for DI-AU₁-FP was comparable to that described for DI-AU₁-FP-AU² (Fig. 2.2C), confirming that the recombination hot spots include the AU¹ sequence and RII. This data supports that one copy of the AU sequence plays an important role in recombination, while the role of the second copy is less obvious.

The isolated DI RNA recombinants from protoplasts appear to be true recombinants, since they can be detected by gel analysis of the total RNA (Fig. 2.1D) and by Northern blotting (Fig. 2.1E). In addition, control RT-PCR performed on gel isolated DI RNA transcripts (the same DI RNA transcripts that were used for electroporation to protoplasts)
did not detect recombinant-sized DI RNAs for these constructs (Fig. 2.1F, lanes 5, 10 and 15).

To test if the length of the inserts is an important factor in recombination events, we generated a DI RNA carrying a ~152 nt long sequence from the barstar gene (Fig. 2.1A) (7). The resulting DI-152Bar* RNA was stable in protoplasts after one (Fig. 2.1D-F, lanes 16-19) or two passages (not shown). This result, together with that of DI-FP, supports that DI RNAs with short inserts can be stable in protoplasts (under the conditions used) if they lack recombination-promoting signals.

To examine if recombination might occur in whole plants as well, we tested the accumulation of DI-FP, DI-AU\textsuperscript{1}-FP-AU\textsuperscript{2} and DI-AU\textsuperscript{1}-FP RNAs (in the presence of the CNV gRNA) in \textit{N. benthamiana} plants ~10 days after inoculation with gel-purified transcripts. Northern blot (Fig. 2.1G) and RT-PCR (not shown) analyses demonstrated that the systemically-infected (uninoculated) leaves contained the parental-sized DI RNAs for each DI RNA tested. In contrast, recombinant-like DI RNAs appeared only in DI-AU\textsuperscript{1}-FP-AU\textsuperscript{2} and DI-AU\textsuperscript{1}-FP RNA-containing plants (Fig. 2.1G, lanes 5-12), but not in DI-FP-containing plants (Fig. 2.1G, lanes 1-4). Cloning and sequencing of a representative number of recombinants confirmed that the accumulating recombinants in plants are similar to those observed in protoplasts with most of the junction sites located within the 5’ AU\textsuperscript{1} sequence and RII (Fig. 2.2B and D).

**RII sequence affects the distribution of recombination sites.** Since many recombinants generated with DI-AU\textsuperscript{1}-FP-AU\textsuperscript{2} and DI-AU\textsuperscript{1}-FP RNAs had the recombination sites within a 40 nt stretch close to the 5’ end of RII(+) (we term this region hs40 hot spot), it is possible
that sequences around hs40 might influence the selection of recombination sites. To test if hs40 is required for replication of DI RNA, we generated a series of DI constructs based on various 5' deletions in RII. First, we tested construct DI-FP-RII/Δ40 (Fig. 2.3A), which has the same sequence as the DI-FP vector (Fig. 2.1A), except lacking the 5’ 40 nt from RII (i.e., hs40). DI-FP-RII/Δ40 RNA accumulated efficiently in protoplasts and no recombinants were detected in protoplasts after the first passage (Fig. 2.3B-C, lanes 1-5). This data suggests that hs40 sequence is not essential for DI RNA accumulation in protoplasts.

To test if hs40 is important for RNA recombination, we tested derivatives of DI-AU1-FP-AU2 with various 5’ deletions in RII. Construct DI-RII/Δ20, which lacked 20 nt from the 5’ end of RII(+) in DI-AU1-FP-AU2, generated recombinants as efficiently as DI-AU1-FP-AU2 did after the first passage (Fig. 2.3B-C, lanes 7-11). Interestingly, the distribution of recombination sites in the recombinants obtained with DI-AU1-FP-AU2 (Fig. 2.2A) and DI-RII/Δ20 (Fig. 2.4A) was similar, except ~50% (7 out of 15) of recombinants obtained with DI-RII/Δ20 contained precise recombination sites between the duplicated AU sequences. The second construct tested, DI-RII/Δ40, which lacked 40 nt from the 5’ end of RII(+) in DI-AU1-FP-AU2, also generated recombinants in 100% of samples (Fig. 2.3B-C, lanes 13-17). Interestingly, 87% (26/30) of the 3’ junctions were within the AU2 copy and only 13% within the RII sequence. We also observed precise recombinants between the repeated AU regions in ~36% (11/30) of recombinants (Fig. 2.4B). The occurrence of precise recombinants may suggest that primer re-alignment (base-pairing between the the primer and the acceptor region) assisted by the presence of duplicated sequences may take place during these recombination events (see Discussion).
Construct DI-RII/Δ60, which lacked 60 nt from the 5’ end of RII(+) in DI-AU₁-FP-AU₂, supported recombination as efficiently as DI-AU₁-FP-AU₂ (Fig. 2.3B-C, lanes 19-23). Cloning and sequencing revealed that all 15 recombinants, each derived from separate samples, were precise recombinants (Fig. 2.4C) with junctions between the duplicated AU sequences. In contrast to the above DI RNA constructs, DI-RII/Δ100, which lacked 100 nt from the 5’ end of RII(+), did not support the generation of recombinants in protoplasts (Fig. 2.3B-C, lanes 25-29). Moreover, DI-RII/Δ100 accumulated only inefficiently (tested after the first passage) in protoplast, suggesting that an important cis-acting signal may have been deleted from this construct (see Discussion). Overall, these results suggest that the RII sequence may have two important effects on RNA recombination: first, it could affect the frequency of recombinant formation and, second, it could also influence the selection of recombination sites. In addition, we conclude that hs40 is not required for recombination to take place within the RII sequence.

**Short AU-rich sequences can also support RNA recombination in DI RNA.** To test if shorter AU-rich sequences could also support RNA recombination in tombusviruses, first we deleted the BMV-derived R’ region (40 nt long, Fig. 2.1B) from the 69 nt AU₁ sequence in DI-RII/Δ40. The resulting construct DI-AUs-FP-AU₂ contained only the 29 nt long artificial AU-rich sequence (termed AUs, with 76% AU content), yet it still supported recombination efficiently with most of the junction sites located within the 5’ AUs and 3’ AU₂ regions (Fig. 2.5A). Note that we found that the distribution of the recombination sites was similar to that observed for DI-RII/Δ40 recombinants, with almost 50% (7/17) of recombinants having precise junctions between the repeated AU-rich sequences (Fig. 2.5A).
Second, we deleted the AUs portion from the 5’ copy of the AU sequence (Fig. 2.1B) in construct DI-R’-FP-AU² (Fig. 2.5B), which left the 44 nt long BMV-derived R’ region at the 5’ location. Although the R’ sequence contained only a 21 nt AU-rich stretch (with ~75% AU content, Fig. 2.1B), it was still active in recombination (Fig. 2.5B). Most of the recombinants obtained with DI-R’-FP-AU² (Fig. 2.5B) contained precise junctions between the repeated AU-rich sequences. Overall, this data suggests that both the BMV-derived R’ and the artificial AUs portions of the original AU sequence are capable of supporting RNA recombination efficiently in TBSV. We conclude that as short as 21 to 29 nt AU-rich sequence can promote recombination in TBSV DI RNA.

**GC-rich sequences cannot “silence” RNA recombination in TBSV.** Since RNA recombination promoted by the AU sequence could efficiently be inhibited by the presence of GC-rich sequences located 3’ of the AU sequence (referred here in plus-strand orientation) in the BMV RNAs (Nagy and Bujarski, 1997), we also tested the effect of one of these well-characterized sequences [termed GC2 with 55% GC-content (Nagy and Bujarski, 1997)] on DI RNA recombination. First, the GC2 sequence was inserted 3’ of the R’ sequence in DI-R’-FP-AU² (Fig. 2.5B). The resulting construct, DI-R’/GC2-FP-AU² (Fig. 2.5C) supported RNA recombination efficiently. Most of the junctions were located within the R’ sequence and the 3’ AU² sequence, including more than 60% (9/15) with precise junctions between the repeated AU-rich sequences (Fig. 2.5C). Second, placing the GC2 sequence behind the R’ sequence within the 3’ repeat (construct DI-AU¹-FP-R’/GC2) did not inhibit recombination that took place between the 5’ AU¹ sequence and either R’, GC2 or the RII sequences (Fig. 2.5C). Interestingly, none of the isolated recombinants contained precise junctions between
the repeated AU-rich sequences, suggesting the GC2 sequence interfered with precise recombination, but not with imprecise recombination.

**The minus-stranded RII binds efficiently to the TBSV RdRp.** Previous in vitro studies with the CNV RdRp (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002) predicted that the same AU sequence tested above is a good donor, while it is a poor acceptor during template switching events. Based on these observations, it is possible that RNA recombination might occur during plus-strand synthesis when the RdRp would “jump” from the AU\(^1\)(-) region to the RII(-) region (around hs40) (see Discussion). Thus, this model predicts that RII(-) should contain a binding site(s) for the tombusvirus replicase proteins. In contrast, the heterologous dEGFP-derived FP sequence is expected to lack high-affinity binding site(s), thus serving as a cold spot during the recombination events. This model was tested below in two different in vitro assays.

First, we have tested the ability of the purified recombinant TBSV replicase proteins (i.e., p33 and p92 expressed and purified from *E. coli*)(Rajendran and Nagy, 2003; Rajendran, Pogany, and Nagy, 2002) to bind to RII(-) sequences or the FP sequence (also tested in the complementary orientation). To obtain quantitative results, we used template competition in a gel mobility shift assay (Rajendran and Nagy, 2003; Rajendran, Pogany, and Nagy, 2002). The experiments included the same amount of \(^{32}\)P-labeled template [derived from the RIII(-) replication enhancer] and the same amount of purified recombinant p92 (Fig. 2.6B) or p33 proteins (Fig. 2.6C). The amounts of unlabeled competitors were used in 5x, 15x and 45x excess over the labeled RNA for each competitor RNA. These experiments revealed that the two competitor RNAs that contained overlapping portions of the RII(-)
sequence (namely construct #7 and #18, Fig. 2.6A) were 2 to 3-fold better competitors for binding to p92 or p33 (Fig. 2.6B-C, lanes 4-9) than the FP control sequence (lanes 1-3). Therefore, this data supports the model that RII(-) is a hot spot region due to its increased binding to the replicase proteins (see Discussion).

Since binding to a template does not necessarily result in productive interaction between the particular sequence and the RdRp, we also tested the ability of the partially purified CNV RdRp to use templates containing various portions of RII(-) sequence and of the FP(-) sequence in the presence of a short primer in an in vitro primer extension assay (Fig. 2.7)(Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002). As shown earlier, primer extension by the CNV RdRp in the in vitro assay depends on (i) the ability of the template region to interact with the CNV RdRp, and (ii) on the number of base-pairs formed between the primer and the template (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002). Since all the constructs tested in this work contained the same 5 base-pair primer/template region (termed art5, which supported self-priming of RNA synthesis on the template by the CNV RdRp)(Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002), we predicted that the activity of a particular template in the CNV RdRp reaction should depend on the ability of the template region to interact with the CNV RdRp. These experiments revealed that primer extension was ~3 to 9-fold more efficient when it started within the RII(-) sequence than the primer extension from the FP sequence (compare constructs #16, #17, and #18 with #20 and #21 in Fig. 2.7A-C), while primer extension from the 3’ end of RII(-) was almost ~2-fold more efficient (construct #11). Interestingly, the most efficient constructs were those that contained the primer within the hs40 hot spot region (see constructs #16 and #17) or in the middle of RII(-) (construct #17, Fig. 2.7), suggesting that
this region contains a putative cis-acting element (see Discussion). Overall, data from the primer extension experiments does support the model that RII(-) might be active in recombination due to its enhanced binding to the RdRp.

DISCUSSION

Comparison of the role of AU-rich sequences in BMV, retrovirus and tombusvirus recombination. The emerging picture in RNA recombination is that the nonrandom distribution of RNA recombination sites observed with several plus-strand RNA viruses [reviewed in (Nagy and Simon, 1997)] are due to the presence of recombination promoting signals in the RNA templates. The viral RdRp is postulated to recognize these recombination signals, which then leads to recombination (template switching) events with high frequencies (i.e., forming recombination hot spots) (Nagy and Simon, 1997). It is currently unknown if recombination signals, which have been described for a particular RNA virus, could also be recognized by another RNA virus. We were particularly interested in AU-rich sequences, which are not only common in many RNA viruses, but they are known to promote recombination in BMV (Nagy and Bujarski, 1995; Nagy and Bujarski, 1996; Nagy and Bujarski, 1997; Nagy and Simon, 1997) and retroviruses (DeStefano, Bambara, and Fay, 1994; Wu et al., 1995) and possibly in poliovirus (Pilipenko, Gmyl, and Agol, 1995). Current models on the role of AU-rich sequences in promoting recombination is based on the assumption that the viral RdRp might pause during RNA synthesis within the AU-rich stretch of the template due to the weak A-U base pairing between the growing nascent strand and the donor RNA template inside the RdRp (Nagy and Bujarski, 1996; Nagy and Bujarski, 1997;
Moreover, the weak base pairing between the nascent strand and the donor RNA within the AU-rich region might also facilitate dissociation of the 3’ end of the nascent strand from the complementary donor strand. The free 3’ end of the nascent strand may then anneal to the acceptor RNA at some frequencies, followed by resumption of the RNA synthesis by the viral RdRp. If the weak base pairing between the growing nascent strand and the donor RNA within the AU-rich region is indeed “forcing” RNA recombination, then we predict that different viral RdRps should recognize this type of signal, albeit with various frequencies depending on the processivity of different viral RdRps.

Indeed, we found that an AU-rich sequence, which promoted RNA recombination in BMV, also facilitated RNA recombination in a TBSV-associated DI RNA. The supporting evidence includes the following: (i) the recombination sites were frequently located within the AU-rich region; (ii) recombination frequency of the DI RNAs containing AU-rich sequences were higher than the control DI RNAs carrying dEGFP-derived or barstar sequences (Fig. 2.1). We also obtained data that support a different role for the AU-rich sequence in tombusvirus recombination than in case of BMV. These include: (i) only one copy of the AU-rich sequence was enough for promoting tombusvirus recombination, while two copies were needed for efficient recombination in BMV; (ii) most of the recombination events occurred precisely or semi-precisely between the two copies of AU-rich sequences in BMV, while this type of recombinants were less frequent in case of TBSV DI RNA; (iii) an additional viral sequence (i.e, RII) was also involved in RNA recombination in TBSV DI RNA, while it is currently unknown if sequences other than the AU-rich regions are involved in facilitating recombination to take place between the AU-rich sequences in BMV. Overall,
the above similarities and differences between the two recombination systems suggest that the AU-rich sequences can promote recombination in both viruses, but the selection of recombination sites (which are probably determined during resolution of the putative recombination intermediates) is different for the majority of recombinants (mostly precise for BMV and largely imprecise for TBSV recombinants). This observation is also supported by the different effect of GC-rich sequences on recombination in the two viral systems. While GC-rich sequences located downstream on the acceptor template (based on the progress of the viral replicase on the template) inhibited or silenced recombination in BMV (Nagy and Bujarski, 1996; Nagy and Bujarski, 1997; Nagy and Bujarski, 1998; Nagy and Simon, 1997), we did not observe similar effects by the same GC-rich sequence in case of TBSV DI RNA (Fig. 2.5C-D).

In addition to the in vivo data discussed above, in vitro data obtained with the purified BMV and CNV (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002; Kim and Kao, 2001) replicases also support the role of AU-rich sequences when present in the donor RNA. For example, the end-to-end template switching assay developed by Kim and Kao (Kim and Kao, 2001) demonstrated that the AU-rich stretches when present at the 5’ end of the donor RNA, where the nascent strand must be released from the donor RNA before the template switching occurs, facilitated end-to-end recombination events. Interestingly, the AU-rich sequence tested in this work also supported template switching by the partially purified CNV RdRp in vitro (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002). The template RNA containing the AU-rich sequence served primarily as a donor RNA in the in vitro CNV RdRp assay (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002), giving valuable insight into the significance of sequence context in RNA recombination in vitro, and possibly in vivo.
Further evidence on the possibility that AU-rich sequences could serve as common recombination signals comes from studies with retroviruses. Using an in vitro template switching assay based on purified reverse transcriptase (RT) from *Human immunodeficiency virus*, it has been demonstrated that AU-rich sequences promoted RNA recombination (DeStefano, Bambara, and Fay, 1994). Since the RT favors base pairing between the nascent strand and the acceptor strand prior to resumption of cDNA synthesis, most of the in vitro recombinants were precise (i.e., occurred precisely within homologous regions), although recombinants with extra nucleotides, mismatched nucleotides, or short deleted regions at the recombination sites were also isolated (DeStefano, Bambara, and Fay, 1994). The observation that AU-rich sequences form hot spots in recombinants obtained with BMV, retrovirus and TBSV suggests that the induction of recombination by AU-rich sequences is similar for these viruses. Moreover, the resolution of recombination intermediates may use somewhat similar mechanism for BMV and retroviruses, but it is different in tombusviruses. We propose that the difference is due to two factors: (i) the BMV RdRp and the retrovirus RT favor a somewhat precise annealing step between the nascent strand (primer) and the acceptor strand prior to the resumption of the RNA/DNA synthesis, while the tombusvirus RdRp does not seem to favor this step. Indeed, we have proposed that the CNV RdRp can easily resume RNA synthesis (primer extension) without the need of extensive base pairing between the primer and the acceptor strand (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002). (ii) There might be differences among these viruses in using *cis*-acting elements to guide the jumping viral replicase to a new acceptor site before resumption of primer extension, resulting in recombination hot spots (see below). It is also important to note that we cannot completely rule out that selection for the best fit recombinants might also
affect the types of recombinants isolated in these virus infections. In summary, in vitro and in vivo data suggests that AU-rich sequences might serve as common recombination signals. However, the likely differences in recombination between viruses are caused by (i) differences in template/sequence recognition during the template switching event; and (ii) various roles for base-pairing between the primer and template during the template switching events.

**Model of AU-rich sequence-driven recombination in tombusviruses.** Since both plus and minus-strand synthesis take place during DI RNA replication in protoplasts, it is difficult to establish whether plus or minus-stranded RNAs are used as templates for recombination. However, based on previous in vitro experiments with the partially-purified CNV RdRp, the likely role of the AU-rich sequence is to promote recombination at the donor sites (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002). Moreover, the same AU sequence used in this study was found to be a relatively poor template (in comparison with known cis-acting elements of tombusviruses) in in vitro primer extension and template switching experiments with the CNV RdRp (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002), suggesting that the AU sequence is unlikely to form a recombination hot spot as an acceptor site. If this is the case in protoplast as well, then most of recombination events might involve minus-stranded templates and, thus, template switching would occur during plus-strand synthesis (Fig. 2.8A). This is because the first copy of the repeated AU regions (i.e. AU$^1$), which is observed as a hot spot with DI-AU$^1$-FP-AU$^2$ (Fig. 2.2A), would be at the 3’ proximal location (relative to the deletion junctions) in the minus-stranded DI-AU$^1$-FP-AU$^2$ RNA (Fig. 2.8A). Thus, the AU$^1$ repeat is favorably positioned to serve as a donor site for promoting
jumping events by the tombusvirus replicase to a new location [which is hs40 within RII(-), see below]. The above model (Fig. 2.8A-B) also predicts that the second AU repeat (i.e., AU²), located at a more 5’ position in the minus-stranded DI-AU¹-FP-AU² RNA should be less favorable as a donor site, because this region could only be copied by the RdRp after the first AU¹ repeat had already been copied. Accordingly, the lack of AU² repeat in DI-AU¹-FP RNA did not significantly alter the recombination sites or the frequency of the recombination events (Figs. 2.1 and 2.2). Interestingly, the AU sequences have also been predicted to support recombination during plus-strand synthesis when they were present in the minus-stranded BMV RNAs (Nagy and Bujarski, 1997).

In contrast to the proposed primary role in promoting replicase jump, the AU-rich sequence may have no or only limited role during RdRp landing (i.e., at the acceptor sites) if it is located “far” from the putative RII(-) cis-acting element. This model is supported by (i) primer extension experiments with CNV RdRp, which demonstrated that the AU sequence is a poor template (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002); and (ii) the lack of effect on distribution of recombination sites in the absence of the AU² repeat (compare data obtained with DI-AU¹-FP-AU² and DI-AU¹-FP RNAs, Fig. 2.2A and C). Interestingly, the AU² sequence became a recombination hot spot, when it was located in the vicinity of the putative RII(-) cis-acting element (such as in constructs DI-RII/Δ40 and DI-RII/Δ60, Fig. 2.4). To explain the formation of these precise recombinants, we propose that the RdRp might still bind to the RII(-) region during the jumping event, but annealing (base-pairing) between the primer and the template might take place within the AU² repeat. This annealing step might then promote primer extension to initiate from the AU² repeat, therefore resulting in precise or semi-precise recombination (Fig. 2.8C-D). We propose that similar annealing
step between the primer and the acceptor region might be inhibited in the presence of extra sequences between AU$^2$ and RII(-) (see construct DI-AU$^1$-FP-AU$^2$, Fig. 2.1), thus favoring recombination to take place within the hs40 region that flanks the putative cis-element in RII(-) (Fig. 2.8A-B).

A surprising observation in this work is the discovery of the major role of RII in TBSV recombination. The in vitro binding studies with recombinant TBSV p33 and p92 replicase proteins and the primer extension studies with the CNV RdRp preparation suggest that there is a putative cis-acting element in RII(-) located 5’ of the hs40 sequence. Indeed, construct 18 that contained the 5’ half of RII(-) (between position 100-239, Figs. 2.6-7) competed efficiently for binding to p33/p92 (Fig. 2.6) and it was an efficient template for primer extension in the in vitro CNV RdRp assay (Fig. 2.7). We propose (Fig. 2.8B) that the jumping viral replicase is likely guided by the RII(-) cis-acting element before resumption of primer extension, resulting in recombination hot spots. The role of this cis-element is likely more important than the effect of base-pairing between the primer and the template during tombusvirus recombination events. Accordingly, we observed earlier that the CNV RdRp could perform RNA synthesis (primer extension and template switching) without the need of extensive base pairing between the nascent strand and the acceptor strand (Cheng and Nagy, 2003; Cheng, Pogany, and Nagy, 2002). This observation can explain why there are not long stretches of sequence identity around the recombination sites generated in infections with DI-AU$^1$-FP-AU$^2$ (Fig. 2.2A).

The role of cis-acting replication elements in recombination has been proposed before in several viral systems [reviewed in (Nagy and Simon, 1997)]. For example, a known replication enhancer element (Nagy, Pogany, and Simon, 1999) in a satellite RNA (termed
satC), which is associated with *Turnip crinkle virus* (TCV) infections, promoted recombination between the mutated satC and satD, another satellite RNA (Cascone, Haydar, and Simon, 1993; Nagy, Pogany, and Simon, 1999). It has been proposed that the replication enhancer of satC is involved in binding to the jumping TCV replicase (Nagy, Pogany, and Simon, 1999; Nagy, Pogany, and Simon, 2001), a process similar to the role that was proposed above for the RII(-) in TBSV DI RNA (Fig. 2.8). A different class of *cis*-acting element, namely the subgenomic promoter region has also been proposed to promote RNA recombination in luteoviruses and BMV (Wierzchoslawski et al., 2003). In summary, the discovery of the role of the RII(-) element in RNA recombination, in combination with published data cited above, suggests that *cis*-acting elements might play much wider roles in viral RNA recombination than previously anticipated.
Fig. 2.1. The AU-rich hot spot sequence of BMV promotes recombination in DI RNA of TBSV. (A) Schematic representation of the recombination vector and its derivatives. A 117 nt long sequence from the 3' end of EGFP (termed FP), the 69 nt long AU-rich sequence of BMV (termed AU\textsuperscript{1} or AU\textsuperscript{2}, depending on the location relative to the FP sequence) or a 152 nt long region from the barstar gene were inserted as shown between RI and RII of the prototypical DI-73 RNA of TBSV. The probe used in Northern blots and the primers for RT-PCR are depicted above the constructs. (B) The sequence and the predicted structure of the AU sequence (shown in plus-strand orientation). Note that the AU sequence consists of two
parts: (i) the BMV-derived R’ that contains a 21 nt AU-rich region (underlined with a thick gray line); and (ii) the artificial highly AU-rich sequence termed AUs that is 25 nt long (encircled, note that the four nucleotides marked with asterisks are part of both R’ and AUs regions). (C) Northern blot analysis of the total RNA extracts obtained from protoplasts (zero passage). Protoplasts were electroporated with the shown gel-isolated DI RNAs in the presence of the CNV genomic (g)RNA. The parental DI RNAs are marked. (D) Ethidium-bromide stained gel of the total RNA extracts obtained from protoplasts (first passage) that were electroporated with the total RNA extract obtained from “zero passage” protoplasts. The parental and recombinant DI RNAs are marked. Each experiment was based on twelve or more independent protoplast samples (four representative samples are shown). (E) Northern blot analysis of the samples shown in panel D. An asterisk marks a unique recombinant (lane 6) that had the recombination site close to the 3’ end, away from the inserted sequences (thus, it could not be detected by RT-PCR). (F) RT-PCR analyses of samples shown in panel D. Lanes 5, 10, 15, and 20 contains the control RT-PCR products obtained with the same DI RNA transcripts that were used for electroporation. (G) Northern blot analysis of the total RNA extracts obtained from the uninoculated (systemically-infected) N. benthamiana leaves 10 days after inoculation with the shown gel-isolated DI RNAs in the presence of the CNV gRNA.
### A. DI-AU¹-FP-AU² (protoplasts)

<table>
<thead>
<tr>
<th>AU²</th>
<th>FP</th>
<th>AU¹</th>
<th>RII</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
</tbody>
</table>

### B. DI-AU¹-FP-AU² (plants)

<table>
<thead>
<tr>
<th>AU²</th>
<th>FP</th>
<th>AU¹</th>
<th>RII</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
</tbody>
</table>

### C. DI-AU¹-FP (protoplasts)

<table>
<thead>
<tr>
<th>AU²</th>
<th>FP</th>
<th>AU¹</th>
<th>RII</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/15</td>
</tr>
</tbody>
</table>

### D. DI-AU¹-FP (plants)

<table>
<thead>
<tr>
<th>AU²</th>
<th>FP</th>
<th>AU¹</th>
<th>RII</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/12</td>
</tr>
</tbody>
</table>

35
Fig. 2.2. Sequences of DI RNA recombinants around the junction sites. A representative number of recombinants obtained with construct DI-AU$^1$-FP-AU$^2$ (panels A-B) or with DI-AU$^3$-FP (panels C-D) were amplified by RT-PCR (see Fig. 2.1), cloned and sequenced around the junction sites. The top sequence represents the sequence of the parental DI RNAs. The regions are shown schematically on the top. The underlined nucleotides represent marker mutations. Double slashes indicate not shown sequences, while the lengths of the not shown sequences are indicated with numbers below the parental sequences. The recombinants are shown below the asterisk. The deleted sequences in the recombinants are marked with dotted lines, while nontemplated nucleotides are shown with small letters. +++ indicates the following not shown sequence: UGGACCGUCACCCUGCAGCCUGCUUCUG. Note that positive strand sequences are shown in 5’ to 3’ orientation. Two cDNA clones were sequenced/per protoplast samples, but we counted only one recombinant/per protoplast sample in the frequency column.
**Fig. 2.3.** The role of the RII sequence in DI RNA recombination. (A) Schematic representation of the parental constructs. See Fig. 1A for details. The deleted sequences are marked with broken lines. (B) Northern blot and (C) RT-PCR analyses of the total RNA extracts obtained from first passage protoplasts. Lanes 6, 12, 18, 24, and 30 (marked with T as transcripts) contains the control RT-PCR products obtained with the same DI RNA transcripts that were used for electroporation. See further details in the legend to Fig. 2.1.
Fig. 2.4. Sequences of DI RNA recombinants obtained with RII deletion mutants. A representative number of recombinants (from first passage protoplast) were amplified by RT-PCR (see Fig. 2.3), cloned and sequenced. The duplicated regions that participated in precise recombination are boxed. Note that due to sequence identity in the duplicated regions, the actual recombination sites cannot be determined in these precise recombinants. The actual sequence of the region marked with +++ is shown at the bottom of panel B. See further details in the legend to Fig. 2.2.
### A. DI-AUS-FP-AU²

<table>
<thead>
<tr>
<th>AU²</th>
<th>FP</th>
<th>RII Δ40</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### B. DI-R' - FP-AU²

<table>
<thead>
<tr>
<th>R'</th>
<th>FP</th>
<th>AU²</th>
<th>RII Δ40</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### C. DI-R'/GC2-FP-AU²

<table>
<thead>
<tr>
<th>R'</th>
<th>GC2</th>
<th>FP</th>
<th>AU²</th>
<th>RII Δ40</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### D. DI-AU²-FP-R'/GC2

<table>
<thead>
<tr>
<th>AU²</th>
<th>FP</th>
<th>R'</th>
<th>GC2</th>
<th>RII Δ40</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Fig. 2.5.** The effect of short AU-rich and GC-rich sequences on DI RNA recombination. All these constructs are derivatives of DI-RII/Δ40 (Fig. 2.3A). The sequences of AUs and R’ are shown in Fig. 2.1B, while the complete GC2 sequence is shown in panels C and D. A representative number of recombinants were amplified by RT-PCR (not shown) from first passage protoplast samples, cloned and sequenced around the junction sites. The duplicated regions that participated in precise recombination are boxed. +++ indicates the following not shown sequence: 5’-GCCGGAGGUGGAGGAGCAGGAUGAU. See further details in the legend to Figs. 2.2 and 2.4.
Figure 2.6. Efficient binding of RII(−) to the recombinant p33 and p92 replicase proteins of TBSV in vitro. (A) Schematic representation of the competitor RNAs used in gel mobility shift experiments. Note that the competitors represent negative-strand sequences. See Fig.
2.1A for further details. (B-C) A gel mobility shift assay showing the competition between the $^{32}$P-labeled probe [i.e., RIII(-) sequence fused to the cPR11 plus-strand initiation promoter] and the shown unlabeled competitors. The unbound, free RNA probe and the shifted (bound) RNA/protein complexes are marked on the right. The competitor RNAs were used 5x, 15x and 45x excess over the labeled probe. Quantification of the binding of the probe to p92 and p33 is shown below the gel image. Binding of the 92 nt long $^{32}$P-labeled probe to the recombinant p92 protein (1 μM) in panel B and to the recombinant p33 (1 μM) in panel C in the absence of the competitor (lane 11) was taken as 100%. Each experiment was repeated three times. Note that lane 10 includes the probe in the absence of proteins and competitor RNAs.
Fig. 2.7. Comparison of the level of primer extension obtained with RII(-) and GFP-derived sequences. (A) Schematic representation of the constructs tested in an in vitro CNV RdRp assay. These templates contain the same artificial primer regions at the 3’ end of the templates that form 5 bp primers (termed Art-5, indicated by a solid arrow) and the shown acceptor regions. Note that the acceptor regions represent (-)-strand sequences. (B) A representative 8M urea/PAGE analysis of CNV RdRp products using the series of templates shown in Panel A. The primer extension products initiated from the Art-5 primers are marked
with arrowheads. All these RdRp products are partially RNase sensitive (not shown). (C) The normalized (based on the number of radiolabel incorporated) % values of the levels of primer extension by the partially purified CNV RdRp were compared to that of construct #21.
Fig. 2.8. Models of AU-rich sequence-promoted RNA recombination in tombusviruses. Panel A1 shows DI-AU₁-FP-AU² template (Fig. 2.1A) in minus-strand orientation. hs40 is the recombination hot spot, while a triangle points to the putative cis-acting element in RII(-) that facilitates binding to the replicase proteins. The CNV RdRp (represented by an oval circle), while copying the minus-strand template from 3’ to 5’ direction, is proposed to jump from the 3’ AU₁ sequence to the putative cis-acting element in RII(-), probably still carrying the primer nascent strand (solid line). (A2) After binding to RII(-), the RdRp resume extension of the primer from the nearby hs40 sequence, thus resulting in imprecise recombinants. (B1) This alternative model explains the possible formation of precise recombinants observed with DI-RII/Δ40 (shown schematically here) and its derivatives (Figs. 3-5). See panel A1 for further details. (B2) After binding to the putative cis-acting element in RII(-), the RdRp might resume extension of the primer from the nearby AU² sequence in the template. We propose that annealing between the primer (which is complementary to AU¹)
and the AU\textsuperscript{2} sequence in the template (represented by five short lines) prior to primer extension might facilitate the formation of precise recombinants.

**TABLE 2.1. Primers used for PCR**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Primer</th>
<th>Sequence</th>
<th>Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>DI-AU\textsuperscript{1}-FP-AU\textsuperscript{2}</td>
<td>114</td>
<td>GCAGTCTAGAGGACGCGTGAGAGACCTGTCCAG</td>
<td>AU1,</td>
</tr>
<tr>
<td></td>
<td>115</td>
<td>GCAGCTCGAGCCATGGATCCATAATTGATGG</td>
<td></td>
</tr>
<tr>
<td>DI-AU\textsuperscript{1}-FP</td>
<td>114</td>
<td>GCAGTCTAGAGGACGCGTGAGAGACCTGTCCAG</td>
<td>AU1,</td>
</tr>
<tr>
<td></td>
<td>115</td>
<td>GCAGCTCGAGCCATGGATCCATAATTGATGG</td>
<td></td>
</tr>
<tr>
<td>DI-Bar</td>
<td>922</td>
<td>GGAGTCTAGATTGTCTGACCGGATGGGT</td>
<td>DI-73-BARS</td>
</tr>
<tr>
<td></td>
<td>923</td>
<td>GGAGGGATCCGAAAGTATGATGGTGA</td>
<td></td>
</tr>
<tr>
<td>DI-FP-RII/Δ40</td>
<td>751</td>
<td>GGCAGGGATCTTTTGGAGCTCTTTC</td>
<td>DI-FP</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>CCGCCATGGTCCATAATTGATGGA</td>
<td></td>
</tr>
<tr>
<td>DI-RII/Δ40</td>
<td>751</td>
<td>GGCAGGAAGGCTTTTGGAGCTCTTTC</td>
<td>DI-FP</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>CCGCCATGGTCCATAATTGATGGA</td>
<td></td>
</tr>
<tr>
<td>DI-RII/Δ20</td>
<td>933</td>
<td>CCGGGATACCCACAACCTACCAAAG</td>
<td>DI-FP</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>CCGCCATGGTCCATAATTGATGGA</td>
<td></td>
</tr>
<tr>
<td>DI-RII/Δ60</td>
<td>934</td>
<td>CCGGGATACCCCCAGGTCAAGAAAAG</td>
<td>DI-FP</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>CCGCCATGGTCCATAATTGATGGA</td>
<td></td>
</tr>
<tr>
<td>DI-RII/Δ100</td>
<td>935</td>
<td>CCGGGATACCCCTTGATGAAAGC</td>
<td>DI-FP</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>CCGCCATGGTCCATAATTGATGGA</td>
<td></td>
</tr>
<tr>
<td>DI-AUs-FP-AU\textsuperscript{2}</td>
<td>860</td>
<td>CGACTCTAGAGCTAATTTATAATTCCATCA</td>
<td>AU1,</td>
</tr>
<tr>
<td></td>
<td>753</td>
<td>CCGCCATGGTCCATAATTGATGGA</td>
<td></td>
</tr>
<tr>
<td>DI-R'/GC2-FP-AU\textsuperscript{2}</td>
<td>752</td>
<td>CGACTCTAGACCCCTGTCAGGCTAGG</td>
<td>GC2,</td>
</tr>
<tr>
<td></td>
<td>794</td>
<td>CCGCCATGGTCCGCCCCCCTATTTT</td>
<td></td>
</tr>
<tr>
<td>DI-AU\textsuperscript{1}-FP-R'/GC2</td>
<td>792</td>
<td>CGACACGCGGTCCCTGTCCAGGTTAGG</td>
<td>GC2,</td>
</tr>
</tbody>
</table>
798  CCGGGATCCTCCGCCGCCCTATTG

DI-R’-FP-AU²  752  CGA[CCTAG]ACCTGTCCAGGTA  AU1,  

796  CCG[CATGGT]AGCTTTTAACCTTAGCC

Competitor templates:

#1  1077  CGCTGCCCATGTCTTGT  DI-AU³-FP
  1005  TAATACGACTCACTATAGGGACACATTG  DI-AU³-FP

#7  1009  TAATACGACTCACTATAGGGACTCGTG  DI-AU³-FP
  933  CGGGATCCACACACTACAA

#18  1038  TCGTCTTATTGGACGACACCGACTTGGGTATGA  DI-AU³-FP
  14  GTAATACGACTCACTATAGGGTTCTCTGCTTTTACGAAG

Templates for primer extensions in vitro:

#11  167  TCGTCTTATTGGACGAAAGAAAGGGAAGCTCGC  DI-AU³-FP
  14  GTAATACGACTCACTATAGGGTTCTCTGCTTTTACGAAG

#16  1037  TCGTCTTATTGGACGAGGAGAGCTCTTGGACGCTT  DI-AU³-FP
  1009  TAATACGACTCACTATAGGGACACATTG  DI-AU³-FP

#17  1037  TCGTCTTATTGGACGAGGAGAGCTCTTGGACGCTT  DI-AU³-FP
  14  GTAATACGACTCACTATAGGGTTCTCTGCTTTTACGAAG

#18  1038  TCGTCTTATTGGACGACACCGACTTGGGTATGA  DI-AU³-FP
  14  GTAATACGACTCACTATAGGGTTCTCTGCTTTTACGAAG

#20  1036  TCGTCTTATTGGACGACAGCCTGCCCACGCTT  DI-AU³-FP
  14  GTAATACGACTCACTATAGGGTTCTCTGCTTTTACGAAG

#21  1035  TCGTCTTATTGGACACACCGACCGAAGGATGGA  DI-AU³-FP
  14  GTAATACGACTCACTATAGGGTTCTCTGCTTTTACGAAG

__________________________________________________________________________

Copyright © Natalia Shapka 2006.
Chapter III

Genome-wide screen identifies host genes affecting viral RNA recombination

INTRODUCTION

Rapid evolution of RNA viruses with mRNA-sense genomes, which include SARS coronavirus, hepatitis C virus, and West Nile virus, makes controlling RNA viruses a difficult task. Emergence of new pathogenic RNA viruses is frequently due to RNA recombination (Lai, 1992; Worobey and Holmes, 1999), which can lead to dramatic changes in viral genomes by creating novel combinations of genes, motifs or regulatory RNA sequences. Thus, RNA recombination can change the infectious properties of RNA viruses and render vaccines and other antiviral methods ineffective (Worobey and Holmes, 1999). RNA recombination likely contributed to outbreaks with dengue- (Holmes, Worobey, and Rambaut, 1999; Worobey, Rambaut, and Holmes, 1999), polio- (Marturano and Fiore, 2002), calici- (Jiang et al., 1999), astro- (Walter et al., 2001), entero- (Lukashev et al., 2004; Oprisan et al., 2002), influenza- (Khatchikian, Orlich, and Rott, 1989), pestiviruses (Becher, Orlich, and Thiel, 2001; Fricke, Gunn, and Meyers, 2001), and SARS coronavirus, a newly-emerged viral pathogen of humans (Bosch, 2004; Rest and Mindell, 2003; Stavrinides and Guttman, 2004). RNA recombination is also important in viral RNA repair, which likely
increases the fitness of RNA viruses that lack proofreading polymerases (Allison, Thompson, and Ahlquist, 1990; Guan and Simon, 2000; Lai, 1992; Nagy and Simon, 1997).

Current models of RNA recombination are based on template-switching mechanism driven by the viral replicase (Lai, 1992; Nagy and Simon, 1997) or RNA-breakage and ligation (Chetverin et al., 1997). The more common template-switching RNA recombination is thought to occur as an error during the replication process (Lai, 1992; Nagy and Simon, 1997). Because viral RNA replication depends not only on viral proteins, but on host factors as well (Ahlquist et al., 2003), it is likely that host factors could affect the recombination process, too. However, despite the significance of RNA recombination in viral evolution, the possible roles of host genes in the viral RNA recombination process are currently unknown.

Tombusviruses, including Tomato bushy stunt virus (TBSV) and Cucumber necrosis virus (CNV), are non-segmented, small model positive-strand RNA viruses (White and Nagy, 2004). Due to their robust replication and the ability to generate novel RNA recombinants in whole plants and single cells, tombusviruses are used extensively to dissect the roles of cis-acting RNA elements during virus infections (White and Nagy, 2004). In vivo and in vitro replication/recombination studies with a small replicon RNA, termed defective interfering (DI-72) RNA (White and Morris, 1994b; White and Nagy, 2004), established a role for RNA sequences/structures and viral replicase proteins in RNA recombination. Co-expression of the replicon RNA with the two essential tombusviral replicase proteins (Fig. 3.1A) resulted in robust DI RNA replication in Saccharomyces cerevisiae (Panavas and Nagy, 2003b; Pantaleo, Rubino, and Russo, 2003), which is a model eukaryotic host. Yeast also supported viral RNA recombination, giving rise to recombinants similar to those in
plants and plant protoplasts (Panavas and Nagy, 2003b). Therefore, yeast could be a useful host to study viral RNA recombination and to identify host proteins involved in this process.

In this paper, we have tested the effect of ~80% of all yeast genes on TBSV recombination based on screening the entire yeast single-gene knockout (YKO) library for the occurrence of viral RNA recombinants. Using the TBSV derived replicon RNA, we identified five YKO strains that supported unusually high levels of new recombinant RNAs. We also identified four yeast deletion strains that showed reduced viral recombinant accumulation. Therefore, selected set of host genes could either suppress or accelerate viral RNA recombination, demonstrating for the first time that host genes play significant roles in virus recombination and evolution.

MATERIALS AND METHODS

Yeast strains and expression plasmids:

_Saccharomyces cerevisiae_ strain BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and the haploid deletion series (BY4741 strain background) were from Open Biosystems (Huntville, AL). The expression plasmids pGBK-His33 (carrying CNV p33 gene behind the ADH1 promoter), pGAD-His92 (containing CNV p92 gene behind the ADH1 promoter), and pYC/DI-72 (expressing TBSV DI-72 RNA under the control of GAL1 promoter) have been previously described (Panavas and Nagy, 2003b; Panaviene et al., 2004). Each yeast strain was co-transformed with all three plasmids using LiAc/ssDNA/PEG method (Gietz and Woods, 2002) and transformants were selected by complementation of auxotrophic markers. Out of 4848 strains, we found that 71 were not transformable and 229 strains did not grow on
galactose-containing medium. Therefore, total of 4548 strains were tested for RNA recombination below.

**Yeast cultivation:**

Each transformed yeast strains from the YKO library were cultured under two different conditions during the genome-wide screen for RNA recombinants. The first screen included yeast strains grown in 96-deep-well plates at 23°C in selective media (SC-ULH) with 2% galactose until reaching cell density of 0.8-1.0 (OD$_{600}$). For the second screen, the yeast strains were grown in 96-deep-well plates at 23°C for 6h in selective media (SC-ULH) with 2% galactose, followed by 1:10 dilution with SC-ULH medium containing 5% glucose. Then, the cells were grown for 24h at 23°C, followed by additional dilution (1:10) and subsequent culturing until cell density reached 0.8-1.0 (OD$_{600}$). Yeast cells were harvested by centrifugation at 1,100g for 5 min.

**High-throughput RNA analysis:**

We performed two separate genome-wide screens of the YKO library that included total of 4-6 independent samples per each strain. Total RNA isolation and Northern blot analysis were done as previously described (Panavas and Nagy, 2003b), except using a high throughput approach. Briefly, yeast cells in 96-deep-well plates were resuspended in RNA extraction buffer (50mM sodium acetate, pH 5.2, 10mM EDTA, 1% SDS) and phenol, followed by incubation for 4 min at 65°C. After removal of phenol, the RNA was recovered by ethanol precipitation. Agarose gel electrophoresis (1.5%) and Northern blotting were done as described (Panavas and Nagy, 2003b; Shapka and Nagy, 2004). For negative-strand
detection, total yeast RNA obtained from selected strains was separated in denaturing 5% polyacrylamide/8M urea gel as described previously (Panavas and Nagy, 2003b). The RNA was quantified using a phosphorImager as described (Panavas, Pogany, and Nagy, 2002).

**RT-PCR analysis of the junction sites in the recombinants:**

We have used both total yeast RNA extracts and gel-isolated recombinants for reverse transcription (RT-)PCR reactions to specifically amplify regions covering junction sites. First, the RT reaction included primer #14 (GTAATACGACTCACTATAGG GTTCTCTGCTTTTACGAAG) for cDNA synthesis, followed by PCR with primers #168 (TCGTCTTTATGGACGAATTCCTGTTTACGAAAG) and #270 (TTGGAAATTCTCCTCAGTCTGAGTTTGTGGA). The PCR products were cloned into pGEM-T Easy vector (Promega) and sequenced using M13 Reverse Primer (Cheng and Nagy, 2003).

**5’RACE and 3’RACE of recombinants:**

The 5’ and 3’ sequences of recombinants were determined by using 5’RACE (rapid amplification of complementary ends) and 3’RACE, respectively, as described (Cheng and Nagy, 2003). To enrich for recombinants, RNA bands were gel-isolated as described previously (Cheng and Nagy, 2003). The resulting products were cloned and sequenced.

**In vitro Tombusvirus replicase assay:** The in vitro replicase assay was performed with the co-purified (endogenous) RNA as described (Panavas and Nagy, 2003b).
RESULTS

Systematic analysis of yeast single-gene deletion strains for enhanced level of viral RNA recombination.

To facilitate identification of host genes involved in RNA virus evolution/recombination, we took advantage of the advanced genomics tools available for yeast and the ability of yeast to support TBSV recombination (Panavas and Nagy, 2003b). The recombination assay was based on a replication-competent TBSV DI-72 RNA replicon, which, when co-expressed with the two essential tombusviral replicase proteins (p33 and p92, Fig. 3.1A), undergoes robust replication and it also generates small amount of RNA recombinants (Panavas and Nagy, 2003b). The 621 nt DI-72 RNA replicon contains four noncontiguous segments (Fig. 3.1A), including the cis-acting replication elements, derived from the full-length genomic RNA (White and Nagy, 2004). It is important to note that in this assay replication/evolution of DI-72 RNA and the de novo generated recombinant RNAs takes place in the absence of artificial selection markers in all viral RNAs.

To systemically test the effect of each host gene on viral RNA recombination, we have developed a high throughput method based on the available yeast single-gene deletion (YKO) library. Briefly, the collection of 4,848 YKO yeast strains representing ~80% of yeast genes (those which are nonessential for yeast growth) was co-transformed with three plasmids expressing DI-72 RNA replicon in addition to p33 and p92 replicase proteins (Fig. 3.1A). We successfully transformed 4548 strains that grew on galactose-containing media (see M&M), and cultured them in 96-deepwell plates, followed by total RNA extraction and
agarose gel electrophoresis. Under these conditions, the replication-competent DI-72 RNA is easily detectable in yeast cells and its amount is similar to the yeast ribosomal RNAs (Fig. 3.1B). In contrast, recombinant RNAs carrying rearranged RNA sequences accumulate inefficiently in the parental yeast strain as demonstrated by Northern blotting (~1-2% of the level of the replicating DI-72 RNA) (Fig. 3.1C). Thus, under the above conditions the screening approach is expected to favor the identification of those YKO strains, which support increased levels of viral RNA recombinants when compared to the parental strain.

**Identification of five host genes whose absence leads to increased frequency of viral RNA recombination.**

Using the above high-throughput genome-wide screen, we identified total of five YKO strains that generated 10-50-fold higher levels of recombinant viral RNAs than did the parental yeast strain (Fig. 3.1B-C). Four of these deletion strains, *ctt1Δ, met22/hal1Δ, xrn1Δ* and *ubp3Δ* accumulated one major type of recombinant RNA at levels comparable to that of the wild-type viral replicon (Fig. 3.1B-C), whereas *hur1Δ* generated four recombinant RNAs, which were ~10-50-fold more abundant than the wt replicon (Fig. 3.1B-C). Northern blot analysis with a probe specific for an internal RIII sequence in the DI-72 RNA replicon demonstrated the viral origin of these novel recombinant RNAs (Fig. 3.1C). On the contrary, a probe specific for the 5’ RI sequence, only detected the wt DI-72 replicon in total RNA samples from all five strains, but not the recombinant-like RNAs (Fig. 3.1D), suggesting that RNA recombination might have led to dramatic rearrangement of the viral RNA. These recombinant viral RNAs accumulated in the presence of the wt DI-72 replicon, suggesting
that they were generated efficiently and/or competed efficiently with the wt DI-72 RNA replicon.

To determine the sequence of the novel recombinant-like RNAs, we gel-isolated them followed by RT-PCR, 3’RACE, 5’RACE, cloning and sequencing (Cheng and Nagy, 2003; Shapka and Nagy, 2004). We found that the most common recombinant RNAs obtained from xrn1Δ, ctl1Δ, met22Δ and ubp3Δ strains were similar, partially dimeric RNAs (Fig. 3.2A and not shown). They contained various duplicated 3’ sequences (part of RII, and complete RIII and RIV) and had deletions of 5’ DI-72 RNA sequences (i.e., RI and part of RII, Fig. 3.2A). Most recombinants differed slightly in their junction sequences, a feature shared with TBSV recombinants arising in planta (5). Recombinants in hur1Δ contained 2-to-5 incomplete copies of DI-72 RNA sequences with highly variable 5’-truncations (Supplement Fig. 3.8). The origin of the 1-to-13 extra nucleotides at the 5’ end or at the junctions is currently unknown. Extra nucleotides are also frequently detected at the junctions in tombusvirus recombinants in plant protoplasts (Shapka and Nagy, 2004; White and Morris, 1994b) and in vitro with purified tombusvirus replicase (Cheng and Nagy, 2003), supporting the model that the tombusviral replicase adds extra sequences to the ends of viral RNAs (Nagy and Simon, 1997).

To gain insights into the dynamics of recombinant formation, we performed time-course experiments by analyzing total RNA samples at given time points after induction of RNA transcription from the GAL1 promoter in hur1Δ and xrn1Δ strains. We found that the recombinants emerged as early as 2 hours after induction (Fig. 3.3A) in the absence of artificial selection to facilitate their appearance, suggesting that their formation is an efficient process. The amount of recombinants increased over time due to either new recombination
events and/or replication of the recombinant RNAs (Fig. 3.3A). Moreover, we found that the recombinant RNAs replicated and evolved further in yeast cells over ten serial dilutions in suppressive media (Fig. 3.3B).

To demonstrate that the recombinant DI RNAs are replication-competent, we isolated membrane fractions containing tombusvirus replicase/viral RNA complexes from xrn1Δ and hur1Δ cells. This was followed by in vitro replicase assays in the presence of added ribonucleotides, including 32P-labeled UTP. These experiments led to in vitro labeling of the recombinant-sized RNAs in the replicase assay, suggesting that the recombinant RNAs were part of the replicase complexes (Fig. 3.3C). Their replication competence was also confirmed by detection of minus-stranded replication intermediates for the recombinant RNAs (Fig. 3.3C). Altogether, we conclude that the recombinant RNAs, similar to the wt DI-72 RNA replicon, are replication-competent and they are maintained on suppressive media for extended period of time in yeast.

To test if the deletion of the host gene altered recombination frequency versus recombinant selection, we analyzed the stability of four cloned recombinants and the wt DI-72 replicon RNA in the parental and xrn1Δ strain. Fig. 3.4 demonstrates that the stability of the recombinants and wt replicon was comparable in the parental strain, whereas the recombinants and the wt replicon RNA showed 2-to-3-fold increased stability in xrn1Δ strain (Fig. 3.4). This suggests that viral RNA degradation is hindered in xrn1Δ strain. Importantly, however, all the recombinants and the wt replicon RNA showed similar level of increase in stability in xrn1Δ strain, suggesting that these RNAs have comparable stability. Overall, selective RNA degradation of wt replicon versus recombinants cannot explain the increased accumulation of recombinants over the wt replicon in xrn1Δ strain.
Systematic analysis of yeast single-gene deletion strains for decreased level of viral RNA recombination. Because the above genome-wide screen was only suitable for testing for increased levels of TBSV recombination, we modified the screening approach to allow the identification of YKO strains supporting reduced levels of virus recombinants when compared with the parental strain. To this end, we induced DI-72 RNA transcription in all 4548 YKO strains (also co-expressing p33/p92) for 6 hours, followed by growing them in glucose-containing medium prior to total RNA extraction and analysis by Northern blotting. Under these conditions, detectable amounts of recombinant RNAs accumulated in the parental strain (10-18% of the standard wt DI-72 RNA, Fig. 3.5A). These recombinant RNAs included complete dimers (two copies of full-length DI-72 replicons joined head-to-tail) and incomplete dimeric DI-RNAs (5’ truncated monomers joined head-to-tail, see Fig. 3.5A). Northern-blot analysis of total RNA extracts from all transformants (two-to-four samples per strain) revealed that the ratio of recombinant RNA versus wt DI-72 RNA was three-to-five-fold lower only in four YKO strains (Fig. 3.5A). Note that we did not measure the absolute amounts of RNA recombinants, but instead, estimated the ratio of recombinants versus non-recombinant DI-72 RNA. This is because DI-72 RNA accumulation level could be different in various YKO strains (Panavas et al., 2005c), which could affect the amount of RNA substrates available for recombination.

Characterization of four host genes whose absence leads to decreased frequency of viral RNA recombination.
To further test the recombination deficiency of the above identified 4 YKO strains, we examined if they could support RNA recombination with a modified viral replicon, DI-AU-FP. This replicon contains a 186 nt long heterologous sequence including a 46 nt-long AU-rich stretch (Fig. 3.5B). Previous work in plant protoplasts demonstrated that DI-AU-FP induced recombination with high efficiency in the presence of the wt helper virus (Shapka and Nagy, 2004). Northern blot analysis of total RNA obtained from the four YKO strains co-expressing DI-AU-FP and p33/p92 proteins revealed that RNA recombinants accumulated poorly (3-5-fold decrease) in these strains when compared to the parental strain (Fig. 3.5B). The sequences of the generated recombinants isolated from the parental and the selected YKO strains were comparable (shown schematically in Fig. 3.5B), indicating that the mechanism of their generation was likely similar. Overall, we conclude that the four identified YKO strains supported recombination with reduced frequency and/or accumulation rate of recombinants was lower in these strains than in the parental yeast.

It is worth noting that the above genome-wide screen also identified 5 additional YKO strains that supported two-fold reduced level of recombinant accumulation in comparison with the parental strain (see “weak accelerators” in Table 3.1). In addition, we found that spe3Δ and spt3Δ generated different recombinant profile from the parental strain (Fig. 3.6A). The notable difference was the accumulation of a single dominant recombinant RNA in these strains (see recM, Fig. 3.6A). To confirm that the lack of SPE3 and SPT3 genes indeed affected RNA recombination, we analyzed recombinant formation during DI-AU-FP RNA replication. This experiment demonstrated that (i) spe3Δ and spt3Δ showed 2-4-fold increased levels of recombinants and (ii) the profile of recombinants generated were somewhat different from that observed with the parental strain (Fig. 3.6B).
Major contribution of host genes to viral RNA recombination

To demonstrate the full extent of contribution by host genes to viral RNA recombination, we compared recombinant accumulation in \(xrn1\Delta\) and \(pep7\Delta\) strains carrying DI-AU-FP and p33/p92 expression plasmids. These experiments revealed that \(xrn1\Delta\) strain accumulated viral RNA recombinants up to 80-fold higher level than \(pep7\Delta\) strain did (Fig. 3.7). Because these yeast strains differed only in the two deleted genes, the above experiment demonstrated that host genes could play major roles in viral RNA recombination.

DISCUSSION

Viruses are known to evolve rapidly in selected hosts, yet the roles of host genes in RNA virus recombination/evolution are currently unknown. This work, based on high throughput genetic screen in yeast, a model host, has led to the identification of 11 host genes that significantly affected tombusvirus recombination. We found that single deletion of the identified genes had three types of effect on tombusvirus recombination: (i) five genes increased, while (ii) four genes decreased recombinant accumulation, and (iii) two genes changed the profile of recombinants. Additional five genes had lesser effect (~2-fold) on RNA recombination.

Suppressors of RNA virus recombination

The observation that the accumulation of viral RNA recombinants increased 10-50-fold in the absence of five host genes (Table 3.1) suggests that these genes, when present, can suppress RNA virus recombination. Interestingly, three of the identified genes, namely \(XRNI, CTLJ\) and \(MET22/HAL2\), are involved in RNA metabolism/degradation. It is plausible
that these genes could affect viral recombination by influencing the 5’-to-3’ RNA degradation pathway (Parker and Song, 2004). The proposed connection between RNA degradation and viral RNA recombination is supported by the following findings: (i) the recombinants had deletions within their 5’ sequences (Fig. 3.2); (ii) 5’ truncated viral RNAs accumulated in these yeast strains (Fig. 3.1B-C); and (iii) identification of Xrn1p, which is the key enzyme in the 5’-to-3’ RNA degradation pathway (Parker and Song, 2004; Sheth and Parker, 2003), as one of the viral recombination affecting proteins; (iv) the increased stability of both recombinant and DI-72 RNA replicon in xrn1Δ strain (Fig. 3.4). Moreover, three of the five identified host genes are predicted and/or known to affect the activity of Xrn1p. For example, Met22p/Hal2p has been shown to affect the activity of Xrn1p via regulating the level of pAp, an inhibitor of Xrn1p (Dichtl, Stevens, and Tollervey, 1997). Also, Ctl1p is known to modify the 5’ end of the RNA by removing a phosphate group that could potentially facilitate Xrn1p-driven 5’-to-3’ RNA degradation (Rodriguez et al., 1999). In addition, Ubp3p has been shown to increase stability of Xrn1p in cells (Brew and Huffaker, 2002). Altogether, the 5’-to-3’ exoribonuclease activity of Xrn1p could be inhibited in the absence of one of these genes. On the contrary, the role/function of Hur1p is currently unknown. The profile of recombinants generated in hur1Δ, however, is different (Supplement Fig. 3.8) from the recombinants identified in the other four YKO strains, indicating that it might use a different mechanism during viral RNA recombination. Overall, this genome-wide screen indicates close connection between viral RNA recombination and RNA metabolism/RNA degradation. Interestingly, a 5’-to-3’ exoribonuclease, similar to Xrn1p, is present in Arabidopsis (Kastenmayer and Green, 2000), and Hal2p homolog has been cloned from rice (Peng and Verma, 1995), suggesting that similar genes are functional in plants, too.
**Proposed mechanism of suppression of viral RNA recombination:**

Based on the identified host genes and the profile of generated viral RNA recombinants, we propose that four out of five host genes, including \textit{XRNI}, \textit{CTL1}, \textit{MET22/HAL2} and \textit{UBP3}, could suppress viral recombination via affecting the Xrn1p-dependent rapid and complete degradation of viral RNA. However, in the absence of Xrn1p, or due to inhibition of Xrn1p activity in the absence of Ctl1p, Met22p or Ubp3p, degradation of the viral RNA gets slower (Fig. 3.4). The resulting incompletely degraded viral RNA could then participate in RNA recombination efficiently, facilitating the accumulation of partly dimeric recombinant RNAs. The generated recombinants are also more stable in \textit{xrn1A} strain, further facilitating the accumulation of recombinants. Moreover, abundance of 5’ truncated RNA species in these strains supports the model that these RNAs are intermediates (substrates) in the RNA recombination process. In addition, efficient recombination is likely due to “exposure” of the highly recombinogenic RII sequences (Shapka and Nagy, 2004) at the ends of the viral RNAs after their partial degradation. On the contrary, the parental yeast cells could efficiently and completely degrade viral RNAs, thus reducing the chance for partly degraded RNAs to participate in RNA recombination.

**Protein accelerators of viral RNA recombination:**

The other set of host genes identified during this genome-wide screen includes four genes, \textit{PEP7}, \textit{IPK1}, \textit{CHO2} and \textit{DCII}, whose deletion resulted in reduced level of viral RNA recombination (Fig. 3.5). The viral replicon RNA, either DI-72 or DI-AU-FP, replicates efficiently in these strains, whereas the dimeric recombinant RNAs accumulate 3-to-5-fold less than in the parental strain. Therefore, these genes might directly influence the frequency
of recombination. Based on the known functions of these genes (Table 3.1), we suggest that (i) intracellular transport of viral and/or host proteins (or possibly protein-viral RNA complexes) to the site of recombination (see genes PEP7 and DCII), and/or (ii) the lipid content/structure of the membranous-compartment, which contains the virus- replicase could be altered in the absence of these genes (IPK1, CHO2 and DCII), resulting in reduced RNA recombination efficiency.

Although the current work has not addressed the mechanism of RNA recombination in the selected strains, comparison of sequences at the recombination junctions suggests that the recombinants represent similarity-nonessential (nonhomologous) recombinants (Nagy and Simon, 1997). The observed recombinants are likely generated via viral replicase-driven template-switching mechanism, which has been shown for tombusviruses before (Cheng and Nagy, 2003; White and Morris, 1994b). Also, data presented in the on line material exclude that DNA recombination or RNA recombination during pol II-driven RNA transcription are the mechanisms of viral RNA recombination (Fig. 3.9).

**General conclusions:**

This genome-wide screen of yeast for host genes affecting viral RNA recombination demonstrates for the first time that selected set of host genes can accelerate or suppress viral RNA recombination. We found that the majority of yeast single-deletion strains showed low level of virus recombination, whereas five strains with particular genetic backgrounds were “hotbeds” for recombination, accelerating virus evolution. This implies that mutation(s) in host genes involved in suppression of virus recombination create “favorable” genetic backgrounds for virus RNA recombination, suggesting that such an individual(s) might
contribute to RNA recombination and virus evolution more significantly than other individuals of the same species with less favorable genetic backgrounds. Altogether, our discovery promises to have a major influence on future thinking about the contribution of particular host genes and individual organisms to virus recombination and evolution.
**Fig. 3.1.** Absence of CTL1, MET22/HAL2, HUR1, XRN1 and UBP3 host genes leads to enhanced recombination of TBSV DI-72 RNA replicon in yeast. (A) Plasmid-based expression of p33 and p92 replicase proteins and DI-72 RNA replicon in yeast. (B) Total RNA extracts from the shown yeast strains (two independent samples are shown for each strain to illustrate the reproducibility of recombinant accumulation) was visualized with ethidium-bromide or probed with a radiolabeled RNA that was (C) complementary with RIII; or (D) with RI of DI-72. Arrow points at the replicon, whereas the novel recombinant RNAs (recRNA) are bracketed. Various recombinants in hur1D are depicted with arrowheads. Samples from hur1Δ yeast were overloaded (~5x) to facilitate visualization of viral RNAs. Short, 5’ truncated viral RNAs are marked with asterisks.
Fig. 3.2. Schematic presentation of the DI-72 replicon with four regions and the recombinants with duplicated 3’ sequences (3’ part of RII, RIII and RIV) and 5’ deletions (RI and 5’ part of RII). The actual sequences of the recombinants (shown for $xrn1\Delta$) at the 5’ ends (left panel) and at the junctions are shown. $\Delta$ indicates the number of deleted nucleotides, whereas virus-templated and nonviral sequences are shown in uppercase and lowercase letters, respectively. The 3’ end in RIV (both at the internal and 3’ terminal locations) contained the authentic sequence.
Fig. 3.3. (A) Time-course experiment with hur1Δ and xrn1Δ co-expressing p33 and p92 reveals rapid generation of recombinants (probed with RIII(-) after induction of DI-72 RNA transcription from plasmid pYC/DI-72. (B) Recombinant RNAs are still present after 10 serial dilutions in glucose-containing medium, which suppresses transcription of DI-72 RNA from the GAL1 promoter. (C) The new viral recombinants are replication competent. The in vitro replicase assay is based on the tombusvirus replicase/viral RNA complex present in the isolated membrane-enriched fraction of yeast (left panel). The presence of minus-stranded RNA replication intermediates for the recombinant RNAs was detected in total RNA extracts using a minus-strand-specific probe (right panel).
**Fig. 3.4.** Deletion of Xrn1p increases the stability of recombinant RNAs and DI-72 replicon RNA. Four representative recombinant RNAs containing partially duplicated sequences (first four bars on the left) and the DI-72 replicon RNA (the dark gray bar on the right) were separately expressed in (A) the parental and (B) xrn1Δ strains from GAL1 promoter. After repression of transcription with glucose (time points of 0, 2, 4 and 6 hours), the residual viral RNAs were measured by Northern blotting and quantified by a phosphoImager. The data are shown in % (the amount of viral RNA at 0 time point is 100%) derived from four independent experiments.
Fig. 3.5. (A) Absence of PEP7, IPK1, CHO2 and DCI1 host genes leads to low frequency of recombination of TBSV DI-72 RNA replicon in yeast co-expressing p33 and p92 replicase proteins. The relative amounts of recombinants T, M and B (in comparison with DI-72 RNA
replicon, which is chosen 100%) are shown. The sequences of the dimeric recombinant RNAs are shown schematically at the bottom. Deletion within RII usually included from 65 to 170 5' nucleotides. The junction sites are circled. (B) Similar recombination experiment was performed with DI-AU-FP replicon RNA. The sequences of DI-AU-FP and the recombinant RNAs are shown schematically on the top and bottom, respectively. See further details in the legend to Fig. 3.1.
Fig. 3.6. Absence of \textit{SPE3} and \textit{SPT3} genes results in altered recombination profile with (\textit{A}) DI-72 RNA replicon and (\textit{B}) DI-AU-FP replicon. See further details in the legend to Fig. 3.1.
Fig. 3.7. Comparison of recombination activity of DI-AU-FP replicon in \textit{xrn1}\Delta, \textit{pep7}\Delta, and parental yeast strains. See further details in the legend to Fig. 3.1.
Fig. 3.8. Schematic presentation of the DI-72 replicon-derived recombinants generated in hur1Δ strain expressing the DI-72 RNA replicon. The recombinants contain duplicated 3’ sequences (3’ part of RII, RIII and RIV) and 5’ deletions (RI and 5’ part of RII). The actual sequences of the recombinants at the 5’ ends (left panel) and at the junctions are shown. Δ indicates the number of deleted nucleotides, whereas virus-templated and nonviral sequences are shown in uppercase and lowercase letters, respectively. The 3’ end in RIV (both at the internal and 3’ terminal locations) contained the authentic sequence. Note that due to less variation at the junctions than at the 5’ end, we sequenced smaller number of clones at the junction sites.
Fig. 3.9. Absence of viral recombinants in yeast DNA and RNA transcripts. (A) Total DNA was extracted from yeast strains carrying the three expression plasmids, followed by PCR analysis for recombinants with primers shown schematically above the representative DI-72 RNA and a recombinant. The positive control lane (marked as C) represents cDNA obtained from xrn1D strain containing recombinant viral RNA. (B) Northern blot analysis of DI-72 RNA transcripts from total RNA extracts obtained from the yeast strains transformed with pYC/DI-72 only (no RNA replication could take place in these cells due to lack of p33 and p92). Note the lack of recombinant RNAs and the presence of the original (containing plasmid-borne 5’ and 3’ sequences) and ribozyme-cleaved DI-72 RNA transcripts in all samples.
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Molecular function/biological process</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supressors:</strong></td>
<td></td>
</tr>
<tr>
<td><em>CTL1</em></td>
<td>polynucleotide 5'-phosphatase</td>
</tr>
<tr>
<td><em>MET22/HAL2</em></td>
<td>3'(2'),5'-bisphosphate nucleotidase</td>
</tr>
<tr>
<td><em>HUR1</em></td>
<td>unknown</td>
</tr>
<tr>
<td><em>XRN1</em></td>
<td>5'-3' exoribonuclease</td>
</tr>
<tr>
<td><em>UBP3</em></td>
<td>ubiquitin-specific protease</td>
</tr>
<tr>
<td><strong>Accelerators:</strong></td>
<td></td>
</tr>
<tr>
<td><em>PEP7/VPS19</em></td>
<td>unknown/Golgi to vacuole transport</td>
</tr>
<tr>
<td><em>IPK1</em></td>
<td>inositol/phosphatidylinositol kinase</td>
</tr>
<tr>
<td><em>CHO2/PEM1</em></td>
<td>phosphatidylethanolamine N-methyltransferase</td>
</tr>
<tr>
<td><em>DCI1</em></td>
<td>dodecenoyl-CoA delta-isomerase</td>
</tr>
<tr>
<td><strong>Weak Accelerators</strong>:</td>
<td></td>
</tr>
<tr>
<td><em>VPS43/VAM7</em></td>
<td>Golgi to vacuole transport</td>
</tr>
<tr>
<td><em>PTH1/VAM3</em></td>
<td>Golgi to vacuole transport</td>
</tr>
<tr>
<td><em>VPS29</em></td>
<td>Endosome to Golgi transport</td>
</tr>
<tr>
<td><em>VPS35</em></td>
<td>Endosome to Golgi transport</td>
</tr>
<tr>
<td><em>NGG1</em></td>
<td>Transcription cofactor</td>
</tr>
<tr>
<td><strong>Modifiers:</strong></td>
<td></td>
</tr>
<tr>
<td><em>SPE3</em></td>
<td>spermidine synthase</td>
</tr>
</tbody>
</table>
Deletion of these genes decreased viral recombinant accumulation by only \(~2\)-fold.
Chapter IV

Phosphorylation of the p33 replication protein of Cucumber necrosis tombusvirus adjacent to the RNA binding site affects viral RNA replication

INTRODUCTION

Replication of plus-strand RNA viruses, which are the most abundant among plant, animal and human viruses, is performed by viral replicases (Ahlquist, 2002; Ahlquist et al., 2003; Buck, 1996). The viral replicase consists of viral proteins, including the RNA-dependent RNA polymerase (RdRp), and putative host factors. In spite of their significance in virus replication, our current understanding of the roles of viral- and host proteins is incomplete. These proteins are likely subject to posttranslational modification(s), which could affect/regulate their functions in viral replication. Indeed, reversible phosphorylation and other post-translational modifications are important for the functions of many cellular and viral proteins. For example, protein phosphorylation/dephosphorylation is frequently used as an “on-off switch”, allowing regulation of protein function. While considerable information is available for the role of phosphorylation of replication proteins in minus-strand RNA viruses, which code for P replication protein, a phosphoprotein (Barik and Banerjee, 1992; Lenard, 1999; Takacs et al., 1992), the role and significance of phosphorylation of replication proteins in plus-strand RNA viruses is currently poorly understood. The examples include phosphorylation and ubiquitinylation of Turnip yellow mosaic virus RNA-dependent RNA polymerase (RdRp) (Hericourt et al., 2000) and the Vpg of potyviruses (Puustinen et al.,
2002). Also, phosphorylation of CMV 2a RdRp protein was shown to occur late in infection, affecting its interaction with the 1a replication protein (a methyltransferase and helicase-like protein) (Kim, Palukaitis, and Park, 2002). Phosphorylation of the NS5A replication protein of hepatitis C virus (HCV) has been indicated to affect HCV replication and cellular signaling, likely facilitating establishment of chronic infections (Huang et al., 2004; Macdonald et al., 2004; Street et al., 2004). Phosphorylation of the NS5B RdRp protein of HCV has been proposed to play critical role in HCV replication, based on enhanced HCV replication in cells over-expressing protein kinase C-related kinase 2 (PRK2, a host kinase), which is known to interact with NS5B (Kim et al., 2004). In addition to the proposed role of phosphorylation in virus replication, posttranslational modification also affects the functions of other viral proteins, as demonstrated for movement (MP) and coat proteins (CP) of plus-stranded RNA viruses of plants (Ivanov et al., 2003; Waigmann et al., 2000).

Tombusviruses are important and emerging plant pathogens that are among the best characterized plant viruses (White and Nagy, 2004). They have monopartite, single-stranded RNA genomes of ~4.8 kb with mRNA-sense polarity. They belong to supergroup 2 viruses that include important plant pathogens (luteoviruses, carmoviruses, and others), human and animal pathogens (hepatitis C virus, flaviviruses, pestiviruses), and bacteriophages (Qbeta). Two of the five Tombusvirus-coded proteins, namely p33 and p92 (Fig. 4.1), were shown to be essential for tombusvirus replication (Oster, Wu, and White, 1998a; Panaviene, Baker, and Nagy, 2003; Panaviene, Panavas, and Nagy, 2005; Scholthof, Scholthof, and Jackson, 1995a) and recombination (Panaviene and Nagy, 2003). Both p33 and p92 are translated from the genomic RNA and p92 is the result of translational readthrough of the translation stop codon of p33 (Fig. 4.1). Due to the protein expression strategy of tombusviruses, the N-
terminal portion of p92 overlaps with p33 (Scholthof, Scholthof, and Jackson, 1995a). p92 contains the signature RdRp motifs and together with p33 is part of the tombusvirus replicase (Nagy and Pogany, 2000b; Panaviene et al., 2004; Scholthof, Scholthof, and Jackson, 1995a). The essential p33 replication protein is proposed to function in template selection by specifically recognizing a conserved internally located stem-loop structure within the p92 ORF (Pogany, White, and Nagy, 2005). Altogether, the p33 is likely involved in many essential replication functions, such as template RNA binding (Monkewich et al., 2005; Panavas et al., 2005b; Pogany, White, and Nagy, 2005), assembly/disassembly of the viral replicase complex (Panaviene, Panavas, and Nagy, 2005; Panaviene et al., 2004), targeting and/or attaching the replicase complexes to subcellular membranes (Navarro, Rubino, and Russo, 2004; Panavas et al., 2005b), complementary RNA synthesis, release of the newly made viral RNA progeny from the site of replication, and so on. Because posttranscriptional modifications of the replication proteins might play roles as molecular switches in the replication process by altering the properties of the replication proteins, we tested if the tombusvirus p33 is phosphorylated in infected plants. This study establishes that phosphorylation of the p33 replication protein takes place both in vivo and in vitro. Using kinase-containing plant extract, the phosphorylation sites were mapped to threonine/serine (T/S) residues located in the vicinity of the RNA-binding site, suggesting the possible role of phosphorylation in modulating the RNA-binding activity of p33. Using phosphorylation-mimicking aspartic acid (D) mutants of p33, we demonstrate that phosphorylation of the T/S residues in p33 could influence subgenomic versus genomic RNA synthesis and the ratio between plus- and minus-stranded RNAs. Phosphorylation incompetent alanine (A) mutants of CNV replicated slower than wild type in plant protoplast and showed less severe
symptoms in *Nicotiana benthamiana* plants. In summary, our data promote a regulatory role for phosphorylation of p33 in tombusvirus replication.

**MATERIALS AND METHODS**

*Detection of phosphorylated p33 and p28 replication proteins by Western blotting using phosphothreonine specific antibody or by staining with phosphoprotein-specific dye.*

Leaves (2 g) of *N. benthamiana* plants systemically infected with CNV and TCV, respectively, were used to obtain the membrane-enriched and soluble fractions as described (Panaviene et al., 2004), except we included phosphatase inhibitor mix containing 25 mM NaF, 10 mM benzamidine, 10 mM sodium orthovanadate. 1/10th of the obtained membrane-enriched fraction was loaded on a 10% SDS-PAGE gels, electrophoresed, transferred to PVDF membranes, followed by using anti-phosphothreonine and anti-phosphoserine antibodies as recommended by the manufacturer (Research Diagnostic, Inc).

The second method to detect phosphorylated p33 in CNV-infected *N. benthamiana* was based on using the phosphoprotein specific Pro-Q Diamond dye from Molecular Probes. First, we prepared the membrane-enriched fraction (see above), followed by solubilization of membrane-bound proteins with 2% Triton X-100 as described (Nagy and Pogany, 2000b; Panaviene et al., 2004). Anti-p33 antibody (prepared against CVGYLKNTPENRLIY peptide, Sigma-Genosys) was then added to the sample, followed by incubation in the cold room and then addition of protein A (Sigma) and further incubation for 60 min. The pellet was recovered after centrifugation, washed three times and then loaded on an SDS-PAGE gel. The gel was then fixed overnight with 200 ml 10% acetic acid (glacial) and 50%
methanol, followed by three time wash with 200 ml deionized water for 15 min. Staining of the gel was done overnight in the dark with 50 ml Pro-Q Diamond dye as recommended by the manufacturer (Molecular Probes). De-staining was done with 200 ml 10% methanol and 7% acetic acid for 30 min. Images of gels were taken when illuminated with 302 nm UV light.

**Preparation of host kinase-containing fraction from N. benthamiana and in vitro phosphorylation of p33**

Young leaves (5 g) of uninfected *N. benthamiana* were used to obtain the membrane-enriched fraction (see above), followed by solubilization with 4% Triton X-100 exactly as described for preparation of CNV replicase (Panaviene et al., 2004). The solubilized membrane-enriched preparation was passed through Sephadex G-25 resin (Amersham). To reduce the background in subsequent assays, the obtained host kinase preparation was mixed with phosphorylation buffer (0.1M HEPES pH 7.4, 10 mM EDTA, 5 mM DTT, 50 mM MgCl₂) in the presence of 1mM ATP to allow phosphorylation of any phosphorylatable proteins that might be present in the preparation for 20 min at room temperature (RT). This was followed by removal of ATP on a Sephadex G-25 column.

To test phosphorylation of CNV p33 in vitro by the above kinase-containing plant membrane preparation, we selected two purified recombinant p33-derived peptides, termed C10 (p33₁₈₁₋₂₄₀) and C11 (p33₂₁₁₋₂₄₀) (Rajendran and Nagy, 2003), expressed in E. coli. The cleavage of the proteins was done with Factor Xa protease (NEB) as described (Rajendran and Nagy, 2003). The in vitro phosphorylation reaction was made in 20 µl volume in the phosphorylation buffer (see above), plus 1 µl γ³²P-ATP (7000 Ci/mmol, 0.05 mCi), 2 µl
kinase-containing plant membrane fraction and 3 µg of p33-derived peptide. After incubation for 20 min, the samples were run in SDS PAGE gels, stained with coomassie blue, dried and exposed to a phospho-storage screen and imaged with a PhosphorImager.

**In vitro phosphorylation of p33 with PKC**

The affinity-purified p33 mutants (200 µg/ml, expressed in *E. coli*) (Rajendran, Pogany and Nagy 2002), carrying or lacking the T205, S210, and T211 positions, were phosphorylated in 20 µl volume containing PKC-phosphorylation buffer (20mM HEPES, 0.03% triton X-100, 4mM MOPS, 5mM 1μglycerol phosphate, 1 mM EGTA, 0.2 mM Na-orthovanadate, 0.2 mM DTT), 0.1 mM ATP (unlabeled), 5 µl PKC lipid activator (Upstate), 1 µl 2mM CaCl, 2.5 µl γ32P-ATP (7000 Ci/mmol, 0.125 mCi), and 0.5 µl PKC (Sigma), for 15 min at room temperature. Afterwards, the samples were boiled in SDS-PAGE running buffer for three min and loaded onto a large 10% SDS-PAGE gel (Bio-Rad protean II xi). The electrophoresis was performed at 200 volts for 6 hours, followed by staining with coomassie blue, drying and then scanning by a PhosphorImager.

**Construction and testing of the phosphorylation-mimicking mutants of CNV p33 in N. benthamiana protoplasts and plants**

To make site-directed mutations that replaced S/T residues with A/D residues, we used standard Quick Change XL Site-Directed Mutagenesis Kit (Stratagene) as described before (Panaviene, Baker, and Nagy, 2003) based on an infectious CNV clone (Rochon, 1991). The mutant clones were confirmed by sequencing, using Beckman Coulter DTCS Quickstart Kit and primer #958 (5’-CCGGCACGGGAGCTCAAGGTAAGGA).
Full-length RNA transcripts of CNV (Rochon, 1991), DI-72 and DI-73 (White and Morris, 1994a; White and Morris, 1994e) were prepared using standard in vitro transcription reaction with T7 RNA polymerase on Smal linearized DNA templates. *N. benthamiana* protoplasts were prepared as described before (Panaviene, Baker, and Nagy, 2003). Protoplasts were divided into aliquots (5 x10^5 protoplasts) and electroporated with 5 µg of wt or mutated CNV gRNA and 1 µg of DI RNA (when the goal was to test DI RNA replication). After electroporation, protoplasts were incubated in the dark at 22°C.

*N. benthamiana* plants were rub inoculated with 3 µg of wt or mutated CNV gRNA. Inoculated plants were kept in the green house. Samples were taken after 5, 7 and 14 dpi.

**CNV RNA extraction and Northern blotting**

Total RNA was prepared from protoplast or plants using a standard phenol/chloroform method (Kong, Wang, and Simon, 1997; Nagy, Pogany, and Simon, 2001) and the samples were electrophoresed in 1.2% agarose (in case of gRNA) or 4% polyacrylamide/8 M urea gels (for DI-72 RNA). The RNA was then transferred from the gels to Hybond XL membrane (Amersham-Pharmacia) by electrotransfer and hybridized with CNV-specific probes for the 3’ noncoding region as described (Panaviene, Baker, and Nagy, 2003). Hybridization was done in ULTRAhyb hybridization buffer at 68°C as recommended by the supplier (Ambion). The riboprobes were made in an in vitro transcription reaction with T7 RNA polymerase in the presence of [^32]P]UTP as described. The CNV DNA template was obtained using PCR with pair of primers #16 (GTAATACGACTCACTATAGGGCTGCATTTCTGCAATGTTC) and #312 (GCTGTCAGTCTAGTGGA) for detection of plus-strands, and with #31
(GTAATACGACTCACTATAGGAAATTCTCCAGGATTTC) and #806
(CCGTCTAGAGTCGTCGGTTACTGGAAGTTAC) for detection of minus-strands. The template for detection of DI-72 RNA (+) and (-) strands (corresponding to RIII/RIV) was obtained with PCR using primers #1165 (AGCGAGTAAAGACCAGACTCTTCA) and #22
(GTAATACGACTCACTATAGGAGAAAGCGAGTAAGACG) versus #157
(GGGTTTCATAGAAAGAAAAACAAAAACC) and #18
(GTAATACGACTCACTATAGGAGAAAGCGAGTAAGACG), respectively.

The two component, trans-complementation assay was performed in N. benthamiana protoplasts with selected mutants as described earlier (Panaviene, Baker, and Nagy, 2003). The mutant clones were confirmed by sequencing, as described above.

**Western blot analysis**

Total proteins from N.bentamiana plant or protoplasts infected with CNV were prepared as described (Panaviene et al., 2004). The samples were run in 10% SDS-PAGE, followed by transfer onto PVDF membrane (Bio-Rad). Immunodetection of p33 protein in the Western blot was performed by using anti-p33 antibody (generous gift of H. Scholthof).

**Analysis of replicon RNA accumulation in yeast**

*S. cerevisiae* strain INVSc1 (Invitrogen) transformed with three plasmids [i.e., pGAD-His92, pGBK-His33 and pYC-DI-72(+)Rz (Panaviene et al., 2004)] was grown in SC-ULT medium containing 2% galactose for 24 h at 30°C (Panavas and Nagy, 2003b). To extract total RNA from yeast, equal volumes of RNA extraction buffer (50 mM sodium acetate, pH 5.3, 10 mM EDTA, 1% SDS) and water-saturated phenol were added to the pelleted cells
Agarose gel electrophoresis and Northern blotting was as described above (Pogany, White, and Nagy, 2005).

Western blot to detect accumulation levels of His-tagged p33 was performed as described earlier (Pogany, White, and Nagy, 2005) using monoclonal anti-His antibodies (Amersham) and secondary alkaline phosphatase-conjugated anti-mouse antibody (Sigma). Western blots were developed using BCIP and NBT (Sigma).

RESULTS

Rationale:
Sequence comparison of p33 replication proteins of tombusviruses (Fig. 4.1A) revealed the presence of conserved S and T residues (i.e., T\textsubscript{205}, S\textsubscript{210} and T\textsubscript{211}) adjacent to the previously identified arginine (R) and proline (P)-rich RNA-binding site (termed RPR-motif, R\textsubscript{213}P\textsubscript{214}R\textsubscript{215}R\textsubscript{216}R\textsubscript{217}P\textsubscript{218})(Fig. 4.1B). These S and T residues are likely phosphorylated in cells based on NetPhos analysis (i.e., they showed the probability of 0.8-0.9 relative value on a scale of 0.0-1.0 for phosphorylation) (software: NetPhos, version 2.0, Technical University of Denmark) (Fig. 4.1B). Possible phosphorylation of these S and T residues in p33 would introduce negative charge in the vicinity of the positively-charged RPR-motif, thus likely decreasing the RNA-binding capacity of p33. In support of this model, substitutions of A for arginines (R), which reduced the overall positive charge of the RPR motif, inhibited the ability of p33 to bind to the viral RNA as well as decreased tombusvirus replication in plant protoplasts (Panaviene, Baker, and Nagy, 2003). Based on these observations, we predicted
that phosphorylation of p33 adjacent to the RPR-motif could be a regulatory “switch”, affecting its biological function(s) in tombusvirus-infected cells.

**In vivo phosphorylation of CNV p33 and the related TCV p28 replicase proteins.**

To test if the CNV p33 and the related TCV p28 replication proteins are phosphorylated in plants, first we enriched p33 and p28, respectively, in our samples by isolating membrane-enriched fractions from CNV and TCV-infected *Nicotiana benthamiana* plants [10 days postinoculation (dpi)]. This approach is based on previous findings that the membrane-fraction contains the active CNV and TCV replicases (Nagy and Pogany, 2000b; Song and Simon, 1994) and the viral replication proteins (Panaviene et al., 2004). The obtained fractions were treated with SDS, followed by SDS-PAGE electrophoresis and blotting on PVDF membranes for Western blot analysis. We used antibody specifically recognizing either phosphorylated T or S residues for the Western blot assays (Kim et al., 2004). Data obtained with the anti-phosphothreonine antibody (Fig. 4.2A) showed that membrane-enriched fraction from CNV and TCV-infected plants contained phosphorylated T residues in p33 and p28, whereas healthy plants lacked similar proteins. In contrast, the anti-phosphoserine antibody based assay was not sensitive enough to detect phosphorylated p33 and p28 (not shown). In addition, we also captured p33 with anti-p33 antibody from CNV-infected plants (using isolated membrane-enriched fraction, which was treated with 2% Triton X-100 to solubilize p33) on a protein A column, followed by elution. The recovered proteins were analyzed by SDS-PAGE, followed by staining with Pro-Q Diamond (Molecular Probes), a phosphoprotein-specific dye. This assay showed the presence of phosphorylated CNV p33 in plants (Fig. 4.2C). Overall, both the Western blot with
phosphothreonine-specific antibody and the staining assay with phosphoprotein-specific dye demonstrated that the CNV p33 (and the related TCV p28) replication protein is phosphorylated at T residue(s) in infected cells.

**In vitro phosphorylation of serine and threonine residues adjacent to the RNA-binding domain in p33**

To determine if the conserved T$_{205}$, S$_{210}$ and T$_{211}$ residues located adjacent to the RNA-binding site in p33 (Fig. 4.1A) can be phosphorylated, as predicted by NetPhos program, we developed an in vitro phosphorylation assay based on partially-purified kinase preparation obtained from healthy *N. benthamiana* plants. First, we prepared a membrane-enriched fraction from uninfected plants using the procedure developed for isolation of p33 (see above) and the CNV replicase (Panaviene et al., 2004). Then, we performed a phosphorylation assay including the kinase in the membrane-enriched preparation, $^{32}$P-labeled ATP and purified peptides including or lacking the predicted phosphorylation sites adjacent to the RPR-motif in p33 (Fig. 4.3). After the kinase reaction, the radioactively-labeled peptides were detected using SDS-PAGE/PhosphoImager analysis. This assay demonstrated that the peptide p33C$_{181-240}$ containing the predicted T$_{205}$, S$_{210}$ and T$_{211}$ phosphorylation sites could be efficiently phosphorylated in vitro by the partially-purified plant kinase (Fig. 4.3A). The second peptide (p33C$_{211-240}$), which was not phosphorylated in this assay, lacked T$_{205}$, S$_{210}$ residues and contained only T$_{211}$ residue in an unfavorable context (Fig. 4.3A). Overall, these experiments demonstrated that healthy *N. benthamiana* plants contain a membrane-associated kinase capable of phosphorylating p33 in vitro at one or more of the conserved T$_{205}$, S$_{210}$ and T$_{211}$ residues.
Based on NetPhos analysis, the probable kinase involved in phosphorylation of T\(_{205}\), S\(_{210}\) and T\(_{211}\) residues in p33 is a protein kinase C (PKC)-like kinase. Interestingly, PKC-like kinase activity has been detected in plants (Lee and Lucas, 2001; Sokolova et al., 1997). Therefore, we tested if purified PKC preparation (Sigma) could phosphorylate p33-derived peptides in vitro. The in vitro kinase assay revealed efficient phosphorylation of a 30 aa-long peptide p33C\(_{205-234}\) carrying wt sequences and less efficient phosphorylation of mutant D\(_{210}\), whereas mutants A\(_{210}\)A\(_{211}\) and D\(_{211}\) (in the context of p33C\(_{205-234}\)) were not phosphorylated at a detectable level (Fig. 4.4A). Thus, these experiments demonstrated that T\(_{211}\) residue can be phosphorylated by PKC, but no data for phosphorylation of S\(_{210}\) residue was obtained. The lack of phosphorylation of S\(_{210}\) residue in mutant D\(_{211}\) (Fig. 4.4A) could be due to the sequence change at position 211, because the sequence context of the phosphorylation site is known to affect the efficiency of phosphorylation. To circumvent this problem, we treated the wt p33C\(_{205-234}\) peptide with PKC, followed by detection of phosphorylated serine with anti-phosphoserine antibody (Fig. 4.4C). This experiment demonstrated that S\(_{210}\) residue could be phosphorylated by PKC in vitro. Therefore, we conclude that both S\(_{210}\) and T\(_{211}\) residues can be phosphorylated by PKC in vitro.

**Effect of phosphorylation of CNV p33 on virus replication based on the use of phosphorylation-mimicking mutants**

Because phosphorylation of neutral S and T residues introduces negative charge around these residues, effect of phosphorylation on protein function can be tested by mutating these residues to D, which is negatively-charged. In contrast, the nonphosphorylated status of S and T residues can be mimicked by mutating S and T residues to A, which is neutral. Thus, D
mutation mimics the phosphorylated, while A mutation mimics the nonphosphorylated forms of S and T residues as previously demonstrated in numerous assays (Freye-Minks, Kretsinger, and Creutz, 2003; Jabbur, Huang, and Zhang, 2001; Jia et al., 2004; Kim, Palukaitis, and Park, 2002; Waigmann et al., 2000). Another advantage is that both D and A residues are similar in sizes to S and T residues, thus likely minimizing the sterical/structural changes caused by these mutations. In spite of the above advantages, D mutations only “mimic” phosphorylation, thus results obtained with D mutants could possibly cause problems unrelated to phosphorylation.

A set of twelve single, double, and triple A/D mutations were introduced at T\textsubscript{205}, S\textsubscript{210} and T\textsubscript{211} within the p33/p92 open reading frame in an infectious cDNA clone of CNV (Rochon, 1991). Note that the above mutations changed the sequences of the overlapping p33 and p92 replication proteins (Fig. 4.1A). We then generated CNV RNA transcripts in vitro, followed by electroporation to \textit{N. benthamiana} protoplasts to test the replication features of the mutated CNV proteins. Total RNA was extracted 12 and 24 hours after electroporation and analyzed by Northern blotting. These experiments demonstrated that single and double D for S/T substitutions (at positions S\textsubscript{210} and T\textsubscript{211}) reduced the ability of CNV to replicate to detectable levels (constructs with single mutations, such as D\textsubscript{210}, and D\textsubscript{211}, and double mutation, D\textsubscript{210}D\textsubscript{211}, Fig. 4.5A). In addition, mutant D\textsubscript{205} replicated poorly (<1% at 12, and 7% of wt at 24 hours time points, lanes 18 and 26, Fig. 4.5A). These data suggest that negative charge(s) at positions S\textsubscript{210}, T\textsubscript{211} and, to a less extent, at T\textsubscript{205} inhibits CNV replication.

All single and double A mutants (at T\textsubscript{205}, S\textsubscript{210} and T\textsubscript{211}) mimicking the nonphosphorylated form of p33/p92 supported replication, albeit with reduced efficiency (34-58% of wt CNV gRNA) when compared to wt (Fig. 4.5A). Even the triple A mutant
(construct A$_{205}$A$_{210}$A$_{211}$, accumulation level of 13-21% of wt, lanes 19 and 27, Fig. 4.5A) was replication competent, suggesting that nonphosphorylated residues at positions 205, 210 and 211 are important for the function(s) of replicase proteins in vivo. However, because all these A mutants replicated less efficiently than wt, it is likely that phosphorylation of p33/p92 plays a regulatory role during virus replication (see below).

To demonstrate that the above mutations at T$_{205}$, S$_{210}$ and T$_{211}$ in CNV affected replication due to changes in the p33/p92 replication proteins, but not due to altered CNV RNA structure, we tested the ability of the mutated CNV replication proteins to support replication in trans, using a model template, DI-73 RNA (Fig. 4.5B). Co-electroporation of DI-73 RNA with the CNV helper (wt or D$_{210}$D$_{211}$) into plant protoplasts demonstrated that CNV D$_{210}$D$_{211}$ did not support replication of DI-73 RNA in trans, indicating that the phosphorylation-mimicking replication protein(s) is not functional in vivo.

To test if the above CNV mutants expressed p33 in comparable amounts, we performed Western blotting on total protein extracts obtained from protoplasts 12, 18 and 24 hpi with an antibody specific for the C-terminal portion of p33 (outside the mutagenized domain of p33). We found that D$_{210}$D$_{211}$ p33 was not detectable (Fig. 4.5C), whereas the amount of D$_{210}$A$_{211}$ was only ~10% of wt p33 (Fig. 4.6B). The accumulation of A$_{210}$A$_{211}$ p33 was comparable to wt p33. The reduced accumulation of phosphorylation-mimicking p33 mutants in plant protoplasts can be explained by several alternative models including decreased stability and/or production of these p33 forms.

_Hypophosphorylation-mimicking p33 mutants affected plus-strand RNA and subgenomic RNA synthesis in vivo_
To test for possible regulatory roles of phosphorylation of p33/p92 in replication, we performed time-course experiments in protoplasts with a viable nonphosphorylation-mimicking mutant, a hypophosphorylation mutant (i.e., D_{210}A_{211} carrying only a single D mutation) and the wt (Fig. 4.6). Total RNA samples were prepared from protoplasts and analyzed with Northern blotting as described above. These experiments revealed that nonphosphorylation-mimicking A_{210}A_{211} and the partial (hypo)-phosphorylation-mimicking D_{210}A_{211} mutants generated 5-6 times more subgenomic (sg) RNA2 than sgRNA1 at early time points, whereas sgRNA2 was only ~3-fold more abundant than sgRNA1 in the wt CNV infections (Fig. 4.6A). Altogether, sgRNA2 was the most abundant viral RNA species in A_{210}A_{211} and D_{210}A_{211} infections, in contrast to wt CNV, which produced the gRNA in the largest amount (except at the early time point, Fig. 4.6A). These data suggest that phosphorylation status of S_{210} and T_{211} in p33/p92 could affect the level and ratio of sgRNA1 and sgRNA2 synthesis in CNV infections.

To test if mimicking mutations could also affect minus- versus plus-strand synthesis during CNV replication, we took advantage of DI-72 RNA replicon, which produces minus-strand RNA at easily detectable levels (Monkewich et al., 2005). In this assay, *N. benthamiana* protoplasts were co-electroporated with wt or mutated (A_{210}A_{211}, A_{205}A_{210}A_{211} and D_{210}A_{211}) CNV RNA and DI-72 RNA, followed by detection of plus- and minus-strand DI-72 RNA levels in total RNA samples by Northern blotting (Fig. 4.7). We observed that DI-72 minus-strand synthesis took place at low levels (10-20%) between 4-12 hours post infection (hpi), followed by increased synthesis between 12-24 hours in the presence of wt or any of the mutated helper CNV. On the contrary, plus-strand RNA synthesis showed remarkable differences depending on the wt and mutated helper CNV (Fig. 4.7A). For
example, detectable level of DI-72 plus-strands was observed as early as 6 and 8 hpi in the presence of wt CNV helper, whereas A205A210A211 and D210A211 mutants of CNV accumulated DI-72 plus-strands at a detectable level only at 12 hpi (Fig. 4.7A). Interestingly, the relative level of minus-strand accumulation (when compared to wt level of minus strands representing 100%) supported by A205A210A211 and D210A211 mutants was ~20% higher than the relative level of plus-strand accumulation (Fig. 4.7C). The altered ratio between minus and plus-strand accumulation means that the RNA synthesis was less asymmetrical with the above mutants than in the case of wt CNV replicase. Overall, these observations indicate that the phosphorylation status of p33 at positions 205, 210 and 211 might affect the level of asymmetry between plus- versus minus-strand synthesis.

**Phosphorylation-mimicking mutations inhibit the function of p33 more than p92 in protoplasts**

To test if phosphorylation-mimicking mutations inhibit p33 and/or p92 functions in plant cells, we utilized a two-component complementation-based replication assay developed by Oster et al. (Oster, Wu, and White, 1998a), which allows separate mutagenesis of p33 and p92. Expression of p33 and p92 (the p92 ORF contains a tyrosine mutation to suppress termination at the end of p33 ORF, Fig. 4.8A) from two separate RNAs leads to efficient replication of DI-p33, whereas CNV-Y replicates poorly in *N. benthamiana* protoplasts [Fig 4.8B-C and (Panaviene, Baker, and Nagy, 2003; Panaviene and Nagy, 2003)]. We found that co-expression of the unmodified p92Y with D210D211 mutation in p33 led to undetectable level of DI-p33 replication (Fig. 4.8B), suggesting that phosphorylation renders p33 nonfunctional. On the contrary, the D210D211 mutation in p92Y inhibited wt DI-p33
replication by only 50% (Fig. 4.8C), suggesting that phosphorylation of p92 RdRp protein at 210-211 positions interfered with its replication function to a lesser extent than phosphorylation of p33.

Yeast based tombusvirus replication assay confirms a role for the phosphorylation status of p33 replication protein in tombusvirus replication

To test if the phosphorylation status of p33 affects its stability and/or functions, we took advantage of the availability of TBSV replicon-based replication assay in yeast expressing the CNV p33 and p92 replication proteins, which supports efficient replication of DI-72 RNA (Panavas and Nagy, 2003b). In the yeast system, the expression of p33 and p92 is driven from separate expression vectors via the constitutive ADH1 promoter. The advantage of the yeast replication system is that we can study the replication of the DI-72 RNA without the need of replication and gene expression of the helper virus. This allowed us to re-test if several mimicking mutants of p33 could support tombusvirus replication in yeast.

We found that co-expression of wt p92 together with p33 carrying D for T mutation (see constructs D_{211}, D_{210}D_{211}, A_{210}D_{211}, D_{205}D_{210}D_{211}, Fig. 4.9A-C) did not result in DI-72 RNA replication, based on undetectable levels of DI-72 plus- and minus-strands in yeast. Similar mutations at T_{205} and S_{210}, (see constructs D_{205} and D_{210}, Fig. 4.9A-C) also reduced replication of DI-72 RNA (down to 12 and 8% of wt levels), but to a lesser extent than observed in plant cells. Western-blot analysis of total protein extracts of yeast strains revealed that, unlike in the plant protoplast-based replication assay, the p33 levels in yeast were similar for wt and the mutated p33 (Fig. 4.9D), suggesting that the stability of p33 is not altered by these mutations. Instead, these data indicate that the phosphorylation-mimicking D
mutations make p33 less functional/nonfunctional in tombusvirus replication. On the contrary, the nonphosphorylatable A mutations had less pronounced effects on p33 functions (16-57% level of replication, see constructs A\textsubscript{210}A\textsubscript{211}, A\textsubscript{205}A\textsubscript{210}A\textsubscript{211}, Fig. 4.9A-C).

We also tested the ability of p33 mutants to support plus- versus minus-strand synthesis of the replicon RNA in yeast. Northern-blot analysis of total RNA samples revealed that the relative-levels of minus-strand (as compared to wt level representing 100%) versus plus-strand accumulation were two-to-five fold higher in case of four out of five functional p33 mutants (i.e., A\textsubscript{210}A\textsubscript{211}, A\textsubscript{205}A\textsubscript{210}A\textsubscript{211}, D\textsubscript{210}A\textsubscript{211}; and D\textsubscript{205}, Fig. 4.9B-C) than that observed with wt p33. However, one p33 mutant (i.e., D\textsubscript{210}) showed lower relative level of minus-strand versus plus-strand synthesis (Fig. 4.9B-C). Altogether, these data indicate that nonphosphorylation or partial phosphorylation-mimicking mutations in p33 at positions 205, 210, 211 selectively decrease the level of plus-strand RNA accumulation, likely resulting from less asymmetrical RNA synthesis.

Delayed accumulation of an nonphosphorylation-mimicking mutant in N. benthamiana plants

To test if CNV expressing the nonphosphorylation-mimicking A\textsubscript{210}A\textsubscript{211} p33 mutant could replicate with similar efficiency as the wt CNV (expressing phosphorylation-competent p33) in whole plants, we extracted total RNA from CNV-infected leaves 5, 7 and 14 days after inoculation (dpi). We found that wt CNV accumulated viral-specific RNAs to higher levels than A\textsubscript{210}A\textsubscript{211} did at 5 and 7 dpi (Fig. 4.10A), suggesting that virus replication/accumulation is more robust in wt CNV infections. Interestingly, the difference in accumulation level between wt and A\textsubscript{210}A\textsubscript{211} gRNA was more significant in plants than in protoplasts (2-to-2.5
fold difference in plants at 5 and 7 dpi versus 5% in protoplast at 24 hpi). This might be due
to an indirect effect of phosphorylation of wt p33 on cell-to-cell movement and/or stability of
gRNA. The symptoms caused by the wt CNV infection were also more intensive than in
A_{210A_{211}} infections (Fig. 4.10B), and led to the death of most leaves by 14 dpi. On the
contrary, symptoms caused by A_{210A_{211}} were mild and did not lead to necrosis. Interestingly,
viral RNA accumulation was three-fold higher in A_{210A_{211}}-infected than in wt CNV-infected
plants at 14 dpi, likely due to the sickness of wt CNV-infected plants (Fig. 4.10A).
Altogether, these data demonstrate that the nonphosphorylatable A mutations at positions 210
and 211 in p33 have pronounced effects on the pathogenicity of CNV in *N. benthamiana*.

**DISCUSSION**

Detection of phosphorylated replication proteins in plus-strand RNA virus-infected
hosts is a challenging task due to the low abundance of viral replication proteins, their
localization to less amenable cellular membranes, the dynamic and reversible nature of
phosphorylation and temporally existence of some phosphorylated proteins. By using both
anti-phosphothreonine antibody-based Western blot and phosphoprotein-specific staining
assays, we demonstrated the accumulation of phosphorylated CNV p33 (and the related TCV
p28) replication protein in plants, while we could not detect phosphorylated CNV p92,
possibly due to its very low abundance [~20-fold less than p33, (Scholthof, Scholthof, and
Jackson, 1995a)]. In addition, we obtained a partially-purified kinase preparation from
solubilized membrane-enriched fraction of *N. benthamiana*, which was able to phosphorylate
the CNV p33 replication protein at conserved S/T residues (likely at T\textsubscript{205}, S\textsubscript{210} and T\textsubscript{211}) adjacent to the RPR motif, an RNA-binding site (Panaviene, Baker, and Nagy, 2003; Rajendran and Nagy, 2003). On the contrary, the soluble (cytoplasmic) fraction of cells did not contain kinases capable of phosphorylation of the same region in p33 in vitro (not shown). Because p33 is membrane-associated in cells and tombusvirus replication occurs in membrane-containing compartments (Navarro, Rubino, and Russo, 2004; Panavas et al., 2005b; Panaviene et al., 2004; Rubino and Russo, 1998; Scholthof, Scholthof, and Jackson, 1995a) it is likely that phosphorylation of p33 takes place in cellular membranes [peroxisomal membrane for CNV (Panavas et al., 2005b)] in infected cells. Thus, the kinase present in the membrane-enriched fraction is likely responsible for in vivo phosphorylation of p33. This host kinase is functionally similar to PKC, based on phosphorylation site prediction and demonstration that purified PKC preparation could also phosphorylate recombinant p33 at residues S\textsubscript{210} and T\textsubscript{211} in vitro (Fig. 4.4). Altogether, we conclude that p33 gets phosphorylated in CNV-infected plants by a membrane-associated host kinase.

The effects of p33 phosphorylation during CNV infection were studied using phosphorylation-mimicking mutants (see rationale in the results section). The advantage of using mimicking mutants is that it simplifies the interpretation of data due to the presence of p33 representing either the “nonphosphorylated” (i.e, alanine mutants) or “phosphorylated” (aspartic acid mutants) forms of p33 during the entire infection. Disadvantage is that the dynamics of phosphorylation cannot be fully examined with these mutants. Nevertheless, based on data obtained with the mimicking mutants, the model predicting that the nonphosphorylated form of p33 (we refer only to phosphorylation of T\textsubscript{205}, S\textsubscript{210}, and T\textsubscript{211} in the text below) is the functional form is supported, while the fully phosphorylated form of p33
seems to be nonfunctional in replication. Accordingly, in vitro phosphorylation of p33 and phosphorylation mimicking mutations in p33 (at S$_{210}$ and T$_{211}$ positions) both inhibited the ability of p33 to bind to the viral RNA in vitro (see Stork et al., accompanying paper), suggesting that phosphorylation and mimicking mutations had similar effect on p33 function. However, we cannot completely exclude that results obtained with D mutants could possibly cause changes in the structure of p33 unrelated to phosphorylation.

The mimicking mutants were also useful to demonstrate that similar phosphorylation of p92 RdRp protein (at T$_{205}$, S$_{210}$, and T$_{211}$), which overlaps with p33 within its N-terminus, had only limited effect on its activity in protoplasts. Thus, we suggest that phosphorylation at T$_{205}$, S$_{210}$, and T$_{211}$ mainly affects p33 function(s) during CNV infections. Phosphorylation of p33 and p92 at additional amino acid residues was not addressed in this paper.

In addition to the nonphosphorylated and fully phosphorylated forms of p33, we suggest that hypophosphorylation of p33 at either T$_{205}$, S$_{210}$, or T$_{211}$ positions might affect its activity/function. For example, we observed changes in the ratio of sgRNA1 versus sgRNA2 and of plus- versus minus-strand levels (i.e., strand asymmetry was reduced) with mimicking mutants affecting only one or two of the predicted phosphorylation sites. These observations open up the possibility that hypophosphorylation of p33 might be involved in (i) timing and/or fine tuning of subgenomic RNA synthesis (sgRNA1 versus sgRNA2) and (ii) controlling the timing and/or ratio of plus- versus minus-strand synthesis. For example, hypophosphorylation of p33 might modify the activity of the CNV replicase, affecting the timing of production of different viral RNA species and the ratio of replication products. In contrast, full phosphorylation of p33 (especially phosphorylation of T$_{211}$ seems to be the most critical) might shut down replication, which could be beneficial for viral RNA encapsidation,
cell-to-cell movement and other processes by releasing the viral RNA from replication (see the accompanying paper by Stork et al.). Indeed, the nonphosphorylated A_{210}A_{211} mutant showed delay in gRNA accumulation in systemically-infected leaves when compared to the wt CNV, suggesting less efficient virus spread in plants and/or decreased stability of mutant A_{210}A_{211} gRNA.

The amount of p33 replication protein was not the same in infections with wt and phosphorylation-mimicking CNV mutants in _N. benthamiana_ protoplasts (Fig. 4.5C and 4.6B). For example, protoplasts containing a replication incompetent mutant (D_{210}D_{211}) and an inefficiently replicating mutant (D_{210}A_{211}) accumulated p33 at undetectable and reduced levels, respectively (Fig. 4.5C and 4.6B). This suggests that phosphorylation-mimicking aspartic acid mutations (and analogously phosphorylation) might promote rapid p33 degradation, thus affecting the stability of p33. Although we cannot fully exclude that the above D mutations facilitate p33 degradation, this model is not supported by the observation that the amounts and stability of wt and phosphorylation-mimicking D mutants of p33 were comparable in yeast, expressing these proteins constitutively from plasmids (Fig. 4.9D). Alternatively, it is possible that reduced accumulation of p33 carrying the phosphorylation-mimicking mutations is due to the production of limited amount of new CNV RNA progeny in cells. We propose that some of the new CNV RNA progeny could participate in new rounds of translation to boost p33/p92 production, and thus replication, in infected cells. The proposed translational boost might not take place for CNV RNA expressing p33 with the phosphorylation-mimicking mutations that produce only limited amount of new progeny plus-strands.
The phosphorylation sites studied in this paper are located adjacent to the positively charged RNA-binding domain in p33 (Fig. 4.1B). Therefore, it is possible that the observed effects of the phosphorylation-mimicking mutants on CNV replication are due to altered RNA-binding capacity of p33. Accordingly, we show in the accompanying paper (Stork et al.) the inhibitory effect of phosphorylation of these sites on RNA-binding by the recombinant p33. Thus, phosphorylation of p33 is likely an important step in CNV replication by affecting replicase and template interactions.
**Fig. 4.1.** Location of the predicted phosphorylation sites in the CNV replication proteins. (A) The CNV genomic RNA containing five open reading frames (indicated with boxes) and noncoding sequences (solid lines) is shown schematically. (B) The known functional domains in p33 are shown schematically at the top. The amino acid sequence of the RPR-motif and the predicted phosphorylation sites are shown for selected members of Tombusviridae. The following abbreviations were used: TBSV, *Tomato bushy stunt virus* (cherry strain); CNV, *Cucumber necrosis virus*; AMCV, *Artichoke mottled crinkle virus*; CymRSV, *Cymbidium ringspot virus*; PLV, *Pear latent virus*; PNSV (*Pelargonium necrotic spot virus*); CBV (Cucumber Bulgarian virus); CIRV, *Carnation Italian ringspot virus*; and TCV, *Turnip crinkle virus* (Carmovirus, Tombusviridae).
Fig. 4.2. Detection of in vivo phosphorylated replication proteins of CNV and TCV. (A) Detection of CNV p33 in the enriched membrane fraction by using anti-p33 antibody. (B) Detection of phosphorylated CNV p33 and TCV p28 proteins in membrane-enriched fractions prepared from mock-, CNV and TCV-infected *N. benthamiana* plants by Western blotting with anti-phosphothreonine antibody. Replication proteins p33 of CNV and p28 of TCV are indicated by arrowheads, whereas a band of unknown origin (present only in samples from CNV-infected plants) is marked with an asterisk. (C) Detection of CNV p33 in membrane-enriched fraction of *N. benthamiana* by staining with phosphoprotein-specific Pro-Q Diamond dye. The CNV p33 was captured on protein A column with anti-p33 antibody from the membrane-enriched fraction prior to loading to SDS-PAGE.
**Fig. 4.3.** In vitro phosphorylation of p33 by a plant kinase. (A) Two peptides derived from p33 were expressed as an MBP fusion in *E. coli*, affinity-purified, cleaved off from MBP with *Factor Xa* and treated with a membrane-enriched preparation obtained from healthy *N. benthamiana* plants in the presence of γ[^32]P]ATP. The arrowhead shows the phosphorylated peptide (p33211-240), whereas asterisk pinpoints the nonphosphorylated peptide (p33219-240). (B) Coomassie blue-stained SDS-PAGE shows the migration of the two peptides used in the phosphorylation assay.
**Fig. 4.4.** In vitro phosphorylation at S210 and T211 residues of p33 by PKC. (A) A p33-derived peptide (p33205-234) and its derivatives with the shown mutations were expressed as MBP fusions in *E. coli*, affinity-purified and treated with PKC in the presence of γ-[\(^{32}\)P]ATP. The arrowheads show the phosphorylated peptides, whereas asterisks pinpoint phosphorylated PKC and high molecular weight proteins of unknown origin (B) Coomassie blue-stained SDS-PAGE shows the migration of the peptides (marked with an arrowhead) used in the phosphorylation assay. (C). Detection of PKC-mediated phosphorylation of S210 using anti-phosphoserine antibody. The reagents used in this assay were the same as in panel A. After the in vitro phosphorylation assay, the samples were loaded on an SDS-PAGE gel, blotted to a PVDF membrane and treated with anti-phosphoserine antibody. Additional details are as described in panel (A).
Fig. 4.5. Comparison of accumulation of CNV gRNA carrying phosphorylation-mimicking mutations adjacent to the RPR-motif in p33/p92 gene. The full-length, infectious CNV gRNA (2 µg) was electroporated into *N. benthamiana* protoplasts (5x10⁵ cells per
experiment). Total RNA was isolated, electrophoresed and blotted onto a nylon membrane, followed by probing with $^{32}$P-labeled RNA specific for gCNV (see Materials and Methods). The same amount (5 µg) of total RNA was used for loading onto the gels (based on the estimation of the host ribosomal RNA in ethidium bromide-stained gels). The positions of the CNV gRNA, sgRNA1 and sgRNA2 are depicted on the right. The names of the constructs are shown on the top. The samples were taken 12 and 24 hours after electroporation. The experiment was repeated three times. (B) The ability of CNV mutants to support DI-73 RNA replication in trans. The full-length DI-73 RNA (1 µg) was co-electroporated with the infectious CNV gRNA (2 µg, mutated or wt) into N. benthamiana protoplasts (5x10$^5$ cells per experiment). Sample preparation and Northern-blotting was done as described in the legend to panel A, except the probing was done with $^{32}$P-labeled RNA specific for DI-73 RNA (see Materials and Methods). (C) Accumulation of CNV p33 at 24 hpi. Western blot of total protein extracts was done with anti-p33 antibody.
Fig. 4.6. Effect of phosphorylation-mimicking mutations on accumulation of (A) CNV (+) RNAs and (B) p33 replication protein. The relative accumulation levels of CNV RNAs are compared to wt CNV gRNA accumulation at 24 hpi (100%). The bottom line shows the ratio of sgRNA2 versus sgRNA1 in each sample. The arrowheads indicate the position of p33 in the blot.
Fig. 4.7. Effect of phosphorylation-mimicking mutations on accumulation of DI-72 RNA \textit{in trans}. The relative accumulation levels of (A) plus-stranded and (B) minus-stranded DI-72 RNAs are compared to infections containing the wt CNV gRNA accumulation at 24 hpi.
[100% separately for both (+) and (-)RNAs, indicated by solid and dotted lines, respectively].

(C) Graph showing the relative accumulation of (+) and (-) strands in wt and mutated CNV infections in *N. benthamiana* protoplasts.
**Fig. 4.8.** Phosphorylation-mimicking mutations affect the function of p33, and a lesser extent, p92. (A). Schematic representation of the two constructs used in two component complementation-based replication studies. The stop codon at the end of the p33 gene in gCNV-Y was mutated to a tyrosine codon (Y) to prevent the production of p33 from this RNA. The second RNA (DI-p33) was derived from a DI RNA that carried a translation competent p33 gene. (B) Northern blot analysis of replication of DI-p33 mutants in the two-component system. The gCNV-Y RNA (2 µg) was co-electroporated with DI-p33 (2 µg) carrying a given mutation into *N. benthamiana* protoplasts. Sample preparation and Northern-blotting was done as described in the legend to Fig. 4.5. The position of the DI-p33 is depicted on the left. (C) Northern blot analysis of replication of gCNV-Y mutants in the two-component system. See panel B for details.
Fig. 4.9. Effect of phosphorylation-mimicking mutations on DI RNA replication in yeast. (A) Ethidium-bromide stained agarose gel showing the accumulation levels of DI-72 RNA in yeast total RNA extracts. Yeast strain Sc1 was co-transformed with three plasmids coding for full-length wt or mutated p33, wt p92Y and wt 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. (B) Northern blot analysis of accumulation levels of (+)DI-72 and (C) (-)DI-72 RNA in yeast. The relative level of DI-72 RNA replication in yeast [based on quantification of the Northern blot using a phosphorimager and ImageQuant (v1.2) software from three separate experiments], is shown below the blots (100% equals with DI-72 RNA levels in yeast expressing wt p33/p92). (D) Western blot analysis of the amount of p33 present in yeast with anti-His antibody.
Fig. 4.10. Nonphosphorylation-mimicking mutant of CNV shows delay in gRNA accumulation and symptom formation in *N. benthamiana* plants. (A) Northern blot analysis of accumulation levels of CNV (+)RNAs in *N. benthamiana* plants. Samples were taken from
two systemically-infected leaves/plant and two separate plants were used in each experiment (top and bottom panels). The relative levels of gRNA, sgRNA1 and sgRNA2 are shown below the blots (100% equals with wt CNV gRNA level at 7 dpi). (B) Symptoms of mock- and CNV-infected plants at 7 dpi.

Copyright © Natalia Shapka 2006.
Chapter V

SUMMARY

Rapid evolution of RNA viruses with mRNA-sense genomes is a major concern to health and economic welfare due to the devastating diseases these viruses inflict on humans, animals and plants. RNA recombination can be facilitated by recombination signals present in viral RNAs. Among such signals are short sequences with high AU contents that constitute recombination hot spots in *Brome mosaic virus* (BMV) and retroviruses. In this paper, we demonstrate that a defective interfering (DI) RNA, a model template associated with *Tomato bushy stunt virus* (TBSV), a tombusvirus, undergoes frequent recombination in plants and protoplast cells when it carries the AU-rich hot spot sequence from BMV. Similar to the situation with BMV, most of the recombination junction sites in the DI RNA recombinants were found within the AU-rich region. However, unlike BMV or retroviruses, where recombination usually occurred with precision between duplicated AU-rich sequences, the majority of TBSV DI RNA recombinants were imprecise. In addition, only one copy of the AU-rich sequence was essential to promote recombination in the DI RNA. The selection of junction sites was also influenced by a putative cis-acting element present in the DI RNA. We found that this RNA sequence bound to the TBSV replicase proteins more efficiently than did control non-viral sequences, suggesting that it might be involved in replicase “landing” during the template switching events. In summary, evidence is presented that a tombusvirus can use the recombination signal of BMV. This supports the idea that common
AU-rich recombination signals might promote interviral recombination between unrelated viruses.

To test if host genes can affect the evolution of RNA viruses, we used a *Saccharomyces cerevisiae* single-gene deletion library, which includes ~80% of yeast genes, in RNA recombination studies based on a small viral replicon RNA derived from *Tomato bushy stunt virus*. The genome-wide screen led to the identification of five host genes, whose absence resulted in rapid generation of novel viral RNA recombinants. Thus, these genes normally suppress viral RNA recombination, but in their absence hosts become viral recombination “hotbeds”. Four of the five suppressor genes are likely involved in RNA degradation, suggesting that RNA degradation could play a role in viral RNA recombination. On the contrary, deletion of four host genes inhibited virus recombination, indicating that these genes normally accelerate the RNA recombination process. Comparison of deletion strains with the lowest and the highest recombination rate revealed that host genes could affect recombinant accumulation by up to 80-fold. Overall, our results demonstrate for the first time that a set of host genes have major effect on RNA virus recombination and evolution.

Replication of the non-segmented, plus-stranded RNA genome of *Cucumber necrosis tombusvirus* (CNV) requires two essential overlapping viral-coded replication proteins, the p33 replication co-factor and the p92 RNA-dependent RNA polymerase. In this paper, we demonstrate that p33 is phosphorylated in vivo and in vitro by a membrane-bound plant kinase. Phosphorylation of p33 was also demonstrated in vitro by using purified protein kinase C. The related p28 replication protein of *Turnip crinkle virus* was also found to be phosphorylated in vivo, suggesting that posttranslational modification of replication proteins
is a general feature among members of the large *Tombusviridae* family. Based on in vitro studies with purified recombinant p33, we show evidence for phosphorylation of threonine and serine residues adjacent to the essential RNA-binding site in p33. Phosphorylation-mimicking aspartic acid mutations rendered p33 nonfunctional in plant protoplasts and in yeast, a model host. Comparable mutations within the pre-readthrough portion of p92 did not abolish replication. The nonphosphorylation-mimicking alanine mutants of CNV were able to replicate in plant protoplasts and in yeast, albeit with reduced efficiency when compared to the wild type. These alanine mutants also showed altered subgenomic RNA synthesis and a reduction in the ratio between plus- and minus-strand RNAs produced during CNV infection. These findings suggest that phosphorylation of threonine-serine residues adjacent to the essential RNA-binding site in the auxiliary p33 protein likely plays a role in viral RNA replication and subgenomic RNA synthesis during tombusvirus infections.

Copyright © Natalia Shapka 2006.
REFERENCES


VITA


Education:

Kiev State University, department of Virology. 1987-1992

Master degree in Virology

Kiev State University, 1987-1990

Bachelor degree

Professional positions:

Belozersky Institute of Physico-Chemical Biology, Moscow State University. 1996-1998

Position: Research assistant

Zlabolotny Institute of Microbiology and Virology, Kiev 1992-1999

Position: Research assistant

Publications:

