

GCD10 AND GCD14 PROTEINS FACILITATE TOMATO BUSHY
STUNT VIRUS (TBSV) RNA REPLICATION IN *SACCHAROMYCES*
CEREVISIAE

Hyukho Sheen

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM OF BIOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

August 2018

© Hyukho Sheen, 2018

Abstract

Tomato bushy stunt virus (TBSV) is a positive-sense RNA virus and its genome replication requires two viral proteins, p33 and RNA-dependent RNA polymerase p92. Previous studies showed that the 5' untranslated region (5'UTR) of the genome is important, albeit not essential, for viral RNA replication in plants. Here, my study shows that the 5'UTR is also important for viral RNA replication in *S. cerevisiae*, a model host for TBSV. Using the TBSV 5'UTR as bait RNA and a streptotag-based RNA pull-down assay, GCD10 and GCD14 were identified as specific 5'UTR-binding proteins. Subsequent electrophoretic mobility shift assay (EMSA) shows that GCD10 and GCD14 proteins bind directly to the 5' UTR and both subunits are required for the binding. To investigate the possible importance of GCD10 and GCD14 proteins in TBSV RNA replication, the replication of a TBSV RNA replicon was analyzed in wild type, GCD10 knockdown and GCD14 knockout yeast strains. Reduction of GCD10 protein and deletion of GCD14 protein in the yeast strains significantly inhibited TBSV subviral defective interfering (DI) RNA replication, suggesting that both GCD10 and GCD14 proteins facilitate TBSV DI replication. Importantly, through western analyses, we confirmed that the reduced levels of the replicon in the yeast strains were not due to altered p33 and p93 protein levels. To gain mechanistic insights on TBSV RNA replication, 5'UTR-lacking DI (Δ 5'UTR-DI) was also tested in the wildtype, GCD10 knockdown and GCD14 knockout strains. Since Δ 5'UTR-DI replicates without the 5' UTR, its replication was expected to be independent of GCD10 knockdown or GCD14 knockout. As expected, replication of the Δ 5'UTR-DI was not inhibited by GCD10 knockdown, showing that GCD10 protein functions only through the 5' UTR. However, replication of Δ 5'UTR-DI was inhibited in GCD14 knockout strain, suggesting that GCD14 protein functions through the 5' UTR, but is also able to affect replication

by a different mechanism. Although the functions of the 5' UTR binding proteins appear to be complex, my results show that GCD10 and GCD14 proteins are important for TBSV RNA replication in *S. cerevisiae*. Mechanistic perspectives of the functions of these two proteins are discussed.

Acknowledgements

I would like to thank my supervisor, Dr. K. Andrew White for his guidance during my Ph.D. studies. He is one of the most intelligent persons I have personally met, and he is also a very good person. The kindness and understanding he showed me throughout my studies will never be forgotten, and I believe all his past graduate students he supported feel the same. I also thank my supervisory committee members, Dr. Katalin A. Hudak and Dr. Terrance Kubiseski. Dr. Hudak is the one of the most friendly professors I have ever known. I appreciate Dr. Terrance Kubiseski for his concerns and kind support during my studies.

I would also like to thank my examining committee members, Dr. Philip Johnson, Dr. Michael Scheid and Dr. Aiming Wang for taking their valuable time for my thesis.

I also thank the present and past members in Dr. White laboratory for their help and friendship, especially Dr. Baodong Wu.

I am grateful to my parents, sisters and brother for their support during my studies, and I especially thank my mother. I would have not been able to come this far without her caring and encouragement. I am forever grateful to my mother.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iv
Table of Contents.....	v
List of Figures.....	ix
List of Abbreviations.....	xi
1. Chapter 1. Introduction.....	1
1.1. Virus Classification.....	2
1.1.1. Positive-strand viruses.....	3
1.1.2. Family Tombusviridae.....	4
1.1.3. Genus <i>Tombusvirus</i>	4
1.2. Functions of TBSV proteins.....	6
1.3. TBSV gene expression.....	13
1.4. TBSV RNA replication.....	15
1.4.1. Defective Interfering RNAs.....	16
1.4.2. Yeast (<i>Saccharomyces cerevisiae</i>) as a model host for DI replication.....	21
1.4.3. Principles behind yeast single gene knockout strains.....	23
1.4.4. Host factors important for TBSV viral RNA replication.....	23
1.5. The objectives of my dissertation.....	29

2. Chapter 2. Results.....	31
2.1. PART 1: Identification of the 5' UTR binding protein(s).....	32
2.1.1. Rationale of searching for 5' UTR binding protein(s).....	32
2.1.2. Structures in the 5' UTR RNA are required for DI72 replication in both yeast plant host cells.....	33
2.1.3. The 5' UTR facilitates DI72 minus-strand RNA production in yeast.....	36
2.1.4. Principles of streptotag RNA affinity isolation of RNA-binding proteins and subsequent identification.....	38
2.1.5. Streptotag RNA does not interfere with DI72 replication in plants.....	38
2.1.6. Binding of the streptotagged 5' UTR to the column.....	40
2.1.7. Identification of the 5' UTR-binding protein(s).....	42
2.2. PART 2: Characterization of Binding of GCD10/GCD14 Proteins to the 5' UTR.....	45
2.2.1. General information about GCD10 and GCD14 proteins.....	45
2.2.2. The T-shape domain is required for GCD10 and GCD14 binding to the 5' UTR.....	45
2.2.3. Expression and purification of GCD10 and GCD14 proteins.....	46
2.2.4. Electrophoretic mobility shift assay (EMSA).....	48
2.3. PART 3: Functionality of GCD10 and GCD14 in DI72 replication.....	50
2.3.1. GCD10 protein is important for DI72 replication in yeast.....	50
2.3.2. GCD10 protein is not required for replication of a 5' UTR-lacking DI (Δ 5'UTR-DI)..	55
2.3.3. GCD14 protein is also important for DI72 replication.....	58
2.3.4. GCD14 protein is important for Δ 5' UTR-DI replication.....	61

2.3.5. Knocking down GCD10 or knocking out GCD14 reduces DI72 minus-strand levels.....	63
3. Chapter 3. Discussion.....	66
3.1. The 5' UTRs of plus-strand RNA viruses.....	67
3.2. Pros and cons of the yeast system and the DI RNA replicon.....	69
3.3. Identification of TBSV 5'UTR binding proteins.....	70
3.4. Characterization of 5' UTR-binding proteins.....	72
3.5. The functional importance of GCD10 and GCD14 proteins in TBSV DI replication.....	73
3.6. Plus- and minus-strand synthesis and GCD10/14 proteins.....	75
3.7. A working model for the role of GCD10/14 in DI RNA replication.....	76
3.8. Future directions.....	80
4. Chapter 4. Materials and Methods.....	82
4.1. Plasmids or DNA template constructions.....	83
4.2. Yeast transformation and yeast DI replication assay.....	86
4.3. In vitro transcription.....	87
4.4. Protoplast preparation infection.....	87
4.5. Yeast total RNA isolation and northern blotting.....	87
4.6. Preparation of streptomycin-conjugated column and binding test for streptotag RNAs.....	88
4.7. Preparation of yeast total soluble extract and streptotag affinity purification.....	88
4.8. Silver staining and identification of proteins.....	89

4.9. Protein extraction for western blotting analyses.....	90
4.10. His affinity purification.....	91
4.11. Electrophoresis mobility shift assay (EMSA).....	92
4.12. RT-PCR for DI72 minus-strand detection.....	92
4.13. The list of primers used in this study.....	94
5. References.....	96
6. Appendices.....	108

List of Figures

Figure 1.	Structure of TBSV viral particle and coat protein.....	5
Figure 2.	TBSV ssRNA genome organization.....	5
Figure 3.	Simplified replication strategy for the TBSV RNA genome.....	7
Figure 4.	The TBSV p92 RdRp and structural features of RdRps.....	9
Figure 5.	Translation of TBSV sRNAs.....	14
Figure 6.	A prototypical TBSV defective intefering (DI) RNA.....	17
Figure 7.	Yeast DI72 replication assay in wildtype and knockdown strains.....	22
Figure 8.	Assembly of the tombusvirus virus replication complex (VRC).....	25
Figure 9.	Higher-order RNA structures in the 5' UTR are important for DI72 replication in plant and yeast cells.....	24
Figure 10.	Analysis of DI72 minus-strand levels in yeast.....	37
Figure 11.	Streptotagged affinity purification.....	39
Figure 12.	Streptotagged 5' UTRs.....	41
Figure 13.	Streptotag RNA affinity isolation of host proteins.....	43
Figure 14.	Streptotag affinity isolation using mutant 5'UTRs.....	47
Figure 15.	Expression and purification of GCD10, GCD14 and MBP.....	49
Figure 16.	EMSA with 5' UTR RNA and the purified GCD10/GCD14 protein complex.....	51
Figure 17.	EMSA with individually purified GCD10 and GCD14 proteins.....	52
Figure 18.	Effect of knockdown of GCD10 protein on DI72 levels.....	54
Figure 19.	Effect of knockdown of GCD10 protein on Δ 5'UTR-DI levels.....	57

Figure 20.	Effect of knockout of GCD14 protein on DI72 levels.....	60
Figure 21.	Effect of knocking out GCD14 protein on Δ 5'UTR-DI levels.....	62
Figure 22.	Analysis of DI72 minus-strands in GCD10 and GCD14 mutant strains by semiquantitative RT-PCR.....	65
Figure 23.	A working model.....	77

List of Abbreviations

BMV:	<i>Brome mosaic virus</i>
cPR:	complementary promoter
DI:	defective interfering
ds:	double-stranded
DSD:	downstream domain
EMSA:	electrophoretic mobility shift assay
ESCRT:	endosomal sorting complex required for transport
GCD:	general control derepressed
gPR:	genomic promoter
KD:	knockdown
Kd:	dissociation constant
KO:	knockout
LC:	loading control
nt:	nucleotide
PC:	phosphatidylcholine
PE:	phosphatidylethanolamine
PK:	pseudoknot
RISC:	RNA-induced silencing complex
RdRp:	RNA-dependent RNA polymerase
rRNA:	ribosomal RNA
RT:	reverse transcription
sg:	subgenomic
SL:	stem-loop
ss:	single-stranded
siRNA:	small interfering RNA

TBSV:	<i>Tomato bushy stunt virus</i>
TMV:	<i>Tobacco mosaic virus</i>
TSD:	T-shape domain
UTR:	untranslated region
VRC:	virus replication complex
wt:	wild-type

CHAPTER 1

INTRODUCTION

Plants are essential for humans, because they are a critical part of our food source and provide the oxygen that we breathe. Additionally, plants are also significant for our environment by reducing global warming by recycling carbon dioxide, adding to their importance for our existence. Therefore, as our global population rises in the future, the cultivation and protection of plants will become increasingly important.

My research focuses on how a particular plant pathogen, a plant virus, utilizes plants for its own reproduction. In the course of reproduction, plant viruses can significantly impact plant physiology and reduce crop yield [1-3]. In this regard, understanding plant viral replication at the molecular level may allow us to interfere with its reproduction cycle and develop novel antiviral strategies to protect our crops.

My project focuses on a model plant virus called Tomato bushy stunt virus (TBSV). It contains single-stranded (ss) RNA as its genetic material, which is the most prevalent type of genome present in viruses that infect plants [4]. Consequently, understanding how TBSV reproduces itself may not only provide novel information on this virus' reproduction mechanism, it could also provide general insights on how this dominant class of plant virus replicates [5].

1.1. Virus Classification

Viruses are most commonly organized into related groups by the genomic material that they contain, and this classification is known as the Baltimore classification [6]. Two classes involve DNA viruses that contain double-stranded (ds) (group I) and single-stranded (ss) DNA

(group II) as their genome. Three other classes are RNA viruses, which contain either dsRNA (group III), or ssRNA as their genomic material. The single-stranded RNA viruses are further defined as positive-sense (group IV) or negative-sense ssRNA viruses (group V). The positive-sense RNA viruses are also called plus-strand RNA viruses and are named so because they are of coding sense. Conversely, negative-strand or minus-strand RNA virus genomes are non-coding. Another class of RNA virus are the retroviruses (group VI), which contain RNA as their genome, but replicate by a DNA intermediate. The last virus group are dsDNA viruses, termed pararetroviruses (group VII), which replicate their DNA genomes via an RNA intermediate using reverse transcriptase. The virus that my study focuses on is a positive-strand RNA virus (group IV), and the reproduction cycle of this group is discussed below.

1.1.1. Positive-strand RNA viruses

Positive-strand RNA viruses have a message-sensed ssRNA as their genomic material. Therefore, upon the entry into a cell, their genomes are immediately translated to produce viral protein(s) using the host translation machinery [7]. An important feature of the viruses in this class is the production of a virally-encoded RNA-dependent RNA polymerase (RdRp). The RdRp is responsible for copying the RNA genome into a full-length complementary negative-sense RNA intermediate, which is in turn copied to produce positive-sense progeny RNA genomes [8]. Progeny RNA genomes generated can then be further translated to produce RdRp for more genome replication or associate with viral coat protein(s) and be packaged into progeny viral particles. These general features are common to all known positive-strand RNA viruses,

regardless of whether they infect bacteria, yeast, insects, plants or animals, indicating a common evolutionary root for these viruses.

1.1.2. Family Tombusviridae

TBSV, which is the focus of my study, belongs to the virus family Tombusviridae. This family contains positive-sense RNA viruses infecting a variety of agriculturally important plants such as tomato, lettuce, cucumber, maize and grape [9, 10]. Some common features of this virus family are as follows. Their ssRNA genomes are approximately 4600 - 4800 nucleotides long and lack a 5' cap and a poly(A) tail at their 3' ends [11]. Viral particles have icosahedral structure with diameters in the range of 30 nm (**Fig. 1A & B**) [12]. The family Tombusviridae is currently divided into 16 genera based on their encoded proteins, genome organization, and sequence similarities [13]. This family includes several genera that have been studied extensively, which include tombusviruses, carmoviruses and dianthoviruses [13].

1.1.3. Genus *Tombusvirus*

The genus *Tombusvirus* currently has 17 species, and TBSV is the type species of this genus [13]. All members of this genus have characteristics that are similar to those of TBSV in terms of genome and particle structure [13]. The TBSV ssRNA genome is 4776 nucleotides long and encodes five open reading frames (ORFs) that are flanked by a 166-nt 5'UTR and a 351-nt 3'UTR (**Fig. 2, top**) [11, 14]. The 5' proximally-encoded ORF is the auxiliary replication protein

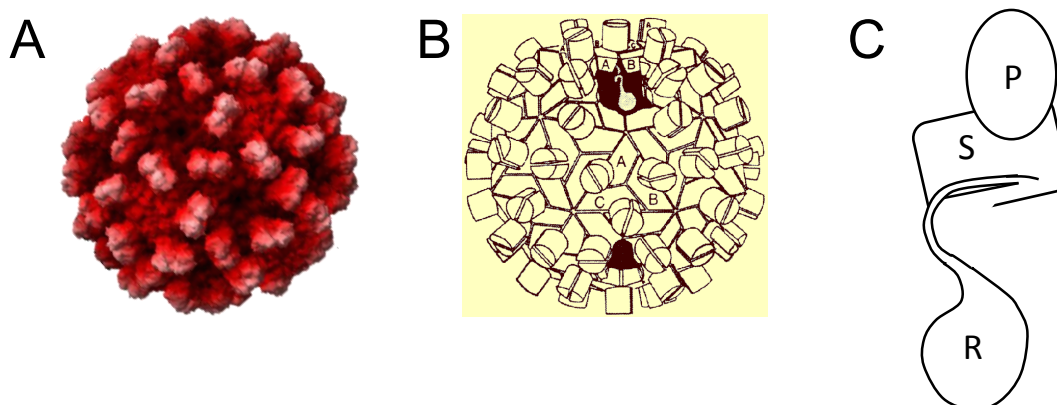


Figure 1. Structure of TBSV viral particle and coat protein. (A) A crystal structure of TBSV virus particle (11). (B) Schematic representation of the TBSV virion (17). Three identical, but conformationally-distinct, subunits, A, B, C form T=3 symmetry. (C) A schematic view of domains in a TBSV coat protein subunit. P denotes protruding domain, while S and R denote shell and RNA binding domains, respectively [modified from (17)].

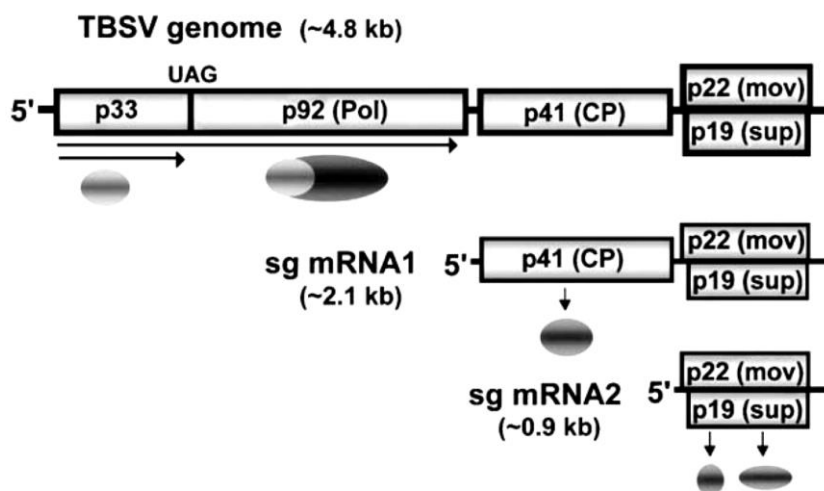


Figure 2. TBSV ssRNA genome organization. Viral proteins p33 and p92 are translated directly from the viral RNA genome, while p41 is translated from sg mRNA1, and p22/p19 are translated from sg mRNA2. p22 and p19 ORFs are not in the same reading frame, thus they are distinct proteins [from (11)].

p33, and its translational readthrough product, p92, is the RdRp. Both are translated directly from the viral genome by host ribosomes and both are essential for TBSV genome replication [15, 16]. These viral proteins complex with specific host proteins to form the TBSV RNA replicase or, more generically termed, the virus replication complex (VRC) [5]. Encoded downstream of p92 ORF are the coat protein, p41, the movement protein, p22, and the suppressor of gene silencing, p19 [11]. These downstream ORFs are translationally silent in the genomic context, and instead are translated from two subgenomic (sg) mRNAs that are transcribed during infections (**Fig. 2, middle and lower**) [11]. Only the full-length viral genome is packaged into viral particles.

1.2. Functions of TBSV proteins

During infections TBSV expresses five different viral proteins and each of these proteins has specific and sometimes multiple functions. The role(s) of each protein is described below.

p33: The first TBSV protein to be expressed upon infection of a plant host cell is p33 [11]. It is a membrane anchored protein that contains a peroxisomal targeting signal [17]. One of the key functions of p33 is to bind to a specific region of TBSV RNA genome termed region II (RII; discussed below) and to recruit the genome to the peroxisome where the viral genome replication occurs (**Fig. 3**) [17-20]. Free p33 is also targeted to peroxisomes, where they form multimers in association with the membrane to form invaginations in the outer peroxisomal membrane [17]. These membranous structures, called spherules, house the VRC that performs viral genome replication and sg mRNA transcription (**Fig. 3**) [17, 21-23]. p33 is the major protein component of spherules, along with at least one molecule of p92. Inside the spherule,

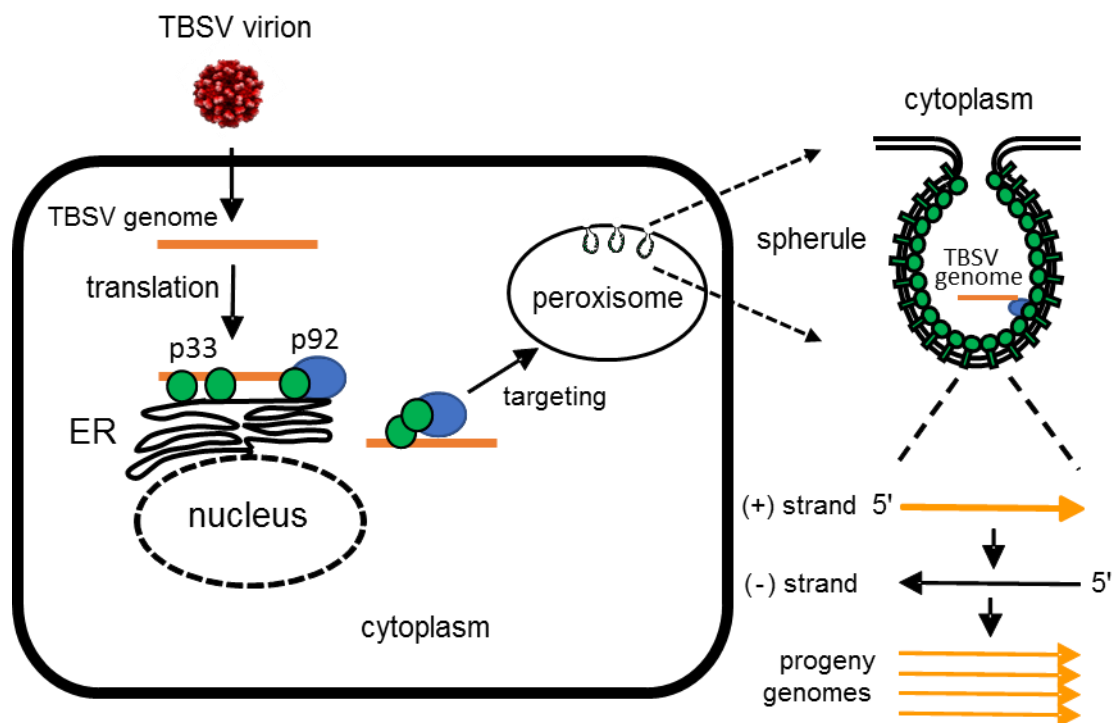


Figure 3. Simplified replication strategy for the TBSV RNA genome. Following translation of p33 and p92, these viral proteins bind to the TBSV genome and the RNA-protein complex is targeted to peroxisomes. The outer peroxisomal membrane is induced to form invaginations, termed spherules. Genome replication that leads to the production of progeny viral genomes occurs within these spherules via synthesis of a (-)-strand RNA genome intermediate.

the viral genome, associated with p33, p92 and host proteins, undergoes viral RNA synthesis, and the progeny genomes and sg mRNAs are released from the spherules via a neck that connects the interior of the spherules with the cytosol [5].

Different regions of p33 have been found to be related to the different roles (**Fig. 4A**) [24]. The N-terminal sequence as well as the two proximal transmembrane domains are found to be critical for peroxisomal targeting function [17]. The RNA binding role of p33 is carried out by arginine- and proline-rich RNA binding motif (₂₁₃RPRRRP) (RPR motif) [16]. Two separate regions downstream of RPR binding motif, termed p33:p33/p92 interaction domains, play an essential role for self oligomerization of p33 and p92, which is required for RNA replication [22].

p92: The hallmark of RNA viruses is a virus-encoded RdRp, which for TBSV is p92 [11]. Notably, p92 is a translational readthrough product of the p33 open reading frame (ORF), and thus is a C-terminally extended version of p33, with the extended portion of p92 containing the core RdRp activity [25]. Since p92 also contains p33 as its N-terminal region, it too is targeted to peroxisomes [17] and is associated with the membrane component of p33-lined spherules (**Fig. 3**) [17, 22]. Once a replication complex is assembled in the spherules, p92 initiates *de novo* RNA synthesis (*i.e.* without a primer) at a promoter element at the 3' end of the genome and synthesizes a complementary negative-strand [11]. After negative-strand synthesis is completed, p92 binds to the 3' end of the negative-strand to produce the complementary positive sense progeny genomes (**Fig. 3**). The RdRp catalytic function of p92 is carried out by the p92 region that does not overlap with p33 (*i.e.* the non-overlapping region) (**Fig. 4A**).

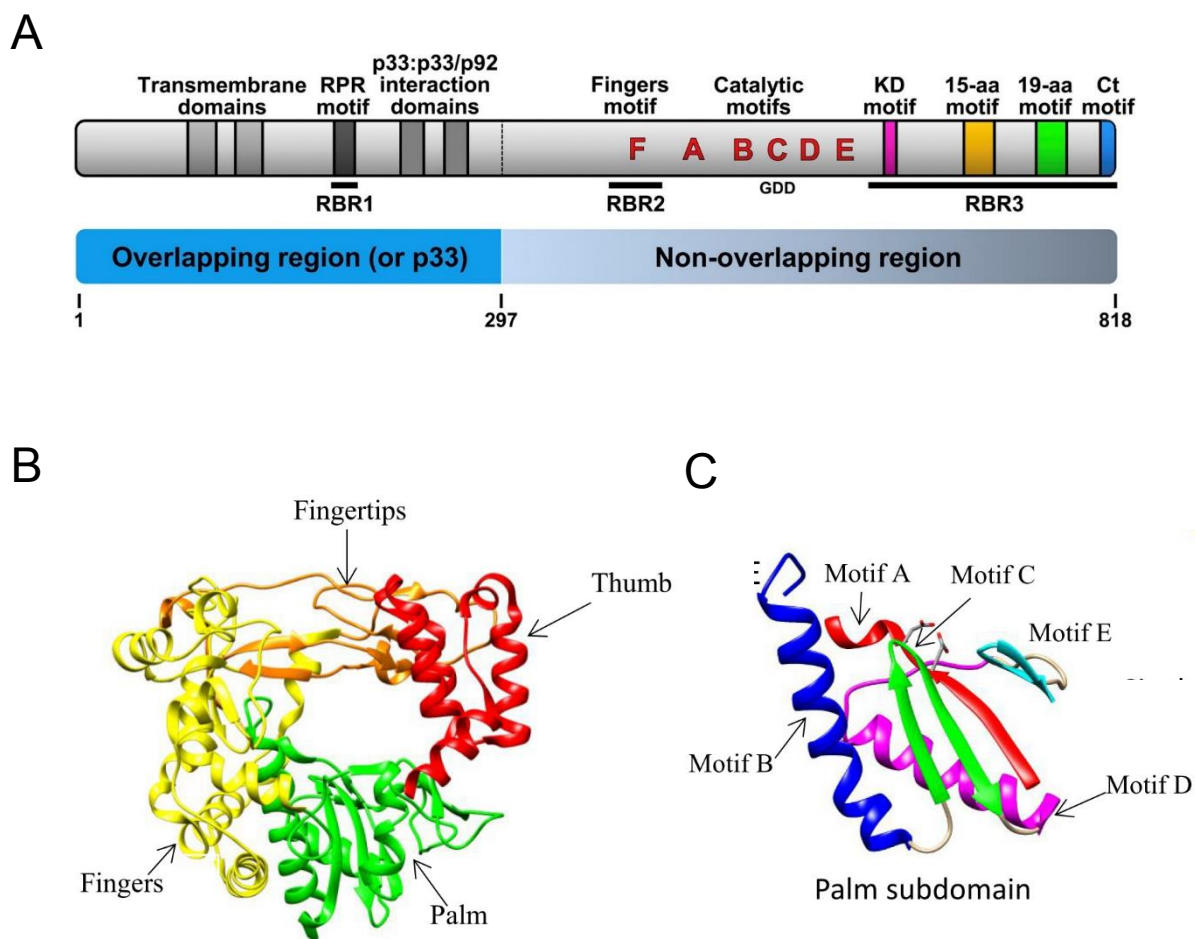


Figure 4. The TBSV p92 RdRp and structural features of RdRps. (A) Linear representation of TBSV p92 depicting functional motifs. The overlapping region corresponds to p33, while p92 contains both overlapping and nonoverlapping regions. The relative locations of RNA binding regions (RBR) are represented with horizontal bars. [from (24)] **(B)** Poliovirus RdRp depicting the conserved fingers, palm and thumb domains, colored in yellow, green and red, respectively. [from (27)]. **(C)** Relative positions and secondary structures of catalytic motifs in the palm subdomain. [from (27)]

All the RdRps structurally characterized to date have a right-hand structure with palm, finger and thumb domains, and have RdRp signature motifs (motifs A-F) (**Fig. 4A**) [24, 26]. Motifs A to D reside in the palm domain, while motif E is located between palm and thumb domain, and motif F (finger motif) is at the finger domain (**Fig. 4B, C**) [27]. The catalytic activity of RdRp requires two magnesium ions [28], and motif A is involved in coordinating both magnesium ions and possibly sugar selection, while motif B participates in binding template and incoming nucleotides [26, 27, 29]. Motif C contains the universally conserved GDD motif. Along with motif A, motif C coordinates the two magnesium ions, which in turn coordinates α -, β -, γ -phosphate groups of the nucleoside triphosphate and 3'-OH of nascent primer for the catalytic reaction [26]. Motif D is proposed to facilitate the RdRp forming an optimal conformation after incorporation of correct nucleotide, and functions in fidelity of nucleotide addition [27]. Motif E has been shown to properly position the 3' end OH of the primer for polymerization [27]. Motif F forms the roof of rNTP entry tunnel and is involved in rNTP binding [27]. In TBSV p92, there is another RNA binding domain in the finger domain (termed RBD2). Based on the similarity in location to the region of hepatitis virus C NS5B, RBD2 has been proposed to guide the RNA template to the active site of RdRp [30]. Since this domain overlaps with motif F, this region is likely multifunctional.

In addition to these signature RdRp motifs, p92 has tombusviridae-specific motifs, termed KD, 15-aa and 19-aa motifs [31]. A di-residue KD (₆₆₃KD), a 15 amino acid (15-aa) motif (₇₁₈GLALSAGIPVVETFY), and 19 amino acid (19-aa) motifs (₇₇₀RASFWAAFGLTGDEQLALE) are highly conserved among Tombusviridae family members and all are important for viral RNA synthesis [31]. However, mutations in the different motifs preferentially affected different viral

RNA species [31]. Specifically, mutations in 19-aa motifs primarily affected minus-strand synthesis, whereas mutations in 15-aa motifs preferentially decreased plus-strand synthesis. Since these three motifs overlap with another RNA binding domain (RBD3), it is plausible that this region of p92 may function in discriminating between different viral RNA species during various stages of viral RNA synthesis [24].

p92 also has a highly conserved C terminal motif (Ct motif), which is conserved only among members of the tombusvirus genus [32]. Mutational and deletional studies of the Ct motif have shown that this motif is required mostly for sg mRNA transcription [32]. Overall, these various motifs at C-terminus proximal region affected different steps of viral RNA synthesis, namely sg mRNA transcription, minus-strand synthesis, and/or plus-strand synthesis [24].

p41: The capsid, or coat, protein for TBSV, p41, packages TBSV progeny genome in a spherical shell (**Fig. 1A**) [12]. The viral particle is nonenveloped and contains 180 coat protein subunits with T=3 icosahedral symmetry (**Fig. 1A**). p41 has three distinct structural domains (**Fig. 1C**) [33]. The shell domain forms a tight covering through subunit-to-subunit interactions, while the protruding domains from two neighbouring subunits project outward from the surface of the particle (**Fig. 1B**) [33]. The RNA binding domain, which is highly basic, is located inward from the shell and interacts with the viral genome in the interior [12, 33]. Depending on their location in the particle matrix, subunits assume one of three slightly different conformations (*i.e.* A, B or C) that help to maximize inter-subunit interactions. While the coat proteins from some viruses in the family Tombusviridae have been reported to function in

suppressing host gene silencing [34, 35], the TBSV coat protein is not involved in this function [34].

p19: Plants have evolved an RNA-based mechanism to defend themselves against viral infections [36]. RNA viruses replicate their genome via negative-strand intermediates, and it is well documented that during infections, a fraction of positive- and negative-strands exist in the form of double-stranded RNA (dsRNA) [37-43]. This viral dsRNA can be detected and targeted for destruction by the cell [44]. The cellular nuclease, dicer, recognizes and cleaves the viral dsRNAs into ~20 to 24 nucleotide (nt) dsRNA fragments, termed small interfering RNAs (siRNAs). One of the strands from the siRNA duplex is then loaded into an RNA-induced silencing complex (RISC), which enables it to specifically target and cleave viral RNAs guided by base pairing complementarity [44]. The TBSV p19 protein is a gene silencing suppressor that inhibits this type of host defence mechanism [45-47]. p19 forms dimers that binds to 21 nt siRNA fragments generated by dicer, thus sequestering them and preventing their loading into RISC [48]. This reduces amounts of siRNAs available for RISC and weakens the plant defense mechanism. This suppression of RNA silencing by p19 is essential for survival of the virus, because without it, plants eventually able to clear the virus and recover from the infection [45, 46].

p22: Unlike animal cell membranes, plant cell walls represent an impassable barrier for viruses. To deal with this challenge, plant viruses express movement proteins to facilitate viral genome transport from one cell to adjacent cells; termed cell-to-cell transport. For this function, TBSV expresses p22 [11]. This protein is membrane-associated and a symptom determinant and is essential for cell-to-cell movement [49, 50]. Although it is not clear how this

protein facilitates genome transmission, studies with Tobacco mosaic virus (TMV) have shown that its movement protein associates with plant plasmodesmata, a channel connecting one cell to a neighbouring cell, and it increases its size exclusion limit allowing genome transport to neighbouring cells [51]. It is possible that TBSV p22 functions similarly, because both are type-I movement proteins, which do not also require coat protein for cell-to-cell movement to occur.

1.3. TBSV gene expression

As mentioned above, the TBSV genome contains five ORFs encoding p33, p92, p41, p22 and p19 (**Fig. 2**), which are expressed through unconventional translational strategies. In general, eukaryotic cellular mRNAs contain 5'-cap and 3'-poly(A) tail, and both are required for efficient translation [52]. mRNAs lacking either of these elements have been shown to have dramatically reduced levels of translations [53]. Strikingly, the viruses in the family Tombusviridae lack both of these terminal elements, however their viral proteins are efficiently translated. Translation of the first two 5'-proximally encoded p33 and p92 occurs directly from the genome and is mediated by an RNA element located at the near 3' end of the genome, termed the 3' cap-independent translational enhancer (3'CITE) [54]. Studies in the tombusvirus Carnation Italian ringspot virus (CIRV), a close relative of TBSV, showed that its 3'CITE recruits eukaryotic translation initiation factors, 4E and 4G, and delivers them to the 5' end of the genome through an RNA-RNA interaction that occurs between the 5' UTR and 3'CITE (**Fig. 5**) [55, 56]. Other translation initiation factors and the 43S ribosomal subunit are then sequentially recruited to the 5' end of the genome to initiate translation of p33 [56]. For

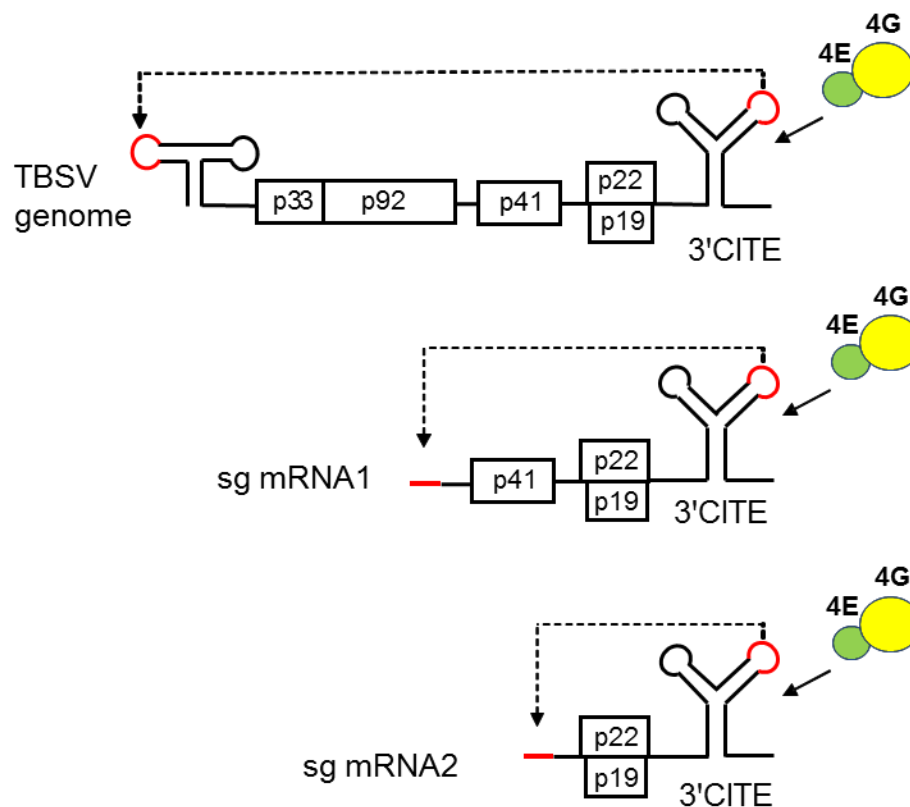


Figure 5. Translation of TBSV sRNAs. Translation initiation factors 4E and 4G bind to the 3'CITE, which is able to simultaneously base pair (red regions) with the 5'UTR via a long-range interaction (depicted by dotted line). This interaction repositions the bound initiation factors at the 5'UTR, where they are able to recruit the 43S ribosome subunit which initiates translation.

translation of p92, translation termination is occasionally suppressed at p33 UAG stop codon by an aminoacylated near-cognate tRNA(s) [11]. This allows ribosomes to bypass this stop codon and continues translation, thereby producing the p92 RdRp [11]. Interestingly, the process of translational readthrough requires a different long-range RNA-RNA interaction between an extended stem loop structure at the readthrough site and the 3'-end of the viral genome [25]. It is proposed that this interaction helps to coordinate translation and replication of the viral genome.

Since eukaryotic ribosomes are generally 5'-end dependent for translation initiation [57], which is the case for TBSV, viral proteins p41, p22 and p19 downstream of p92 ORF are not translated from the genome. Instead, the virus transcribes two sg mRNAs during the infection, which encode the ORFs for these proteins [11]. p41 is at the 5'-end of sg mRNA1 and is translated from this message. Sg mRNA2 has two overlapping ORFs in different reading frames, one for p22 and the other for p19, and both of these proteins are translated from this message (**Fig. 5**). The initiation of translation from both sg mRNAs is believed to be similar to what occurs for viral genome translation. That is, the 3' CITE near 3' end of each sg mRNA binds and delivers the eIF4E and eIF4G to the 5' ends of the sg mRNAs via RNA-RNA interaction between the 5' UTRs of sg mRNAs and 3'CITE (**Fig. 5**) [56].

1.4. TBSV RNA replication

As mentioned above, p33 and p92 RdRp are essential for TBSV genome replication [15, 16]. Once enough p33 and p92 proteins are translated, the genomic RNA is then used as a

template for genome replication. p33 and p92 associate with each other and also with the genomic RNA [18, 19, 22]. The complex is transported to the replication centre comprised of p33-based spherules derived from the peroxisomal outer membrane (**Fig. 3**) [17, 21].

Four regions in the viral RNA genome have been found to be critical for the genome replication [58, 59]. These elements have been identified and characterized mainly through subviral RNA replicons termed defective interfering (DI) RNAs [58, 59]. In order to understand how TBSV genome replicates, it is relevant here to understand the basic features of DI RNAs. The important features of the DI RNAs as well as the four regions of the genome important for RNA replication elements are discussed below.

1.4.1. Defective interfering RNAs

Defective interfering (DI) RNAs were discovered after serial passage of tombusvirus infections in plants at high titer [60]. They are small, noncoding RNAs derived from four different parts of the viral genome. They are able to replicate in infections when accompanied by the “helper” wt TBSV genome, which provides necessary RNA replication proteins. DI RNAs represent highly deleted forms of the viral RNA genome and are about 400-800 nts long. Typical TBSV DI RNAs are composed of four noncontiguous segments of the viral genome (**Fig. 6A**) [58, 59]. Formation of DI RNAs is believed to occur when the replicase occasionally dissociates from the template and then resumes RNA synthesis at an upstream site, resulting in a deletion of the intervening region. This polymerase jumping mechanism generates small viral RNAs composed of different non-contiguous segments of the viral genome (**Fig. 6A**) [59].

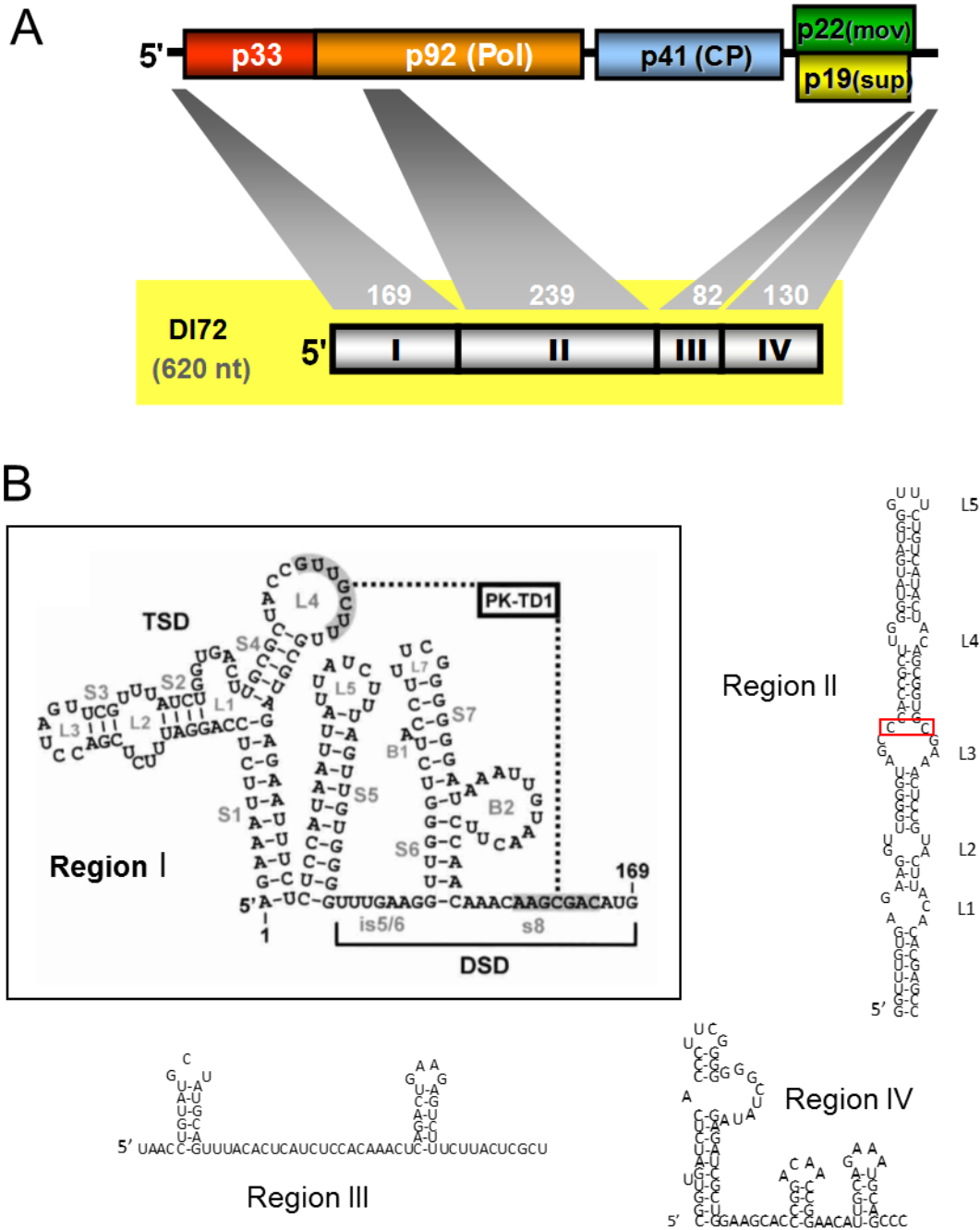


Figure 6. A prototypical TBSV defective interfering (DI) RNA. (A) Schematic of DI72 is shown along with the relative locations (and lengths) from which regions I to IV are derived from the genome. **(B)** The secondary structure of each region is shown. For region I, stem-loop 5 (SL5) corresponds to the regions of S5 and L5 in the diagram. The critical C-C mismatch in the L3 of region II is boxed in red. The structure of region III is from minus-strand.

Because these viral RNAs do not encode viral proteins, they are unable to replicate on their own [58]. As a result, DI RNAs are only amplified in coinfections with a complete viral genome, which provides p33 and p92 for its replication [58, 59]. Therefore, these small non-coding viral replicons require a “helper” virus for their replication. Nonetheless, their ability to replicate indicates that they contain the minimal RNA replication elements required for TBSV genome replication. Also, because DI RNAs are very efficient viral RNA replicons and are free from encoding viral proteins, they are highly useful in studying *cis*-acting RNA elements required for TBSV genome replication.

For TBSV, DI72 is the DI RNA that has been studied most extensively (**Fig. 6A**). It contains four segments of the TBSV genome, namely, regions I, II, III and IV [58, 59]. Region I is identical to the 5' untranslated region (UTR) of TBSV genome. Regions II and III are from internal segments in the genome, while region IV corresponds to the 3' end of the genomic RNA. The basic characteristics of each region is discussed below in the order that they are required for TBSV RNA replication.

Region II. Once enough p33 and p92 are translated, they bind specifically to region II in the genome and target the viral RNA (e.g., TBSV genome or DI72) to peroxisomes for assembly of the VRC and the initiation of RNA replication [18-21]. Region II is located within p92 ORF, and it forms a long stem-loop RNA structure with multiple internal loops (**Fig. 6A**) [18]. p33, in complex with p92, binds as a multimer to one of the internal loops (termed L3) in the RNA structure [19]. This RNA-protein interaction has been shown to be essential for the viral replication, and the sequence identity of the internal loop, L3, is highly critical for p33/p92 binding, with the C-C mismatch being essential (red box in **Fig. 6B**) [18, 19]. Notably, p33/p92

were not found to interact with any of the other regions (i.e. RI, RIII or RIV) of the plus-strand of the DI RNA [19].

Region IV. Region IV is 130 nucleotides long and corresponds to the 3' end of the TBSV genome (**Fig. 6A**). After DI72 or TBSV genome is recruited to the VRC, complementary stand RNA synthesis begins at the 3' end (or region IV) of the RNA templates [11]. *In vitro* assays for minus-strand synthesis using partially purified viral replicase from virus-infected plants showed that the 19 nt sequence (5'-CAUUGCAGAAAUGCAGCCC-OH) at the 3' end functions as a promoter (termed genomic promoter, gPR) for de novo initiation of minus-strand synthesis [61]. Solution structure probing and mutational analyses has shown that region IV folds into three hairpin structures with a single-stranded tail, -CCC-OH at the end (**Fig. 6B**) [62-64]. Recognition of region IV by RdRp and minus-strand synthesis is greatly influenced by the promoter sequence identity, as well as the structures and sequences within the three RNA hairpins [61-64]. Importantly, within the context of the viral genome, the p92 that binds to RII is able to access the 3'-promoter in RIV by assistance of yet another long-range RNA-RNA interaction that unites these two regions [65].

Region III. The viral replicase copies DI72 RNA or the TBSV genome from their 3' ends to generate complementary minus-strands. After this, the RdRp recognizes the 3' end of the minus-strand to initiate the plus-strand synthesis. In DI72, the binding of the replicase to the minus-strand is assisted by its binding to region III in the minus-strand (**Fig. 6B**). Region III assists in the recruitment of the replicase to the minus-strand template and facilitate its access to the promoter for plus-strand synthesis [66, 67]. Consequently, regions III has been termed

an enhancer element, because it promotes efficient plus-strand synthesis from minus-strand templates [66].

Region I. Region I is 166 nts long and identical to the 5' UTR of TBSV genome (**Fig. 6A**). To initiate the plus-strand synthesis, the RdRp recognises the 3' end of the minus-strand, which is complementary to the TBSV 5' UTR and region I. The minimal promoter sequence for the initiation of the plus-strand synthesis (termed complementary promoter, cPR) has been determined as the 3' end 11 nts of the minus-strand [_{H₀}-CCUUAAGAGG..5'(-)] [68]. Thus, the sequence complementary to region I functions as cPR to facilitate plus-strand synthesis.

Relevant to my study is the fact that the plus-strand sequence of region I is also important for efficient viral RNA replication. Conserved sequences and secondary and tertiary structures RNA structures in the plus-strand of region I mediate its function (**Fig. 6B**). Solution structure probing analysis showed that region I has two distinct RNA structural domains separated by a stable hairpin termed stem loop 5 (SL5) (**Fig. 6B**) [69, 70]. The 5' proximal structure forms a T-shaped domain (TSD) with a three-way junction [69]. The 3' domain, termed downstream domain (DSD), has both helical and unpaired segments, with high sequence conservation in is5/6 and B2 (**Fig. 6B**) [70]. Additionally, there is a pseudoknot, PK-TD1, that forms between a loop in the TSD and the 3' portion of the DSD [70]. Mutational analyses of the various structural features of region I in DI72 revealed that the predicted double stranded regions, the conserved sequence segments, and the pseudoknot all contributed to efficient DI72 replication. Additional experiments established that these RNA elements functioned in the plus-strand of region I and that their disruption inhibited minus-strand RNA production [70].

1.4.2. Yeast (*Saccharomyces cerevisiae*) as a model host for DI replication

Since viral genomes contain only limited numbers of genes, viruses require host proteins to assist in their replication [71]. An important area in viral research is determining which cellular proteins are used and how they mediate viral RNA replication. Due to difficulties associated with identifying host proteins in plant host system (*e.g.*, complete plant protein databanks for many host plants may be unavailable), this area of research did not progress rapidly until a pioneering surrogate yeast host system was developed. Using *Saccharomyces cerevisiae* (*S. cerevisiae*) as a heterologous host, the Ahlquist laboratory was able to show authentic genome replication of brome mosaic virus (BMV) and flock house virus in yeast cells [72, 73]. Importantly, due to the availability of genome-wide single gene knockout and knockdown yeast strains, the yeast systems allowed for facile identification of many host proteins important for the viral genome replication [74-78]. Although yeast does not support replication of the full-length TBSV genome, the Nagy lab determined that it can support TBSV DI RNA replication very efficiently when p33 and p92 replicase proteins are provided in trans from separate expression plasmids (**Fig. 7**) [79]. In this DI RNA replication system, p33 and p92 are expressed from constitutive ADH1 promoter, while the RNA replicon, DI72 RNA, is transcribed under the control of an inducible GAL1 promoter. Upon the addition of galactose to the yeast culture medium, DI72 transcription is activated and provides a viral RNA for replication. The p33/92 replicase proteins, in cooperation with host factors, then mediate robust DI72 replication (**Fig. 7**) [79-81]. To identify and elucidate the functions of host proteins that contribute to TBSV RNA replication, genome-wide single gene knockout strains (or knockdown strains, if the genes are essential) were employed. DI72 replication in both wild type and the

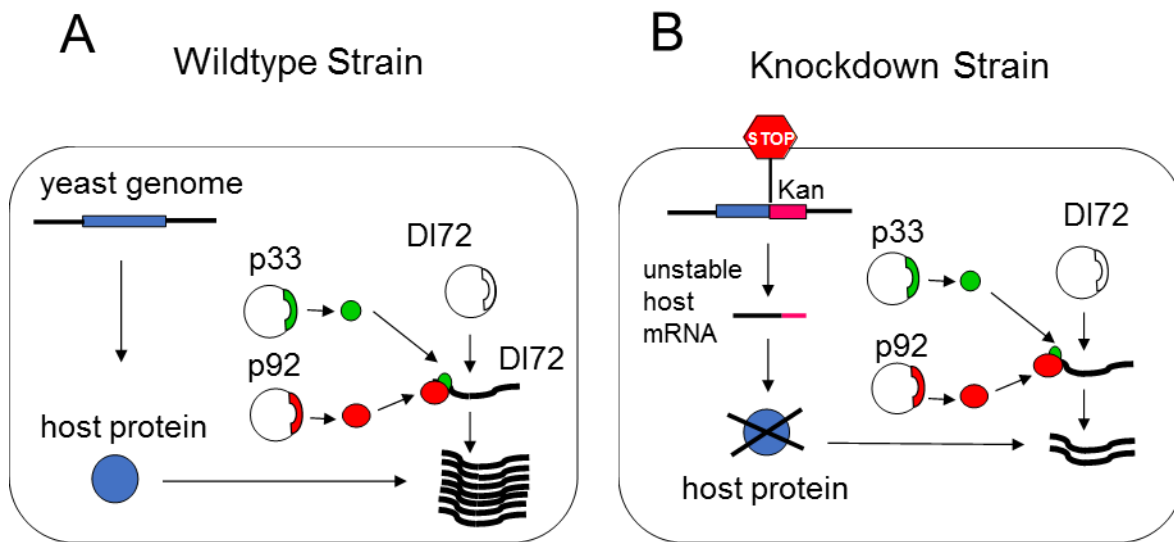


Figure 7. Yeast DI72 replication assay in wildtype and knockdown strains. (A) In wt yeast cells, p33 and p92 are expressed constitutively, while DI72 RNA transcription is induced upon the addition of galactose. p33 and p92 facilitate DI72 replication with the assistance of host proteins expressed from the yeast genome (blue). **(B)** In DAmP knockdown strains, a kanamycin resistance gene (pink) is inserted to the 3' UTR of a host gene, which destabilises the host mRNA. This in turn reduces amounts of a host protein. If the host gene that is knocked down is important for DI RNA replication, DI72 levels will be reduced.

single gene knockout (or knockdown) strains were compared, uncovering many host proteins important for DI72 replication [82, 83].

1.4.3. Principles behind yeast single gene knockdown strains

Most yeast single-gene knockdown and knockout strains are commercially available. For essential genes, knocking out the genes renders the yeast nonviable, and thus, in most cases only knockdowns are possible (but exceptions exist, discussed later). The knockdown strains relevant to my study is termed DAmP (Decreased Abundance by mRNA perturbation) [84]. In these strains, essential genes are disrupted by an insertion of kanamycin antibiotic resistance cassette into the gene region corresponding to the 3' UTR of the corresponding mRNA (**Fig. 7**). This insertional modification makes the mRNA transcripts unstable, and their levels can be reduced 4- to 10-folds, with reduced levels of their encoded proteins [84]. When present in haploid yeast, the knockdown of the target proteins is more profound than for diploid strains, in which only one allele is knocked down.

1.4.4. Host factors important for TBSV viral RNA replication

Many host proteins affecting DI72 replication have been identified using genome-wide yeast single-gene knockout or knockdown libraries [82, 83]. Moreover, temperature sensitive mutant yeast library and individual overexpression of ~5500 yeast genes in wildtype yeast also led to the identification of more host proteins affecting DI72 replication [85, 86]. Altogether,

over 500 host proteins were identified to affect DI72 replication or viral RNA recombination [82, 83, 85, 86]. As mentioned above, the replication of TBSV viral RNA is carried out by VRC in the sequestered spherules. The host-tombusvirus interactions can be grouped into five stages based on how VRC is formed and how viral RNA replication proceeds (**Fig. 8**). The current knowledge of host-tombusvirus interactions are presented below.

Stage 1: Recruitment of host factors and preassembly of VRC. Upon their translation, the viral replicase proteins, p33 and p92 interact with region II of the viral genome and are transported to the peroxisome via their peroxisomal targeting signal at near N-terminal region [17, 20]. Many host proteins are also suggested to be recruited at this stage, and both the replicase proteins and the viral RNA play active roles in recruiting the host factors [5, 24]. p33 binds to Pex19 and is targeted to peroxisomes [20]. Pex19 delivers most of the cellular peroxisome biogenesis proteins to the peroxisome boundary membrane, and thus, Pex19 assists in transport of p33 to peroxisomes.

p33 also directly interact with Heat shock protein 70 (Hsp70) and translation elongation factor 1A (eEF1A) and recruit them to the replication sites [87-89]. These two proteins are found to be part of core replication complex, and they play critical roles in stabilizing and properly folding p33 protein [87-89]. Downregulation of these two proteins have led to rapid degradation of both p33 and p92 [87-89]. p92 directly binds to and recruits regulatory particle non-ATPase 11 (Rpn 11) [90]. Rpn11 is generally known to play important roles in various protein complex formations such as proteasome lid and proteasome storage granules [91, 92]. It is suggested that Rpn11 may contribute to VRC formation [90]. Downregulation of Rpn11 reduced TBSV viral RNA replication in yeast, plants and in vitro [90]. TBSV viral RNA is also

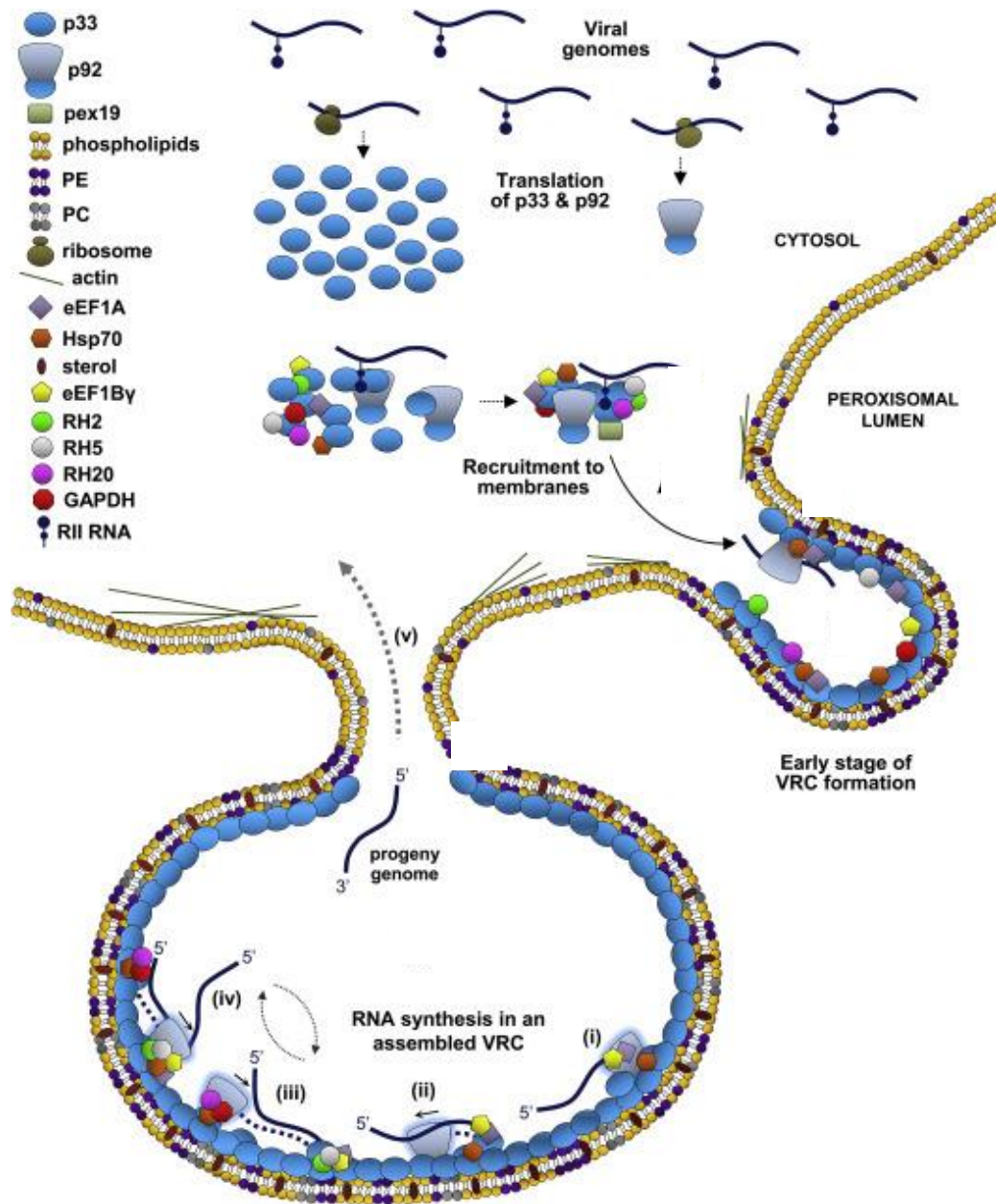


Figure 8. Assembly of the tombusvirus virus replication complex (VRC). Upon the translation of p33 and p92 from the viral RNA, various host factors (discussed in the Introduction) are recruited, via p33, p92 and the viral RNA, and direct preassembly of VRC at the peroxisomal outer membrane. Viral RNA synthesis that occur in the spherules in four distinct steps. (i) Initiation of minus-strand synthesis. (ii) Elongation of minus-strand synthesis. (iii) Initiation of plus-strand synthesis. (iv) Initiation of plus-strand synthesis displaces the first progeny genome. (v) The displaced progeny genomes exit via the neck of the spherules [from (24)].

suggested to recruit host factors such as eukaryotic translation elongation factor 1B gamma (eEF1By) and 1A (eEF1A) [88, 93]. Both eEF1By and eEF1A are found to bind region IV of TBSV RNA in vitro [88, 93]. Also, some host DEAD-box helicases are shown to be part of VRC and thought to be recruited possibly by binding to TBSV RNA [94, 95]. However, how they are recruited is unknown.

Stage 2: Regulation of actin network and lipid composition to build VRC. The actin network has been shown to play an important role in the formation of VRC. Specifically, p33 directly interacts with an actin depolymerization factor, Cof1 [96]. Cof1 normally functions in severing existing actin filament to recycle actin monomers [97]. The interaction between p33 and Cof1 inhibited the normal function of Cof1, resulting in more stabilization of actin filaments and enhanced delivery of p33/p92 and sterol at the tombusvirus replication sites [96]. Subsequently, this led to more efficient formation of VRC. On the other hand, over-expression of Cof1 resulted in sequestration of p33/p92 away from VRC assembly and inhibition of sterol enrichment at the spherules [96].

Several studies on TBSV have shown that lipid composition of VRC is also important. Sterol is known to tightly pack phospholipids (a major component of membranes) and reduce fluidity of membranes [98]. Inhibition of sterol biosynthesis by either genetic or pharmacological means showed drastic inhibition of TBSV replication [99], whereas addition of sterol to artificial phospholipid vesicles enhanced TBSV replication in vitro [100]. These two studies suggest that sterol may stabilize VRCs.

Another lipid has also been shown to be important for TBSV genome replication. In plant cells infected with TBSV, phosphatidylethanolamine (PE), one of the phospholipids in the membranes increased significantly compared to other phospholipids [101]. Inhibition of PE biosynthesis decreased TBSV replication, while increased level of PE in yeast defective in conversion of PE to phosphatidylcholine (PC) increased TBSV replication significantly [101]. PE has a conical structure and is thought to help membrane deformation, thus this feature may be important for spherule formation [98].

Stage 3: Spherule formation. The replicase proteins, TBSV RNA, and recruited host factors are targeted and concentrated at the cytosolic surface of peroxisomal membrane and induce invagination resulting in spherules with narrow openings to the cytosol (**Fig. 3**) [5, 24]. Spherules are believed to protect the viral RNA and proteins from host antiviral defense mechanism, such as RNA interference [5, 24]. The spherule formation is carried out by ubiquitinated p33, which recruits host membrane-deforming proteins, termed endosomal sorting complex required for transport (ESCRT). ESCRT proteins are normally involved in membrane bending and budding for transporting ubiquitinated membrane proteins to lysosomes for their degradation [102]. p33 directly recruits Vps23 and Bro1 ESCRT proteins, and it is thought that recruitment of these two host proteins attracts other ESCRT proteins, such as Vps2, Vps20, Vps24 and Snf7, leading to the spherule formation [21]. Another closely related tombusvirus, Carnation Italian ringspot virus also recruits Vps23 and other ESCRT proteins to facilitate spherule formation [103], and expression of dominant-negative mutants of these ESCRT proteins reduced tombusvirus genome replication, supporting essential roles of ESCRT proteins in spherule formation [21, 103, 104].

Stage 4: Minus-strand RNA synthesis. It is currently unknown whether minus-strand synthesis begins during the VRC assembly process or after the completion of spherule formation [5]. Regardless, once a VRC is functional, p92 begins minus-strand synthesis at the 3' end of the viral RNA. However, the 3' end of the RNA is highly folded (termed closed conformation) and it must be converted to open conformation in order to be accessed by p92 for minus-strand initiation [64, 105]. This process is thought to be facilitated by two host factors, eEF1A and eEF1By [87, 88, 93]. eEF1A binds to region IV and stimulated minus-strand synthesis in vitro. Since it also binds to p92, it is suggested that it may help position p92 properly to initiate minus-strand synthesis [88]. eEF1By also binds to region IV, and it opens the closed 3' end structure to facilitate minus-strand synthesis [93]. Along with these two host proteins, p33 is also found to be critical in initiation of minus-strand synthesis from the 3' terminal sequence [106]. Altogether, these two host proteins as well as p33 are shown to synergistically function in the minus-strand synthesis [87, 88, 93, 106].

Stage 5: Plus-strand RNA synthesis. Studies on plus-strand RNA viruses including TBSV have shown that completion of minus-strand synthesis results in the formation of double-stranded (ds) RNA intermediate consisting of the minus-strand and the plus-strand genome template [37-43, 107]. However, viral RdRp cannot efficiently initiate plus-strand synthesis from the dsRNA structure [108, 109]. Using yeast in vitro replication assay, several host factors (namely, Ded1, GAPDH, Fal1, Dbp2 and Dbp3) have been shown to assist p92 to initiate plus-strand synthesis [94, 110-113]. Yeast RNA helicase, Ded1 (RH20 in plant), which was found to be a component of VRC, preferentially bound to the region near the 3'-end of the minus-strand and opened dsRNA structure locally, making the 3' end of the viral RNA accessible to p92 [110].

Dbp2 played a similar role to Ded1 in vitro assay [94]. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was found to bind to the 3' end of the minus-strand and enhance plus-strand synthesis both in yeast and plant [111, 112]. Since GAPDH also interacts with p92, this protein was proposed to bring p92 in proximity to the end of the minus-strand for plus-strand synthesis [111, 112]. Both Ded1 and GAPDH synergistically enhanced plus-strand synthesis [5].

Two other yeast RNA helicases, Fal1 and Dbp3 (RH2 and RH5 in plants, respectively) have also been shown to enhance plus-strand synthesis [113]. They bind to the dsRNA intermediate via region III in the minus-strand and open the dsRNA in the region. Region III is located near the 5' end of the minus-strand, and this opened single-stranded region then interacts with the 3' end of minus-strand through long-distance RNA-RNA interaction [113, 114]. As a result, this brings the 5'- and 3'-ends of the minus-strand close to each other, helping the RdRp reinitiate plus-strand synthesis for multiple rounds. The generated progeny viral genomes are then released into the cytosol through the neck of spherules.

1.5. The objectives of my dissertation.

TBSV DI RNAs have served as very useful viral RNA replicons to study RNA elements that are required for viral RNA replication, in a context that is free from the constraints of translation. Similarly, the yeast system and its powerful genetics has been valuable for identifying cellular proteins that contribute to TBSV RNA replication. Indeed, using the DI RNA with the yeast system has identified numerous host proteins important for this process [82, 83].

The 5' UTR of the TBSV RNA genome, which corresponds to region I in DI RNAs, represent an interesting RNA segment that is important for both minus- and plus-strand viral RNA synthesis. Prior studies indicated that structural features of the plus-strand of region I are important for DI RNA replication, however, it was not known how this was accomplished. Although many host proteins have been found to be important for TBSV viral RNA replication, no host proteins have yet been shown to assist the function of the 5' UTR. Since the viral replication proteins p33/92 were not found to interact with the plus strand of region I, I **hypothesized that region I is bound by a host protein(s) that facilitates viral RNA replication.**

To test this hypothesis, I devised a plan to:

- (i) Determine if yeast was an appropriate system to study the function of region I.
- (ii) Identify protein(s) that bound specifically to the plus-strand of region I RNA.
- (iii) Establish if the RNA-protein interaction was direct or indirect.
- (iv) Investigate if the protein influenced DI RNA replication.
- (v) Propose a model to explain how the protein could contribute to viral RNA replication.

Additionally, during the course of my graduate studies I: (i) developed a quick and effective method for de-salting and removing detergents from protein samples [115], (ii) investigated the utility of a T7 polymerase-based system for heterologous protein expression in plants [116], and (iii) contributed to a study investigating intracellular targeting of tombusvirus proteins in plant cells [103]. These additional contributions are provided in the Appendices.

CHAPTER 2

RESULTS

2.1. PART 1: Identification of the 5'UTR binding protein(s)

2.1.1. Rationale for searching for 5'UTR binding protein(s)

The 5'UTR of tombusviruses is important for viral genome replication and contains conserved primary, secondary and tertiary structures that can act as binding sites for proteins involved in this process. Only p33 and p92 are required for TBSV RNA replication, therefore these two proteins are the only viral candidates for 5'UTR-binding proteins. However, previous RNA-protein binding analyses of all four regions of a TBSV DI RNA involved in RNA replication determined that these viral proteins bind efficiently to region II, but do not bind to any of the other three regions, including region I, which corresponds to the 5'UTR [19]. Considering this result, my study focused on host proteins as the primary candidates for TBSV 5'UTR-binding proteins.

Ideally, the best approach to identify the 5'UTR-binding protein(s) is to search for proteins that originate from a natural host plant of the virus. However, there is a technical difficulty with this approach, as none of the genomes of TBSV's hosts were sequenced to any significant levels when my Ph.D. thesis project was initiated. Additionally, it is difficult to prepare functional protein extracts from plants, especially extracts where added RNA would be stable. These issues presented a significant hurdle to identifying proteins, because my tactic to find 5'UTR-binding protein(s) was to use the 5'UTR as bait for RNA affinity purification from protein extracts, and then use mass spectrometry to identify the bound proteins. The mass spectrometry and downstream analyses would determine mass-to-charge ratios (m/c) of the tryptic peptides from the candidate protein(s) and this information would then be used to find

matching values of the known peptides from a protein databank. However, if the genome of the host plant of interest was not sequenced and available in a databank, then this approach would not be feasible. Additionally, obtaining RNase-free functional protein extracts from plants is very difficult to achieve.

Therefore, to circumvent these problems, my study utilized yeast as a surrogate model host, because (i) the preparation of RNase-free protein extracts is possible in this system, (ii) yeast supports the replication of TBSV DI RNAs [79] (iii) all of its protein open reading frames are annotated [117], and (iv) genome-wide single gene knockdown or knockout strains are available. Thus, yeast contains the proteins required for DI72 replication, and is a tractable system for protein analysis in extracts and their identification by mass spectrometry.

2.1.2. Structures in the TBSV 5'UTR are required for DI72 replication in both yeast and plant host cells

Previously it was shown that the DI72 TBSV RNA replicon can replicate in yeast when p33 and p92 are provided by co-transfected plasmids encoding these proteins [79]. However, before initiating the study, I wanted to determine if DI72 replication in yeast had the same structural requirements in the 5'UTR (*i.e.* region I) as it did in plant cells. To test this idea, the 5'UTR secondary and tertiary RNA structures known to be important for DI72 replication in a plant cells were disrupted and tested for their effects on DI72 replication in yeast (**Fig. 9**).

Specifically, in mutants M2 and S51b, base pairing in stem 1 of the TSD and the centrally-positioned stem-loop 5 (SL5) were disrupted, respectively, whereas in mutant TD1b,

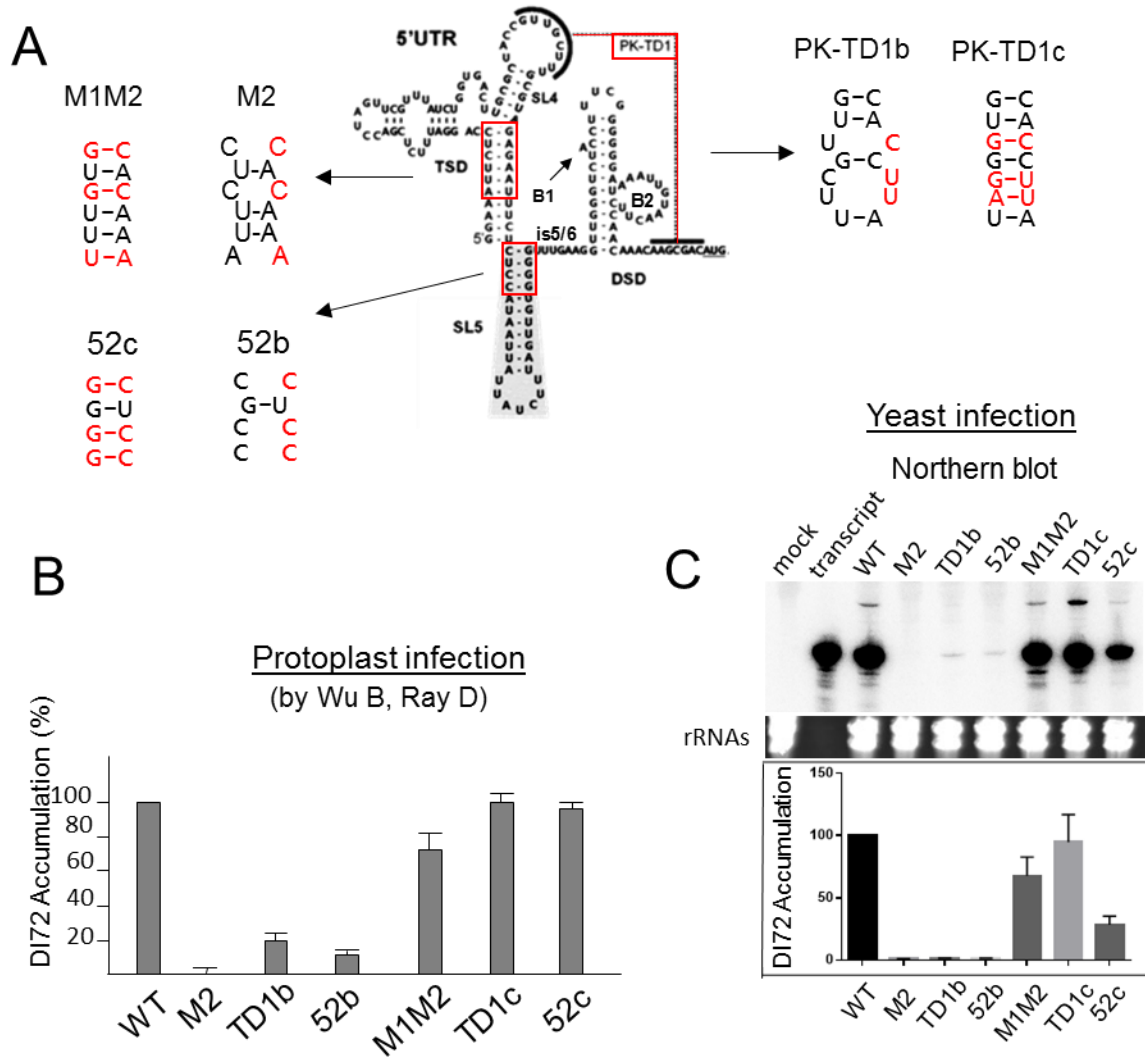


Figure 9. Higher-order RNA structures in the 5' UTR are important for DI72 replication in plant and yeast cells. (A) The three different RNA structures targeted for disruption and restoration are boxed. Nucleotide substitutions in each mutant are shown in red. (B) Summarized published results for replication of DI72 mutants in cucumber protoplasts; conducted by Wu B, Ray D & White KA (69, 70). (C) Replication analysis of the same DI72 mutants in yeast. The upper panel is a Northern blot detecting DI RNAs in total RNA extracts from transformed yeast. The lower graph is quantification of DI RNAs from Northern blots with averages with standard deviations derived from three separate infections. Yeast ribosomal RNAs (rRNAs) levels are shown as loading controls. The error bars represent standard deviations (SD) from three trials. The SD values are 0.23, 0.43, 0.090, 15, 22, 7.0 for M2, TD1b, 52b, M1M2, TC1c and 52c, respectively.

the PK-TD1 pseudoknot interaction was destabilized (**Fig. 9A**). Additional corresponding mutants, M1M2, 52C and TD1c, were also made which had additional substitutions that restored base pairing (**Fig. 9A**). Importantly, the formation of each of these substructures in the 5'UTR/region I (*i.e.* stems 1 and 5, and PK-TD1) was shown previously to be important for efficient DI72 replication in plant protoplasts (**Fig. 9B**) [69, 70]. Specifically, the disruptive mutants notably decreased DI72 levels in plant cells, while those with pairing restored were at near wt levels (**Fig. 9B**).

To test yeast counterparts, the wild-type and mutant DI72s were cloned into pYC, a yeast plasmid that allows for transcription of DI72 from a galactose (GAL1) promoter. Yeast BY4741 strain was transformed with pHisGBK-His33 and pGAD-His92 plasmids, encoding p33 and p92, respectively, along with one of the pYC-DI72 constructs. The transformed yeast was then cultured to assess the levels of replication of the modified DI72s. As shown by Northern blotting in **Fig. 9C**, disruptions of the 5'UTR substructures in M2, 52b and TD1b notably inhibited DI72 replication in yeast. However, when the disrupted structures were restored through compensatory mutations in M1M2, 52c and TD1c, DI72 replication was partially or fully restored (**Fig. 9C**). These results indicate that the RNA substructures assessed are also important in yeast. In conclusion, despite some differences in recovery levels, the results from yeast were generally consistent with those from plant protoplasts, indicating that similar structures of the 5'UTR are required for DI72 replication in both the plant host and yeast. It follows that, if host proteins do interact with the 5'UTR, these protein(s) are likely similar in yeast and plant hosts.

2.1.3. The 5'UTR facilitates DI72 minus-strand RNA production of in yeast.

It was previously shown that disrupting the substructures, TSD, SL5 and PK-TD1 in the 5'UTR led to a reduction of DI72 minus-strand levels in a natural plant host [70, 118]. To further determine whether DI72 replicates similarly in both plants and yeast, the minus-strand levels of the wildtype and mutant DI72 in yeast were assessed. To this end, the yeast total nucleic acid samples obtained from wt and mutant DI72 yeast replication assays (**Fig. 9C**) were first treated with DNase I to remove the pYC DI72 DNA templates. Subsequently, the relative levels of DI72 minus-strand RNA levels were assessed using semi-quantitative RT-PCR. As shown in **Fig. 10**, the RT-PCR results showed that disrupting the substructures in the 5'UTR reduced levels of DI72 minus-strand (**panel 1, Fig. 10**). No visible PCR products were observed when reverse transcriptase was not added (**panel 2, Fig. 10**), confirming that the PCR products in the panel 1 were the results of the amplification of DI72 minus-strand RNA, and not derived from pYC DI72 plasmid templates. RT-PCR results targeting actin mRNA (**panel 3, Fig. 10**) showed that similar amount of total yeast RNAs were used for these RT-PCR analyses. Thus, disrupting the substructures, TSD, SL5 and PK-TD1 in the 5'UTR also reduced the DI72 minus-strand levels in yeast model host, similar to the observations from the plant host [70, 118]. This suggests that, with respect to 5'UTR function, DI72 replicates similarly in both natural plant hosts and yeast.

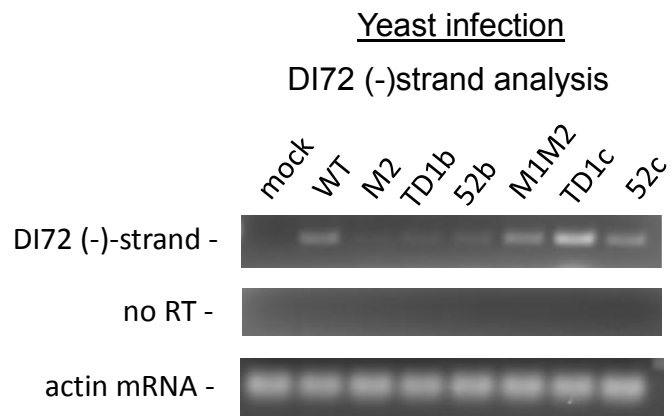


Figure 10. Analysis of DI72 minus-strand levels in yeast. Total yeast nucleic acids were treated with DNase I, and the relative levels of DI72 minus-strands from each sample were determined by semiquantitative RT-PCR (upper panel). PCR was also performed without prior reverse transcription as a control for DNA-derived amplification (middle panel) and the relative levels of actin mRNA levels determined to serve as internal controls (lower panel).

2.1.4. Principles of streptotag RNA affinity isolation of RNA-binding proteins and subsequent identification

To isolate interacting proteins in this approach the 5'UTR RNA is immobilized to a matrix in a column. This is accomplished by using a streptotag RNA aptamer, simply termed streptotag. The streptotag is a 46 nt long RNA that folds into a hairpin-based structure which binds specifically to the antibiotic, streptomycin (i.e. $K_d \approx 1 \mu\text{M}$) in the presence of magnesium ions (**Fig. 11A**) [119]. The streptotag can be incorporated at the 5' or 3' end of a bait RNA or, if the structure of the bait RNA is well characterized, it can be inserted internally in a way that does not disrupt the structure of the bait RNA [120]. To create a hybrid RNA, conventional overlapping PCR is used to generate a dsDNA template containing a T7 promoter. This PCR product is used for in vitro transcription of the streptotag-bait RNA hybrid. The generated chimeric RNA is then immobilized to a streptomycin-conjugated sepharose column (**Fig. 11B**) [120, 121], and total protein extract is added to the column to allow for protein-RNA interactions to occur. After washing the column to remove non-specifically bound proteins, the streptotagged RNA and its bound protein(s) are eluted from the column using excess streptomycin [120-122]. The eluted proteins are then separated by SDS-PAGE and proteins of interest are cut from the gel and submitted for identification via mass spectrometry analysis.

2.1.5. Streptotag RNA does not interfere with DI72 replication in plant.

In order to immobilize the 5'UTR to a column for RNA affinity purification, the streptotag has to be fused with the 5'UTR. However, this addition should not cause misfolding that could

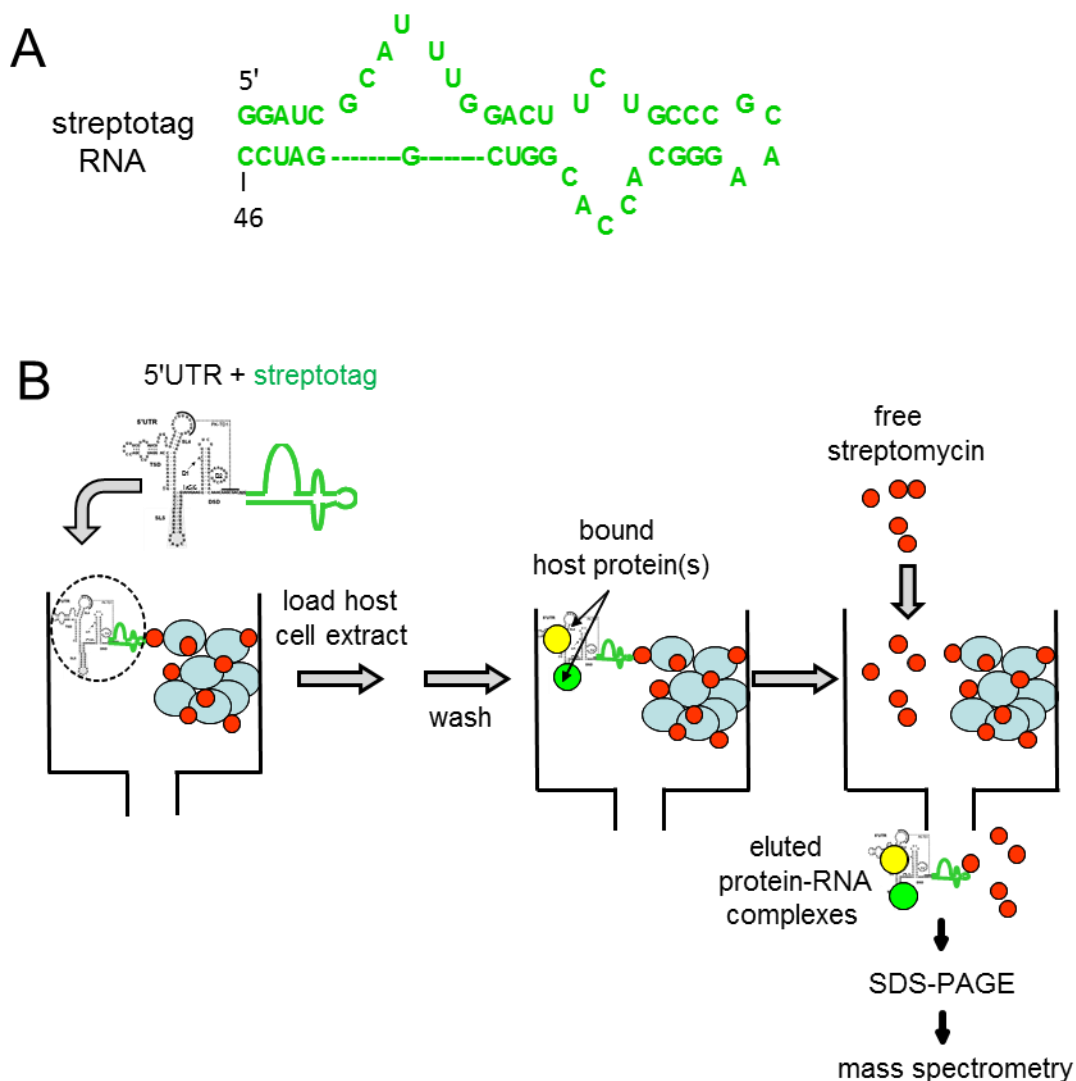


Figure 11. Streptotag Affinity Purification (A) The sequence and secondary structure of the streptotag RNA aptamer. The streptotag binds to streptomycin in the presence of magnesium ions (119). **(B)** The steps required for streptotag RNA affinity purification are illustrated. The hybrid 5'UTR-streptotag RNA is first immobilized to a streptomycin-conjugated column (streptomycin shown in red and sepharose matrix in pale blue). Next, a host cell extract is loaded, and after washing the column, the proteins bound to the RNA are eluted using an excess amount of streptomycin. The eluted protein-RNA complexes were then analyzed by SDS-PAGE and subjected to mass-spectrometry to identify 5'UTR-binding proteins.

interfere with the normal function of the 5'UTR sequence. To test the affect of insertion within or next to the 5'UTR, the streptotag sequence was inserted in different positions in several mutant DI72 RNAs (**Fig. 12A**). In ST1, the streptotag was fused as an extension of SL3. In ST2 and ST3, the streptotag was introduced internally as an extension of SL5, with different added junction nucleotides. In ST4, the streptotag was added at the 3' end of the 5'UTR. *In silico* analysis of folding of these mutant DI72 RNAs using the secondary structure predicting program *mFold* suggested that the modified 5'UTRs would fold properly (data not shown) [123].

After generating the mutant streptotagged DI72 RNAs by *in vitro* transcription, they were co-transfected along with TBSV genome into cucumber protoplasts. The TBSV genome provides the p33 and p92 proteins that facilitate DI72 replication in the protoplasts. After incubating the protoplasts for 22 hrs, the total RNAs were isolated and Northern blot analysis was carried out. As shown in **Fig. 12B**, addition of the streptotag at any of the chosen positions did not inhibit DI72 replication, and instead augmented accumulation to different degrees. Since the additional sequence in SL5 was well tolerated in ST3, this streptotagged 5'UTR, termed ST3-5'UTR, was chosen as a bait RNA to affinity-purify 5'UTR-binding protein(s).

2.1.6. Binding of the streptotagged 5'UTR to the column

Prior to attempting affinity purification of 5'UTR binding protein(s), an additional test was performed to confirm that ST3-5'UTR was able to bind to the streptomycin-coupled sepharose column. For this, *in vitro* transcribed ST3-5'UTR was loaded onto a streptomycin-conjugated column and, after washing the column, the bound RNA was eluted with excess

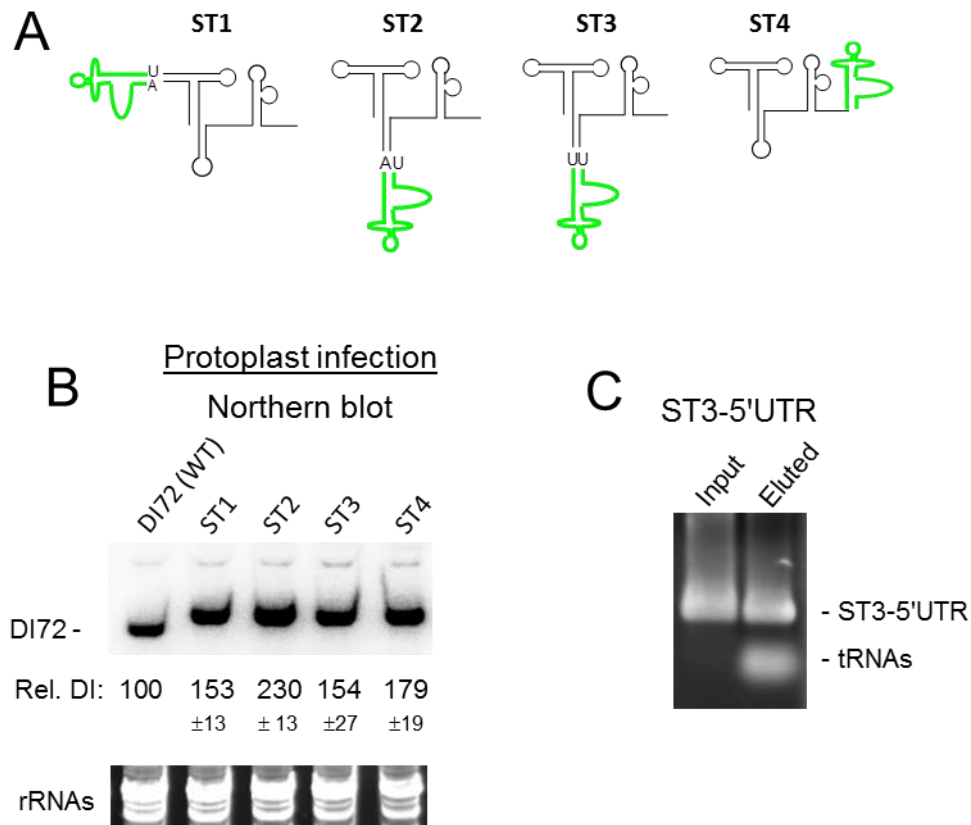


Figure 12. Streptotagged 5'UTRs. (A) The locations where the streptotag (green) was introduced within the 5'UTR are shown for each mutant. Two extra nucleotides at the junction sites of ST2 and ST3 are also shown. (B) Northern blot analysis to assess replication of wildtype DI72 and streptotagged DI72 mutants in plant protoplasts. DI RNA levels were determined with standard deviations from three repeats. Plant rRNAs levels are shown as loading controls. (C) Binding efficiency of ST3-5'UTR to streptomycin-conjugated column was evaluated by comparing the input and eluted amounts of ST3-5'UTR RNA via ethidium bromide-stained agarose gel analysis. The results for column loading for ST3-5'UTR is shown. The input lane contains 5% of the loaded RNA and the eluted lane contains 5% of the RNA eluted after binding RNA. tRNAs were added as a carrier RNA for the EtOH precipitation step of recovered eluted RNA to maximize yield.

streptomycin. To gauge the binding efficiency of ST3-5'UTR to the column, 5% of the initial amount of the RNA applied to the column was compared to 5% of the eluted amount. As shown in the agarose gel (**Fig. 12C**), the input and eluted amounts of streptotagged RNA were similar, showing that most of the streptotagged RNA added was efficiently bound to the column and retained during the washing step. The maximum amount of streptotagged RNA that could be loaded onto the column was calculated to be ~30 µg of streptotagged RNA per 0.5 ml of streptomycin-conjugated column. These binding results showed that ST3-5'UTR binds efficiently to the column and validated its use for isolation of 5'UTR-binding protein(s).

2.1.7. Identification of the 5'UTR-binding protein(s)

For affinity isolation of 5'UTR-binding protein(s), total yeast soluble protein extract was added to the affinity column containing ST3-5'UTR as bait RNA. As a negative control, another column was prepared with a streptotagged ~200 nt long nonviral RNA that was in vitro transcribed using a portion of the pUC19 backbone as a template. Following addition of the yeast protein lysate to both columns, the eluted protein-RNA complexes were separated by 10% SDS-PAGE, and the gel was subsequently silver-stained (**Fig. 13A**). The co-eluted streptotagged RNAs were visible as yellow regions in the silver-stained SDS-PAGE (**Fig. 13A**, lanes 2 and 3). Thus, to improve the view in the areas covered by the RNAs, another aliquot of eluted sample was treated with RNase A before SDS-PAGE analysis (**Fig. 13A**, lanes 4 and 5). RNase A alone was also loaded in the same gel to identify the position of this protein (**Fig. 13A**, lane 6). Both ST3-5'UTR and nonviral RNA samples showed many similar proteins bound to

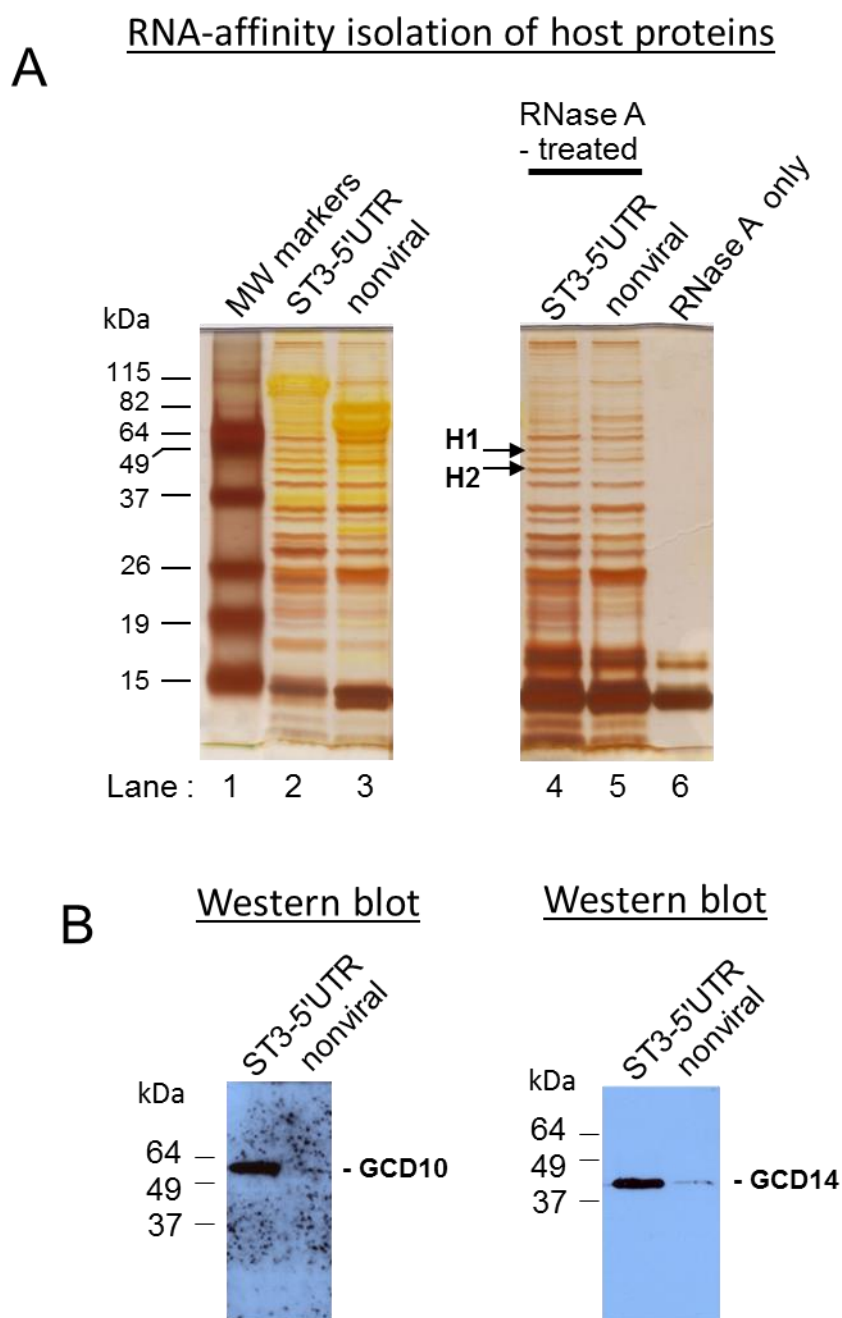


Figure 13. Streptotag RNA affinity isolation of host proteins. (A) A silver-stained SDS-PAGE gel from streptotag affinity isolation. The eluted protein-RNA complexes from ST3-5'UTR or nonviral RNA affinity isolation were analyzed. The yellow regions in the panel on the left correspond to eluted RNAs, which were removed by RNase A treatment in the right panel. H1 and H2 bands correspond to host factors 1 and 2 that were enriched in streptotag-based isolations. Mass spectrometry analysis showed that the H1 and H2 bands contained GCD10 and GCD14 proteins, respectively. **(B)** Western blot analysis with GCD10 and GCD14 antisera confirming the identity of the candidate proteins.

their bait RNAs. Even though extensive column-washing was performed, the majority of the proteins were nonspecifically bound. However, two proteins (H1 and H2) interacted preferentially with ST3-5'UTR (**Fig. 13A**, compare lanes 4 and 5). Throughout repeated trials, these two bands were consistently associated with ST3-5'UTR bait RNA.

In order to obtain a sufficient amount of the candidate proteins for subsequent mass spectrometry analysis, the eluted samples from three trials were pooled and acetone-precipitated. Next, these concentrated samples were separated by SDS-PAGE, stained with Coomassie blue, and the two specific protein bands were excised and sent for a mass spectrometry analysis. In parallel, as controls, the corresponding zones in the nonviral RNA lane were also excised and sent to the mass spectrometry facility. Mass spectrometry analysis was performed through liquid chromatography MS/MS analysis by the mass spectrometry facility at the Hospital for Sick Children at Toronto. The results revealed that the major proteins present in the excised bands from ST3-5'UTR were yeast GCD10 and GCD14 proteins. In contrast, these two proteins were largely absent in the control nonviral RNA sample, as judged by the mass spectrometry analysis (data not shown).

To confirm that the proteins in the two bands were indeed GCD10 and GCD14 proteins, antibodies against yeast GCD10 and GCD14 proteins were obtained from Dr. A. Hinnebusch (National Institute of Health, USA) and western blot analysis was performed on the eluted samples from both ST3-5'UTR and nonviral affinity isolation. As shown in **Fig. 13B**, the samples from ST3-5'UTR reacted efficiently with the GCD10 and GCD14 antibodies, while the control nonviral derived sample did not. Taken together, from the mass spectrometry and western blot

analyses, it was concluded that GCD10 and GCD14 proteins bind preferentially to the TBSV 5'UTR.

2.2. PART2: Characterization of Binding of GCD10/GCD14 Proteins to the 5'UTR

2.2.1. General Information about GCD10 and GCD14 proteins

The GCD10 and GCD14 proteins, formally known as tRNA modifying enzymes 6 and 61 (TRM6 and TRM61), respectively, are essential in yeast and are likely also vital in other eukaryotes [124, 125]. Importantly, orthologs of GCD10 and GCD14 also exist in plants. Therefore, in addition to yeast, the interaction of these two proteins with the viral 5'UTR may be relevant in natural hosts. As a functional enzyme, GCD10 and GCD14 proteins form a heterotetramer complex (GCD10₂ GCD14₂) [126-129]. This complex binds to most eukaryotic tRNAs and methylates N1 of adenosine at position 58. This particular methylation event has been observed in all three domains of life (Bacteria, Archaea, and Eukaryota), suggesting that evolutionarily this modification is very old [124]. Methylation of initiator tRNA at A₅₈ is essential in yeast and without the methylation of tRNA^{i_{met}}, the tRNA is unstable and degraded rapidly [126, 129].

2.2.2. The T-shaped domain is required for GCD10 and GCD14 binding to the 5'UTR

To further test whether GCD10 and GCD14 proteins interact specifically with the 5'UTR, the binding of these proteins to two mutated 5'UTRs was tested. Streptotagged 5'UTR transcripts, ST-M2-5'UTR and ST-TD1b-5'UTR, were generated in which pairing in either stem 1

of the TSD or the pseudoknot was disrupted, respectively (**Fig. 14A**). Corresponding D172 mutants with these disruptions are highly defective for replication in both plant and yeast cells (refer to **Fig. 9**). ST-M2-5'UTR and ST-TD1b-5'UTR were applied separately to columns and used as bait RNAs to test whether GCD10 and GCD14 proteins could bind to these mutant 5'UTR RNAs. After performing streptotag affinity isolation, the eluted proteins from ST-M2-5'UTR and ST-TD1b-5'UTR were separated by SDS-PAGE and silver-stained (**Fig. 14B**). As a positive control, streptotag affinity purification was also performed with the wildtype ST-5'UTR. The results indicated that GCD10 and GCD14 proteins bound efficiently to the wildtype ST-5'UTR and ST-TD1b-5'UTR, but not to ST-M2-5'UTR or the nonviral control, and these results were confirmed by western blotting (**Fig. 14B**, lower panel). Accordingly, GCD10 and GCD14 proteins were unable to bind to the 5'UTR when stem 1 in the T-shaped domain structure was disrupted, but were not affected by disruption of the pseudoknot formation. Thus, GCD10 and GCD14 binding requires formation of the T-shape domain in the 5'UTR.

2.2.3. Expression and purification of GCD10 and GCD14 proteins

The next question addressed was whether GCD10 and GCD14 proteins bind directly to the 5'UTR. Although the streptotag RNA affinity isolation showed that GCD10 and GCD14 proteins interact with the 5'UTR, this approach does not distinguish if the two proteins bind directly or indirectly. Therefore, it is possible that an unknown protein(s) acts as a bridge between the TBSV 5'UTR and GCD10/14 proteins.

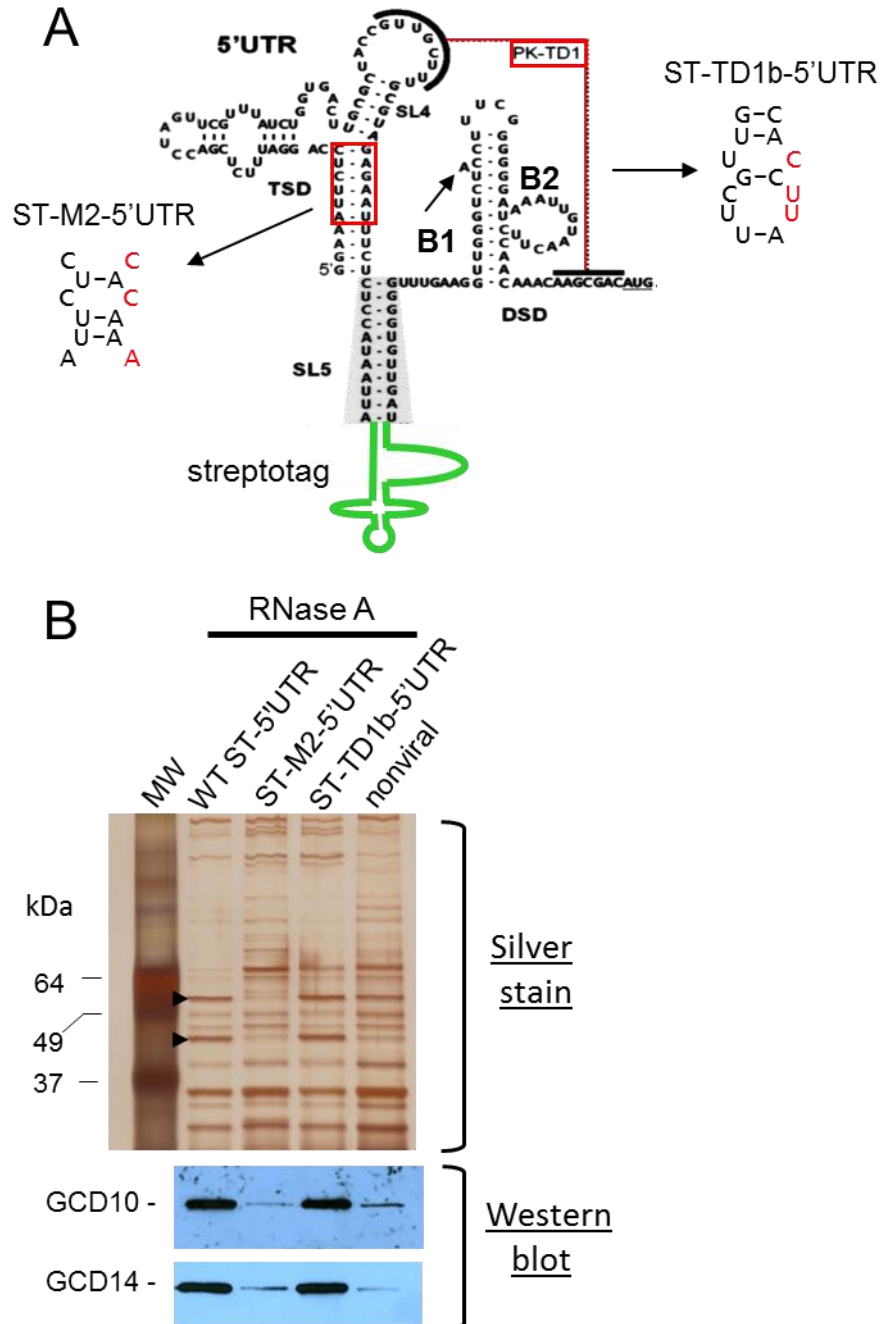


Figure 14. Streptotag affinity isolation using mutant 5'UTRs. (A) The secondary structures of the streptotagged 5'UTRs are shown. The nucleotide substitutions are shown in red with the corresponding mutant names. **(B)** A silver-stained SDS-PAGE gel from streptotag affinity isolation. The eluted protein-RNA complexes isolated with wt and mutant streptotagged 5'UTRs were analyzed after RNase A treatment. The black arrowheads indicate the locations of GCD10 (upper) and GCD14 (lower) proteins. The lower panels show the results from the western blot analysis of the samples using anti-GCD10 and anti-GCD14 antisera.

To address this possibility, GCD10 and GCD14 proteins were his-tagged at their C-termini (GCD10His and GCD14His, respectively), expressed individually in *E. coli*, and purified through polyhistidine affinity. In addition, both proteins were also simultaneously expressed in *E. coli* from a bicistronic plasmid, in which case only GCD14 protein was C-terminally his-tagged. Recombinant protein expression in *E. coli* was confirmed by western blot analysis of total protein extracts using anti-polyhistidine antibody (**Fig. 15A**). As a control, an unrelated protein, maltose-binding protein (MBP) was also his-tagged, expressed, and purified using the same isolation scheme. Aliquots of nickel column-purified samples were separated by SDS-PAGE and Coomassie-stained, revealing highly enriched levels of the proteins of interest (**Fig. 15B**). Co-expression of GCD10 and GCD14His proteins and the subsequent purification via the GCD14His protein resulted in co-purification of GCD10 as well (**Fig. 15B**), indicating that GCD10 and GCD14 proteins form a complex in *E. coli* [127, 128]. GCD10 and GCD14 co-purification yielded approximately 1:1 molar ratio of these proteins, as would be anticipated for stoichiometrically correct assembly of the complex.

2.2.4. Electrophoretic Mobility Shift Assay (EMSA)

To determine whether GCD10 and GCD14 proteins bind directly to the 5'UTR, the 5'UTR RNA was in vitro transcribed and radiolabelled internally with α -³²P-UTP. As a negative control, radiolabelled mutant M2 5'UTR RNA was also generated, since these proteins were shown previously to not bind efficiently to a 5'UTR containing a disrupted stem 1 (refer to ST-M2-5'UTR in **Fig. 14B**). The purified GCD10/GCD14 complex was then incubated with radiolabelled

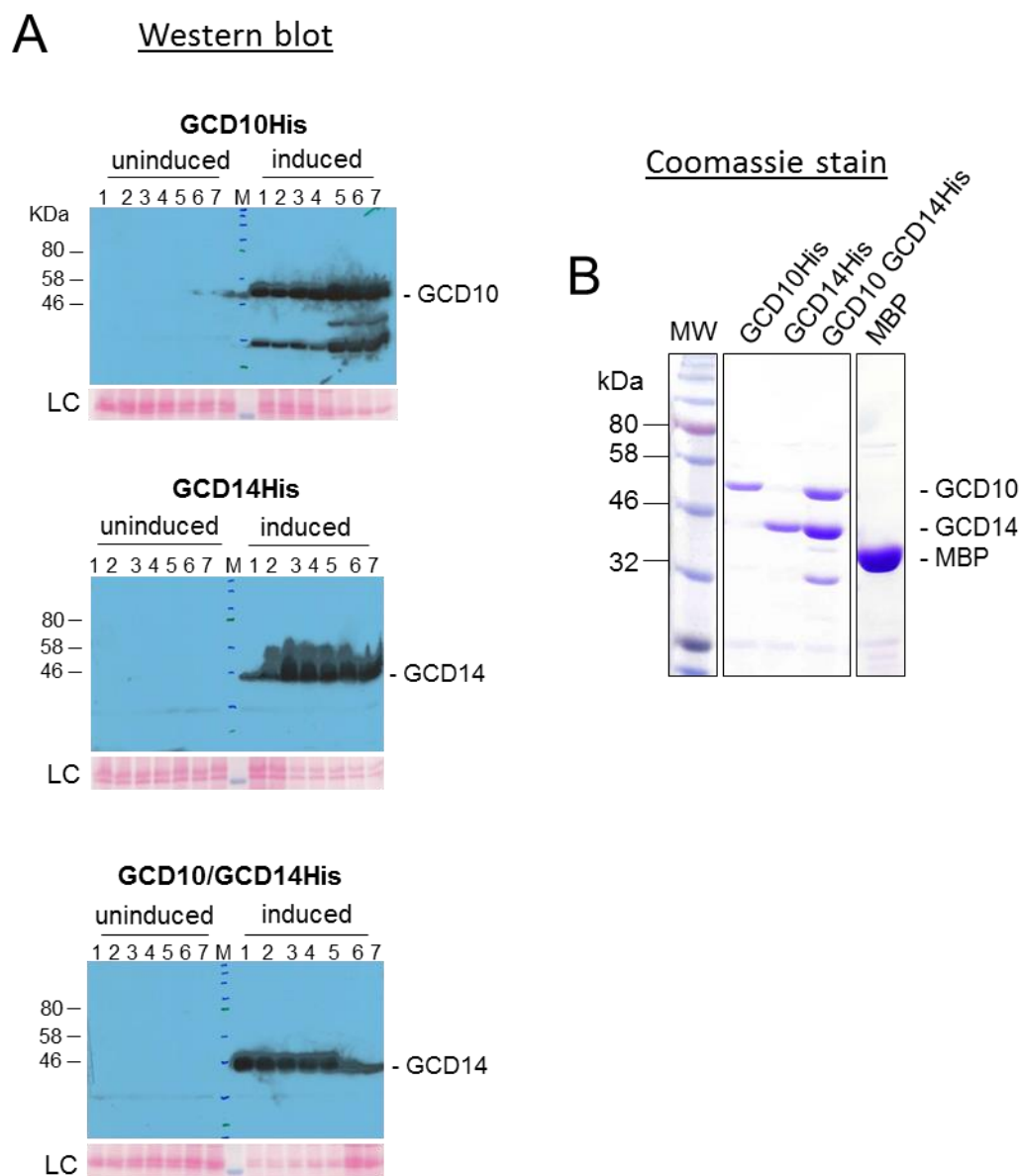


Figure 15. Expression and purification of GCD10, GCD14 and MBP. (A) Expression of GCD10His, GCD14His and GCD10/GCD14His proteins in *E. coli* was confirmed by western blotting using anti-polyhistidine antibody. The top, middle and lower panels show the expression of GCD10His, GCD14His and GCD10/GCD14His, respectively. The numbers above each lane represent different colonies that were analyzed from uninduced or induced cells. Lane M carried molecular size marker. Ponceau S staining of the membranes serve as loading controls (LC). **(B)** Coomassie stained SDS-PAGE of the proteins isolated using polyhistidine tag affinity purification. MBP, maltose binding protein.

wt 5'UTR and M2 5'UTR RNAs and analysed by EMSA for direct protein-RNA interactions. As shown in **Fig. 16A**, GCD10/GCD14 proteins bound to the wt 5'UTR more efficiently than to M2 5'UTR RNA. Quantification of the binding revealed that the dissociation constants for the 5'UTR and M2 5'UTR were 55 nM and 181 nM, respectively (**Fig. 16B**). Thus, wt 5'UTR exhibited about 3.3-fold greater affinity under these in vitro assay conditions.

The EMSA analysis showed that GCD10 and GCD14 proteins bind to the 5'UTR directly, but did not reveal if they do so individually or as a complex. To address this question, the individually purified GCD10His and GCD14His proteins were incubated with the radiolabelled wt 5'UTR, and EMSA analysis was carried out. As a control, MBP was also incubated with the wildtype 5'UTR RNA, and it did not bind to the 5'UTR (**Fig. 17B, C**). Compared to the GCD10/GCD14 complex, individual GCD10 or GCD14 proteins did not bind efficiently to the 5'UTR (**Fig. 17A, C**). The collective results indicate that, GCD10 and GCD14 proteins bind as a complex and interact more efficiently with the wt 5'UTR than with the M2 mutant.

2.3. PART 3: Functionality of GCD10 and GCD14 in DI72 replication

2.3.1. GCD10 protein is important for DI72 replication in yeast

In standard yeast replication assays, TBSV replicase proteins p33 and p92 are expressed from a constitutive ADH1 promoter in their respective plasmids, pHisGBK-His33 and pGAD-His92, while DI72 transcription is under the control of GAL1 promoter from the plasmid, pYC-DI72 (refer to **Fig. 7**). As a consequence, the addition of galactose to the yeast medium launches DI72 transcription, allowing the constitutively expressed p33 and p92 to engage the

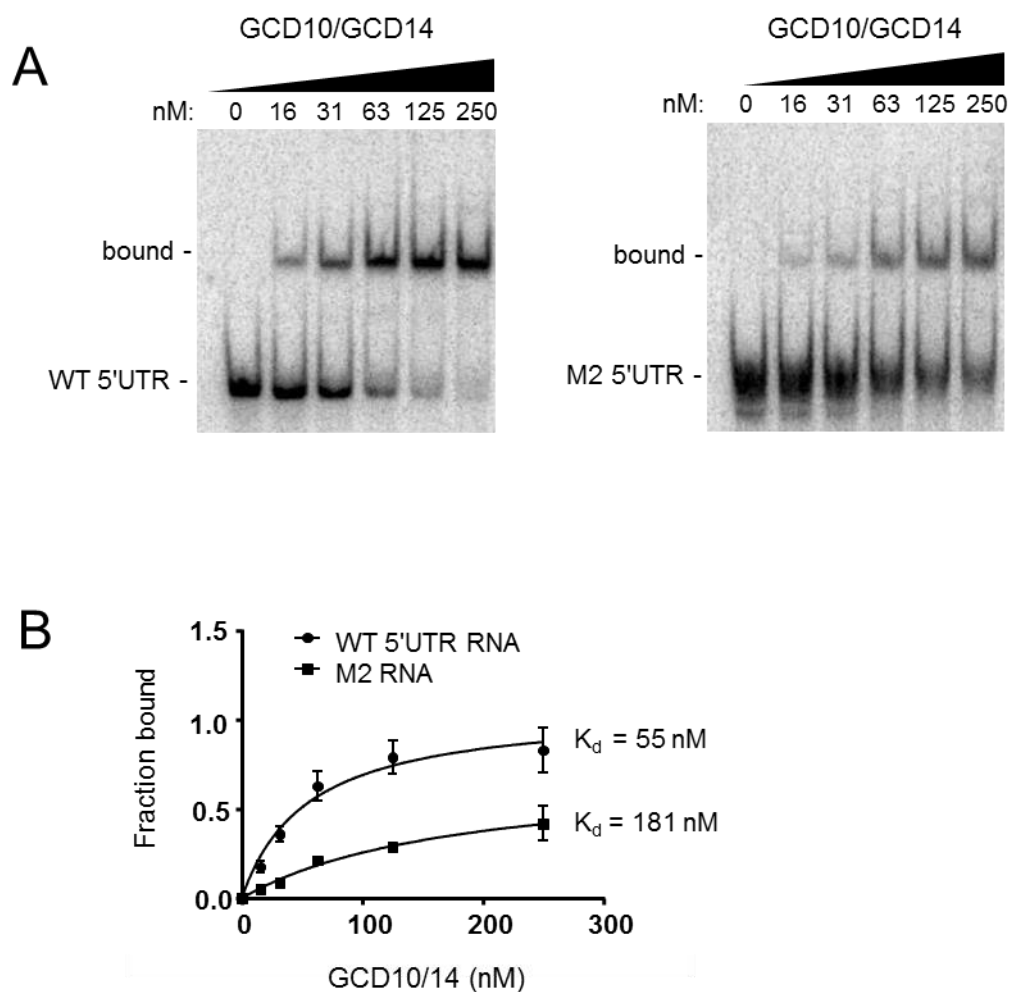


Figure 16. EMSA with 5'UTR RNA and purified GCD10/GCD14 protein complex. (A) Increasing amounts of GCD10/14 protein complex were mixed with labelled WT 5'UTR or mutant M2 5'UTR and separated by non-denaturing PAGE. **(B)** The fractions bound from panel A were quantified and binding curves were generated with the bars indicating standard deviation (SD) from three repeated trials. The SD values for WT 5' UTR RNA are 0.037 (16 nM), 0.043 (31 nM), 0.081 (63 nM), 0.091 (125 nM), 0.12 (250 nM), while the SD values for M2 RNA are 0.012 (16 nM), 0.025 (31 nM), 0.0068 (125 nM), 0.023 and 0.10 (250 nM).

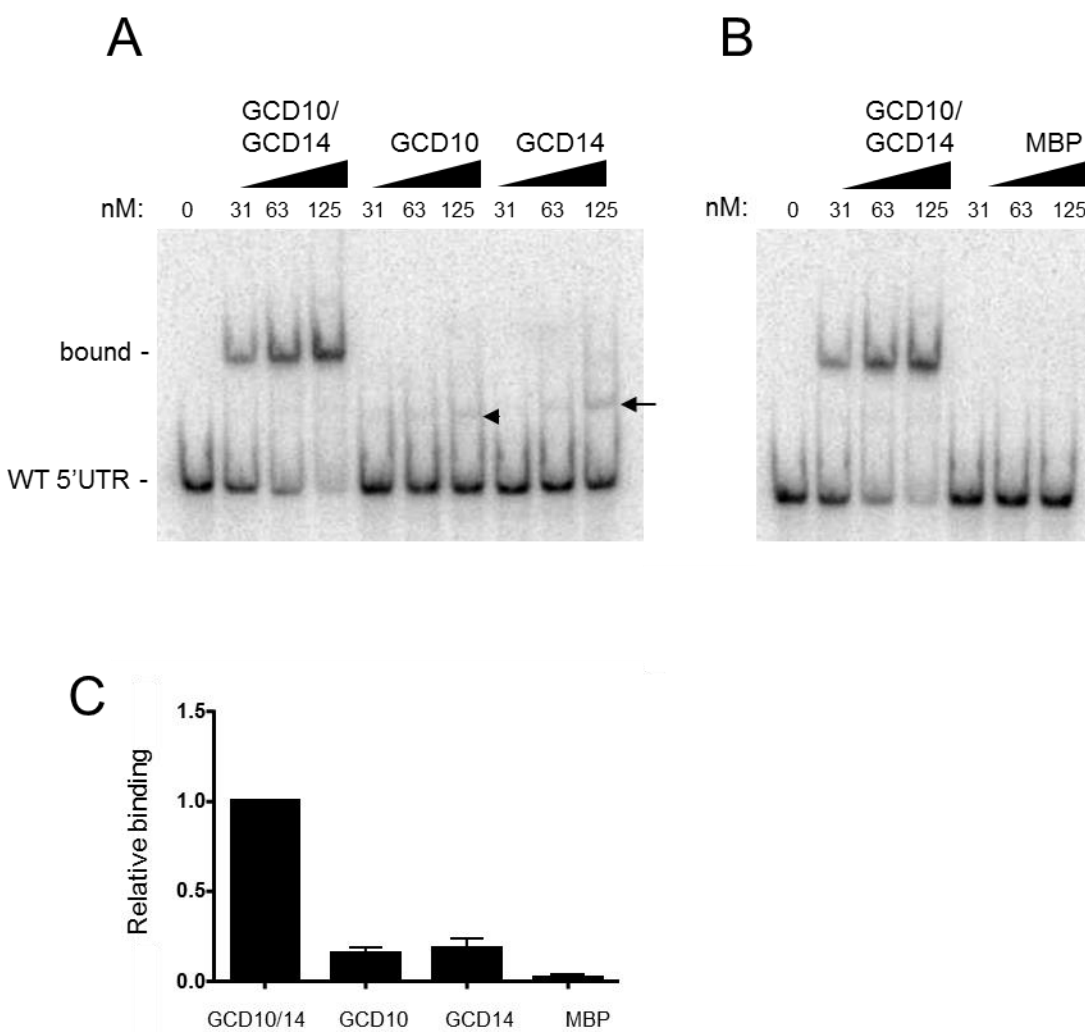


Figure 17. EMSA with individually purified GCD10 and GCD14 proteins. (A) Increasing amounts of either GCD10 protein or GCD14 protein were mixed with the WT 5' UTR. The arrowhead and arrow to the right shows the likely positions of bound GCD10 or GCD14, respectively. **(B)** EMSA analysis comparing the GCD10/14 complex with MBP. **(C)** Quantification of fractions bound at 63 nM for samples in panels A and B. The bar represent standard deviation (SD) from three repeated trials, and their values are 0.032, 0.055 and 0.013 for GCD10, GCD14 and MBP, respectively.

transcript and carry out DI72 replication. In order to test whether a host protein of interest is important for DI72 replication, DI72 replication in wild-type and a knockdown yeast strain of the protein of interest are compared (**Fig. 7**). This type of assay was used to determine whether GCD10 protein is important for DI72 replication.

The wild-type or GCD10 knockdown strains were transformed with pHisGBK-His33, pGAD-His92 and pYC-DI72 plasmids. After growing the transformed strains overnight in glucose medium, the strains were converted to galactose medium for 24 hrs to induce DI72 transcription and allow for its replication by p33 and p92. Subsequently, total yeast RNAs were isolated and Northern blot analysis was carried out. At least three different colonies were examined for wild-type and for GCD10 knockdown strains, and 5 trials were performed. As shown in **Fig. 18A**, DI72 RNA replicated robustly in the wild-type strain. In contrast, in GCD10 knockdown strain, DI72 replication was significantly reduced to 16% of that of the wild-type strain. As a loading control, the same northern blot was also probed for 25S ribosomal RNA (25S rRNA) and untransformed wild-type and GCD10 strains (mock) were used as negative controls for DI72 replication (**Fig. 18A**).

In order to confirm that the reduced DI72 replication correlated with reduced levels of GCD10 protein in the GCD10 knockdown strain, the extracted yeast total proteins were probed for GCD10 protein using anti-GCD10 antibody. Equal loading of the protein samples was verified by Ponceau S staining of yeast total proteins transferred onto the nitrocellulose membrane (**Fig. 18B** middle panel). As shown in the **Fig. 18B** lower panel, the levels of GCD10 protein was noticeably reduced in the GCD10 knockdown strain compared to the wild-type.

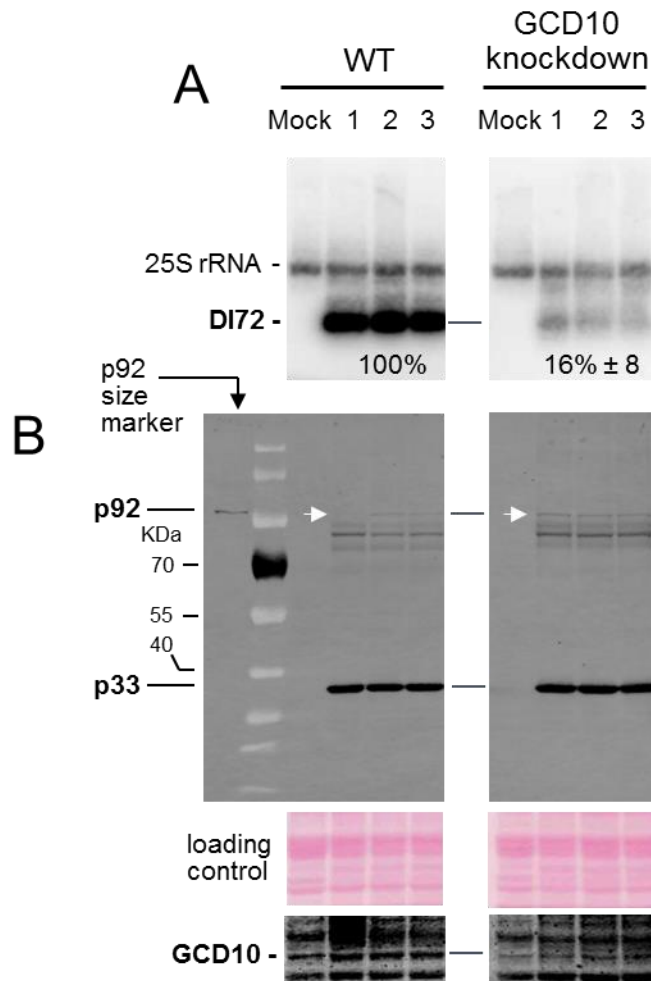


Figure 18. Effect of knockdown of GCD10 protein on DI72 levels. (A) Northern blot analysis. DI72 replication assay was performed in both wildtype and GCD10 knockdown strains. DI72 was probed and quantified. 25S rRNA was also probed as a loading control. The quantified values are shown with standard deviation from three repeated trials. The numbers above the panel means different colonies used. **(B)** Western blot analyses. The upper panel shows the levels of p33 and p92 proteins. Lane 1 serves as a p92 size marker, and lane 2 is a molecular weight marker. The white arrows point out p92 in the blot. The middle panel is a loading control. The lower panel shows the levels of GCD10 protein in both wildtype and GCD10 knockdown strains.

To assess that the reduced DI72 replication in GCD10 knockdown strain was not due to the reduced levels of p33 or p92 replicase proteins, total yeast proteins were extracted from each sample and western blot analysis was carried out using anti-p33 antibody provided by Dr. R. Mullen (University of Guelph, Canada) (**Fig. 18B**). This antibody recognizes both p33 and p92 proteins, as the epitope used to generate it resides in p33, which is also present in p92 [17]. To determine precisely the location of the weakly accumulating p92 protein on the Western blot, total proteins were extracted from the wild-type yeast strain transformed only with pGAD-His92, and the extract was included in the SDS-PAGE analysis as a p92 size marker (**Fig. 18B**, upper panel, first lane). As shown in **Fig. 18B**, p33 levels were similar in both strains, while p92 protein levels were notably higher in the GCD10 knockdown strain. Thus, it can be concluded that the reduced level of DI72 replication in GCD10 knockdown strain is not due to reduced p33/p92 replicase levels and that reducing GCD10 protein impedes DI72 replication.

2.3.2. GCD10 protein is not required for replication of a 5'UTR-lacking DI (Δ 5'UTR-DI)

Although DI72 replication was significantly reduced in GCD10 knockdown strain, it was not known whether the reduced DI72 replication was due to the reduced availability and binding of GCD10 protein to the 5'UTR. To address this question, another DI RNA was constructed in pYC plasmid, termed pYC- Δ 5'UTR-DI. Δ 5'UTR-DI is identical to DI72, except that it lacks the entire 5'UTR. Although the 5'UTR has been shown to be important for replication of both DI72 and the TBSV genome [69, 70], the 5'UTR is not essential [130], and Δ 5'UTR-DI can replicate at a low level in plant cells (~10% that of wt DI72). Theoretically, if GCD10 protein

binds to the 5'UTR to facilitate DI72 replication, then $\Delta 5'$ UTR-DI replication should be independent of GCD10 protein, and DI RNA levels should not be reduced in the GCD10 knockdown strain.

To address this possibility, the wildtype and GCD10 knockdown strains were transformed with pHisGBK-His33, pGAD-His92 and pYC- $\Delta 5'$ UTR-DI, and the replication levels of $\Delta 5'$ UTR-DI was assessed in both wt and GCD10 knock down strains. As shown in **Fig. 19A**, the replication level of the $\Delta 5'$ UTR-DI in the GCD10 knockdown strain was not lower, and instead was ~5-fold higher than that of wt. To determine p33 and p92 levels in both strains, western blot analysis was done with anti-p33 antibody. **Fig. 19B** upper panel shows that p33 levels were similar but p92 levels were higher in the GCD10 knockdown strain, as observed previously (**Fig. 18B, upper panel**). The increased levels of p92 may be, at least partially, responsible for the higher replication of $\Delta 5'$ UTR-DI in GCD10 knockdown strain.

It should be noted that replication levels of $\Delta 5'$ UTR-DI are generally low. To confirm that the bands observed from the Northern blot is a result of its replication, not the result of residual transcription from pYC $\Delta 5'$ UTR-DI plasmid, the wild-type and GCD10 knockdown strains were also transformed with pYC- $\Delta 5'$ UTR-DI and pHisGAD-His92 (without p33), and the isolated total yeast RNAs were also analyzed for the Northern blot (**Fig. 19A**, labelled transcript). The transcription levels of the $\Delta 5'$ UTR-DI without replication was too low to be observed on the blot, therefore the $\Delta 5'$ UTR-DI levels observed are indeed due to replication. Collectively, it can be concluded that, in GCD10 knockdown strain, wt DI72 replication is reduced whereas $\Delta 5'$ UTR-DI replication is enhanced. This supports the idea that GCD10 protein functions through an interaction with the wt 5'UTR to regulate DI72 replication.

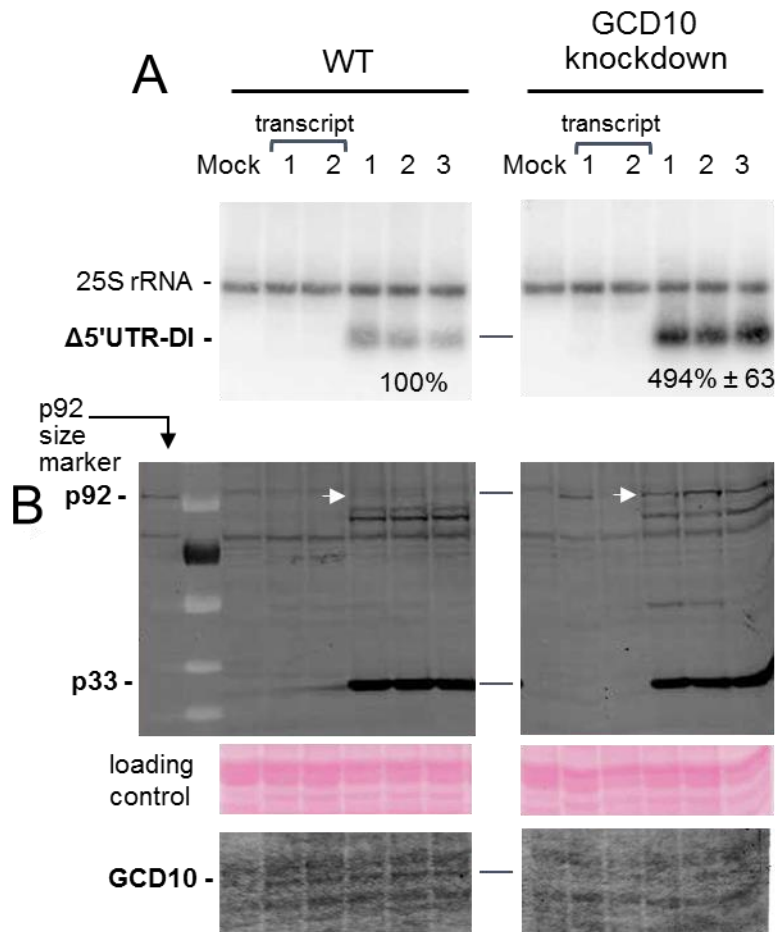


Figure 19. Effect of knockdown of GCD10 protein on $\Delta 5'$ UTR-DI levels. (A) Northern blot analysis. $\Delta 5'$ UTR-DI replication assay was performed in both wildtype and GCD10 knockdown strains. $\Delta 5'$ UTR-DI was probed and quantified. 25S rRNA was also probed as a loading control. The numbers above the panel means different colonies used. Both wildtype and GCD10 knockdown strains were also transformed without pHisGBK-His33 in order to distinguish $\Delta 5'$ UTR-DI transcription from its replication (lanes labeled transcript). **(B)** Western blot analyses. The upper panel shows the levels of p33 and p92 proteins. Lane 1 serves as a p92 size marker, and lane 2 is a molecular weight marker. The white arrows point out p92 in the blot. The middle panel is a loading control. The lower shows the levels of GCD10 protein in both wildtype and GCD10 knockdown strains.

2.3.3. GCD14 protein is also important for DI72 replication

Next, it was investigated whether GCD14 is also important for DI72 replication. For this, wt and GCD14-DAmpp knockdown strain [84] were transformed with pHisGBK-His33, pGAD-His92 and pYC-DI72, and the replication levels of DI72 in both strains were assessed. Because GCD14 is an essential protein, it was expected that the yeast strain would grow more slowly if this protein were significantly knocked down, as was observed for the GCD10-DAmpp knockdown strain (data not shown). However, the growth rate of the GCD14 knockdown strain was undistinguishable from that of the wild-type strain, and DI72 replication levels were similar in both strains (data not shown). At the time of this yeast analysis, there was no anti-GCD14 antibody available, so it was difficult to confirm whether GCD14 protein was indeed knocked down in this strain.

Faced with this uncertain situation, an alternative strategy was used that involved a GCD14 “knockout” provided by Dr. A. Hinnebusch (National Institute of Health, USA). Since GCD14 protein is an essential protein [126], knocking out this gene renders the yeast nonviable. However, Hinnebusch and his colleagues were able to create a viable knockout strain in a two-step process. First, a yeast plasmid overexpressing initiator tRNA (termed hclMT4, from a high copy plasmid) was introduced into wild-type yeast strain, and subsequently, the GCD14 gene was inactivated by its replacement with a URA3 selection marker [131]. This strain requires over-expression of initiator tRNA from hclMT4 for its survival and this plasmid is maintained using a leucine auxotrophic selection marker.

So far, pHisGBK-His33, pGAD-His92 and pYC-DI72 were used for yeast transformation and DI RNA replication assays in my study. However, since pGAD-His92 plasmid is also maintained through leucine selection marker, hclMT4 and pGAD-His92 plasmids could not be used simultaneously to transform the GCD14 knockout strain. In order to solve this problem, the cDNA encoding p92 and its ADH1 promoter were moved from pGAD-His92 into pYC-DI72, creating the plasmid pYC-DI72/p92rev. After transformation of wildtype and GCD14 knockout strains with pGBKHis33 and pYC-DI72/p92rev, the levels of DI72 replication were assessed in both strains. It should be noted here that in order to have the identical genetic background for both strains, the wild-type strain was also transformed additionally with hclMT4 plasmid to overexpress initiator tRNA.

As shown in **Fig. 20A**, the northern analysis revealed that DI72 replication in GCD14 knockout strain was reduced to ~29% of that in the wt strain. Western blot analysis using anti-p33 antibody confirmed that p33 and p92 protein levels were not reduced in the GCD14 knockout strain (**Fig. 20B**, upper panel). In fact, as observed for the GCD10 strain (**Fig 18B**, upper panel), p92 levels in GCD14 knockout strain were notably higher than those from the wt strain (**Fig. 20B**, upper panel). Thus, if the p92 levels in GCD14 knockout strain were reduced to match those of the wildtype strain, an even greater reduction of DI72 replication would be expected. Western blot analysis of subsequently acquired anti-GCD14 antibody was also performed for both wildtype and GCD14 knockout strains, and it confirmed that GCD14 protein was not expressed in GCD14 knockout strain (**Fig. 20B**, lower panel). Taken together, the lack of GCD14 protein correlated with the reduced level of DI72 replication in the knockout strain.

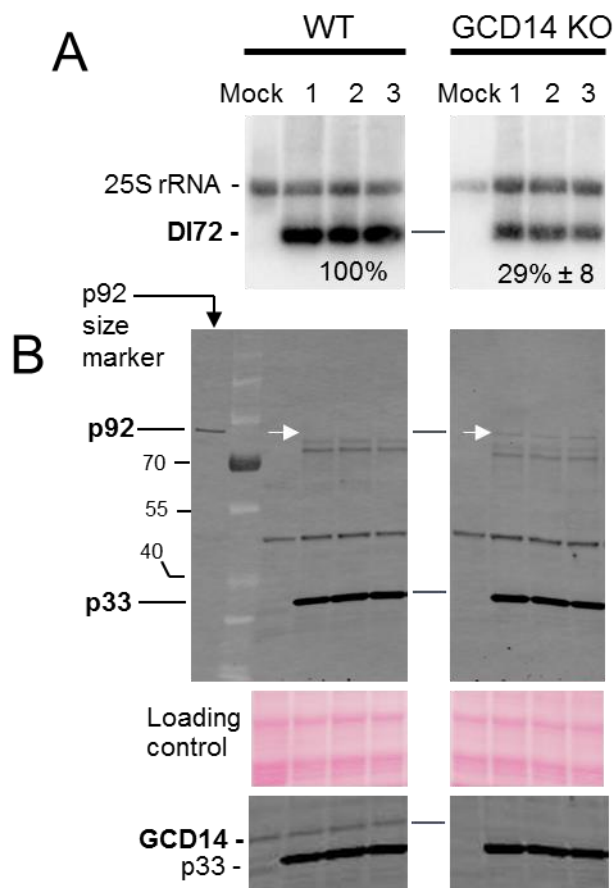


Figure 20. Effect of knockout of GCD14 protein on DI72 levels. (A) Northern blot analysis. DI72 replication assay was performed in both wildtype and GCD14 knockout strains. DI72 was probed and quantified. 25S rRNA was also probed as a loading control. The numbers above the panel means different colonies used. **(B)** Western blot analyses. The upper panel shows the levels of p33 and p92 proteins. Lane 1 serves as a p92 size marker, and lane 2 is a molecular weight marker. The white arrows point out p92 in the blot. The middle panel is a loading control. The lower panel shows the levels of GCD14 protein in both wildtype and GCD14 knockout strains.

2.3.4. GCD14 protein is important for $\Delta 5'$ UTR-DI replication

The reduced DI72 replication seen in the GCD14 knockout strain could be due to the absence of GCD14, which would prevent formation of the GCD10/14 complex that binds to the 5'UTR. For the GCD10 knockdown strain, replication of $\Delta 5'$ UTR-DI was not negatively effected, consistent with the positive effect observed on wt DI72 replication being mediated through its 5'UTR. To determine if the same was true for the GCD14 knockout strain, replication of $\Delta 5'$ UTR-DI was tested.

The wt and GCD14 knockout strains were transformed with pHisGBK-His33 and pYC- $\Delta 5'$ UTR-DI/p92rev plasmids, and the replication of $\Delta 5'$ UTR-DI was assessed through northern blot analysis. As shown in **Fig. 21A**, the replication of $\Delta 5'$ UTR-DI was reduced in GCD14 knockout strain compared to the wildtype strain. Further analysis also revealed that the p33 and p92 levels were not lower in the GCD14 knockout strain (**Fig. 21B**, upper panel) and that GCD14 was indeed knocked out (**Fig. 21B**, lower panel). Thus, these results suggest that GCD14 protein is important for $\Delta 5'$ UTR-DI replication.

The $\Delta 5'$ UTR-DI RNA detected on the northern blot was the result of its replication, because the accumulation of transcripts in the absence of p33 replication protein was negligible (**Fig. 21A**, lanes 2-3 and 8-9). Also, the levels of $\Delta 5'$ UTR-DI transcripts was more closely examined in wt and GCD14 knockout strains in an overexposure of the northern blot to determine if the knockout strain was defective in transcription (**Fig. 21C** lanes 2-3 and 8-9). The $\Delta 5'$ UTR-DI plasmid-derived transcript appeared more abundant in the GCD14 knockout strain

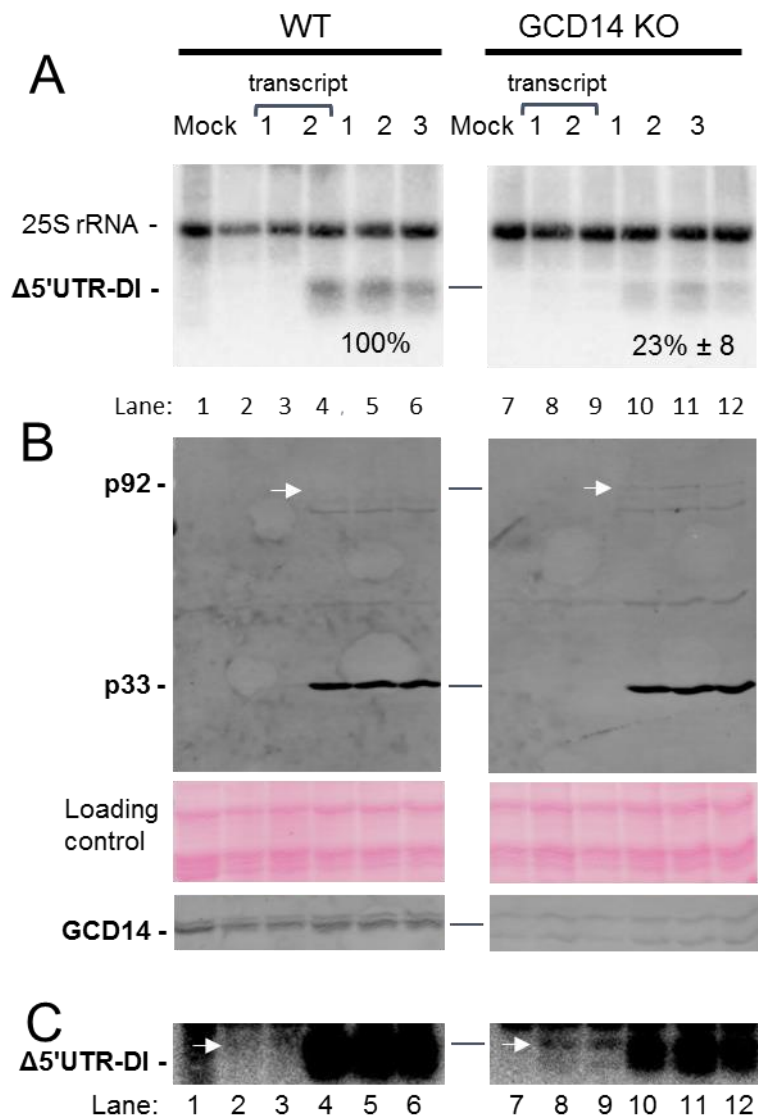


Figure 21. Effect of knocking out GCD14 protein on $\Delta 5'$ UTR-DI levels. (A) Northern blot analysis. $\Delta 5'$ UTR-DI replication assay was performed in both wildtype and GCD14 knockout strains. $\Delta 5'$ UTR-DI was probed and quantified. 25S rRNA was also probed as a loading control. The numbers above the panel means different colonies used. Both wildtype and GCD14 knockout strains were also transformed without pHisGBK-His33 in order to distinguish $\Delta 5'$ UTR-DI transcription from its replication (lanes labeled transcript). **(B)** Western blot analyses. The upper panel shows the levels of p33 and p92 proteins, white arrows point out p92 in the blot. The middle panel is a loading control. The lower panel shows the levels of GCD14 protein in both wildtype and GCD14 knockout strains. **(C)** The northern blot in Fig. 21(A) was over-exposed to visualize the $\Delta 5'$ UTR-DI transcript (see the lanes 2, 3 vs. 8, 9). The white arrows show the location of the transcript.

than the wt strain, therefore the lower level of replication in the GCD14 knockout strain was not due to less efficient launching of the $\Delta 5'$ UTR-DI. The collective results indicate that GCD14 protein is not only important for DI72 replication but also important for replication of $\Delta 5'$ UTR-DI.

2.3.5. Knocking down GCD10 or knocking out GCD14 reduces DI72 minus-strand levels

As shown previously in **Figs. 9** and **10**, disrupting the substructures in the 5'UTR not only reduced the levels of the DI72 plus-strand, minus-strands accumulation was also diminished. Knocking down GCD10 protein or knocking out GCD14 proteins significantly reduced the levels of DI72 plus-strand (**Figs. 18** and **20**). To determine whether GCD10 and GCD14 proteins also affect the levels of DI72 minus-strands, GCD10 knockdown and GCD14 knockout strains were assessed. Total yeast nucleic acid samples in **Figs. 18** and **20** were first treated with DNase I to remove pYC-DI72 DNA template, and a semi-quantitative RT-PCR was performed to determine the relative levels of DI72 minus-strands. As shown in **Fig. 22A**, the minus-strand levels were also reduced in GCD10 knockdown strain compared to those from the wild-type strain. Also, the minus-strand levels in GCD14 knockout strain was lower than those of the wild-type strain (**Fig. 22B**). Taken together, similar to the effect of disrupting the 5'UTR substructures, knocking down GCD10 protein or knocking out GCD14 protein reduced the levels of DI72 plus- and minus-strands. These results show that the effect of disrupting the 5'UTR substructures is similar to the effect of downregulation of GCD10 protein and GCD14 proteins. This, in turn,

supports the idea that yeast GCD10 and GCD14 are important 5'UTR-binding proteins for DI RNA replication.

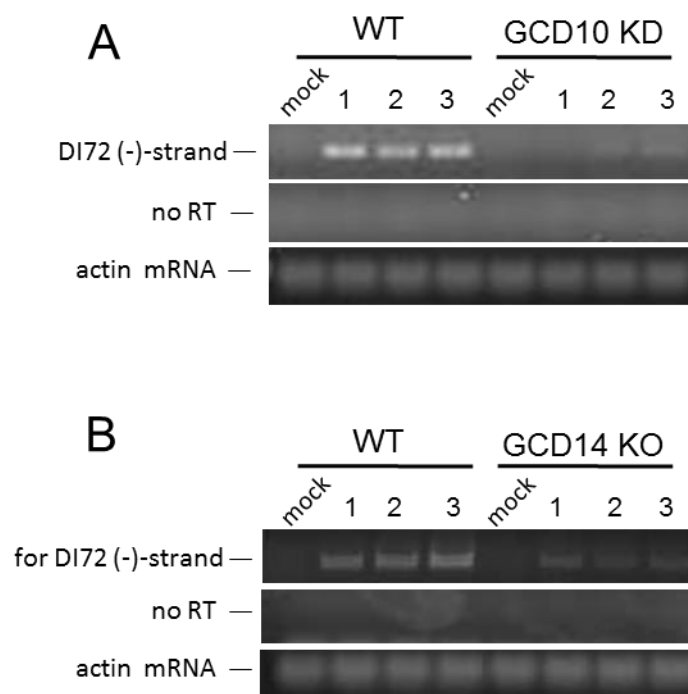


Figure 22. Analysis of DI72 minus-strands in GCD10 and GCD14 mutant strains by semiquantitative RT-PCR. (A) Analysis of GCD10 knockdown mutants and **(B)** analysis of GCD14 knockout mutants. Total yeast nucleic acids were treated with DNase I, and the relative levels of DI72 minus-strands from each sample were determined by semiquantitative RT-PCR (upper panel). PCR was also performed without prior reverse transcription as a control for DNA-derived amplification (middle panel) and the relative levels of actin mRNA levels determined to serve as internal controls (lower panel).

CHAPTER 3

DISCUSSION

Plants are essential for our daily lives, and plant viruses are significant plant pathogens affecting crop yield and economic loss [1-3]. Notably, positive-strand RNA viruses are the major plant viruses found in nature [4], and understanding the molecular mechanism of their reproductions will undoubtedly be useful in developing antiviral approaches for better crop yields. Moreover, insights learned from understanding positive-strand plant RNA viruses will likely expand our general understanding of positive-strand RNA viruses infecting animal, insect, yeast or bacteria, because they all likely share the same evolutionary origin [132].

In this study I investigated the hypothesis that a host protein(s) binds to the 5'UTR of TBSV RNA genome and contributes to viral genome replication. The study used yeast as a surrogate host and a DI RNA as a translation-independent viral RNA replicon. A protein complex consisting of GCD10/14 proteins was identified to preferentially bind to the 5'UTR of TBSV. Further analysis revealed that the TSD sub-structure in the 5'UTR was required for GCD10/14 binding. Also, decreasing the amount of the components of this complex negatively affected DI RNA levels in yeast cells. Lower levels of GCD10/14 proteins lowered both minus and plus-strand accumulation of the DI RNA, suggesting that the defect originated at the level of minus-strand synthesis. Below, I discuss these results and their relevance to TBSV genome replication and possible modes of action.

3.1. The 5'UTRs of plus-strand RNA viruses

Positive-strand RNA viruses replicate their genome through a negative-strand intermediate, which in turn is copied into plus-strand progeny genome. In different plus-strand RNA viruses, 5'UTRs have been found to effect either plus- or minus-strand synthesis, or both

[133-142]. For plus-strand RNA synthesis, a promoter at the 3'-end of the RNA minus-strand is required for initiation. Since the complement of this sequence is present in the plus-strand, modifications in the plus-strand, *i.e.* the 5'UTR, would be transferred on to the complementary RNA minus-strands. Therefore, mutations to the sequence in the 5'UTR that correspond to promoter sequences in the minus strand would be expected to adversely affect plus-strand production. However, there are cases where plus-strand inhibition is observed with mutation in the 5'UTR that do not map within the corresponding minus-strand promoter, *e.g.* poliovirus and Brome mosaic virus [139, 140]. Similarly, in the full-length TBSV genome, some mutation in the 5'UTR that map outside of the minus-strand promoter region, cause defects in plus-strand genome production [133, unpublished works from S. Wang, O. Chernysheva and K. A. White]. Interestingly, none of the TBSV 5'UTR mutants tested showed defects in genome minus-strand synthesis.

In some viruses, however, *e.g.* dengue virus and Sindbis virus, it was found that mutations in the 5'UTR caused defects in minus-strand genome production [137, 141,143]. This suggests that, in some cases, the 5'UTR of a plus-strand RNA virus genome is able to control the minus-strand initiation process at the 3'-end of the genome. For Dengue virus, it was shown that for minus-strand genome synthesis, the viral RdRp actually binds to the 5'UTR in the genome and then accesses the 3'-promoter by a long-range RNA-RNA interaction [143]. The binding of the RdRp to the 5'UTR explains why this sequence is important for minus-strand genome synthesis. It was hypothesised that the requirement for both ends is a strategy that ensures that only complete genomes are replicated. As mentioned above, for TBSV, none of

the 5'UTR mutants tested caused a defect in minus-strand genome synthesis [133], therefore, in the genome context, the 5'UTR appears to control only plus-strand genome synthesis.

3.2 Pros and cons of the yeast system and the DI RNA replicon

My study used a yeast-based system to investigate a TBSV RNA replicon. This system has both advantages and disadvantages. The well established yeast system provides excellent proteomic and genetic tools for studying its proteins. Fortunately, it was determined by the Nagy lab that yeast cells are able to support TBSV viral RNA replication [79]. This finding established that yeast can provide the proteins necessary for this viral process to occur and, thus, was an acceptable surrogate to study TBSV RNA replication. However, it is possible that this system does not completely reflect what occurs in plant cells. Although sufficient critical proteins are present to allow the TBSV DI RNA to replicate in yeast, it is possible that these cells may lack some proteins that are present in plant cells that contribute to modulating the replication cycle. Consequently, the results from yeast cell may not be completely mirror what occurs in plant cells.

The DI RNA has been used in many studies to identify and characterize RNA elements that contribute to its replication [62-64, 69, 70]. A key advantage of this small replicon is that it is not translated, and this uncoupling allows one to focus on the RNA replication process. However, a disadvantage is that the RNA elements are not in their natural genomic context, which could influence how they work. For example, as mentioned above, the 5' UTR mutants in the genome context did not inhibit minus-strand synthesis, but the same mutations in a DI context caused significant defects in DI72 minus-strand synthesis. It is therefore important to

realize that differences exist and that results with a DI RNA may not always be the same for the full-length genome. My study found that downregulation of the 5' UTR binding proteins shows a similar effect as the mutations introduced to the 5' UTR in the DI context. In both cases, the minus-strand levels were reduced unlike the TBSV genome context. Until we determine how the 5'UTR binding proteins function in the genome context, the function of these proteins may be discussed only in the DI context.

3.3. Identification of TBSV 5'UTR binding proteins

Since dicot plants such as tomatoes are a natural host for TBSV, plant proteins would be ideally for the identification of the 5'UTR binding proteins. However, since the genomes of TBSV host plants were not sequenced to any significant levels, *S. cerevisiae*, which supports DI72 replication, was chosen as an alternative choice to isolate the 5'UTR binding proteins.

Although yeast supports DI72 replication, it was not clear if DI72 replication would be mechanistically similar in both yeast and plants. Specifically, it was uncertain whether the 5'UTR functions similarly in yeast and plants. Therefore, I tested whether the same 5'UTR sub-structures were required for DI RNA replication in both cell types. **Fig. 9** shows that the T-shape domain, SL5, and PK-TD in the 5'UTR are required for efficient DI72 replication in both plant and yeast hosts, suggesting that similar hypothetical 5'UTR binding proteins would likely exist in both hosts. Thus, as a result, yeast proteins were used to identify a 5'UTR binding protein(s).

In order to immobilize the 5'UTR, a streptotag RNA was fused to the 5'UTR, and construct ST3-5'UTR was chosen as the bait RNA. Importantly, insertion of streptotag to SL5 did not interfere with DI72 replication, suggesting that the streptotagged RNA was properly

designed (**Fig. 12**). The streptotag RNA affinity method isolated many yeast proteins, however, within the group, two proteins were found to specifically and consistently bind to the 5'UTR bait. The two proteins were isolated through SDS-PAGE and Commassie staining, and the downstream mass spectrometry identified them as GCD10 and GCD14 proteins (formally, Trm6 and Trm61 proteins). A positive aspect of the results of this co-isolation is that the two proteins normally exist as a complex in natural conditions, suggesting that yeast proteins prepared for the affinity purification were maintained in properly folded states and complexes remained intact throughout the affinity purification.

Previous studies have shown that all sub-structures in the TBSV 5'UTR are important for DI RNA replication [69, 70]. The streptotag-based RNA binding assays indicated that formation of the TSD, but not PK-TD1, was important for GCD10/14 binding (**Fig. 14**). The formation of PK-TD1 involves a segment pairing with SL4 of the TSD, and this interaction may stabilize or alter the TSD structure. Since the PK-TD1 structure is not required for binding, the proteins likely bind to a region of the TSD that is not involved in PK-TD1 formation. Also, because both the PK-TD1 and TSD are important for replication, but the PK-TD1 is not required for binding, the functions of the TSD and PK-TD1, although structurally connected via the PK, may be distinct. It should be noted that the requirement of SL5 for binding was not tested because its disruption would have impeded the folding of the streptotag, which would have interfered with its binding to the column. However, two other regions other than the 5'UTR of DI72 were tested for binding, and it was determined that the proteins do not bind to region II or region IV (data not shown). Therefore, GCD10/GCD14 appear to be specific for the 5'UTR (*i.e.* region I), and the TSD is an important determinant for binding.

3.4. Characterization of the 5'UTR-binding proteins

As a next step to characterize their binding to the 5'UTR, GCD10 and GCD14 proteins were expressed either individually or as a complex in *E. coli* and subsequently purified using his-tag affinity purification. The subsequent EMSA studies using the purified proteins showed that GCD10 and GCD14 proteins can bind directly to the 5'UTR. Also, they bind to the wild-type 5'UTR more efficiently than to functionally-defective M2-5'UTR (**Fig. 16**, up to 3.3-fold). The analyses also revealed that these proteins bind as a complex (**Fig. 17**). Since the active GCD complex is known to consist of two GCD10 molecules and two GCD14 molecules, the EMSA results are consistent with the original streptotag isolation in which both proteins were present in approximately equimolar amounts (**Fig. 13**).

The crystal structure of human GCD10/14 proteins bound to tRNA^{Lys} was recently determined [127]. Also, the crystal structure of yeast GCD10/14 proteins with SAM substrate (methyl donor) was recently solved [128]. The two proteins form a heterotetramer (GCD10₂GCD14₂), and the protein complex contacts most of the tRNA backbone. In addition, the T-loop of tRNA is embedded into the catalytic core of the complex [127]. It is not known how the GCD10/14 complex interacts with TBSV 5'UTR. However, considering that tRNAs and the TSD of the 5'UTR are similar in size (70-90 nts) and have multi-stem junctions, the TSD may be able to adopt a structure that mimics tRNAs and is recognized by the GCD complex.

Notably, neither GCD10 nor GCD14 proteins in the crude yeast extract bound to the streptotagged M2-5'UTR (**Fig. 14**). In contrast, the purified GCD10/14 proteins bound weakly to M2-5'UTR (without streptotag) (**Fig. 16**). This discrepancy between these two results may be explained by the nature of the protein samples used. In the streptotag affinity purification, a

total yeast protein extract containing GCD10 and GCD14 proteins was used, therefore other yeast proteins in the extract may have non-specifically outcompeted GCD10 and GCD14 proteins for binding to the M2-5'UTR mutant. Indeed, many additional RNA-binding proteins were isolated in streptotag pulldowns (**Fig. 14**). Thus, the conditions for the interaction between GCD10/14 proteins and M2-5'UTR was likely more stringent when yeast total extract was used to test binding. As a result, in the streptotag RNA binding assay, GCD10/14 was likely unable to nonspecifically interact with M2-5'UTR RNA to a notable degree due to competition from other nonspecific RNA-binding proteins. In contrast, with the EMSA, purified GCD10 and GCD14 were allowed to more readily bind to M2-5'UTR nonspecifically.

3.5. The functional importance of GCD10 and GCD14 proteins in TBSV DI replication

After characterizing the interaction between GCD10/14 proteins and the 5'UTR through the streptotag affinity assay and EMSA, the functional importance of GCD10 and GCD14 proteins in TBSV viral RNA replication was investigated using a yeast DI72 replication assay. As shown in **Fig. 18A**, knocking down GCD10 protein resulted in a significant reduction of DI72 replication. Importantly, this reduction was not due to altered p33 or p92 levels as shown in the western blot (**Fig. 18B**). Notably, p92 level in GCD10 knockdown strain was higher than that in wild-type strain. If p92 level in GCD10 knockdown strain were reduced to match the p92 level in the wild-type, DI72 replication in GCD10 knockdown strain would have been further reduced, suggesting that knocking down GCD10 protein is highly detrimental to DI72 replication.

To address whether the reduction of DI72 replication in GCD10 knockdown strain is due to the reduced availability and binding of GCD10 protein to the 5'UTR, Δ 5'UTR-DI replication was tested in yeast DI replication assay. In principle, the Δ 5'UTR-DI should replicate independently of the knockdown of GCD10, if GCD10 protein functions only through the 5'UTR. More specifically, replication of the Δ 5'UTR-DI should "not be reduced", unlike reduced DI72 replication in GCD10 knockdown strain, and this is what was observed (**Fig. 19A**). The replication of the Δ 5'UTR-DI was not reduced, but rather enhanced in GCD10 knockdown strain. This may be due to the higher p92 levels in GCD10 knockdown strain (**Fig 19B**). Taken together, the results are consistent with GCD10 protein functioning through the 5'UTR, where reduction of GCD10 protein resulted in reduced DI72 replication levels.

Next, the functional importance of GCD14 protein was studied using GCD14 knockout strain. As shown in **Fig. 20A**, knocking out GCD14 protein inhibited DI72 replication to 28% of wild-type strain. This is an expected result considering that GCD10 and GCD14 proteins bind to the 5'UTR as a complex. By removing the GCD14 subunit, no complex will form, causing reduced DI72 replication. This result also suggests that the GCD10/14 complex is not essential for DI RNA replication, as GCD14 is absent in these assays, yet a low level of replication occurred. Also, as observed for the GCD10 KD (**Fig. 18A**), the reduced replication was not the result of reduced p33 and p92 levels; since p92 levels were actually higher in the GCD14 knockout strain (**Fig. 19B**).

In order to test whether the reduced DI72 replication was due to the lack of the binding of GCD14 protein to the 5'UTR, the replication of Δ 5'UTR-DI was tested in both wild-type and GCD14 knockout strain (**Fig. 21A**). Unexpectedly, and unlike for results with GCD10 protein

knockdown, knocking out GCD14 protein also reduced replication of $\Delta 5'$ UTR-DI. This suggests that GCD14 protein not only functions through the 5'UTR to mediate DI72 replication, but it also contributes to replication by some other direct or indirect mechanism (A mechanistic model discussed below).

3.6. Plus- and minus-strand synthesis and GCD10/14 proteins

The production of progeny plus-strand RNA virus genomes involves the synthesis of a minus-strand RNA intermediate. This process is asymmetrical, with much more plus-strands being produced over minus-strands, i.e. 10-100-fold more plus than minus [80]. This strategy maximizes the use of the cellular ribonucleotide pool for production of the active form of the viral genome and limits the presence of double-stranded RNA intermediates that are triggers for antiviral RNAi in plants [5]. Exactly how this imbalance is maintained is not fully understood, but both RNA elements in the genome and proteins factors likely function in the regulation. Results from previous studies in plant cells showed that modifications in the 5'UTR in the TBSV genome caused defects only in plus-strand synthesis [133, unpublished works from S. Wang , O. Chernysheva and K. A. White]. In contrast, the same mutations introduced into DI RNAs and tested in plant cells resulted in defects in minus-strand synthesis [70, 118]. These observations confirm that different viral contexts for the 5'UTR can lead to contrasting results. Indeed, the genome is a much more complex molecule which, in addition to RNA replication, is involved in translation of viral proteins, transcription of viral sg mRNAs, and packaging of progeny genomes. As all these processes involve the genome, a high level of coordination and regulation is required to allow for these events to occur without cross-interference.

My results with 5'UTR DI RNA mutants in yeast cells are consistent with those observed with 5'UTR DI RNA mutants in plant cells. That is, the accumulation defect manifested at the minus-strand levels (**Fig. 10**). This suggests that the mutations in the 5'UTR DI RNA result in inhibition of minus-strand synthesis. In turn, this would indicate that the 5'-end of the DI RNA is somehow able to communicate with the 3'-end of the DI RNA, where minus-strand initiation takes place. Interestingly, when DI72 minus-strands levels were investigated for GCD10 knockdown and GCD14 knockout stains, a similar defect in minus-strand levels was also observed (**Fig. 22**). Since the same minus-strand defect was caused by a mutation to DI72's 5'UTR in M2-5'UTR (**Fig. 10**) and this mutant 5'UTR was unable to bind to GCD10/14 in crude yeast extracts (**Fig. 14**) or bound more weakly to purified GCD10/14 (**Fig. 16**), the combined results are consistent with the binding of GCD10/14 to the 5'UTR mediating minus-strand RNA synthesis of DI72.

3.7. A working model for the role of GCD10/14 in DI RNA replication

Taken together, based on the overall results from this study, a working model can be built for how GCD10 and GCD14 proteins facilitate DI72 replication in yeast cells (**Fig. 23**). Knocking down GCD10 protein reduced DI72 replication (**Fig. 18**), but not $\Delta 5'$ UTR-DI replication (**Fig. 19**), suggesting that GCD10 protein functions "only through the 5' UTR". Also, GCD10 and GCD14 proteins bind to the 5' UTR as a complex (**Fig. 14**). These results indicate that GCD10 protein binds to the 5' UTR as a GCD10/GCD14 complex to facilitate DI72 replication. This binding is mediated by the TSD sub-structure, but is independent of PK-TD1 formation. As the complex normally binds to tRNAs, it is possible that the TSD mimics the portion of the tRNA to

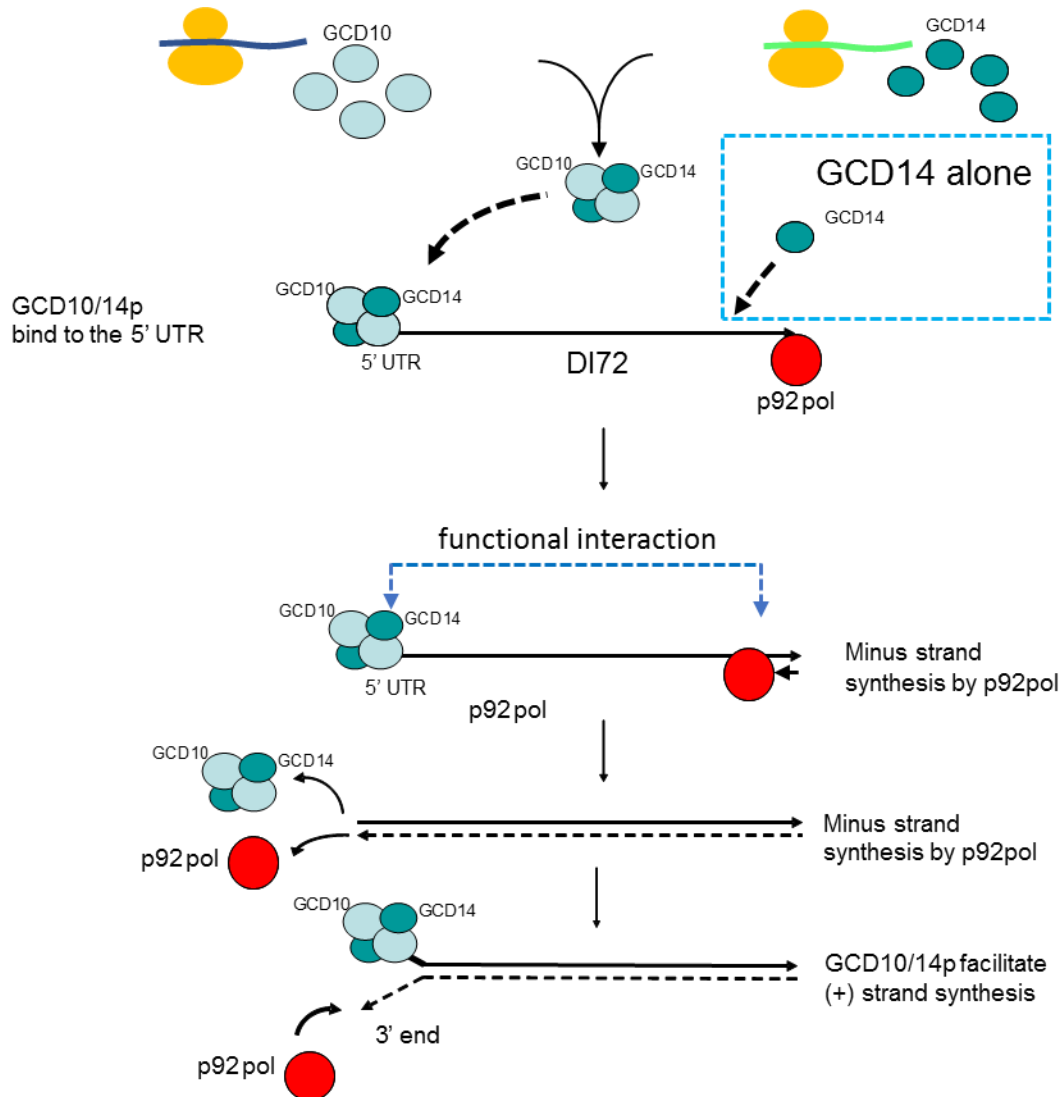


Figure 23. A working model. GCD10 and GCD14 proteins bind to the 5' UTR, and they may be recruited to the replication centre. Their binding to the 5' UTR may affect the minus-strand synthesis. After the minus-strand synthesis, they may bind back to the 5' UTR to allow p92 to initiate the plus-strand synthesis. After its translation, GCD14 protein alone may also be associated with regions other than the 5' UTR to facilitate the viral RNA synthesis.

which GCD10/14 binds. Although the TSD is clearly important for binding, the analysis does not preclude the possibility that other portions of the 5'UTR also contribute to binding. The interaction of GCD10/14 with the 5'UTR is somehow able to promote initiation of minus-strand synthesis at the 3'-end of DI72. As DI72 is relatively small and has a high level of secondary structure, this could provide access of the 5'UTR-bound GCD10/14 to the 3'-end.

Minus-strand synthesis is mediated by p92, the viral RdRp. This polymerase functions in a complex with p33 and host proteins [5]. Precisely how GCD10/14 could affect the activity of this polymerase complex is unknown, but some possibilities could include: (i) recruiting the polymerase complex to the DI RNA template, (ii) becoming part of the polymerase complex or (iii) attracting other components to the polymerase complex. Regardless of the detailed mechanism, the GCD10/14 complex is somehow able to increase the efficiency of minus-strand synthesis. Interestingly, as minus-strand synthesis proceeds toward the 5'-end of the DI RNA, it would have to displace any bound GCD10/14 when copying this region (**Fig. 23**). This suggests that the complex interacts with the 5'UTR in such a way that does not impede polymerase copying of this segment. The displaced complex or a different GCD10/14 may then be able to bind to the 5'UTR of the original genome and facilitate plus-strand synthesis by making the 3' end of the minus-strand more accessible to p92 (**Fig. 23**).

Knocking out GCD14 protein reduced “both” DI72 and $\Delta 5'$ UTR-DI replication (**Figs. 20 & 21**). In the case of a GCD14 knockout, no GCD14 or GCD10/14 complexes would be present in cells. This finding suggests that some GCD14 is needed for efficient replication of $\Delta 5'$ UTR-DI. Although GCD10 and GCD14 proteins normally function as a complex, it has been shown that GCD10 and GCD14 proteins do not always function together. The idea that GCD14 protein can

have some separate functions from GCD10 protein is not unprecedented. It has previously been shown in a human cell line that GCD14 protein alone interacts with protein kinase C α in the cytoplasm, while most of GCD10 protein is localized in the nucleus [146]. Additionally, there are about 3-fold more number of GCD14 molecules than GCD10 in yeast cells [147], suggesting that GCD14 protein may have some independent functions apart from GCD10 protein.

The results with 5'UTR-lacking DI RNA indicate that the role of GCD10 and GCD14 in TBSV viral RNA replication is not straightforward. These proteins bind the 5'UTR as a complex, however this does not preclude other activities for the individual components. Such activities may be the result of direct interaction with the DI RNA or be indirect effects mediated through other components with which they interact. Overall, the results of this study implicate GCD10 and GCD14 as playing important roles in modulating DI RNA replication in yeast.

Interestingly, GCD10 has been implicated in RNA replication in another plus-strand RNA plant virus. The replicase complex purified from tobacco mosaic virus (TMV)-infected tobacco plants contained two viral 183- and 126 kDa replicase proteins and GCD10 protein, but not GCD14 protein [144]. Also, in a yeast two-hybrid screen, GCD10 protein was able to bind to the methyltransferase domain of the viral 183- and 126 kDa replicase proteins [145]. The presence of GCD10 in a viral replicase complex strongly suggests a role in virus genome replication and indirectly supports my finding that TBSV viral RNA replication also involves this protein.

3.8. Future Directions

The functional assay used for GCD10 and GCD14 proteins were based on DI72 replication in yeast heterologous host. An important next step would be to test the functional roles of GCD10/GCD14 proteins in TBSV genome replication in a plant host, such as cucumber protoplasts or *N. benthamiana* plants. Double-stranded RNAs generated from GCD10 and GCD14 ORFs can then used to induce RNA interference and knock down the two proteins in both cucumber protoplasts and *N. benthamiana* plants. The effect of their knockdown on TBSV genome replication and DI RNA replication would then be monitored by northern blotting.

Additional studies could also localize the binding sites for the GCD10/14 complex in the 5'UTR. My results showed that the TSD is important, however further deletions in the 5'UTR combined with EMSAs could allow for the minimal RNA regions to be identified.

Also of interest would be to see if the GCD10/14 complex actually modifies the 5'UTR by methylating an adenine. In Brome mosaic virus, a viral RNA element involved in viral RNA replication, box B motif, was found to be a tRNA mimic of the TpsiC-stem loop in tRNA and two uridines in this sequence was modified to 5mU and pseudouridine [148]. Although it was not determined if the modifications were necessary for function of the viral replication element, it is possible that the changes are important for its activity. Similarly, covalent modifications to the TSD by GCD10/14 could possibly effect its function.

Since GCD10 was identified in the TMV RNA replication complex, it would be worth seeing if either GCD10 or GCD14 are components of the TBSV replication complex. This could be achieved by immunoprecipitating p92 from virus infected cells and determining the host

components that co-immunoprecipitate. The presence of one or more of the GCD proteins would suggest that one mode of their function is to become part of the RNA replicase complex.

CHAPTER 4

MATERIALS and METHODS

4.1. Plasmid or DNA template constructions

The primers used in this study are listed at the end of the Materials and Methods.

pYC DI72 mutants:

pYC DI72 is a gift from Dr. P. D. Nagy (University of Kentucky, USA). The plasmids pYC M2 DI72, pYC 52b DI72 and pYC TD1b DI72 were kindly provided by Dr. B. Wu from our laboratory. For the generation of pYC M2 DI72, a standard overlapping PCR was performed. First, PCR product 1 was amplified using pUC19 M2 DI72 [69] and the primer set, BA58 and BA59. PCR product 2 was generated using pYC DI72 and the primers, BA57 and BA60. PCR product 3 was made using PCR products 1 and 2 and primers BA58 and BA60, and it was digested with HindIII and SacI, and inserted to pYC DI72 digested with Hind III and SacI. For the generations of pYC 52b DI72, pYC TD1b, pYC M1M2 DI72, pYC 52c DI72 and pYC TD1c DI72, the same procedure as above was performed except that pUC19 52b DI72, pUC19 TD1b DI72, pUC19 M1M2 DI72, pUC19 52c DI72, pUC19TD1c [69, 70] were used as a template in place of pUC19 DI72.

pUC19 ST1-DI72, pUC19 ST2-DI72, pUC19 ST3-DI72 and pUC19 ST4-DI72:

For the generation of the streptotagged DI72 *in vitro* transcripts, (i.e., ST1, ST2, ST3 and ST4), pU19 ST1-DI72 to pUC19 ST4-DI72 plasmids were constructed and linearized with SmaI. For the generation of pUC19 ST1-DI72, PCR product 1 was made using pYC DI72 and the primers, BA58 and pHH15. PCR product 2 was generated using pYC DI72 and the primers, pHH16 and BA60. An overlapping PCR was then performed using the PCR products 1 and 2 with the primers BA58 and BA60, and the subsequent PCR product was digested with HindIII and SacI and inserted to pYC digested with HindIII and SacI, making pYC ST1-DI72. The inserted DNA is then amplified using

pHH38 and pHH39, and the resultant PCR product was then digested with *SacI* and *SmaI*, and inserted to pUC19 cut with *SacI* and *SmaI*, generating pUC19 ST1-DI72.

pUC19 ST2-DI72, pUC19 ST3-DI72, pUC19 ST4-DI72 were made similarly except that pHH15 and pHH16 were replaced by pHH17 and pHH18 for pUC19 ST2-DI72. For pUC19 ST3-DI72, pHH19 and pHH20 replaced pHH15 and pHH16, while pHH21 and pHH22 replaced pHH15 and pHH16 for pUC19 ST4-DI72.

To generate ST1-DI72, ST2-DI72, ST3-DI72 and ST4-DI72 in vitro transcript RNAs, pUC19 ST1-DI72, pUC19 ST2-DI72, pUC19 ST3-DI72 and pUC19 ST4-DI72 were linearized with *SmaI* and used as a template for in vitro transcription.

The DNA templates for in vitro transcription of ST-5' UTR, ST-M2-5' UTR and ST-TD1b-5'UTR

RNAs:

For the generation of the DNA templates for the in vitro transcription of the wild-type and mutant streptotagged 5' UTR RNAs, the following PCR products were directly used as transcription templates. For ST-5' UTR RNA, PCR reaction was performed using pUC19 ST3 DI72 and a primer set (pHH57 and pHH89). For ST-M2-5' UTR RNA, PCR product 1 was made using pUC19 ST3 DI72 and a primer set (pHH57 and pHH91), while PCR product 2 was made using pUC19 ST3 DI72 and a primer set (pHH92 and pHH89). Then an overlapping PCR was performed using the two PCR products and the primer set (pHH57 and pHH89), generating the template for ST-M2-5' UTR RNA. For the template for ST-TD1b-5' UTR RNA, PCR was carried out using pUC19 ST3 DI72 and a primer set (pHH57 and pHH93).

pYC DI72/p92rev:

To make pYC DI72/p92rev, ADH1 promoter, p92 ORF, Cyc terminator were digested out from pGAD His92 CEN with SpeI and XmnI and blunt-ended using T4 DNA polymerase (New England Biolab) following the manufacturer's protocol. Then, it was ligated to pYC DI72 digested with NaeI. p92 ORF in the opposite direction to DI72 RNA was chosen and named pYC DI72/p92rev. To generate pGAD His92 CEN, a standard overlapping PCR was performed. PCR product 1 was made using pYC DI72, the primers pCEN92-5 and pCEN92-6, while PCR product 2 was generated using pGAD His92, the primers pCEN92-7, pHH98. PCR product 3 was then made using these two PCR products with pCNE92-5 and pCEN92-8 primers. The digested PCR product (BmgBI and BamHI) and p92 ORF from pGAD His92 (BamHI and XhoI) were triple-ligated to pGAD His92 cut with BmgBI and XhoI.

pET GCD10His:

Yeast GCD10 ORF was amplified from pCUP GCD10 using a primer set (pHS51 and pHS52), digested with PciI and BamHI, and subcloned into pET15 digested with NcoI and BamHI.

pET GCD14His:

Yeast GCD14 ORF was amplified from pCUP GCD14 using a primer set (pHS55 and pHS56), digested with NcoI and BamHI, and subcloned into pET15 digested with NcoI and BamHI.

pET GCD10 GCD14His:

GCD10 ORF was amplified from pCUP GCD10 using a primer set (pHS51 and pHS50), digested with PciI and BamHI, and subcloned into pET15 digested with NcoI and BamHI, generating pET

GCD10. GCD14His ORF was cut out by digesting pET GCD14His with XbaI and BamHI and subcloned into pET GCD10 digested with NheI and BamHI.

pET MBP:

MBP ORF was amplified from pRK1043 [149] using a primer set (pHS67, and pHS68), digested with NcoI and BamHI, and subcloned into pET15 digested with NcoI and BamHI.

4.2. Yeast transformation and DI replication assay

Transformation of *S. cerevisiae* was carried out as previously described [150]. Specifically, wildtype BY4741 yeast strain (Open biosystem, USA) was simultaneously transformed with three plasmids, pHisGBK-His33 (-His), pGAD-His92 (-Leu) and pYC DI72 (-Ura) (gifts from Dr. PD Nagy) [82]. For transformation of mutant DI72, yeast was transformed pHisGBK-His33, pGAD-His92 and the pYC plasmids expressing mutant DI72, such as pYC-M2 DI72 and pYC-52b DI72. Appropriate auxotrophic nutrient markers (histidine, leucine and uracil) were omitted in yeast plates to select the properly transformed yeast colonies. After growing the colonies for 4 days at 30°C, at least three to four colonies were chosen to grow in 3 ml SC(-HUL) with 2% glucose overnight at 30°C. Next morning, the culture was washed with SC(-HUL) plus 2% galactose, and OD₆₀₀ of the yeast was adjusted to ~0.1 and cultured in 3 ml of SC(-HUL) with 2% galactose for 24 hours at 30°C to induce DI72 replication.

GCD10 knockdown strain (DAmp strain, Open biosystem, USA) was similarly transformed as described above. To maintain growth of the proper GCD10 knockdown strain, its growth media

also contained 200 µg/ml of geneticin (Sigma, Canada). The transformed GCD10 knockdown colonies grew significantly more slowly than the wildtype strain, so that they needed to be grown on plates for 6 days at 30°C before using them for DI72 replication assay, while the wild-type colonies were grown on plates for 4 days at 30°C to reach similar colony sizes.

GCD14 knockout strain (a gift from Dr. Hinnebusch) was transformed with pHisGBK-His33 (-His), pYC DI72/p92 rev (-Ura), while its corresponding wild-type strain (W303α, Open biosystem) was transformed with pHisGBK-His33(-His), pYC DI72/p92 rev (-Ura) and hclMT4 (-Leu) [126].

4.3. In vitro transcription

The transcription was performed using Ampliscribe T7 high yield transcription kit and its provided protocol (Epicentre, Canada).

4.4. Protoplast preparation and infection

Isolation and infection of protoplasts were previously described [69].

4.5. Yeast total RNA isolation and northern blotting

Yeast total RNAs were isolated as previously described [151], and 1 to 2 µg of yeast total RNAs were used for northern blotting analysis as previously carried out [69]. To detect TBSV viral RNAs,

p9 oligonucleotide was radio-labeled as previously described [69]. As a loading control, an oligonucleotide p25S was similarly radio-labeled to detect yeast 25S rRNA.

4.6. Preparation of streptomycin-conjugated column and binding test for streptotag RNAs

Conjugation of streptomycin to sepharose beads (GE healthcare, USA) was done as previously described [121]. To test whether streptotagged RNAs bind to streptomycin-conjugated sepharose, 30 µg of streptotagged RNAs in column buffer (50 mM Tris HCl, pH7.5, 300 mM NaCl, 5 mM MgCl₂) were loaded to 0.5 ml of the streptomycin conjugated sepharose, and incubated at room temperature for 10 min. The resin was washed with 5 column volumes of column buffer, and the bound streptotagged RNAs were eluted with 2 ml of column buffer containing 1 mM streptomycin. One tenth of the eluted RNA was precipitated with 0.3 M NaOAc [pH5.6], 20 µg of yeast tRNAs and 70 % ethanol. The precipitated RNA was then dissolved in water and one half of the RNA was loaded on agarose gel. To assess the binding efficiency of the streptotagged RNAs to the column, 1/20 of the streptotagged RNAs before loading to the column was also loaded to the same agarose gel as an initial input.

4.7. Preparation of yeast total soluble extract and streptotag affinity purification

A freshly grown colony of BY4741 strain on YPD plate was cultured in 30 ml YPD medium overnight at 30 °C, and it was transferred to 600 ml YPD medium next morning, and further grown until its OD₆₀₀ reached between 0.7-1.0. The culture was then split into two containers and harvested at 6000 g for 20 min, and the pellet was washed with cold water twice and then with column buffer. Each pellet was then resuspended in 10 ml column buffer plus 2 mM DTT and 1x

protease inhibitors (Roche, USA). The yeast cells were then lysed using beadbeater (Biospec Products, USA) following the manufacturer's protocol. The soluble total protein extracts were obtained by spinning the lysed yeast cells at 2200 g for 5 min at 4 °C. The supernatant was further spun at 1,5000 g for 10 min and the subsequent supernatant was spun again at 1,5000 g for 5 min. The final supernatant was frozen at -80°C until use.

Streptotag affinity purification was performed as previously described [56] with some modifications. Specifically, ~200 µg of streptotag RNA in 2 ml column buffer were first heated to 100 °C and cooled slowly to renature the RNA. The RNA was then loaded to the column and incubated for 30 min at room temperature. tRNAs was not additionally loaded to the streptomycin-conjugated column to block nonspecific binding of yeast total proteins to the column. For loading protein extract, 2 ml of yeast protein extract prepared above (~10 mg/ml) was loaded onto the column. For the washing step, washing was done with 20 column volumes of column buffer and the bound proteins were eluted in 1 ml fractions. The fractions 1 to 7 were pooled and concentrated using Amicon (Millipore, Canada) following its provided protocol until the sample volume becomes ~200 µl.

4.8. Silver staining and identification of proteins

40 µl of the eluted protein sample from streptotag affinity column was mixed with 5x SDS sample buffer, loaded to 10% SDS-PAGE (1 mm thick gel with 10 well combs), and SDS-PAGE gel was silver-stained as described previously [152]. For RNase A treatment of the eluted samples, 1ul of protease free RNase A (10 mg/ml) (Fermentas, Canada) was added to 40 µl of eluted samples and

incubated at 37°C for 5 min. For mass spectrometry analysis, the eluted proteins from three independent trials were pooled and acetone-precipitated (80% v/v) overnight at -20°C, and spun at maximum speed at a benchtop centrifuge at 4°C for 10 min. After washing with 90% cold acetone, the dried pellet was dissolved in 1x SDS sample buffer and boiled at 100°C for 3 min before loading onto SDS-PAGE. After staining the SDS-PAGE gel with Coomassie blue overnight, the gel was destained according to the previously described method [153]. The bands of interest were excised and submitted to the mass spectrometry facility at the Hospital for Sick Children at Toronto for identification of the proteins in the bands.

4.9. Protein extraction for western blotting analyses

Extraction of yeast total proteins were performed as previously described [154], and similar amounts from each sample were loaded onto SDS-PAGE as previously described [115]. The proteins in the SDS-PAGE gels was transferred to nitrocellulose membrane (GE healthcare) at 30V at 4°C overnight, and the membrane was stained with Ponceau S (Sigma Aldrich, Canada) using the manufacture's protocol. After removing the stain, the membrane was then blocked with 5% skim milk in TBS-T (Tris buffered saline, 0.1% Tween 20) for at least 1 hr at RT. For detection of p33/p92, rabbit anti-p33 antibody (1:1000 dilution) (a gift from Dr. R Mullen, University of Guelph, Canada) [17] was incubated with the membrane for 1 hr at RT. For detection of GCD10 and GCD14 proteins, rabbit anti-GCD10 and anti-GCD14 antibodies (gifts from Dr. Hinnebusch) [126] were diluted in 1:500 and used for western blotting. For the detection of his tagged proteins, mouse anti-his antibody (GE Healthcare, USA) was used in 1:1000 dilution. As a secondary antibody,

either anti-rabbit antibody conjugated with horse radish peroxidase (Sigma Aldrich, Canada) or anti-rabbit antibody conjugated with Cy3 (Jackson ImmunoResearch, USA) was used in 1:5000 or 1: 400 dilution, respectively. For the mouse primary antibody, anti-mouse antibody conjugated with horse radish peroxidase (Jackson ImmunoResearch, USA) was used in 1: 5000 dilution.

4.10. His affinity purification

After transforming BL21 strain with pET GCD10His, pET GCD14His, pET GCD10/GCD14His or pET MBP, a freshly grown colony was grown in 10 ml LB medium at 30°C overnight. The entire culture was then transferred to 500 ml LB and cultured at 30°C until its OD₆₀₀ reached ~0.5. 0.5 mM IPTG was added to the culture to induce the recombinant protein expression, and the culture was grown for additional 3 hrs before the cells were harvested. The pelleted cells were frozen at -80°C until use.

The cells were resuspended in 10 ml lysis buffer (20 mM Tris, pH8.0, 300 mM NaCl, 1 mM DTT, 1x protease inhibitors, 10 mM imidazole) and sonicated on ice at 40% amplitude (FischerScientific 150E digital sonic dismembrator) for 10 seconds with 50 second pause. The procedure was repeated for 5 times. The lysed samples were then spun at 12,000 g for 30 min at 4°C, and the supernatant was mixed with 1 ml nickel resin (50% slurry, equilibrated with his affinity column buffer (the same as the lysis buffer but without protease inhibitors)) for 1 hour at 4°C. After washing the resin with 20 column volumes of column buffer, the resin was further washed with 0.5 ml of increasing imidazole concentrations (20 mM, 30 mM and 40 mM imidazole). Significant amounts of GCD10/GCD14His was lost during this step. The bound his tagged proteins were

eluted with 7 column volumes of column buffer containing 500 mM imidazole. The eluate fractions containing most of the recombinant proteins were pooled and concentrated using amicon column (Millipore, USA) and the company's protocol. Then the column buffer without imidazole was added and concentrated again. This buffer exchange step was repeated at least four times. To avoid protein aggregation, the centrifugation was stopped frequently, and the samples were mixed by pipetting up and down.

The concentrations of purified proteins were generally within the range of 0.5 mg/ml to 1 mg/ml. Then glycerol was added to the samples (10% v/v) before storing at -80 °C.

4.11. Electrophoretic mobility shift assay (EMSA)

The 5' UTR or M2 5' UTR RNAs was labeled with P³² alpha UTP as described previously [56], and purified using 5% polyacrylamide gel as described before [155]. The RNAs were then mixed with purified proteins (GCD10His, GCD14His, GCD10/GCD14His proteins or MBP) [156] and incubated for 10 min at room temperature before loading onto 5% polyacrylamide gel. The gel was run at 100 V at 4°C and dried for 1 hour at 80°C. Autoradiography was developed using Typhoon Trio (GE Healthcare, USA), the bands of interests were quantified using ImageQuant TL software (GE healthcare, USA). The binding curves were generated using Prism software.

4.12. RT-PCR for DI72 minus-strand detection

1 - 1.5 µg of total yeast nucleic acids in 1x transcription buffer (Epicentre, Canada) were first incubated with DNase I for 10 min at 37°C (Epicentre, Canada) to remove DNA plasmids. To remove DNase I, the RNA samples were cleaned using a gel extraction kit (BioBasics, Canada). Then, the samples were heated to 100°C for 2 min to separate the minus-strands from the plus-strands and cooled immediately on ice. Reverse transcription was performed using a primer pHH200 (or pHH5 for the 5' UTR mutants) and superscript IV. The generated cDNA was amplified using Phusion polymerase (New England BioLab, USA) and a primer set (pHH79-2 and pHH200 (or pHH5 and pHH200 for 5' UTR mutants)). To detect yeast actin mRNA, RT-PCR was similarly performed, but with a primer set (pHH158, pHH159).

4.13. The list of the primers used in this study

BA58: 5'- GTAATAAAAGTAAGCTTGGAAATTCTCCAGGATTTCTCG

BA59: 5'- GGGCTGCATTTCTGCAATGTTCCGGTTGTC

BA57: 5'- GCAGAAATGCAGCCCAGTCCTGTTTCTTGCCAAACAG

BA60: 5'- CTAGTGGATCCGAGCTCTACCAGGTAATATAC

pHH15: 5'- GCCCTTGCGGGCAGAAGTCCAAATGCGATCCAGGTCGAGAAATCCTGGAG

pHH16: 5'- CTTCTGCCC GCAAGGGCACCACGGTCGGATCCAGTTCGTTTATCTGGTGAC

pHH17: 5'- GCCCTTGCGGGCAGAAGTCCAAATGCGATCCTAATAATTATGGAGAGAA

pHH18: 5'- CTTCTGCCC GCAAGGGCACCACGGTCGGATCCTCTTAGTTGTGGGGTTTG

pHH19: 5'- GCCCTTGCGGGCAGAAGTCCAAATGCGATCCATAATTATGGAGAGAAATTC

pHH20: 5'- CTTCTGCCC GCAAGGGCACCACGGTCGGATCCTTAGTTGTGGGGTTTGAAG

pHH21: 5'- GCCCTTGCGGGCAGAAGTCCAAATGCGATCCTCT AGACATGTCGCTTG

pHH22: 5'- CTTCTGCCC GCAAGGGCACCACGGTCGGATCCAGAAACGGGAAGCTCGCTC

pHH38: 5'- AATTCGAGCTCTAATACGACTCACTATAGGAAATTCTCCAGGATTTCTC

pHH39: 5'- AATTCGAGCTCTAATACGACTCACTATAGGAAATTCTCCAGGATTTCTC

pHH57: 5'- GAATTCGAGCTCTAATACGATCACTATA

pHH89: 5'- CATGTCGCTTGTGTTGTTGGAAGTTACAATTTATCC

pHH91: 5'- GGAGAGAATTTGTGTACGCAAAGCAACGG

pHH92: 5'- CCGTTGCTTTGCGTACACAAATTCTCTCC

pHH93: 5'- CATGTGGAATGTTTGTGGAAGTTACAATTTATCC

pHH98: 5'- CAACCCTACACACTAATTGGAAGATGTCC

pCEN92-5: 5'- GTACGACGTGGGTCCTTTTCATCACGTGCTATAA

pCEN92-6: 5'- ACTAGTGGATCATCCCCACGC

pCEN92-7: 5'- GCGTGGGGATGATCCACTAGTTGCATGCCTGCAGGTCGAGATCC

pHS50: 5'- CATGGAGCTAAGAA ACAAAAAATATAATGAGCTAGCATCAAGTGGATCCAGTATC

pHS51: 5'- GATCAGAACATGTCAGGTG GTGGTATGAATGCTTTGACAACCATAG

pHS52: 5'- GATACTGGTCTCATTAGTGATGATGATGATGATGATGACCACCACCTATTTTTTGTTCCTT
CTGGATCCTCATTAGTGATGATGATGATGATGATGATGACCACCACCTATTTTTTGTTCCTTAGCT
CCATG

pHS55: 5'- GATCAGACCATGGGTATGTCAACAAATTGTTTTTCCGGTTAC

pHS56: 5'- GATACTGGATCCAACAGATGCTAGCTCATTAGTGATGATGATGATGATG ACCACC
ACTTTTTCCGTGGATCGAAGAATTC

pHS67: 5'- ATCAGACCATGGGTCATCATCATCATCACGGTGGTGGTATGAAAACCTGAAGAA
GGTAAACTG

pHS68: 5'- CTAGTTGGATCCTCATTACGAGCTCGAATTAGTCTGCGCGTC

p25S: 5'- CGGTTATCAGTACGACCTGGCATGAAAACCTATTCCTTCC

pHH79-2: 5'- GGGCTGCATTTCTGCAATGTTCCGGTTGTC

pHH158: 5'- AGAAGCCCGTGCATTA AAAAG

pHH159: 5'- CTTA CACTTCGTTATAACATTAAC

pHH200: 5'- CATTGAGGCAAATAAAAATATGAAAAAAG

p9: 5'- GGGCTGCATTTCTGCAATGTTCCGGTTGTC

References

1. Pontis RE, Garcia O, Feldman JM. 1968. Tomato bushy stunt virus on tomato crops in Argentina. *Plant Dis. Rep.* 52: 676-677.
2. Fischer HU, Lockhart BLE. 1977. Identification and comparison of two Isolates of tomato bushy stunt virus from pepper and tomato in Morocco. *Phytopath.* 11: 1352-1355.
3. Gerik JS, Duffus JE, Perry R, Stenger DC, Van Maren AF. 1990. Etiology of tomato plants in the California desert. *Phytopath.* 12: 1352-1356.
4. Roossinck MJ. 2003. Plant RNA virus evolution. *Virology* 6(4): 406-409.
5. Nagy PD. 2016. Tombusvirus-Host interactions: co-opted evolutionarily conserved host factors take center court. *Annu. Rev. Virol.* 3(1): 491-515.
6. Baltimore D. 1971. Expression of animal virus genomes. *Bacteriol. Rev.* 35(3): 235-241.
7. Nicholson BL, White KA. 2014. Functional long-range RNA-RNA interactions in positive-strand RNA viruses. *Nat. Rev. Microbiol.* 12: 493-504.
8. Ahlquist P. 2002. RNA-dependent RNA polymerases, viruses and RNA silencing. *Science* 296 (5571): 1270-1273.
9. Kegler G, Kegler H. 1980. Untersuchungen zur natürlichen Übertragung des tomato bushy stunt virus bei Obstgehölzen. *Tagungsberichte der Akademie der Wissenschaften der DDR.* 184: 297-302.
10. Schmelzer K, 1958. Wirtspflanzen des tomatenzwergbusch-virus (marmor odecahedron holmes). *Zeitschrift für PflanzenKrankheit und PflanzenSchutz* 65: 80-89.
11. White KA, Nagy PD. 2004. Advances in molecular biology of tombusviruses: gene expression, genome replication, recombination. *Prog. Nucleic Acid Res. Mol. Biol.* 78: 187-226.
12. Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G. 1978. Tomato bushy stunt virus at 2.9 Å resolution. *Nature* 276(5686): 368-373.
13. King AMQ, Lefkowitz E, Adams MJ, Carstens EB. 2012. *Virus Taxonomy Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses.* Elsevier Academic Press.
14. Hearne PQ, Knorr DA, Hillman BI, Morris TJ. 1990. The complete genome structure and synthesis of infectious RNA from clones of tomato bushy stunt virus. *Virology* 177(1): 141-151.
15. Oster SK, Wu B, White KA. 1998. Uncoupled expression of p33 and p92 permits amplification of tomato bushy stunt virus RNAs. *J. Virol.* 72(7): 5845-5851.

16. Panaviene Z, Baker JM, Nagy PD. 2003. The overlapping RNA-binding domains of p33 and p92 replicase proteins are essential for tombusvirus replication. *Virology* 308: 191–205.
17. McCartney AW, Greenwood JS, Fabian MR, White KA, Mullen RT. 2005. Localization of tomato bushy stunt virus p33 replication protein reveals a peroxisome-endoplasmic reticulum sorting pathway. *Plant Cell*. 17: 3513-3531.
18. Monkewich S, Lin H, Fabian MR, Xu W, Na H, Ray D, Chernysheva OA, Nagy PD, White KA. 2005. The p92 polymerase coding region contains an internal RNA element required at an early step in tombusvirus genome replication. *J. Virol.* 79(8): 4848-4858.
19. Pogany J, White KA, Nagy PD. 2005. Specific binding of tombusvirus replication protein p33 to an internal replication element in the viral RNA is essential for replication. *J. Virol.* 79(8): 4859-4869.
20. Pathak KB, Sasvari Z, Nagy PD. 2008. The host Pex19p plays a role in peroxisomal localization of tombusvirus replication proteins. *Virology* 379(2): 294-305.
21. Kovalev N, Martin IFC, Pogany J, Barajas D, Pathak K, Risco C, Nagy PD. 2016. Role of viral RNA and co-opted cellular ESCRT-1 and ESCRT-III factors in formation of tombusvirus spherules harboring tombusvirus replicase. *J. Virol.* 90(7): 3611-3626.
22. Rajendran KS, Nagy PD. 2004. Interaction between replicase proteins of tomato bushy stunt virus in vitro and in vivo. *Virology*. 326: 250-261.
23. Kopek BG, Perkins G, Miller DJ, Ellisman MH, Ahlquist P. 2007. Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biol.* 5(9): 2202- 2034.
24. Gunawardene CD, Donalson LW, White KA. 2017. Tombusvirus polymerase: Structure and function. *Virus Res.* 234: 74-86.
25. Cimino PA, Nicholson BL, Wu B, Xu W, White KA. 2012. Multifaceted regulation of translational readthrough by RNA replication elements in a tombusvirus. *PLoS Pathog.* 7(12): e1002423.
26. Ng KK, Arnold JJ, Cameron CE. 2008. Structure-function relationships among RNA-dependent RNA polymerases. *Curr. Top. Microbiol. Immunol.* 320: 137-156.
27. Jácome R, Becerra A, de León SP, Lazcano A. 2015. Structural analysis of monomeric RNA-dependent polymerases: evolutionary and therapeutic implications. *PLoS ONE* 10(9): e0139001.
28. Steitz TA. 1998. A mechanism for all polymerases. *Nature.* 391:231-232.

29. O'Reilly EK, Kao CC. 1998. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology* 252: 287-303.
30. Rajendran KS, Nagy PD. 2003. Characterization of the RNA-binding domains in the replicase proteins of Tomato bushy stunt virus. *J. Virol.* 77: 9244-9258.
31. Gunawardene D, Jaluba K, White KA. 2015. Conserved motifs in a tombusvirus polymerase modulate genome replication, subgenomic transcription, and amplification of defective interfering RNAs. *89(6): 3236-3246.*
32. Wu B, White KA. 2007. Uncoupling RNA virus replication from transcription via the polymerase: functional and evolutionary insights. *EMBO J.* 26: 5120-5130.
33. Hopper P, Harrison SC, Sauer RT. 1984. Structure of tomato bushy stunt virus. V. Coat protein sequence determination and its structural implication. *J. Mol. Biol.* 177(4): 701-713.
34. Qu F, Ren T, Morrison TJ. 2003. The coat protein of turnip crinkle virus suppresses post-transcriptional gene silencing at an early initiation step. *J Virol.* 77(1): 511-522.
35. Voinnet O, Pinto YM, Baulcombe DC. 1999. Suppression of gene silencing: A general strategy used by diverse DNA and RNA viruses in plants. *Proc. Natl. Acad. Sci. USA.* 96: 14147-14152.
36. Baulcombe D. 2002. RNA silencing. *Curr. Biol.* 12(3): R82-R84.
37. Targett-Adams P, Boulant S, MacLauchlan J. 2008. Visualization of double-stranded RNA in cells supporting hepatitis virus C RNA replication. *J. Virol.* 82(5): 2182-2195.
38. Weber F, Wagner V, Rasmussen SB, Hartmann R, Paludan SR. 2006. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses, but not in detectable amounts by negative-strand RNA viruses. *J. Virol.* 80(10): 5059-5064.
39. Hagemeyer MC, Vonk AM, Monastyrska I, Rottier PJM, de Haan CAM. 2012. Visualizing coronavirus RNA synthesis in time by click chemistry. *J. Virol.* 86(10): 5808-5816.
40. Gamarnik AQ, Andino R. 1998. Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev.* 12: 2293-2304.
41. Chu PW, Westaway EG. 1985. Replication strategy of Kunjin virus: evidence for recycling role of replicative form RNA as template in semiconservative and asymmetric replication. *Virology* 140: 68-79.
42. Cleaves GR, Ryan TE, Schlesinger RW. 1981. Identification and characterization of type 2 dengue virus replicative intermediate and replicative form RNAs. *Virology* 111: 73-83.

43. Gong Y, Shannon A, Westaway EG, Gowans EJ. 1998. The replicative intermediate molecule of bovine viral diarrhoea virus contains multiple nascent strands. *Arch. Virol.* 143: 399–404.
44. Herr AJ, Baulcoubm DC. 2004. RNA silencing pathways in plants. *Cold Spring Harb. Symp. Quant. Biol.* 69: 363-370.
45. Qu F, Morris TJ. 2002. Efficient infection of *Nicotiana benthamiana* by tomato bushy stunt virus is facilitated by the coat protein and maintained by p19 through suppression of gene silencing. *Mol. Plant-Microbe Interact.* 15: 193–202.
46. Chu N, Desvoyes B, Turina M, Noad R, Scholthof HB. 2000. Genetic dissection of tomato bushy stunt virus p19 protein mediated host-dependent symptom induction and systemic invasion. *Virology* 266(1): 79-87.
47. Qiu W, Park JW, Scholthof HB. 2002. Tombusvirus p19-mediated suppression of virus-induced gene silencing is controlled by genetic and dosage features that influence pathogenicity. *Mol. Plant-Microbe Interact.* 15: 269–280.
48. Silhavy D, Molnar A, Lucioli A, Szittyta G, Hornyik C, Tavazza M, Burgyan J. 2002. A viral protein suppresses RNA silencing and binds silencing-generated 21- to 25-nucleotide double-stranded RNAs. *EMBO J.* 21: 3070–3080.
49. Scholthof HB, Scholthof KB, Kikkert M, Jackson AO. 1995. Tomato bushy stunt virus spread is regulated by two nested genes that function in cell-to-cell movement and host dependent systemic invasion. *Virology* 213: 425–438.
50. Dalmay T, Rubino L, Burgyan J, Kollar A, Russo M. 1993. Functional analysis of cymbidium ringspot virus genome. *Virology* 194: 697–704.
51. Oparka KJ, Prior DA, Santa Cruz S, Padgett HS, Beachy RN. 1997. Gating of epidermal plasmodesmata is restricted to the leading edge of expanding infection sites of tobacco mosaic virus (TMV). *Plant J.* 12(4): 781-789.
52. Gallie DR. 1998. A tale of two termini: A functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene* 216: 1-11.
53. Gallie DR. 2001. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev.* 5:2108-2116.
54. Fabian MR, White KA. 5'-3' RNA-RNA interaction facilitates cap- and poly(A) tail-independent translation of tomato bushy stunt virus mRNA. *J. Biol. Chem.* 279(28): 28862-28872.
55. Nicholson BL, White KA. 2008. Context-influenced cap independent translation of Tombusvirus mRNAs in vitro. *Virology* 380: 203-212.
56. Nicholson BL, Wu B, Chevtchenko I, White KA. 2010. Tombusvirus recruitment of host machinery via the 3' UTR. *RNA* 16: 1402-1419.

57. Kozak M. 2002. Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299(1-2): 1-34.
58. White KA, Morris TJ. 1999. Defective interfering RNAs of monopartite plus-strand RNA plant viruses. *Curr. Top. Microbiol. Immunol.* 239: 1–17.
59. White KA. 1996. Formation and evolution of tombusvirus defective interfering RNAs. *Semin. Virol.* 7: 409–416.
60. Hillman BI, Carrington JC, Morris TJ. 1987. A defective interfering RNA that contains a mosaic of a plant virus genome. *Cell* 51: 427–433.
61. Nagy PD, Pogany J. 2000. Partial purification and characterization of cucumber necrosis virus and tomato bushy stunt virus RNA-dependent RNA polymerases: Similarities and differences in template usage between tombusvirus and carmovirus RNA-dependent RNA polymerases. *Virology* 276: 279–288.
62. Fabian MR, Na H, Ray D, White KA. 2003. 3'-terminal RNA secondary structures are important for accumulation of tomato bushy stunt virus DI RNAs. *Virology* 312: 567-580.
63. Na H, White KA. 2006. Structure and prevalence of replication silencer- 3' terminus RNA interactions in Tombusviridae. *Virology* 345: 305-316.
64. Na H, Fabian MR, White KA. 2006. Conformational organization of the 3' untranslated region in the tomato bushy stunt virus genome. *RNA* 12(12): 2199-2210.
65. Wu B, Pogany J, Na H, Nicholson BL, Nagy PD, White KA. 2009. A discontinuous RNA platform mediates RNA virus replication: Building an integrated model for RNA-based regulation of viral processes. *PLoS Pathog.* 5(3): e1000323.
66. Ray D, White KA. 2003. An internally located RNA hairpin enhances replication of tomato bushy stunt virus RNAs. *J. Virol.* 77(1): 245–257.
67. Panavas T, Nagy PD. 2003. The RNA replication enhancer element of tombusviruses contains two interchangeable hairpins that are functional during plus-strand synthesis. *J. Virol.* 77(1): 258–269.
68. Panavas T, Pogany J, Nagy PD. 2002. Analysis of minimal promoter sequences for plus-strand synthesis by the cucumber necrosis virus RNA-dependent RNA polymerase. *Virology* 296: 263–274.
69. Wu B, Vanti WB, White KA. 2003. An RNA domain within the 5' untranslated region of the tomato bushy stunt virus genome modulates viral RNA replication. *J. Mol. Biol.* 305: 741-756.
70. Ray D, Wu B, White KA. 2003. The second functional RNA domain in the 5' UTR of tomato bushy stunt virus genome: Intro- and interdomain interactions mediate viral RNA replication. *RNA* 9: 1232-1245.

71. Ahlquist P, Noueiry AO, Lee W, Kushner DB, Dye BT. 2003. Host factors in positive-strand RNA genome replication. *J. Virol.* 77(15): 8181-8186.
72. Janda M, Ahlquist P. 1993. RNA-dependent replication, transcription, and persistence of brome mosaic virus RNA replicons in *S. cerevisiae*. *Cell* 72: 961-970.
73. Price BD, Roeder M, Ahlquist P. 2000. DNA-directed expression of functional flock house virus RNA1 derivatives in *Saccharomyces cerevisiae*, heterologous gene expression, and selective effects on subgenomic mRNA synthesis. *J. Virol.* 74: 11724-11733.
74. Ishikawa M, Diez J, Restrepo-Hartwig M, Ahlquist P. 1997. Yeast mutations in multiple complementation groups inhibit brome mosaic virus RNA replication and transcription and perturb regulated expression of the viral polymerase-like gene. *Proc. Natl. Acad. Sci. USA* 94(25): 13810-13815.
75. Diez J, Ishikawa M, Kaido M, Ahlquist P. Identification and characterization of a host protein required for efficient template selection in viral RNA replication. *Proc. Natl. Acad. Sci. USA* 97(8): 3913-3918.
76. Kushner DB, Lindenbach BD, Grdzelishvili VZ, Noueiry AO, Paul SM, Ahlquist P. 2003. Systemic genome-wide identification of host gene affecting replication of a positive-strand RNA replication. *Proc. Natl. Acad. Sci. USA* 100(26): 15764-15769.
77. Gancarz BL, Hao L, He Q, Newton MA, Ahlquist P. 2011. Systematic identification of novel, essential host genes affecting bromovirus RNA replication. *PLoS One* 6(8): e23988.
78. Hao L, Lindenbach B, Wang X, Dye B, Kushner D, He Q, Newton M, Ahlquist P. 2014. Genome-wide analysis of host factors in nodavirus RNA replication. *PLoS One* 9(4): e95799.
79. Panavas T, Nagy PD. 2003. Yeast as a model host to study replication and recombination of defective interfering RNA of tomato bushy stunt virus. *Virology* 314: 315-325
80. Wang RY, Nagy PD. 2008. Tomato bushy stunt virus co-opts the RNA-binding function of a host metabolic enzyme for viral genomic RNA synthesis. *Cell Host Microbe.* 3(3): 178-187.
81. Barajas D, Li Z, Nagy PD. 2009. The Nedd4-type Rsp5p ubiquitin ligase inhibits tombusvirus replication by regulating degradation of the p92 replication protein and decreasing the activity of the tombusvirus replicase. *J. Virol.* 83(22): 11751-11764.
82. Panavas T, Serviene E, Brasher J, Nagy PD. 2005. Yeast genome-wide screen reveals dissimilar sets of host genes affecting replication of RNA viruses. *Proc. Natl. Acad. Sci. USA* 102(20): 7326-7331.
83. Jiang Y, Serviene E, Gal J, Panavas T, Nagy PD. 2006. Identification of essential host factors affecting tombusvirus RNA replication based on the yeast Tet-promoters Hughes collection. *J. Virol.* 80(1): 7394-7404.

84. Breslow DK, Cameron DM, Collins SR, Schuldiner M, Stewart-Ornstein J, Newman HW, Braun S, Madhani HD, Krogan NJ, Weissman JS. 2008. A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat. Methods* 5(8): 711-718.
85. Nawaz-ul-Rehman MS, Prasanth KR, Baker J, Nagy PD. 2013. Yeast screens for host factors in positive-strand RNA virus replication based on a library of temperature-sensitive mutants. *Methods* 59:207-216.
86. Nawaz-ul-Rehamna MS, Martinez-Ochoa N, Pascal H, Sasvari Z, Herbst C, Nagy PD. 2012. Proteome-wide overexpression of host proteins for identification of factors affecting tombusvirus RNA replication: an inhibitory role of protein kinase C. *J. Virol.* 86: 9384-9395.
87. Li Z, Pogany J, Panava T, Xu K, Esposito AM, Kinzy TG, Nagy PD. 2009. Translation elongation factor 1A is a component of the tombusvirus replicase complex and affects the stability of the p33 replication co-factor. *Virology* 385: 245-260.
88. Li Z, Pogany J, Tupman S, Esposito AM, Kinzy TG, Nagy PD. 2010. Translation elongation factor 1A facilitates the assembly of the tombusvirus replicase and stimulates minus-strand synthesis. *PLoS Pathog.* 6:e1001175.
89. Wang RY, Stork J, Pogany J, Nagy PD. 2009. A temperature sensitive mutant of heat shock protein 70 reveals an essential role during the early steps of tombusvirus replication. *Virology* 394: 28-38.
90. Prasanth KR, Barajas D, Nagy PD. 2015. The proteasomal Rpn11 metalloprotease suppresses tombusvirus RNA recombination and promotes viral replication via facilitating assembly of the viral replicase complex. *J. Virol.* 89: 2750-2763.
91. Verma R, Aravind L, Oania R, McDonald WH, Yates JR 3rd, Koonin EV, Deshaies RJ. 2002. Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* 298(5593): 611-615.
92. Saunier R, Esposito M, Dassa EP, Delahodde A. 2013. Integrity of the *Saccharomyces cerevisiae* Rpn11 protein is critical for formation of proteasome storage granules (PSG) and survival in stationary phase *PLoS One* 8:e70357.
93. Sasvari Z, Izotova L, Kinzy TG, Nagy PD. 2011. Synergistic roles of eukaryotic translation elongation factors 1By and 1A in stimulation of tombusvirus minus-strand synthesis. *PLoS Pathog.* 7: e1002438.
94. Kovalev N, Barajas D, Nagy PD. 2012. Similar roles for yeast Dbp2 and *Arabidopsis* RH20 DEAD-box RNA helicases to Ded1 helicase in tombusvirus plus-strand synthesis. *Virology* 432: 470-484.
95. Kovalev N, Pogany J, Nagy PD. 2012. A co-opted DEAD-box RNA helicase enhances tombusvirus plus-strand synthesis. *PLoS Pathog.* 8: e1002537.
96. Nawaz-ul-Rehman MS, Prasanth KR, Xu K, Sasvari Z, Kovalev N, Martin IFC, Barajas D, Risco C, Nagy PD. 2016. Viral replication protein inhibits cellular cofilin actin depolymerization

- factor to regulate the actin network and promote viral replicase assembly. *PLoS Pathog.* 12(2): e1005440.
97. Theriot JA. 1997. Accelerating on a treadmill: ADF/cofilin promotes rapid actin filament turnover in the dynamic cytoskeleton. *J. Cell Biol.* 136(6):1165-1168.
 98. Lorizate M, Krausslich HG. 2011. Role of lipids in virus replication. *Cold Spring Harb. Perspect. Biol.* 3: a004820.
 99. Sharma M, Sasvari Z, Nagy PD. 2010. Inhibition of sterol biosynthesis reduces tombusvirus replication in yeast and plants. *J. Virol.* 84: 2270-2281.
 100. Barajas D, Xu K, Martin IFC, Sasvari Z, Brandizzi F, Risco C, Nagy PD. 2014. Co-opted oxysterol-binding ORP and VAP proteins channel sterols to RNA virus replication sites via membrane contact sites. *PLoS Pathog.* 10(10): e1004388.
 101. Xu K, Nagy PD. RNA virus replication depends on enrichment of phosphatidylethanolamine at replication sites in subcellular membranes. *Proc. Natl. Acad. Sci. USA* 112: E1782-1791.
 102. Schmidt O, Teis D. 2012. The ESCRT machinery. *Curr. Biol.* 22(4): R116-R120.
 103. Richardson LG, Clendening EA, Sheen H, Gidda SK, White KA, Mullen RT. 2014. A unique N-terminal sequence in the Carnation Italian ringspot virus p36 replicase-associated protein interacts with the host cell ESCRT-1 component Vps23. *J. Virol.* 88: 6329-2344..
 104. Barajas D, Jiang Y, Nagy PD. 2009. A unique role for the host ESCRT proteins in replication of Tomato bushy stunt virus. *PLoS Pathog.* 5(12): e1000705.
 105. Pogany J, Fabian MR, white KA, Nagy PD. 2003. A replication silencer element in a plus-strand RNA virus. *EMBO J.* 22: 5602-5611.
 106. Stork J, Kovalev N, Sasvari Z, Nagy PD. 2011. RNA chaperone activity of the tombusviral p33 replication protein facilitates initiation of RNA synthesis by viral RdRp in vitro. *Virology* 409 (2): 338-347.
 107. Kovalev N, Pogany J, Nagy PD. 2011. Template role of double-stranded RNA in tombusvirus replication. *J. Virol.* 88(10): 5638-5651.
 108. Kao CC, Singh P, Ecker DJ. 2001. De novo initiation of viral RNA-dependent RNA synthesis. *Virology* 287: 251–260.
 109. Panavas T, Stork J, Nagy PD. 2006. Use of double-stranded RNA templates by the tombusvirus replication in vitro: implication for the mechanism of plus-strand initiation. *Virology* 352: 110-120.
 110. Kovalev N, Pogany J, Nagy PD. 2012. A co-opted DEAD-box RNA helicase enhances tombusvirus plus-strand synthesis. *PLoS Pathog.* 8: e1002537.
 111. Wang RY, Nagy PD. 2008 Tomato bushy stunt virus co-opt the RNA-binding function of a host metabolic enzyme for viral genomic RNA synthesis. *Cell Host Microbe* 3:178-187.

112. Huang TS, Nagy PD. 2011. Direct inhibition of tombusvirus plus-strand RNA synthesis by a dominant negative mutant of a host metabolic enzyme, glyceraldehyde-3-phosphate dehydrogenase, in yeast and plants. *J. Virol.* 85:9090-9102.
113. Kovalev N, Nagy PD. 2014. The expanding functions of cellular helicases: the tombusvirus RNA replication enhancer co-opts the plant eIF4AIII-like AtRH2 and the DDX5-like AtRH5 DEAD-box RNA helicases to promote viral asymmetric RNA replication. *PLoS Pathog.* 10:e1004051.
114. Panavas T, Nagy PD. 2005. Mechanism of stimulation of plus-strand synthesis by an RNA replication enhancer in a tombusvirus. *J. Virol.* 79: 9777-9785.
115. Sheen H. 2016. Quantitation of yeast total proteins in sodium dodecyl sulfate-polyacrylamide gel electrophoresis sample buffer for uniform loading. *Anal Biochem.* 498: 95-97.
116. Sheen H, White KA. 2018. Expression of T7-based constructs in tobacco cells. *Biochem. Biophys. Res. Commun.* 499(2): 196-201.
117. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Phillippsen P, Tettelin H, Oliver SG. 1996. Life with 6000 genes. *Science* 274(5287): 563-567.
118. Debashish R, Na H, White KA. 2004. Structural properties of a multifunctional T-shaped RNA domain that mediate efficient Tomato bushy stunt virus RNA replication. *J. Virol.* 78(19): 10490-10500
119. Wallace ST, Schroeder R. 1998. In vitro selection and characterization of streptomycin-binding RNAs: Recognition discrimination between antibiotics. *RNA* 4(1): 112–123.
120. Bachler M, Schroeder R, Ahsen UV. 1999. StreptoTag: a novel method for isolation of RNA binding proteins. *RNA* 5(11): 1509-1516.
121. Windbichler N, Schroeder R. 2006. Isolation of specific RNA-binding proteins using the streptomycin-binding RNA aptamer. *Nat. Protoc.* 1(2): 638-641.
122. Dangerfield JA, Windbichler N, Salmons B, Günzburg WH, Schröder R. 2006. Enhancement of StreptoTag method for isolation of endogenously expressed proteins with complex RNA binding proteins. *Electrophoresis* 27: 1874-1877.
123. Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31: 3406-3415.
124. Anderson JT, Droogmans L. 2005. Biosynthesis and function of 1-methyladenosine in transfer RNA. *Top. Curr. Gen.* 12: 121-139.
125. Saikia M, Fu Y, Pavon-Eternod, He C, Pan T. 2010. Genome-wide analysis of N¹-methyladenosine-modification in human tRNAs. *RNA* 16: 1317-1327.

126. Anderson J, Phan L, Cuesta R, Carlson BA, Pak M, Asano K, Bjork GR, Tamame M, Hinnebusch AG. 1998. The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev.* 12: 3650-3662.
127. Finer-Moore J, Czudnochowski N, O'Connell JD, Wang AL, Stroud RM. 2015. Crystal structure of the human tRNA m¹A58 methyltransferase-tRNA₃^{LYS} complex: refolding of substrate tRNA allows access to the methylation target. *J. Mol. Biol.* 427: 3862-3876.
128. Wang M, Zhu Y, Wang C, Fan X, Jiang X, Ebrahimi M, Qiao Z, Niu L, Teng M, Li X. 2016. Crystal structure of the two-subunit tRNA m(1)A58 methyltransferase Trm6-Trm61 from *Saccharomyces cerevisiae*. *Sci. Rep.* 1(6): 32562.
129. Anderson J, Phan L, Hinnebusch AG. 2000. The Gcd10p/Gcd14p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 97: 5173-5178.
130. Wu B, White KA. 1998. Formation and amplification of a novel tombusvirus defective RNA which lacks the 5' nontranslated region of the viral genome. *J. Virol.* 72(12): 9897-9905.
131. Calvo O, Cuesta R, Anderson J, Gutierrez N, Gacia-Barrio MT, Hinnebusch AG. 1999. GCD14p, a repressor of GCN4 translation, cooperates with Gcd10p and Lhp1p in the maturation of initiator methionyl-tRNA in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 19(6): 4167-4181.
132. Koonin EV, Dolja VV and Kupovic M. 2015. Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology* 479: 2-25.
133. Wang S, White KA. 2007. Riboswitching on RNA virus replication. *Proc. Natl. Acad. Sci. USA* 104(25): 10406-10411.
134. Vlot AC, Bol JF. 2003. The 5' untranslated region of alfalfa mosaic virus RNA1 is involved in negative-strand RNA synthesis. *J. Virol.* 77 (20): 11284-11289.
136. Herold J, Andino R. 2001. Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol. Cell* 7: 581-591.
137. Iglesias NG, Gamarnik AV. 2011. Dynamic RNA structures in the dengue virus genome. *RNA Biol.* 8(2): 249-257.
138. Andino R, Rieckhof GE, Baltimore D. 1990. A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell* 63: 369-380.
139. Vogt DA, Andino R. 2010. An RNA element at the 5'-end of the poliovirus genome functions as a general promoter for RNA synthesis. *PLoS Pathog* 6(6): e1000936. doi:10.1371/journal.ppat.1000936

140. Pogue GP, March LE, Hall TC. 1990. Point mutations in ICR2 motif of brome mosaic virus RNAs debilitate (+)-strand replication. *Virology* 178: 152-160.
141. Frolov I, Hardy IFR, Rice CM. 2001. cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. *RNA* 7: 1638-1651.
142. Yu H, Isken O, Grassmann CW, Behrens S. 2000. A stem-loop motif formed by the immediate 5' terminus of the bovine viral genome modulates translation as well as replication of the viral RNA. *J. Virol.* 74(13): 5825-5835.
143. Filomatori CV, Lodeiro MF, Alvarez DE, Samsa MM, Pietrasanta L, Gamarnik AV. A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes Dev.* 20(16): 2238-2249.
144. Osman TAM, Buck KW. 1997. The tobacco mosaic virus RNA polymerase complex contains a plant protein related to the RNA-binding subunit of yeast eIF3. *J. Virol.* 71(8): 6075-6082.
145. Taylor DN, Carr JP. 2000. The GCD10 subunit of yeast eIF3 binds the methyltransferase-like domain of the 124 and 183 kDa replicase proteins of tobacco mosaic virus in the yeast two-hybrid system. *J. Gen. Virol.* 81: 1587-1591.
146. Macari F, El-houfi Y, Boldina G, Xu H, Khoury-Hanna S, Ollier J, Yazdani L, Zheng G, Bièche I, Legrand N, Paulet D, Durrieu S, Byström A, Delbecq S, Lapeyre B, Bauchet L, Pannequin J, Hollande F, Pan T, Teichmann M, Vagner S, David A, Choquet A, Joubert D. 2016. TRM6/61 connects with PKC with translational control through tRNA_i^{Met} stabilization: impact on tumorigenesis. *Oncogene* 35: 1785-1796.
147. Ghaemmaghami S, Huh WK, bower K, Howson RW, Belle A, Dephoure N, O'shea EK, Weissman JS. 2003. Global analysis of protein expression in yeast. *Nature* 42: 727-741.
148. Baumstark T, Ahlquist P. 2001. The brome mosaic virus RNA3 intergenic replication enhancer folds to mimic a tRNA TpsiC-stem loop and is modified in vivo. *RNA* 7(11): 1652-1670.
149. Kapust RB, Waugh DS. 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 8: 1668-1674.
150. Gietz RD, Woods RA. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* 350: 87-96.
151. Schmitt ME, Brown TA, Trumpower BL. 1990. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucl. Acids Res.* 18(10): 3091-3092.
152. Shevchenko A, Wilm M, Vorm O, Mann M. 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* 68: 850-858.

153. Sasse J, Gallagher SR. 2009. Staining proteins in gel. *Curr. Protoc. Mol. Biol.* 85: 10.6.1-10.6.27.
154. Kushnirov VV. 2000. Rapid and reliable protein extraction from yeast. *Yeast.* 16: 857-860.
155. Gagnon KT, Maxwell ES. 2011. Electrophoretic mobility shift assay for characterizing RNA-protein interaction. *Methods Mol. Biol.* 703: 275-291.
156. Black DL, Chan R, Min H, Wang J, Bell L. 2002. The electrophoretic mobility shift assay for RNA binding proteins. *RNA-protein interactions: A practical approach.* Smith CWJ. Oxford University Press. New York. p109-136.

Appendices



Expression of T7-based constructs in tobacco cells

Hyukho Sheen, K. Andrew White*

Department of Biology, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada



ARTICLE INFO

Article history:

Received 13 March 2018

Accepted 15 March 2018

Available online 22 March 2018

Keywords:

T7 promoter

T7 RNA polymerase

Recombinant protein expression

Plant

ABSTRACT

Bacteriophage T7 promoter and RNA polymerase (T7-Pol) are widely used for recombinant protein expression in bacteria. In plants, there exists conflicting results regarding the efficacy of protein expression from T7-Pol-derived mRNAs. To reconcile these contradictory observations, the expression of green fluorescent protein (GFP) from T7 constructs was evaluated in tobacco protoplasts. T7 constructs transcribed by a nuclearly targeted T7-Pol did not express GFP in plant protoplasts, however T7-Pol lacking a nuclear targeting signal was able to translate cytosolically transcribed mRNAs, but only if the messages contained a viral translation enhancer. GFP expression was further evaluated at the plant level by using agroinfiltration-mediated transient expression system. Unlike for cytosolic expression, nuclear T7 transcripts containing a viral translation enhancer element did not express GFP, and modifications designed to stabilize and facilitate export of T7 transcripts to the cytosol did not improve the expression. We conclude that expression of nuclear T7 constructs is not feasible in tobacco cells, but cytosolic transcription provides an alternative means to over-express RNAs directly in the cytosol.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Bacteriophage T7 promoter and T7 RNA polymerase (T7-Pol) have been used widely to express recombinant proteins in bacteria and are one of the most important systems for protein production [1]. This T7 promoter/Pol combination has also been assessed for recombinant protein expression in yeast and mammalian cells. In both cases, a reporter gene construct under the control of a T7 promoter was efficiently transcribed by T7-Pol that was targeted to the nucleus. However, efficient reporter protein expression was observed in mouse cells, but not yeast cells [2,3]. The latter null result was further supported by other studies of mammalian cells which determined that nuclear T7 transcripts are uncapped and retained in the nucleus, thereby preventing translation [4–6]. To overcome the nuclear retention barrier, cytosolically localized T7-Pol was used to transcribe reporter genes from DNA plasmids residing in the cytosol [4,5]. Additionally, T7 transcripts were designed to include a viral internal ribosomal entry site (IRES) from encephalomyocarditis virus to facilitate cap-independent translation [4,5]. These modifications significantly improved reporter protein expression in these [4,5] and other T7 promoter/Pol-based mammalian systems [7,8].

Two laboratories have previously investigated reporter protein expression from mRNA transcribed by T7-Pol in the nucleus of tobacco cells, with contrasting results. The earlier study showed that reporter protein from nuclearly targeted T7 constructs were inefficiently expressed in tobacco plants [9], while the latter study found efficient expression of nuclearly localized T7 constructs in tobacco plants [10]. Based on these conflicting results, the utility of T7-based expression of heterologous proteins in plants remains uncertain.

In this report, the expression of T7-based constructs was initially evaluated in plant cells, and the results revealed that nuclearly transcribed T7 constructs do not express GFP. In contrast, cytosolically transcribed T7-Pol mRNAs that also contained a plant virus translational enhancer (TE) efficiently mediated GFP expression. Expression from T7 constructs containing a TE was also investigated at the whole plant level, via transient agroinfiltration. No reporter protein expression was observed from TE-containing mRNAs, further supporting the concept that expression of nuclear T7 constructs is not feasible in plant systems.

2. Materials and methods

2.1. Plasmid construction

Plasmids containing poly-adenosine tracts were maintained in DH5 α strains below 30 °C to avoid truncation of A-tracts.

* Corresponding author.

E-mail addresses: hyukhos@yorku.ca (H. Sheen), kawhite@yorku.ca (K.A. White).

pT7-GFP: PCR product 1 was generated by using a primer set (pT7-11-1 and pT711-2) without a template, and PCR product 2 was made by using pESC-Leu-GFP-VHL [11] and a primer set (pCaMV4 and pT7-12). PCR product 1 (digested with SbfI and NcoI) and PCR product 2 (digested with NcoI and SbfI) were triple-ligated into pRTL2 [12] cut with SbfI.

pT7-GFP-T ϕ : PCR product 1 was generated by amplifying T7 termination signal sequence with SacI and HindIII sites at each end, and PCR product 2 was generated by amplifying T7 termination signal sequence with Hind III and EcoRI sites at each end. These two PCR products, digested with SacI/HindIII and HindIII/EcoRI, respectively, were triple-ligated to pHST20 TBSV [13] cut with SacI and EcoRI, generating pHST20 TBSV Tt. Using this plasmid as a template, PCR product 3 was made using the primers pT7-14 and pT7-15.

PCR product 4 containing GFP ORF was amplified using pT7 GFP and a primer set (pT7-11-1 and pT7-13). The PCR product 3 above (digested with SacI and SbfI) and PCR product 4 (digested with SbfI and SacI) were triple-ligated to pRTL2 cut with SbfI.

pCaMV-GFP: 35S CaMV promoter sequence is amplified using pRTL2 as a template and a primer set (pCaMV1 and pCaMV2), and digested with SbfI and NcoI. Also, GFP ORF was cut out from pT7 GFP using NcoI and SbfI. 35S CaMV promoter and GFP ORF were triple ligated to pRTL2 cut with SbfI. For the generation of p2xCaMV GFP, dual 35S CaMV promoter was amplified by PCR and was similarly triple ligated to pRTL2 cut with SbfI.

pRTL2-nT7-Pol or pRTL2-ctT7-Pol: T7-Pol cDNA with or without nuclear localization signal was amplified using a standard PCR, and inserted to pRTL2 cut with Sall and NheI.

pTBSV-sg2-GFP: For pTBSV-sg2-GFP, T7 promoter, sg2 5' UTR (nucleotides, 3842–3889), GFP ORF, and the 3' UTR of TBSV genome were fused. The overlapping PCR products were ligated to pHST20 TBSV Tt cut with SacII and SmaI.

pRTL2-GFP: GFP ORF was amplified using pESC-Leu-GFP-VHL and a primer sets (pTrans12-1 and pTrans13-1), digested with XbaI and NheI, and inserted to pRTL2 (cut with the same enzymes).

pTEV-GFP: T7 promoter and TEV 5' UTR were amplified using pRTL2 GFP and a primer set (pTEV1 and pTEV2), and digested with SbfI and NcoI. GFP ORF and CaMV poly A signal sequence were cut out from pRTL2 GFP using NcoI and SbfI. The PCR product, GFP ORF and CaMV poly A signal sequence were triple ligated to pRTL2 cut with SbfI, creating pT7 TEV GFP CaMV polyA sig. Then T7 promoter, TEV 5' UTR and GFP ORF were amplified using this plasmid and a primer set (pTEV3 and pTEV 4-1). After digesting the PCR product with SacII and SspI, it was inserted to pHST20 TBSV Tt cut with SacII and SmaI.

T7-TEV-GFP-A50: T7 promoter, TEV 5' UTR and GFP ORF were digested out from pT7 TEV GFP CaMV poly A sig using SbfI and NheI. Poly adenosine residues, the ribozyme sequence, T7 termination signal were digested out from pTEV GFP using NheI and PvuII. These two digested DNAs were triple ligated to pCambia 0305 cut with SbfI and AfeI.

T7-CTE-TEV-GFP-A50: CTE sequence [14] was amplified and inserted to T7 TEV GFP A50 cut with ApaI.

2.2. Protoplast preparation and transfection

Tobacco protoplasts were isolated from tobacco plants grown for 2–3 months in green house at ambient room temperature (RT), and they were transfected with the plasmids according to the previously described method [15] with slight modifications. 1×10^5 protoplasts in 200 μ l of MMG buffer were transfected with a total of 40 μ g of plasmid(s) and 225 μ l of 40% PEG solution. For the transfections with two plasmids (e.g., pT7 GFP and pRTL2 nT7 pol), 20 μ g of each plasmid were used. The transfected protoplasts were

incubated at 25 °C for ~24 h in the dark before the fluorescence microscopy analyses (Zeiss AxioCam IC fluorescence microscope, Zen software).

2.3. Agroinfiltration

Tobacco plants were agroinfiltrated as previously described [16] except that OD₆₀₀ of agrobacterium was adjusted to 0.5. The agrobacterium used was AGL1. After agroinfiltration, the plants were grown for four days in a growth chamber (20 °C for 8 h in the dark and 23 °C for 16 h with light), before the fluorescence microscopy analyses.

2.4. Isolation of RNA and extraction of proteins

The agroinfiltrated areas of tobacco leaves were cut out and ground using liquid nitrogen, and the ground power was stored immediately at –80 °C until use. For isolation of RNA, about 1/5 full of the ground power in eppendorf tubes was mixed with 1 ml of Trizol solution and left for 15 min at RT with occasional vortexing. 200 μ l of chloroform was then added and left for additional 10 min. The samples were centrifuged at 12,000 g for 20 min at 4 °C, and the upper phase was precipitated with 1 ml of isopropanol overnight at –20 °C. After centrifugation at 16,000 g for 20 min, the pellet was dried, dissolved in 100 μ l water, and extracted with PCI (phenol: chloroform: isoamyl alcohol) three times. The upper phase was then ethanol-precipitated, washed and dissolved in 50 μ l of water. The isolated total nucleic acids containing $1 \times$ T7 transcription buffer (Epicentre) were digested with 2 μ l of DNase I (Epicentre) for 20 min at 37 °C, and the samples were PCI-extracted three times and ethanol-precipitated overnight. After centrifugation and washing the pellet with 70% ethanol, the pellet was dried and resuspended in 20 μ l of water.

For extraction of proteins, equal volumes of the ground power and $2 \times$ SDS sample buffer were mixed and heated to 100 °C for 3 min before loading to SDS-PAGE. Western blotting was carried out using rabbit anti-GFP antibody (Cell Signalling) and secondary antibody conjugated with Cy3 (Jackson ImmunoResearch).

2.5. RT-PCR

Reverse transcription was carried out following the manufacturer's protocol (Superscript IV, ThermoFisher Scientific). PCR reactions were performed using fusion polymerase and its protocol (New England Biolabs). For the detection of GFP transcripts, PCR was done using the primers, pTrans 12-1 and pTrans 13-1. For the detection of T7 transcripts, the primers pT7 downstr and pTrans13-1 were used. To detect actin mRNA, the primers pToAct1 and pToAct2 were used.

2.6. The primers used in this study

pT7-11-1: gagaaacctgcaggtaatagcactcactatagacacgctgaagctagtcgactc
 pT7-11-2: cttcagaatcaactttgttccatggctcggctgagctgactgcttcagcg
 pT7-12: cttgcatgctcaggtcactggattttg
 pT7-13: gttcctgagctcaatagagagatagattgtagag
 pT7-15: gttcttctgcaggcagctatgacatgattacgaattc
 pCaMV1: gaacaacctgcaggtgagactttcaacaagggtaatac
 pCaMV2: gttcttccatggctcaggtgagctgactgactgactgcttccctc
 ccaaatgaaatgaac
 pCaMV4: gatcaacctggtgagcaaggcgaggagctg
 pTrans 12-1: gatcaatctagaatggtgagcaaggcgaggagctgttc
 pTrans 13-1: ttgatcctgacttactgtacagctgcttccatgc

pTEV1: gagaaacctgcaggaatacactgactcactataggaataacaactctcaacac
 aacatatacaaaaacacg
 pTEV2: ttttgtccatggctatcgttcgtaaatg
 pTEV3: gagaaaccgcggaatacactgactcactataggaataacaactctcaacac
 pTEV4-1: cttgttaataatTTT
 tagagagagatagattttagag
 pT7MCS-2: gagaaacctgcaggaatacactgactcactatagggcccactagtgatc
 ggaataacaactctcaacacacatatac
 pT7 downstr: cccactagtgatcggaaataac
 pToAct1: gctactcttaccactactgctgaacgag
 pToAct2: ttagaagcattttctgtgcacaatggatg

3. Results and discussion

3.1. T7-Promoter GFP constructs with nuclearly targeted T7-Pol do not express GFP in protoplasts

As an initial test of reporter protein yield, plant protoplasts were co-transfected with a T7-promoter-driven GFP construct and a nuclearly targeted T7-Pol. Tobacco (*Nicotiana tabacum*) protoplasts were transfected with DNA plasmid pT7-GFP, where GFP is under the control of a T7 promoter and a cauliflower mosaic virus (CaMV) poly(A) signal sequence (Fig. 1A). This construct was co-transfected with another plasmid, pRTL2-nT7-Pol, which provided nuclearly targeted T7-Pol (nT7-Pol); due to its fusion to the nuclear localization signal (NLS) of SV40 T antigen. SV40 NLS was previously shown to import T7-Pol to the nucleus in tobacco cells in two independent studies [9,17]. pRTL2-nT7-Pol was also tested in co-transfections with another T7-promoter-driven GFP construct, pT7-GFP-T ϕ , that contained tandem T7 termination signals in place of the CaMV poly(A) signal sequence in the pT7-GFP plasmid, to allow for T7-Pol-based termination of transcription (Fig. 1A). For positive controls, protoplasts were transfected with constructs pCaMV-GFP or p2xCaMV-GFP, which are identical to pT7-GFP, except that they contain single or double 35S cauliflower mosaic

virus (CaMV) promoters, respectively (Fig. 1A). However, unlike for pT7-GFP and pT7 GFP T ϕ , these control transcripts are transcribed by plant RNA polymerase II and thus capped and poly(A) tailed. At 24 h post transfection, similar numbers of protoplasts from each sample were analyzed by bright field and fluorescence microscopy. The tobacco cells co-transfected with pRTL2-nT7-Pol and pT7-GFP or pT7-GFP-T ϕ did not show detectable presence of GFP (Fig. 1B). In contrast, the GFP constructs under single or dual 35S promoters, pCaMV-GFP or p2x-CaMV-GFP, exhibited readily observable levels of GFP (Fig. 1B). Thus, the combination of T7-promoter-driven GFP constructs and nuclearly targeted T7-Pol did not yield detectable levels of reporter protein in plant protoplasts.

3.2. GFP expression from cytosolically transcribed T7-Pol-derived mRNAs containing a viral TE

The poor expression from T7 GFP constructs provided with nuclear T7-Pol may be due to the absence of a cap and poly(A) tail, causing either nuclear retention and/or weak translation. Such deficiencies could be remedied by transcribing the messages in the cytosol and adding a viral TE to the mRNAs. It was previously shown that T7-Pol lacking the nuclear localization signal is cytosolic [17]. Thus, tobacco protoplasts were co-transfected with pT7-GFP or pT7-GFP-T ϕ along with T7-Pol-expressing pRTL2-cT7, which lacked the NLS. As seen in Fig. 2B, neither transfection displayed detectable GFP, indicating that the T7 constructs cannot express GFP by directing their transcription in the cytosol. Moreover, this result indicates that even if nuclear T7-Pol-transcribed messages were to be exported to the cytosol, they would not be translatable.

The lack of translation observed for cytosolically-expressed reporter mRNAs may be due to the absence of a cap and poly(A) tail. If this is correct, then the deficiency could potentially be overcome by adding alternative translation enhancing elements. To this end, plant virus TEs from two different viruses, Tomato bushy stunt virus (TBSV) or Tobacco etch virus (TEV), were added to pT7-GFP-T ϕ ,

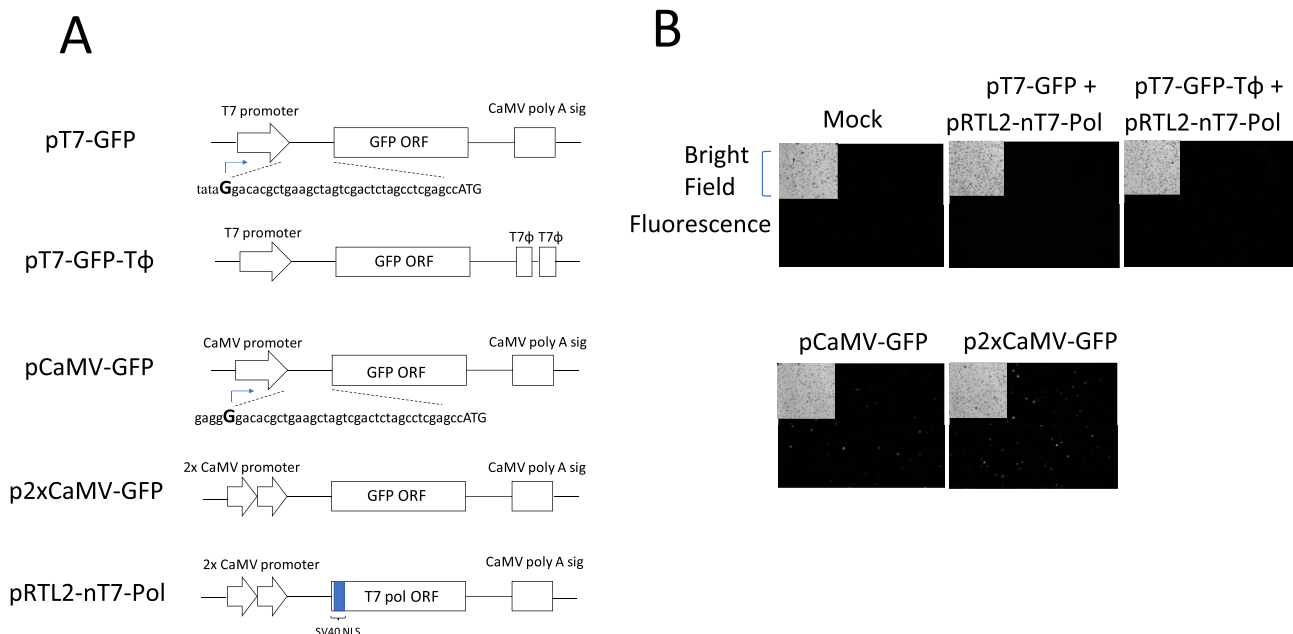


Fig. 1. Expression of GFP from T7 constructs provided with nuclear T7-Pol. **A.** Schematic representation of plasmids used for protoplast transfection. Transcription initiation sites from both T7 promoter and 35S CaMV promoters are shown with an arrow. The start codon of GFP ORF is shown in upper case. **B.** Expression of GFP from T7 constructs provided with nuclear T7-Pol. The plasmids used for transfections were indicated above the pictures. For transfection with two plasmids (e.g., pT7-GFP and pRTL2-nT7-pol), 20 μ g of each plasmid were used. For pCaMV-GFP and p2x-CaMV-GFP, 40 μ g of plasmids were used. The pictures were taken 24 h post transfection. Bright field microscopy was used to confirm that similar numbers of protoplasts were analyzed. pCaMV-GFP and p2x-CaMV-GFP serve as transfection positive controls.

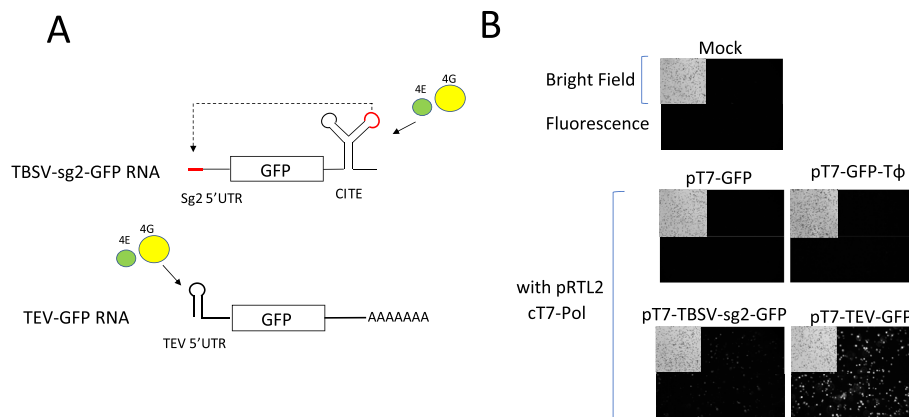


Fig. 2. Expression of GFP from T7 constructs provided with cytosolic T7-Pol. **A.** Schematics of T7-GFP RNAs containing different viral translation enhancers (TE) along with translational initiation factors (eIFs) 4E and 4G. The base-pairing regions in the 5' UTR and TE are shown by a dotted line. **B.** Expression of the T7 constructs with or without the TE and cytosolic T7-Pol.

generating pTBSV-sg2-GFP and pTEV-GFP (Fig. 2A). Both of these viral TEs bind translation initiation factors (eIFs) [18,19]. The TEV TE is located and functions in the 5'UTR, while the TBSV TE is located in the 3'UTR and requires a complementary sequence in the 5'UTR, to which it base pairs and repositions it, and its bound eIFs, proximal to the 5'-proximal start codon (Fig. 2A) [18–20]. As shown in Fig. 2B, when T7 constructs containing TBSV or TEV TEs were transcribed by cytosolic T7 pol, they showed readily detectable levels of GFP expression in the tobacco protoplasts. The ability to rescue translation of the cytosolically-expressed reporter mRNAs by adding viral TEs is consistent with the concept that T7-Pol-synthesized transcripts lack a cap and poly(A) tail.

3.3. No translation of nuclear-expressed, T7-Pol-transcribed mRNAs in plants

Expression of T7 constructs was also evaluated at the plant level. Previously, Zeitouni et al. [9] showed that their T7 construct, consisting of T7 promoter, GUS reporter gene and Nos terminator (Nos T), did not express GUS in their transgenic tobacco plants. They also found that T7 transcription occurred actively, but that the transcript was post-transcriptionally silenced [9]. Thus, the lack of T7 expression in that study [10], and the efficient T7 expression in another study [10], may be due to different degree of gene-silencing in tobacco plants. If this was the case, then some reporter protein expression from T7 constructs should be transiently observable if introduced by agroinfiltration. The construct, T7-TEV-GFP-NosT, similar to the construct from Zeitouni et al. [9], consisted of a T7 promoter, a TE element from TEV, a GFP ORF and a Nos terminator. When this construct was introduced via agroinfiltration along with the nT7-Pol construct to tobacco plants, no GFP expression in the agroinfiltrated leaves was observed (data not shown), suggesting that protein expression from T7 constructs at the plant level is unachievable, as suggested by Zeitouni et al. [9].

It is known that capping and polyadenylation require host RNA polymerase II in mammalian cells [21,22], and that polyadenylation is critical for mRNA export [23,24]. Also, transcripts from transgene that are abnormally processed at the 3' end can induce post-transcriptional silencing [25]. Therefore, transcripts made by T7 polymerase lacking a poly(A) tail could promote silencing and/or inhibit export. In an attempt to improve GFP expression from T7-TEV-GFP-NosT, the Nos terminator was replaced by a 50 nt long poly(A) tract and a Hepatitis delta virus ribozyme, creating construct T7-TEV-GFP-A50 (Fig. 3A). Tobacco leaves agroinfiltrated

with T7-TEV-GFP-A50 along with nT7-Pol construct, did not generate any observable GFP expression (Fig. 3B).

It was reported previously that T7 transcripts containing a stable RNA stem-loop at the 5'-end significantly improved stability of T7 transcript [26]. Thus, a well-characterized stem-loop, termed constitutive transport element (CTE), from Mason Pfizer monkey virus (MPMV) [14] was added to T7-TEV-GFP-A50, creating T7-CTE-TEV-GFP-A50 (Fig. 3A). The idea being that the CTE could potentially facilitate nuclear export of the message and its secondary structure would protect the 5'-terminus. However, as shown in Fig. 3B, neither the T7-CTE-TEV-GFP-A50 + nT7-Pol combination nor the negative control, T7-CTE-TEV-GFP-A50 alone, showed any GFP expression. GFP levels from these T7 constructs, as well as a positive control, CaMV-GFP (Fig. 3B), was also confirmed by western blotting using anti-GFP antibody (Fig. 3C). Thus, T7 constructs do not express GFP even transiently in tobacco plants, even with modifications designed to enhance expression.

3.4. Low levels of GFP transcripts are synthesized by T7 and plant RNA polymerases

Previously, Zeithoune et al. [9] performed Northern blotting in an attempt to detect T7 transcripts from their tobacco transgenic plants, but they were unable to observe any T7 transcripts due to post-transcriptional gene silencing. Studies from drosophila and mammalian cells showed truncated messages indicating that T7-Pol was not very processive, likely due to interference by the chromatin structures in higher eukaryotes [27,28]. It was therefore of interest to determine if T7 transcripts corresponding to a full GFP ORF were produced in our agroinfiltrated plants.

To address this question, semi-quantitative RT-PCR was performed on isolated total nucleic acids from the agroinfiltrated tobacco leaves in Fig. 3B. To ensure that RT-PCR products were the result of amplification of T7 RNA transcripts, and not derived from the T7 construct DNAs, all the isolated nucleic acids were treated with DNase I prior to RT-PCR to remove genomic DNA and RT-PCR reactions minus reverse transcriptase were carried out in parallel (Fig. 4). Also, RT-PCR primers spanned the GFP coding region, to test for the presence of the entire ORF. RT-PCR without RT did not amplify any T7 construct DNAs (Fig. 4, panel 1), while in RT-containing reactions, GFP transcripts were detected from the tobacco leaves agroinfiltrated with T7-TEV-GFP-A50 + nT7-Pol and T7-CTE-TEV-GFP-A50 + nT7-Pol (Fig. 4, lanes 5 and 6 in panel 2). Unexpectedly, GFP transcripts were also detected from the leaves

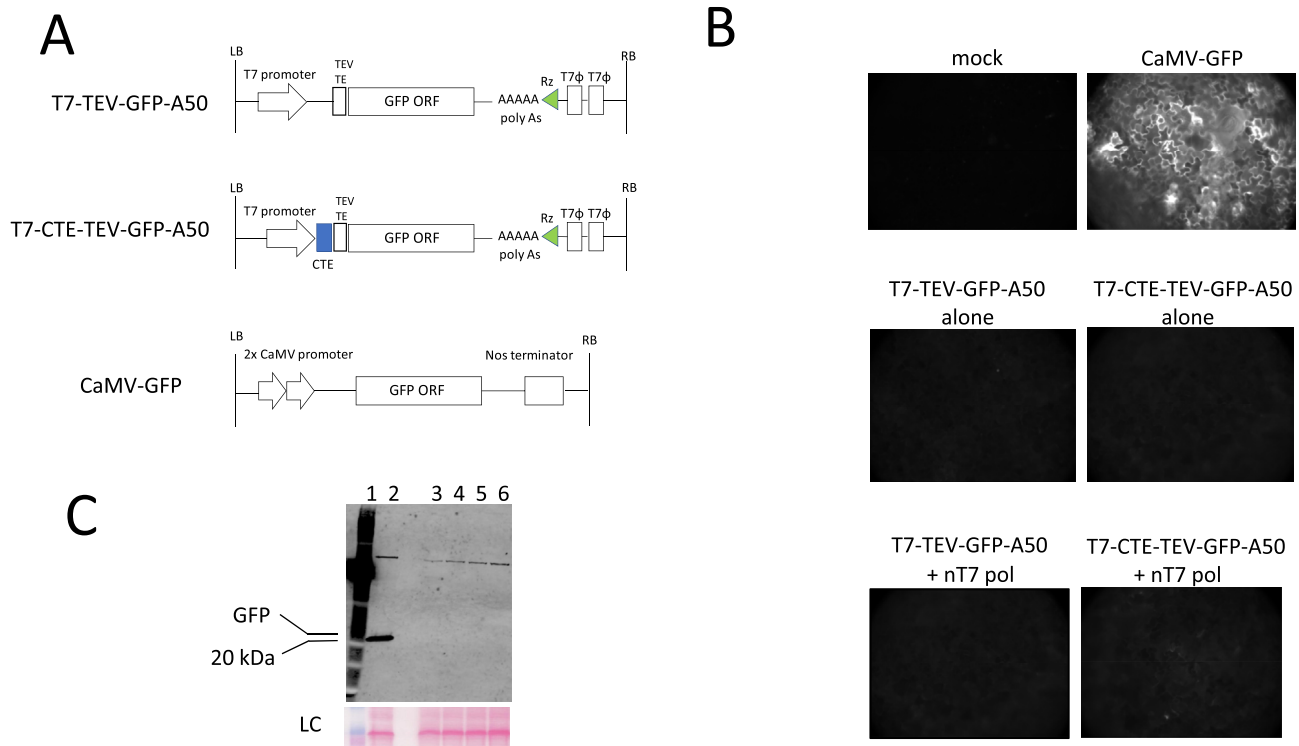


Fig. 3. Agroinfiltration of tobacco leaves with T7 constructs. **A.** Schematic representation of the DNA constructs used for agroinfiltration. LB and RB are left border and right border, respectively. Rz is the ribozyme from Hepatitis delta virus antigenome. T7 ϕ is T7 termination signal, and fifty adenosines are attached to the 3' end of the GFP 3'UTR in the T7 constructs. **B.** GFP expression from the agroinfiltrated leaves were shown under fluorescence microscopy 4 days after agroinfiltration. **C.** Proteins were extracted from the agroinfiltrated leaves in B and analyzed by western blotting to confirm the absence of GFP expression from the T7 constructs. Ponceau S staining is provided as loading controls (LC). Lane 1 contains a molecular weight marker with the 20 kDa band indicated. Lanes 2–6 are proteins extracted from the tobacco leaves agroinfiltrated with pCaMV-GFP (lane 2), T7-TEV-GFP-A50 alone (lane 3), T7-CTE-TEV-GFP-A50 alone (lane 4), T7-TEV-GFP-A50 + nT7-Pol (lane 5), T7-CTE-TEV-GFP-A50 + nT7-Pol (lane 6).

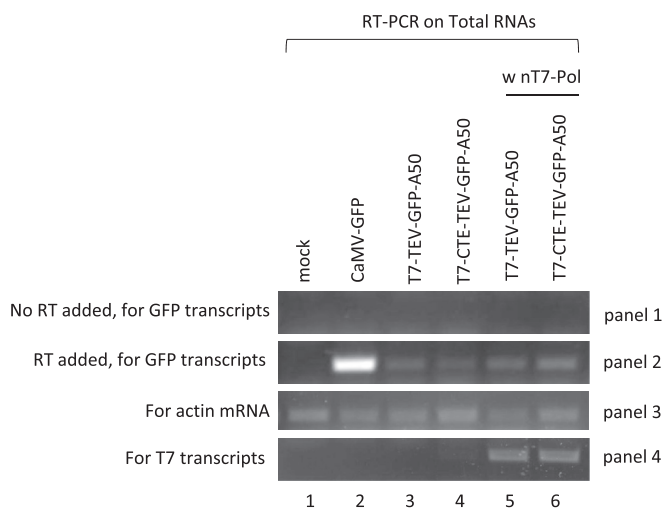


Fig. 4. RT-PCR on the total RNAs isolated from the agroinfiltrated leaves. The first panel is the RT-PCR results performed without the addition of reverse transcriptase, while the second panel is RT-PCR performed with the addition of reverse transcriptase. The first and second panels were done to detect GFP transcripts, and they were performed under the same RT-PCR conditions in parallel. The third panel is the RT-PCR results to detect actin mRNA from the samples. The last panel is the RT-PCR results to detect T7 transcripts made only by T7-Pol.

agroinfiltrated with the T7 constructs without nT7 pol (Fig. 4, lanes 3 and 4 in panel 2), suggesting that host RNA polymerase fortuitously transcribed the T7 constructs in tobacco cells. Importantly, GFP transcript levels from the T7 constructs were much lower than

from the positive control CaMV-GFP construct (Fig. 4, lane 2 in panel 2), suggesting that the lack of GFP expression from our T7 constructs is likely due to the very low levels of GFP transcripts.

Because host RNA polymerase fortuitously transcribed the T7 constructs, the RT-PCR results did not clearly show whether T7-Pol was also able to transcribe the GFP ORF. If the transcription initiation site used by the host RNA polymerase was downstream of the T7 transcription initiation site, performing RT-PCR using a different more upstream primer could exclude its amplification. An upstream primer between CTE and TE sequences was selected for use with the original downstream primer. RT-PCR results using the new upstream primer amplified T7 transcripts only from the leaves agroinfiltrated with T7-TEV-GFP-A50 + nT7-Pol or T7-CTE-TEV-GFP-A50 + nT7-Pol (Fig. 4, lanes 5 and 6 in panel 4), showing that T7-Pol does transcribe the GFP ORF, albeit at low levels, based on the results from the panel 1.

In conclusion, two different laboratories previously studied expression of T7 GUS constructs in tobacco transgenic plants, but with contrasting results. The earlier study showed a lack of GUS expression [9], while the more recent report showed its efficient expression [10]. Our results are consistent with the early study, because our equivalent T7 GFP constructs did not express GFP to detectable levels in protoplasts, regardless if they were provided with nuclear- or cytosolically-targeted T7-Pol. Only when a viral TE was added, and transcription was targeted to the cytosol, did notable GFP expression occur, suggesting that nuclear-based T7-Pol expression is not viable in this system. Expression of T7 constructs was further evaluated using agroinfiltration-mediated transient expression system. However, a construct consisting a T7 promoter, a viral TE, GFP ORF and Nos terminator showed no transient GFP

expression, and adding potentially beneficial terminal modifications did not alter this result. Accordingly, it appears that nuclear T7 constructs in both protoplasts and whole plants are unable to express their reporter proteins. Nonetheless, cytosolic expression of T7 constructs in plant cells could be useful for the production and analysis of non-mRNAs, such as regulatory noncoding plant RNA transcripts.

Conflicts of interest

The authors declare that potential conflicting interests do not exist.

Acknowledgements

We are grateful to Irina Shuralyova and Maria Mazzurco for the use of fluorescence microscope. We appreciate Dr. Kathalin Hudak for the kind gifts of pCambia 0305 and agrobacterium AGL1. We also thank Glen Marlok for growing the tobacco plants. This work was supported by NSERC.

References

- [1] T.W. Overton, Recombinant protein production in bacterial host, *Drug Discov. Today* 19 (5) (2014) 590–601.
- [2] B.M. Benton, W. Eng, J.J. Dun, F.W. Studier, R. Sternglanz, P.A. Fisher, Signal mediated import of bacteriophage T7 RNA polymerase into *Saccharomyces cerevisiae* nucleus and specific transcription of a target genes, *Mol. Cell Biol.* 10 (1) (1990) 353–360.
- [3] A. Lieber, U. Kiessling, M. Strauss, High level gene expression in mammalian cells by a nuclear T7 phage RNA polymerase, *Nucleic Acids Res.* 17 (21) (1989) 8485–8593.
- [4] O. Elroy-Stein, M. Fuerst, B. Moss, Cap-independent translation conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system, *Proc. Natl. Acad. Sci. U.S.A.* 86 (16) (1989) 6126–6130.
- [5] O. Elroy-Stein, B. Moss, Cytoplasmic expression system based on constitutive synthesis of bacteriophage T7 RNA polymerase in mammalian cells, *Proc. Natl. Acad. Sci. U.S.A.* 87 (1990) 6743–6747.
- [6] H. Cheng, K. Dufu, C. Lee, J.L. Hsu, A. Dias, R. Reed, Human mRNA export machinery is recruited to the 5' end of mRNA, *Cell* 127 (2006) 1389–1400.
- [7] M. Brisson, Y. He, S. Li, J.P. Yang, L. Huang, A novel T7 RNA polymerase autogene for efficient cytoplasmic expression of target genes, *Gene Ther.* 6 (2) (1999) 263–270.
- [8] J. Finn, A.C.H. Lee, I. MacLachlan, P. Cullis, An enhanced autogene-based dual promoter cytoplasmic expression system yields increased gene expression, *Gene Ther.* 11 (2004) 276–283.
- [9] S. Zeitoune, O. Livneh, L. Kuzuetzova, Y. Stram, I. Sela, T7 RNA polymerase drives transcription of a reporter gene from T7 promoter but engenders post-transcriptional silencing of expression, *Plant Sci.* 141 (1999) 59–65.
- [10] H.T. Nguyen, S. Leelavathi, V.S. Reddy, Bacteriophage T7 RNA polymerase directed, inducible and tissue-specific over-expression of foreign genes in transgenic plants, *Plant Biotechnol. J.* 2 (2004) 301–310.
- [11] D. Kaganovich, R. Kopito, J. Frydman, Misfolded proteins partition between two distinct quality control compartments, *Nature* 454 (7208) (2008) 1088–1095, 2008 Aug 28.
- [12] J.C. Carrington, D.D. Freed, C.S. Oh, Expression of potyviral polyproteins in transgenic plants reveals proteolytic activities required for complete processing, *EMBO J.* 9 (5) (1990) 1347–1353.
- [13] H.B. Scholthof, Rapid delivery of foreign genes into plants by direct rub-inoculation with intact plasmid DNA of a tomato bushy stunt virus gene vector, *J. Virol.* 73 (9) (1999) 7823–7829.
- [14] R.T. Ernst, M. Bray, D. Rekosh, M.L. Hammarskjold, Secondary structure and mutational analysis of the Mason Pfizer monkey virus RNA constitutive transport element, *RNA* 3 (1998) 210–222.
- [15] S.D. Yoo, Y.H. Cho, J. Sheen, Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis, *Nat. Protoc.* 2 (2007) 1565–1575.
- [16] S. Bachan, S.P. Dinesh-Kumar, Tobacco rattle virus (TRV)-base virus-induced gene-silencing, *Meth. Mol. Biol.* 894 (2012) 83–92.
- [17] M.W. Lassner, A. Jones, S. Daubert, L. Comai, Targeting of T7 RNA polymerase to tobacco nuclei mediated by an SV40 nuclear location signal, *Plant Mol. Biol.* 17 (1991) 229–234.
- [18] B.L. Nicholson, B. Wu, I. Chevtchenko, K.A. White, Tombusvirus recruitment of host translational machinery via the 3' UTR, *RNA* 16 (2006) 1402–1419.
- [19] D.R. Gallie, Cap-independent translation conferred by the 5' leader of tobacco etch virus is eukaryotic initiation factor 4G dependent, *J. Virol.* 75 (24) (2001) 12141–12152.
- [20] B.L. Nicholson, K.A. White, Functional long-range RNA-RNA interactions in positive-strand RNA viruses, *Nat. Rev. Microbiol.* 12 (2014) 493–504.
- [21] Y. Hirose, J.L. Manley, RNA polymerase II is an essential mRNA polyadenylation factor, *Nature* 413 (1997) 538–542.
- [22] S.S. Sisodia, B. Sollner-Webb, D.W. Cleveland, Specificity of RNA maturation pathways: RNAs transcribed by RNA polymerase III are not substrates for splicing or polyadenylation, *Mol. Cell Biol.* 7 (1987) 3602–3612.
- [23] K. Dower, M. Rosbash, T7 RNA polymerase directed T7 transcripts are processed in yeast and link 3' end formation to mRNA export, *RNA* 8 (2002) 686–697.
- [24] H. Fuke, M. Ohno, Role of poly(A) tail as an identity element for mRNA nuclear export, *Nucleic Acids Res.* 36 (2008) 1037–1049.
- [25] Z. Luo, Z. Chen, Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in Arabidopsis, *Plant Cell* 19 (3) (2007) 943–958.
- [26] T.R. Fuerst, B. Moss, Structure and stability of mRNA synthesized by vaccinia virus-encoded bacteriophage T7 RNA polymerase in mammalian cells. Importance of the 5' untranslated leader, *J. Mol. Biol.* 206 (2) (1989) 333–348.
- [27] K. McCall, W. Bender, Probes of chromatin accessibility in the drosophila bithorax complex respond differently to polycomb mediated response, *EMBO J.* 15 (3) (1996) 569–580.
- [28] T. Jenuwein, W.C. Forrester, R.G. Qui, R. Grosschedl, The immunoglobulin mu enhancer establishes local factor access nuclear chromatin independent of transcription stimulation, *Gene Dev.* 7 (10) (1993) 2016–2032.



Notes & Tips

Quantitation of yeast total proteins in sodium dodecyl sulfate–polyacrylamide gel electrophoresis sample buffer for uniform loading



Hyukho Sheen

Department of Biology, York University, Toronto, Ontario M3J 1P3, Canada

ARTICLE INFO

Article history:

Received 12 July 2015

Received in revised form

29 December 2015

Accepted 3 January 2016

Available online 13 January 2016

Keywords:

Protein quantitation

SDS sample buffer

Laemmli buffer

Protein precipitation

Uniform loading

Acetone precipitation

ABSTRACT

Proteins in sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) sample buffer are difficult to quantitate due to SDS and reducing agents being in the buffer. Although acetone precipitation has long been used to clean up proteins from detergents and salts, previous studies showed that protein recovery from acetone precipitation varies from 50 to 100% depending on the samples tested. Here, this article shows that acetone precipitates proteins highly efficiently from SDS–PAGE sample buffer and that quantitative recovery is achieved in 5 min at room temperature. Moreover, precipitated proteins are resolubilized with urea/guanidine, rather than with SDS. Thus, the resolubilized samples are readily quantifiable with Bradford reagent without using SDS-compatible assays.

© 2016 Elsevier Inc. All rights reserved.

Uniform loading of total proteins for sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) is an important step for many Western blot analyses, and for this total proteins are typically extracted from cells or tissues using a buffer that is compatible with subsequent protein assays. However, in some instances, proteins are extracted directly using SDS–PAGE sample buffer (hereafter SDS sample buffer). For example, efficient extraction of yeast total proteins is achieved by incubating cells in 0.1 N NaOH and then boiling directly with SDS sample buffer [1]. However, a drawback to such extractions is that quantitation of proteins in SDS sample buffer is difficult because the presence of SDS and a reducing agent in the buffer interferes with normal protein quantitation assays [2]. Neither detergent-compatible nor reducing agent-compatible assays are compatible with SDS sample buffer because of the simultaneous presence of both SDS and reducing agents in the sample buffer. Thus, for protein samples extracted with SDS sample buffer, their quantitation and subsequent uniform loading for SDS–PAGE are difficult.

One option for quantitating proteins in SDS sample buffer is to measure turbidity of proteins in SDS sample buffer after the addition of trichloroacetic acid (TCA); however, turbidity can change dramatically over a short period of time, and some substances in the sample can interfere with quantitation (e.g., DNA) [3]. Another option is to use proprietary assay kits from commercial vendors (e.g., Pierce 660 nm, Bio-Rad RC DC, or GE Healthcare 2-D Quant kits), but these are more expensive than normal protein assays such as Bradford assay. Furthermore, 660 nm protein assay requires diluting samples from SDS sample buffer at least 10- to 20-fold before quantitation to minimize interference. Thus, the starting protein samples need to be at a high concentration, and at least 0.5–1 $\mu\text{g}/\mu\text{l}$ is required. RC DC and 2-D Quant kits are based on precipitation of proteins while leaving interfering substances in the supernatant.

Here, acetone precipitation is tested to quantitatively recover proteins from SDS sample buffer for subsequent protein quantitation. Conventionally, acetone precipitation is used to clean up protein samples from detergents and salts for downstream applications such as mass spectrometry and two-dimensional gel analyses. However, the recovery of proteins from acetone precipitation varies significantly from 50 to 100% [4–8] depending on the samples tested. However, Crowell and coworkers recently showed that the ionic strength of samples is one of the key factors for recovery of

Abbreviations: SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TCA, trichloroacetic acid; BSA, bovine serum albumin; RT, room temperature.

E-mail address: hyukhos@yorku.ca.

proteins from acetone precipitation [9]. According to their ion-pairing model, ionic substances such as NaCl can neutralize protein surface, and this helps aggregation and precipitation of proteins on the addition of acetone. In their study, acetone precipitated bovine serum albumin (BSA) completely with the addition of 1–100 mM NaCl or 0.1% SDS, whereas acetone precipitated BSA poorly without these additions [9]. These results clearly showed the importance of the ionic strength of the samples for efficient recovery of proteins. Even so, for myoglobin and cytochrome *c* prepared in 0.25–1.25% SDS solution, their recovery from acetone precipitation was still approximately 80% [4]. In addition, for ribosomal proteins prepared in 0.5% SDS, the recovery varied from 80 to 95% [5]. Notably, recovery of these proteins becomes poor (i.e., 10–30% recovery) when the precipitation condition was changed from -20°C overnight to 20°C for 5 min [5]. Normally, $1\times$ SDS sample buffer contains 2% SDS, and due to the variable and incomplete protein recoveries in the previous studies, it is questionable whether acetone precipitation can quantitatively recover total proteins from SDS sample buffer for subsequent quantitation.

The aim of this study was to quantitate total proteins in SDS sample buffer by removing SDS and reducing agent through acetone precipitation. For this, it was first tested whether acetone precipitation at ambient room temperature (RT) for 5 min could fully recover total proteins from SDS sample buffer. Second, it was tested whether acetone-precipitated proteins can be efficiently resolubilized without SDS to use Bradford assay. In general, precipitated proteins from TCA or acetone precipitation are difficult to redissolve without SDS [10], and resolubilization with SDS inevitably requires SDS-compatible protein quantitation assays. However, Bradford assay, the most widely used assay, is not compatible with SDS-solubilized samples.

The results presented here show that acetone is highly efficient at precipitating BSA and yeast total proteins from SDS sample buffer. Near full recovery was achieved in 5 min at RT, in contrast to a previous study [5]. Furthermore, precipitated proteins could be resolubilized with urea and guanidine and could be accurately quantitated with Bradford assay without using SDS-compatible assays.

The first aim of this study was to determine whether acetone can quantitatively recover proteins from SDS sample buffer. To this end, the recovery of proteins from SDS sample buffer was monitored using BSA and yeast total proteins. First, the known amounts of BSA (0.1–2.0 $\mu\text{g}/\mu\text{l}$) were prepared in 20 μl of $1\times$ SDS sample buffer (50 mM Tris [pH 6.8], 10% [v/v] glycerol, 2% [w/v] SDS, 100 mM dithiothreitol, and 0.02% [w/v] bromophenol blue) and precipitated with 4 volumes of 100% acetone (final acetone concentration, 80%, v/v) at ambient RT for 5 min. The samples were then spun at 16,000 g for 2 min at RT, and the supernatant was removed slowly and thoroughly in a single operation. After drying the pellets for 2 min, BSA was resolubilized in 20–40 μl of 6 M urea/3 M guanidine-HCl for 10 min with occasional tappings. All or a portion of the samples was then used for the Bradford assay (Bio-Rad, Canada). As shown in Fig. 1A, acetone recovered BSA nearly completely from SDS sample buffer in a quantitative manner throughout the concentrations tested, whereas acetone did not recover BSA well from H_2O alone when its concentrations were below 0.4 $\mu\text{g}/\mu\text{l}$. These results suggest that acetone and SDS sample buffer together facilitate recovery of BSA. Moreover, the recovery results in Fig. 1A are consistent with the recovery results determined by densitometric analysis of precipitated BSA, which was separated by SDS-PAGE (cf. Fig. 1A vs. Fig. 1B). This indicates that resolubilization of BSA with urea/guanidine and subsequent use of Bradford assay accurately quantitated precipitated BSA.

Second, recovery of yeast total proteins from SDS sample buffer was monitored by comparing the amounts of the total proteins

before and after acetone precipitation. Specifically, total proteins from *Saccharomyces cerevisiae* were extracted with SDS sample buffer as described previously [1], and cell debris was removed by a quick spin (5–10 s) at 16,000 g. Then, the supernatants were collected and serially diluted by 2-fold with SDS sample buffer (up to 16-fold) (Fig. 1C). The first half of each sample was loaded to SDS-PAGE (lanes 1–5 of Fig. 1C), whereas the second half of each sample was acetone precipitated (5 min at RT), resolubilized with $1\times$ SDS sample buffer, and loaded to SDS-PAGE (lanes 6–10). As shown in Fig. 1C, acetone precipitated each sample without significant loss of proteins (cf. lanes 1–5 vs. lanes 6–10). Recovery efficiency determined by the densitometric analysis indicated that near full recovery (97–100%) was possible up to approximately 0.2 $\mu\text{g}/\mu\text{l}$ total proteins in 10 μl (lanes 6–9 in Fig. 1C). Taken together, these results indicate that acetone can efficiently precipitate BSA and yeast total proteins from SDS sample buffer and that these proteins can be accurately measured by the Bradford assay following acetone precipitation and resolubilization with urea/guanidine.

To further test the accuracy and usefulness of this quantitation method, samples containing various unknown amounts of yeast total proteins were prepared in SDS sample buffer. After their quantitation through acetone precipitation, calculated equal amounts were loaded to SDS-PAGE. Uniform loading of these samples was then used to assess the accuracy of this quantitation method. Specifically, to prepare these samples, wild-type and eight single-gene knockout *S. cerevisiae* strains were cultured in yeast synthetic medium overnight. An equal volume from each culture (1 ml) was harvested, and the total proteins were extracted with SDS sample buffer as described previously [1]. Due to different growth rates of these strains, extraction of proteins from these samples results in different amounts of extracted proteins in each sample. As shown in Fig. 2A, loading equal volumes of these samples showed varied levels of proteins. However, after quantitation of the samples through acetone precipitation, the loading of an estimated equal amount led to uniform loading of the samples (Fig. 2B). These results support that quantitation of total proteins by acetone precipitation and resolubilization with urea/guanidine are accurate and useful for uniform loading of total proteins.

A few technical aspects of acetone precipitation of proteins from SDS sample buffer are discussed below. First, acetone coprecipitates trace amounts of bromophenol blue along with proteins from SDS sample buffer. However, these amounts of bromophenol blue are negligible for protein quantitation because only a small volume of resolubilized samples (20 μl or less, although the final volumes were made up to 20 μl with 6 M urea/3 M guanidine) is mixed with a significantly greater amount of Bradford reagent (800 μl) for protein quantitation. Second, for accurate quantitation, it is recommended that the optical density (OD_{595}) of the Bradford assay be below 0.5 due to the nonlinear nature of Bradford assay at high protein concentrations. Third, it should be noted here that although near full recovery of yeast total proteins is achievable, recovery efficiency of specific protein types (e.g., membrane proteins) from SDS sample buffer was not tested here. Lastly, Bradford assay is sensitive to SDS higher than 0.125% [2]. However, acetone precipitation efficiently removes SDS from SDS sample buffer, and no interference with Bradford assay was observed for the samples analyzed in this report. Consistent with this observation, it has been previously shown that acetone precipitation used for mass spectrometry analysis removed SDS by 40- to 80-fold without a washing step (e.g., from initial 2% to final 0.05% SDS) [11].

In conclusion, acetone can quantitatively precipitate BSA and yeast total proteins from SDS sample buffer. Furthermore, the conventional prolonged cold acetone precipitation procedure (e.g., overnight at -20°C) is not required for protein recovery from SDS

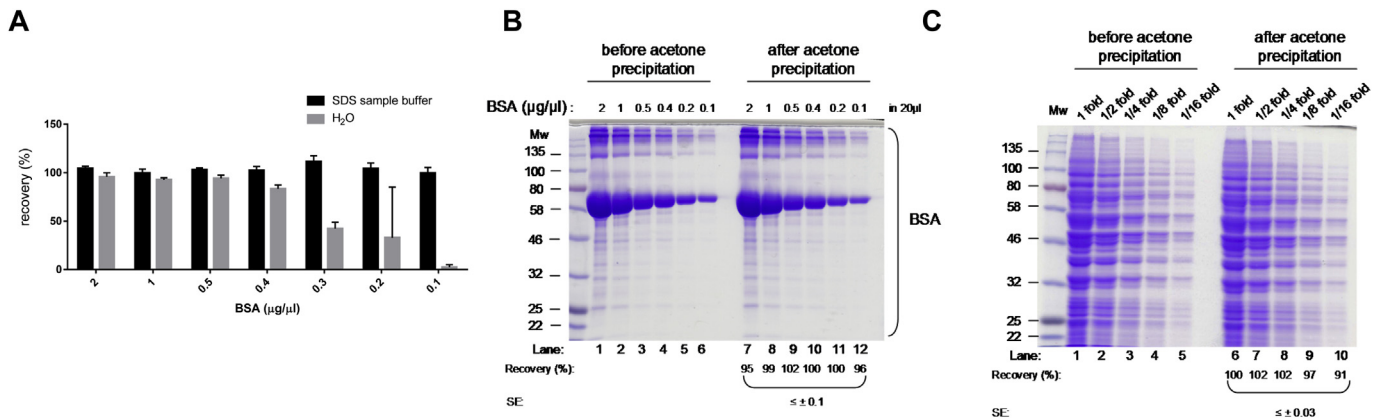


Fig. 1. Acetone quantitatively precipitates BSA and yeast total proteins from SDS sample buffer. (A) Recovery of BSA from SDS sample buffer or H₂O was measured after acetone precipitation. Specifically, 20 µl of BSA in 0.1–2.0 µg/µl was precipitated with 80% (v/v) acetone and dissolved in 20–40 µl of 6 M urea/3 M guanidine-HCl. All or a portion of the samples (final volume, 20 µl) were mixed with 800 µl of 1× Bradford reagent (Bio-Rad), and the amounts of BSA were quantitated according to the manufacturer's protocol. (B) Recovery of BSA in SDS sample buffer was measured after acetone precipitation. Specifically, 20 µl of BSA in 0.1–2.0 µg/µl was precipitated with 80% (v/v) acetone and resolubilized in 1× SDS sample buffer (not in 6 M urea/3 M guanidine-HCl). The entire samples (2–40 µg) were run on 10% SDS-PAGE. The recovery efficiency of BSA from acetone precipitation was quantitated using ImageJ densitometric analysis. Recovery efficiency is indicated below the gel image. The standard error of the mean (SE) was less than 0.1. (C) Recovery of yeast total proteins from SDS sample buffer was monitored by comparing protein amounts before and after acetone precipitation. Specifically, the yeast total proteins in SDS sample buffer were 2-fold serially diluted up to 16-fold with SDS sample buffer. The first half of each sample (10 µl) was loaded to 10% SDS-PAGE (lanes 1–5). The second half of each sample was acetone precipitated, dissolved in 10 µl of SDS sample buffer, and loaded to SDS-PAGE (lanes 6–10). All of the samples analyzed here were heated to 100 °C for 3 min before loading. The gel was stained with Coomassie blue R-250. The concentration of the 16-fold diluted sample (lane 5) was estimated to be 0.11 µg/µl by the method described here. Recovery efficiency is indicated below the gel image. The SE was less than 0.03. Mw corresponds to color prestained protein standard from New England Biolabs, and the sizes are shown in kilodaltons (kDa). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

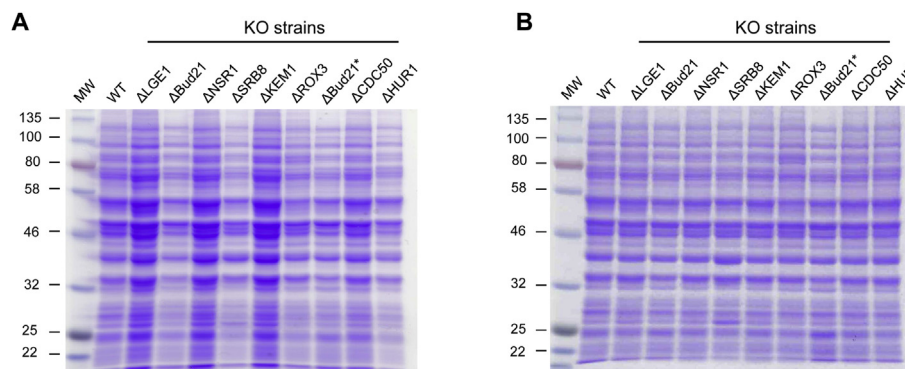


Fig. 2. Uniform loading of yeast total proteins. (A) *S. cerevisiae* total proteins were extracted from BY4741 parental (wild-type, WT) and single-gene knockout strains (KO strains). The names of the deleted genes are shown above the gel. An equal volume (10 µl) of the samples was loaded. ΔBud21 and ΔBud21* are from two different sources. (B) Aliquots of the protein samples were quantitated using the method described in this article, and equal amounts (8.2 µg) of total proteins were loaded in each lane for uniform loading.

sample buffer. In addition, solubilization of precipitated proteins by urea and guanidine makes samples compatible with Bradford assay, and SDS-compatible assays are not required. Thus, quantitation of total proteins in SDS sample buffer can be done quickly and economically using simple acetone precipitation.

Acknowledgments

This work is funded by Natural Sciences and Engineering Research Council of Canada (NSERC).

References

- [1] V.V. Kushnirov, Rapid and reliable protein extraction from yeast, *Yeast* 16 (2000) 857–860.
- [2] J.E. Noble, M.J.A. Bailey, Quantitation of proteins, *Methods Enzymol.* 463 (2009) 73–95.
- [3] J. Karlsson, K. Ostwald, C. Kabjorn, M. Anderson, A method for protein assay in Laemmli buffer, *Anal. Biochem.* 219 (1994) 144–146.
- [4] M. Puchades, A. Westman, K. Blennow, P. Davidsson, Removal of sodium dodecyl sulfate from protein samples prior to matrix-assisted laser desorption/ionization mass spectrometry, *Rapid Commun. Mass Spectrom.* 13 (1999) 344–349.
- [5] D. Barritault, A. Expert-Bezançon, M.F. Guérin, D. Hayes, The use of acetone precipitation in the isolation of ribosomal proteins, *Eur. J. Biochem.* 63 (1976) 131–135.
- [6] V. Thongboonkerd, K.R. McLeish, J.M. Arthur, J.B. Klein, Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation, *Kidney Int.* 62 (2002) 1461–1469.
- [7] O.P. Srivastava, K. Srivastava, Purification of γ-crystallin from human lenses by acetone precipitation method, *Curr. Eye Res.* 17 (1998) 1074–1081.
- [8] J. von Hagen, *Proteomics Sample Preparation*, John Wiley, Hoboken, NJ, 2011.
- [9] A.M.J. Crowell, M.J. Wall, A.A. Doucette, Maximum recovery of water soluble proteins through acetone precipitation, *Anal. Chim. Acta* 796 (2013) 48–54.
- [10] D.I. Jacobs, M. van Rijssen, R. van der Heijden, R. Verpoorte, Sequential solubilization of proteins precipitated with trichloroacetic acid in acetone from cultured *Catharanthus roseus* cells yields 52% more spots after two dimensional gel electrophoresis, *Proteomics* 1 (2001) 1345–1350.
- [11] D. Botelho, M.J. Wall, D.B. Vieira, S. Fitzsimmons, F. Liu, A. Doucette, Top-down and bottom-up proteomics of SDS-containing solutions following mass-based separation, *J. Proteome Res.* 9 (2010) 2863–2870.

A Unique N-Terminal Sequence in the *Carnation Italian Ringspot Virus* p36 Replicase-Associated Protein Interacts with the Host Cell ESCRT-I Component Vps23

Lynn G. L. Richardson,^{a*} Eric A. Clendening,^a Hyukho Sheen,^b Satinder K. Gidda,^a K. Andrew White,^b Robert T. Mullen^a

Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada^a; Department of Biology, York University, Toronto, Ontario, Canada^b

ABSTRACT

Like most positive-strand RNA viruses, infection by plant tombusviruses results in extensive rearrangement of specific host cell organelle membranes that serve as the sites of viral replication. The tombusvirus *Tomato bushy stunt virus* (TBSV) replicates within spherules derived from the peroxisomal boundary membrane, a process that involves the coordinated action of various viral and cellular factors, including constituents of the endosomal sorting complex required for transport (ESCRT). ESCRT is comprised of a series of protein subcomplexes (i.e., ESCRT-0 -I, -II, and -III) that normally participate in late endosome biogenesis and some of which are also hijacked by certain enveloped retroviruses (e.g., HIV) for viral budding from the plasma membrane. Here we show that the replication of *Carnation Italian ringspot virus* (CIRV), a tombusvirus that replicates at mitochondrial membranes also relies on ESCRT. In plant cells, CIRV recruits the ESCRT-I protein, Vps23, to mitochondria through an interaction that involves a unique region in the N terminus of the p36 replicase-associated protein that is not conserved in TBSV or other peroxisome-targeted tombusviruses. The interaction between p36 and Vps23 also involves the Vps23 C-terminal steadiness box domain and not its N-terminal ubiquitin E2 variant domain, which in the case of TBSV (and enveloped retroviruses) mediates the interaction with ESCRT. Overall, these results provide evidence that CIRV uses a unique N-terminal sequence for the recruitment of Vps23 that is distinct from those used by TBSV and certain mammalian viruses for ESCRT recruitment. Characterization of this novel interaction with Vps23 contributes to our understanding of how CIRV may have evolved to exploit key differences in the plant ESCRT machinery.

IMPORTANCE

Positive-strand RNA viruses replicate their genomes in association with specific host cell membranes. To accomplish this, cellular components responsible for membrane biogenesis and modeling are appropriated by viral proteins and redirected to assemble membrane-bound viral replicase complexes. The diverse pathways leading to the formation of these replication structures are poorly understood. We have determined that the cellular ESCRT system that is normally responsible for mediating late endosome biogenesis is also involved in the replication of the tombusvirus *Carnation Italian ringspot virus* (CIRV) at mitochondria. Notably, CIRV recruits ESCRT to the mitochondrial outer membrane via an interaction between a unique motif in the viral protein p36 and the ESCRT component Vps23. Our findings provide new insights into tombusvirus replication and the virus-induced remodeling of plant intracellular membranes, as well as normal ESCRT assembly in plants.

Tombusviruses are positive-strand RNA [(+)RNA] viruses that infect a wide range of plant species and replicate at host cell membranes derived specifically from either peroxisomes (e.g., *Tomato bushy stunt virus* [TBSV]) or mitochondria (e.g., *Carnation Italian ringspot virus* [CIRV]) (1). Upon infection and depending on the tombusvirus, the peroxisomal or mitochondrial (outer) membranes progressively proliferate and invaginate, resulting in the formation of hundreds of spherules that serve to concentrate viral and host cell factors required for synthesis of the viral RNA genome and to protect nascent viral RNAs from degradation by host cell defenses (2, 3). Concomitant with these morphological changes, the modified organelles also form large appendages and coalesce, yielding aggregated structures that no longer resemble the organelles from which they were derived (1, 4).

The morphological transformation of peroxisomes or mitochondria in tombusvirus-infected cells involves two viral replication proteins: an auxiliary viral RNA-binding protein and an RNA-dependent RNA polymerase, referred to as p33 and p92, respectively, in TBSV, or p36 and p95, respectively, in CIRV (5).

Both sets of replicase proteins are essential for viral genome replication (6, 7) and are encoded by overlapping open reading frames (ORFs), and p92 and p95 are products of translational read-through of an amber stop codon in p33 and p36, respectively (8, 9). Consequently, the N-terminal portion of p92/p95 is identical to p33/p36. Both sets of replicase proteins are also integral membrane proteins, each possessing two transmembrane domains (TMDs), as well as unique targeting signals that mediate

Received 27 December 2013 Accepted 18 March 2014

Published ahead of print 26 March 2014

Editor: A. Simon

Address correspondence to Robert T. Mullen, rtmullen@uoguelph.ca.

* Present address: Lynn G. L. Richardson, Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, Massachusetts, USA.

Copyright © 2014, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.03840-13

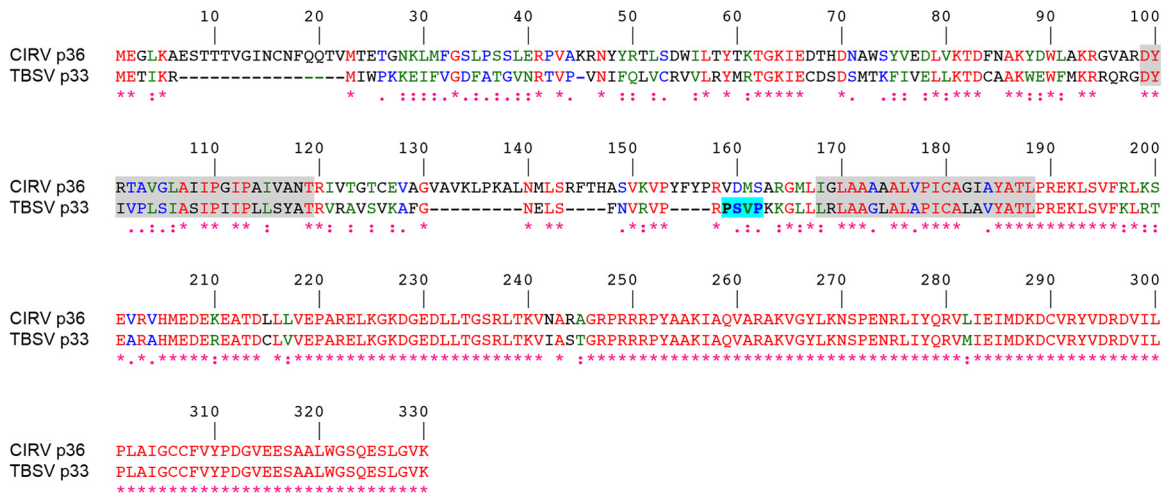


FIG 1 Alignment of the deduced amino acid sequences of CIRV p36 and TBSV p33. Sequences were obtained from GenBank (accession numbers CAA59477.2 and NP_062898.1) and aligned using ClustalW. Identical and similar amino acids in each protein are colored red and green or blue, respectively, and indicated also with asterisks and colons or periods, respectively. The numbers represent specific amino acid residues in full-length p36 (330 residues). Putative TMDs (shown on a gray background) were determined using TOPCONS and visual inspection, and the late-domain-like motif identified in the intervening loop sequence of p33 (23) but absent in p36 is shown on a light blue background.

their specific sorting to either peroxisomes or mitochondria (4, 10, 11) and thus dictate the intracellular site for viral replication.

Numerous host cell factors involved in tombusvirus replication have been identified as part of several large-scale genomic and proteomic studies performed with TBSV and *Saccharomyces cerevisiae* as a model host (12). Among these factors are several components of endosomal sorting complex required for transport (ESCRT). ESCRT is a network of ~20 soluble proteins that, in noninfected cells, are sequentially recruited from the cytosol and assembled into several multiprotein subcomplexes (ESCRT-0, -I, -II, and -III) at the late endosomal surface, where they participate in sorting of ubiquitinated membrane-bound cargo proteins into intraluminal vesicles derived from the endosomal boundary membrane during multivesicular body (MVB) biogenesis. According to models based primarily on studies with yeasts and mammalian cells (13), ESCRT assembly begins with ESCRT-0 recognizing ubiquitinated cargo in the endosomal membrane and recruiting ESCRT-I to the endosomal surface. ESCRT-I also participates in ubiquitinated cargo sorting and recruits ESCRT-II, which subsequently recruits ESCRT-III, which forms polymeric filaments that drive membrane vesiculation. Thereafter, the AAA-ATPase Vps4 (vacuolar protein sorting 4) catalyzes the disassembly of ESCRT-III, a process that is coupled with membrane fission and results in the recycling of ESCRT(III) subunits to the cytosol, while the mature MVB fuses with the lysosome/vacuole, where its contents are degraded. Interestingly, while the ESCRT machinery and its interaction network are relatively well conserved and have similar cellular functions in plants compared to yeasts and mammals (14, 15), some key differences exist, which are not well understood. For example, homologs of mammalian and yeast ESCRT-0 do not exist in plants (16), and instead, recent work has identified the TOL family of proteins to be involved in early recognition of ubiquitinated cargo and their sorting to MVBs (17). However, the relationship between cargo recognition by TOL proteins and subsequent ESCRT recruitment and assembly has yet to be investigated.

In addition to MVB biogenesis, certain ESCRT components

also participate in other important cellular processes involving membrane deformation, including scission of the midbody during cytokinesis or budding of enveloped retroviruses (18, 19). In cells infected with human immunodeficiency virus (HIV), for example, ESCRT is initially redirected or “hijacked” to the plasma membrane via interactions between peptide late-domain motifs in the HIV structural protein, Gag p6, and the ESCRT-I protein Tsg101 and/or the ESCRT accessory protein Alix (20). Among the most common and best characterized of these late-domain motifs is a proline-rich sequence (i.e., P[S/T]AP), which interacts with the ubiquitin E2 variant (UEV) domain of Tsg101 via a binding mechanism that mimics the interaction of certain ESCRT-0 proteins with Tsg101 (21, 22).

Analogous to retroviral budding, the proposed model for the role of ESCRT during TBSV infection is that ESCRT is recruited to the surface of the peroxisome to facilitate invagination of the peroxisomal membrane, as well as to concentrate and assemble the viral replicase complexes within nascent spherules (23). Evidence in support of this model includes the identification of several ESCRT proteins in a genome-wide screen for factors involved in TBSV replication in yeast (24) and the inhibition of TBSV replicase activity in yeast and plant cells upon overexpression of mutant versions of various ESCRT proteins (23). In addition, coexpression of the TBSV replicase protein p33 and the yeast ortholog of mammalian Tsg101, Vps23p, in yeast cells results in the relocalization of Vps23p to peroxisomes (23). TBSV p33 also interacts with yeast Vps23p in a manner that depends on the N-terminal UEV domain of Vps23p and two (mono)ubiquitinated lysine residues in p33, as well as a peptide sequence in p33 (i.e., PSVP) that resembles the PSAP late-domain motif of HIV Gag p6 (25).

Given the ultrastructural similarities of the modified peroxisomal and mitochondrial membranes in TBSV- and CIRV-infected cells, respectively, as well as the ability of chimeric versions of these two viruses to produce the corresponding organelle-specific membrane rearrangements (5, 11, 26), it seems likely that CIRV relies on ESCRT in a manner similar to TBSV. However,

CIRV p36 does not possess a late-domain-like motif resembling that identified in TBSV p33 (Fig. 1) (23) (and in mammalian enveloped retroviruses [20]), implying that CIRV uses a distinct mechanism to recruit ESCRT to mitochondrial membranes. Consistent with this premise, in this study, we show that CIRV p36 binds to and recruits Vps23 to mitochondria in plant cells via a unique N-terminal sequence in p36 that is not present in TBSV p33. We show also that, in contrast to TBSV p33 (and mammalian enveloped retroviruses), the interaction between p36 and Vps23 does not require the UEV domain of Vps23 but, instead, requires the C-terminal steadiness box domain (StBox) of Vps23, which in yeasts and mammals is important for the assembly of ESCRT-I (27–30).

Collectively, our results highlight a unique mechanism for ESCRT recruitment by CIRV that utilizes a Vps23-interacting motif that appears to be distinct from other viruses that co-opt ESCRT, such as TBSV and HIV. This novel mechanism may be reflective of the absence of an ESCRT-0 complex, which raises new mechanistic questions about general features of ESCRT assembly in plants.

MATERIALS AND METHODS

Recombinant DNA procedures and reagents. Molecular biology reagents were purchased either from New England BioLabs, Promega, PerkinElmer Life Sciences Inc., Stratagene, or Invitrogen. Custom oligonucleotides were synthesized by Sigma-Aldrich Ltd. All DNA constructs were verified using automated sequencing performed at the University of Guelph Genomics Facility. PCR-based mutagenesis was carried out using the QuikChange site-directed mutagenesis kit (Stratagene).

Plasmid construction. The majority of the plasmids used in this study have been described previously, including the following binary plasmids used in *Agrobacterium* infiltration and rub inoculation of *Nicotiana benthamiana* leaves: pHST20/TBSV and pHST20/CIRV, encoding the entire TBSV and CIRV genomes, respectively (8, 31); pTRV1 and pTRV2, encoding RNA1 and RNA2 of the bipartite *Tobacco rattle virus* (TRV) genome (32), which were obtained from the *Arabidopsis* Biological Resource Center (ABRC); pMO4-AtSKD1[E232Q], encoding the green fluorescent protein (GFP) fused to the N terminus of a mutant version (i.e., a glutamic acid at position 232 replaced with glutamine) of *Arabidopsis thaliana* Vps4 (also referred to as SKD1 [33]), which was kindly provided by M. Otegui (University of Wisconsin); and pRCS2, serving as an “empty” binary vector control (34). All binary plasmids, as well as all other plant expression plasmids based on pRTL2, pSAT, and pUC18 vectors (see below), contain the cauliflower mosaic virus 35S promoter (CaMV 35S).

Plant expression plasmids used in transient transformations of *Nicotiana tabacum* Bright Yellow-2 (BY-2) cells include the following plasmids that have been described previously (11, 35): pRTL2/p36, encoding CIRV p36; pRTL2/Myc-p36, encoding p36 with an N-terminal Myc epitope tag; pRTL2/p95, encoding CIRV p95, in which the p36 amber stop codon was mutated to a tyrosine codon; pRTL2/Rep, encoding the overlapping CIRV open reading frame 1 (ORF1) and ORF2 that encode both p36 and p95; p36¹⁻⁹⁰-CAT and p36⁹⁰⁻¹⁹⁰-CAT, encoding the N-terminal 90 or 90 to 190 amino acid residues of p36 fused to the N terminus of chloramphenicol acetyltransferase (CAT), respectively; pRTL2/Myc-Vps23, pRTL2/GFP-Vps23, pUC18/Vps23-GFP, and pRTL2/HA-Vps23, encoding either the Myc or hemagglutinin (HA) epitope tag or the green fluorescent protein fused to the N or C terminus of isoform A of *Arabidopsis* Vps23 (referred to as Vps23 in this study); pRTL2/Myc-Vps28 and pRTL2/Myc-Vps25, encoding N-terminal Myc-tagged *Arabidopsis* Vps28 isoform A and *Arabidopsis* Vps25, respectively. Other previously described plant expression vectors include the following plasmids: pRTL2/RFP, encoding the red fluorescent protein (RFP) (36); pUC18/GFP-Syp21 and pUC18/GFP-Syp52, which encode GFP fused to the N termini of the *Arabidopsis* membrane-bound Qa-SNAREs (soluble *N*-ethylmaleimide-sensitive factor at-

tachment protein receptors) Syp21 (Syntaxin of plants 21) and Syp52 (37) proteins (kindly provided by M. Sato [Kyoto University]); pSAT2/Cherry-PTS1, encoding the monomeric cherry fluorescent protein fused to the C-terminal 10 amino acid residues of pumpkin hydroxypyruvate reductase, including its type 1 peroxisomal targeting signal (38); pRTL2/TIC40-RFP, encoding the *Arabidopsis* 40-kDa component of the translocon at the inner membrane of chloroplasts fused to the N terminus of RFP (39); and pSAT4A/AtPAP26-mCherry, encoding the *Arabidopsis* purple acid phosphatase isoform 26 fused to the N terminus of the monomeric cherry fluorescent protein (40). Yeast two-hybrid expression vectors, pGADT7 and pGBKT7 (Clontech) encoding *Arabidopsis* Vps23, Vps25, Vps37, and Vps28 were also previously described (35).

Plasmids used for bimolecular fluorescence complementation (BiFC) assays were based on pSAT4/nEYFP and pSAT4/cEYFP, which encode the N-terminal and C-terminal halves of the enhanced yellow fluorescent protein (nEYFP and cEYFP), respectively (kindly provided by S. Gelvin [Purdue University]) (41). Briefly, pSAT4/p36-cYFP, p36⁹¹⁻³³⁰-cYFP, and pSAT4/nYFP-Vps23 were constructed by PCR amplifying the p36 or Vps23 open reading frames (ORFs), or portions thereof, and cloned into pSAT4/nEYFP or pSAT4/cEYFP. Complete details on the oligonucleotide primers used for generating these and other plasmids are available upon request.

Plant expression vectors encoding modified (truncated) versions of p36 (i.e., p36⁹¹⁻³³⁰, p36⁶⁸⁻³³⁰, and p36²³⁻³³⁰), Vps23 (GFP-UEV, Myc-CC+StBox [CC stands for coiled-coil, and StBox stands for steadiness box], Myc-CC, Myc-StBox, and Myc-UEV+CC) were generated by PCR amplification of the sequence corresponding to the indicated proteins or protein domains and ligated into either pRTL2/MCS (containing a multiple cloning site [MCS]) (42), pRTL2/Myc-MCS (encoding an initiation methionine, followed by the Myc epitope tag and then an MCS), pRTL2/GFP-MCS (encoding GFP followed by an MCS) (14), or pUC18/NheI-GFP (encoding GFP with a 5′ unique NheI restriction site) (43). Similarly, yeast two-hybrid expression vectors encoding p36 or p36¹⁻⁹⁰ or modified (truncated) versions of Vps23 (UEV, CC+StBox, CC, StBox, and UEV+CC) were generated by PCR amplification and subcloned into pGBKT7 (encoding the GAL4 DNA-binding domain [BD] and a Myc epitope tag, followed by an MCS) or pGADT7 (encoding the GAL4 activation domain [AD] and an HA epitope tag, followed by an MCS), respectively.

For construction of pRTL2/Myc-NtVps23, the 3′ sequence of VPS23 from *Nicotiana tabacum* was obtained by 3′ rapid amplification of cDNA ends (RACE)-PCR using cDNA synthesized from *N. tabacum* suspension cell mRNA, and a gene-specific forward primer based on the available *N. tabacum* VPS23 expressed sequence tag (EST) (NCBI accession number EB680173). Subsequently, full-length NtVPS23 was amplified from BY-2 cDNA using gene-specific primers and ligated into pRTL2/Myc-MCS (via pCR2.1TOPO serving as a shuttle vector), yielding pRTL2/Myc-NtVps23. pRTL2/Myc-NtVps23 was then used as the template DNA for constructing pRTL2-Myc/NtUEV, which includes the entire UEV domain of *N. tabacum* Vps23.

pRTL2/p36Δ7-22 was constructed using PCR-based site-directed mutagenesis, whereby sequences encoding amino acid residues 7 to 22 in the p36 ORF were deleted using the appropriate forward and reverse mutagenic primers and pRTL2/p36 as the template DNA. pRTL2/p33¹⁻⁷⁴-p36⁹¹⁻³³⁰ and pRTL2/p36¹⁻²⁸-p33¹³⁻⁷⁴-p36⁹¹⁻³³⁰ were constructed in the following manner. First, sequences encoding the N-terminal 74 amino acid residues of p33 were amplified from pRTL2/p33 (4), and the resulting PCR products were ligated into pRTL2/p36⁹¹⁻³³⁰, yielding pRTL2/p33¹⁻⁷⁴-p36⁹¹⁻³³⁰. Next, sequences encoding the N-terminal 12 amino acid residues of p33 were replaced (via successive site-directed mutagenesis reactions and pRTL2/p33¹⁻⁷⁴-p36⁹¹⁻³³⁰ as the initial template DNA) with sequences encoding the N-terminal 28 residues of p36, yielding pRTL2/p36¹⁻²⁸-p33¹³⁻⁷⁴-p36⁹¹⁻³³⁰. pRTL2/Myc-p36¹⁻⁹⁰-TraB and pRTL2/Myc-p36¹⁻²⁸-TraB, encoding an N-terminal Myc epitope tag fused to the N-terminal 90 or 28 amino acids of p36 followed by the C-terminal 34 or 75 amino acid residues of *Arabidopsis* TraB (At1g05270),

respectively, were constructed by first introducing (using PCR-based site-directed mutagenesis and pRTL2/Myc-TraB as the template DNA) an NheI site immediately upstream of either codon 338 or codon 297 (of 371) in the TraB ORF. The modified plasmids were then digested with NheI (removing sequences coding for residues 1 to 337 or 1 to 296 of TraB and the N-terminal Myc epitope tag) followed by ligation with NheI-digested PCR products encoding the N-terminal 90 or 28 residues of p36, along with an appended Myc epitope tag from pRTL2/Myc-p36 (see above). pRTL2/Myc-TraB was constructed by ligating the full-length TraB ORF (cDNA provided by the ABRC) into pRTL2/Myc-MCS.

Agrobacterium infiltration and rub inoculation of *N. benthamiana* and RNA gel blot analysis. *N. benthamiana* plants were grown in chambers at 21°C with a 16-h/8-h light/dark cycle. The leaves of approximately 3-week-old plants were infiltrated or, for experiments involving TRV, triple infiltrated with cultures of *Agrobacterium tumefaciens* (LBA4404) carrying the appropriate binary vectors. Procedures involving *Agrobacterium* have been described previously (4). Rub inoculations were performed 2 days after infiltration with 10 µg of plasmid DNA encoding full-length infectious TBSV or CIRV cDNA diluted in 30 µl of RNA inoculation buffer (11, 44). Approximately 4 to 6 days after inoculation or, for experiments involving TRV, 2 or 4 days after infiltration, the leaves were flash frozen in liquid nitrogen followed by total RNA extraction (45). Aliquots of isolated RNA were separated in nondenaturing 1.2% (wt/vol) agarose gels, and viral RNAs were detected by electrophoretic transfer to nylon (Hybond-N; Amersham Biosciences) followed by incubation with ³²P-end-labeled oligonucleotide probes complementary to the CIRV and TBSV genome or TRV RNA1 genome (45, 46). Complete details of oligonucleotides used for RNA gel blot analysis are available upon request. Labeled RNAs were visualized using a phosphorimager. Results presented are representative of at least two separate experiments.

Biolistic bombardment and fluorescence microscopy of BY-2 cells. *N. tabacum* Bright Yellow-2 suspension cell cultures were maintained and prepared for biolistic bombardment with a Biolistic PDS-1000/He particle delivery system (Bio-Rad Laboratories) as described previously (47). Bombarded cells were incubated for ~4 to 8 h to allow for expression and sorting of the introduced gene product(s). Cells were fixed in 4% (wt/vol) formaldehyde, followed by permeabilization with 0.01% (wt/vol) pectolyase Y-23 (Kyowa Chemical Products) and either 0.3% (vol/vol) Triton X-100, which permeabilizes the plasma membrane and all organellar membranes, or 25 µg/ml digitonin, which permeabilizes only the plasma membrane. Details on the differential detergent permeabilization of BY-2 cells have been previously described (48).

Primary and dye-conjugated secondary antibodies used for immunofluorescence staining of cells and their sources were as follows: rabbit anti-Myc IgGs and mouse anti-hemagglutinin (anti-HA) IgGs (Bethyl Laboratories); mouse anti-Myc antibodies in hybridoma medium (Princeton University, Monoclonal Antibody Facility); mouse anti- α -tubulin (Sigma-Aldrich Ltd.); mouse anti-CAT antibodies in hybridoma medium (provided by S. Subramani); mouse anti-maize β -ATPase antibodies in hybridoma medium (49); rabbit anti-cytochrome *c* oxidase subunit II (CoxII) IgGs (50); rabbit anti-p36 IgGs raised against a synthetic peptide corresponding to an amino acid sequence either in the C terminus of p36 (residues 218 to 237 [4] or the intervening loop [residues 147 to 160]) of p36 (Cedarlane Laboratories); goat anti-mouse and goat anti-rabbit Alexa Fluor 488 IgGs (Molecular Probes); and goat anti-rabbit rhodamine red-X IgGs (Jackson ImmunoResearch Laboratories).

Microscopic images of cells were acquired using an AxioScope 2 MOT epifluorescence microscope (Carl Zeiss Inc.) or a Leica DM RBE microscope. Figure compositions and merged images were generated using Openlab (Improvision) or Northern Eclipse (Empix Imaging Inc.) software and Adobe Photoshop CS (Adobe Systems). Images presented in all figures are representative of the results obtained from analyzing ≥ 25 independently transformed cells from at least three separate experiments. Colocalization of proteins was quantified using the ImageJ plugin “Co-localization Finder” and methods as described previously (51). Pear-

son’s correlation coefficient *r* values of -1.0 to 1.0 are considered to be equivalent to all of the pixels from the regions of interest within the individual red and green channels of the images being 100% noncolocalized to 100% colocalized, respectively.

BiFC assays. Bimolecular fluorescence complementation assays were performed as previously described (11). BY-2 cells were transformed via biolistic bombardment with plasmid DNA encoding RFP serving as a transformation control and an internal reference for assessing any cell-to-cell variability in RFP/YFP fluorescence values due to differences in protein expression, together with nYFP-Vps23, and p36-cYFP or p36⁹¹⁻³³⁰-cYFP. Transformed cells were visualized (via epifluorescence microscopy) based on RFP fluorescence, and both RFP and reconstituted YFP fluorescence intensities were collected with identical image acquisition settings (e.g., gain, offset, and exposure). Acquisition settings, amounts of plasmid DNA bombarded, and postbombardment cell incubation times employed in BiFC assays were chosen based on preliminary optimization experiments aimed at minimizing the possibility of nonspecific interactions. Likewise, p36⁹¹⁻³³⁰-cYFP, rather than empty cYFP vector, was chosen as a potential negative control based on guidelines for assessing membrane-bound protein interactions using the BiFC assay (52). The mean intensity of RFP and YFP fluorescence in transformed cells was calculated by defining the boundary of each cell followed by quantification of the mean pixel intensity using ImageJ software. The raw data for at least 25 cells were then expressed as a mean YFP-to-RFP ratio and a Student two-tailed *t* test assuming unequal variance between samples was used to determine statistical significance. Results shown are representative of three independent experiments.

Yeast two-hybrid analysis. Yeast two-hybrid assays were carried out as described previously (53) with some modifications (35). Yeast cells (PJ69-4A) containing pGADT7 (activation domain fusions) and pGBKT7 (DNA-binding domain fusions) plasmids were cultured in synthetic dextrose medium (2% [wt/vol] dextrose, 0.67% [wt/vol] yeast nitrogen base without amino acids, 2 g/liter synthetic mix of amino acid supplements [SD-Leu,Trp; Bufferad]), diluted in a 1:5 dilution series, and then replica plated on agar plates containing SD-Leu,Trp or SD-Leu,Trp,His,Ade. Results of growth assays presented in figures are representative of the results obtained from analyzing three isolated yeast colonies from at least two separate cotransformations. In addition, all fusion proteins described in this study were confirmed to be properly expressed based on Western blot analysis of protein lysates obtained from yeast (co)transformed with two-hybrid plasmids, as described elsewhere (35).

In vitro coimmunoprecipitations. Myc-tagged versions of p36¹⁻⁹⁰, p36, and Vps28 were synthesized *in vitro* using the TNT T7 coupled reticulocyte lysate system (Promega), with the corresponding pGBKT7-based plasmids serving as the template DNA. S-epitope-tagged Vps23 (S-Vps23) was expressed in *Escherichia coli* BL21 Codon Plus (Stratagene). Cultures were grown to an optical density at 600 nm (OD₆₀₀) of ~0.6 to 0.8, and protein expression was induced with 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) (Sigma-Aldrich Ltd.) at 30°C for 3 h. For experiments involving “mock” lysate, *E. coli* cells were transformed with empty vector (pET29a). Total soluble proteins were isolated in extraction buffer using a French press, and lysate was cleared by centrifugation as described elsewhere (54). Total soluble protein from cleared S-Vps23-containing or mock lysates were then separated using SDS-PAGE and stained with Coomassie blue R250 to ensure approximately equal total protein. Coimmunoprecipitations were carried out as described elsewhere (54), and proteins were resolved by SDS-PAGE, subjected to Western blotting using rabbit anti-Myc or mouse anti-S-tag (Novagen) antibodies, and detected using chemiluminescence.

RESULTS

CIRV replication in *N. benthamiana* is disrupted by a mutant of Vps4. To begin to test whether ESCRT plays a role in CIRV replication, a dominant-negative version of the *Arabidopsis* ESCRT protein Vps4 (Vps4^{E232Q}) was expressed in CIRV-infected *N. ben-*

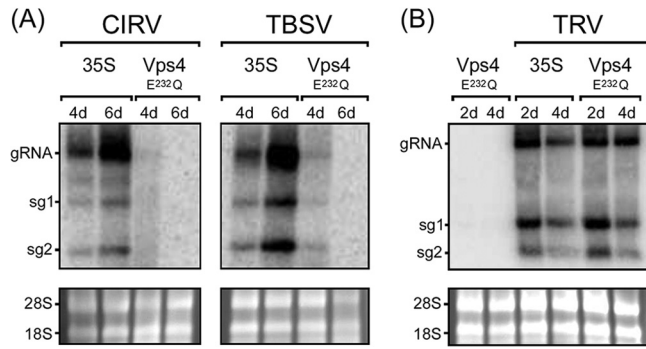


FIG 2 Inhibition of CIRV and TBSV replication in *N. benthamiana* by Vps4^{E232Q}. (A) RNA blot analysis of *N. benthamiana* leaves infiltrated with *Agrobacterium* harboring either an empty vector (35S) or a vector encoding *Arabidopsis* Vps4^{E232Q}, followed by rub inoculation with full-length infectious CIRV or TBSV cDNA. The relative positions of the CIRV and TBSV genomic (gRNA) and the two subgenomic mRNAs (sg1 and sg2) are shown to the left of the gel. The numbers of days after infiltration or inoculation are indicated above the lanes. (B) RNA blot analysis of *N. benthamiana* leaves agro-infiltrated with either Vps4^{E232Q} alone, empty vector (35S), and two other vectors (TRV1 and TRV2) encoding the full-length infectious TRV genome (TRV), or Vps4^{E232Q}, TRV1, and TRV2. In panels A and B, the 28S and 18S rRNAs in the corresponding ethidium bromide-stained agarose gels are shown as a loading control.

thamiana leaves. Vps4 is an AAA-ATPase, and Vps4^{E232Q} contains a Glu-to-Gln mutation in the ATPase domain that blocks ATP hydrolysis and causes defects in endosomal protein sorting (33, 55), presumably by preventing disassembly of the endogenous ESCRT-III machinery at the late endosomal surface, as it does in yeast (56). Leaves of *N. benthamiana*, which is a host of tobamoviruses, including CIRV (57), were infiltrated with *Agrobacterium* harboring a plasmid expressing Vps4^{E232Q} (33) (agro-infiltrated). The same leaves were then rub inoculated 2 days later with a CaMV 35S promoter-containing plasmid encoding the full-length infectious CIRV cDNA (11), and 4 or 6 days thereafter, viral genomic and subgenomic RNAs were analyzed by Northern blotting.

As shown in Fig. 2A, compared to CIRV-infected *N. benthamiana* leaves that were agro-infiltrated with an empty vector (35S), expression of Vps4^{E232Q} dramatically reduced the amounts of CIRV genomic RNA and both subgenomic mRNAs (sg1 and sg2 mRNA). Consistent with the results reported previously on the inhibition of TBSV replication in *N. benthamiana* by overexpression of another dominant-negative, ATPase-deficient version of *Arabidopsis* Vps4, i.e., Vps4^{K178A} (23), we found that overexpression of Vps4^{E232Q} also inhibited the accumulation of TBSV genomic and subgenomic mRNAs (Fig. 2A). Overexpression of Vps4^{E232Q} did not, however, inhibit the replication of the *Tobacco rattle virus* (TRV) in *N. benthamiana* leaves (Fig. 2B), which is a (+)RNA tobamovirus (46) and does not rely on ESCRT for its replication (23). This confirms that the inhibition of CIRV (and TBSV) replication by Vps4^{E232Q} is not simply a consequence of an indirect, inhibitory effect(s) due to the overexpression of this ESCRT mutant in plant cells. As expected, no viral RNAs were detectable in mock-infected *N. benthamiana* leaves expressing Vps4^{E232Q} (results are presented for TRV only [Fig. 2B]).

CIRV p36 recruits Vps23 to mitochondria in plant cells. Given the well-documented role of Tsg101 (the mammalian homolog of Vps23) in HIV budding from the plasma membrane in

mammalian cells (21, 22) and that yeast Vps23p is redirected to peroxisomes in yeast cells coexpressing TBSV p33 (23), we tested whether, in an analogous manner, CIRV causes the relocalization of Vps23 to mitochondria in plant cells. Toward this end, *N. tabacum* BY-2 suspension-cultured cells—which serve as a well-characterized system for studying protein targeting in plant cells (57), including viral proteins (58, 59)—were transiently (co)transformed (via biolistic bombardment) with plasmids encoding full-length CIRV and/or an N-terminal Myc epitope-tagged version of *Arabidopsis* Vps23 (Myc-Vps23) and then processed for immunofluorescence microscopy.

As shown in Fig. 3A and consistent with previously published results (11), both the p36 and p95 replicase proteins in CIRV-transformed BY-2 cells localized to endogenous β -ATPase-containing mitochondria, which as a consequence of the expression of the viral proteins, were conspicuously altered (i.e., aggregated) in terms of their intracellular distribution. For comparison purposes, refer to the normal appearance of the β -ATPase-containing mitochondria that are distributed throughout the cytosol in a nontransformed BY-2 cell (Fig. 3A). Also consistent with previously published results (35), Myc-Vps23 localized in BY-2 cells to the cytosol and late endosomes, shown by its partial colocalization with GFP-Syp21 (Fig. 3A), which is a well-known late endosomal marker protein (38). In contrast, coexpression of CIRV and Myc-Vps23 in the same cell resulted in relocalization of Myc-Vps23 to CIRV-induced aggregated mitochondria (Fig. 3B), which was confirmed as such based on immunostaining of both Myc-Vps23 and endogenous mitochondrial CoxII in the same batch of CIRV- and Myc-Vps23-cotransformed cells (Fig. 3B). These results support the premise that, similar to HIV and TBSV, CIRV causes a change in the intracellular localization of Vps23.

We determined next whether the replicase proteins p36 and/or p95 expressed outside of the context of full-length CIRV were capable of recruiting Vps23 to mitochondria in plant cells. As shown in Fig. 3C, coexpression of Myc-Vps23 with either or both replicase proteins resulted in the relocalization of Myc-Vps23 to aggregated mitochondria (Fig. 3C), similar to when Myc-Vps23 was coexpressed with full-length CIRV (Fig. 3B). Colocalization between p36 and Myc-Vps23 is also shown in Fig. 3D using confocal microscopy, and quantification of colocalization in images obtained from medial (midcell) optical sections using the mean Pearson's correlation coefficient, r , confirmed a high degree of fluorescence signal overlap for both sets of proteins (see the legend to Fig. 3D for r values). In contrast, confocal microscopy and quantification of colocalization confirmed that Myc-Vps23 was not associated with mitochondria in BY-2 cells in the absence of coexpressed p36, as expected (see Fig. 3D and the legend for r value).

We also performed a number of additional control experiments to confirm that the recruitment of Vps23 to mitochondria by p36 was not dependent on the appended peptide sequence or fluorescent reporter protein (Fig. 3E) or the heterologous expression of *Arabidopsis* Vps23 in tobacco cells (Fig. 3F) and that it was specific for Vps23, since at least two other ESCRT proteins, Vps28 (ESCRT-I) and Vps25 (ESCRT-II), were not recruited by p36 (Fig. 3G). Moreover, we demonstrated that the appearance (i.e., distribution) of organelles other than mitochondria was unaffected in p36 (co)expressing cells (Fig. 3H), indicating that colocalization of Vps23 and p36 at mitochondria is not simply due to a general aggregation of organelles.

Taken together, the results presented in Fig. 3 indicate that

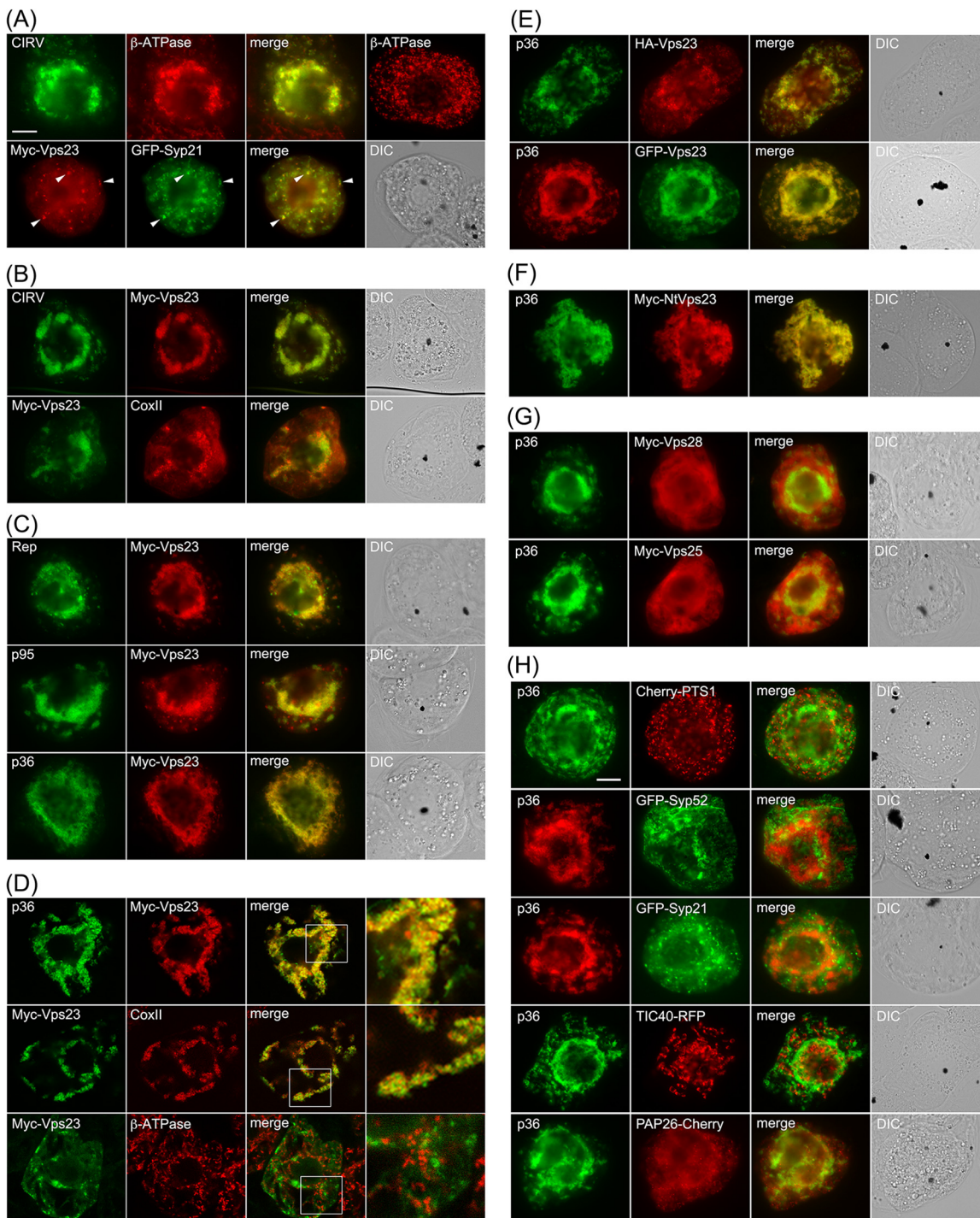


FIG 3 Relocalization of Vps23 from the cytosol and late endosomes to mitochondria in BY-2 cells coexpressing CIRV p95 and/or p36. (A to C) Representative epifluorescence micrographs of BY-2 cells (co)transformed (as indicated by panel labels) with either full-length CIRV, Rep (both p95 and p36 together), p95, p36, Myc-Vps23, or GFP-Syp21, or mock transformed with empty plasmid DNA and then immunostained for the endogenous mitochondrial protein β -ATPase (top right-hand photo in panel A). All cells shown in Fig. 3 were formaldehyde fixed and processed for immunofluorescence microscopy as described in Materials and Methods. Note that p95 and p36 were immunodetected using primary antibodies raised against a synthetic peptide that corresponds to an amino acid sequence in both proteins. p95 is produced by the translational read-through of the p36 amber stop codon (8). Selected transformed cells were also immunostained (as indicated) for endogenous mitochondrial β -ATPase or CoxII. Also shown are the corresponding merge and differential interference contrast (DIC) images. The yellow color in the merged images indicates colocalization, and the white arrowheads in panel A indicate obvious examples of Myc-Vp23 and GFP-Syp21 colocalization at late endosomes. Bar = 10 μ m. (D) Representative confocal micrographs of BY-2 cells cotransformed with either p36 and Myc-Vps23 (top and middle rows), or GFP-Syp21 and Myc-Vps23 (bottom row), which were immunostained for Myc-Vps23 and endogenous CoxII (GFP fluorescence is not shown). All images are medial (midcell) optical sections of cells, and boxes represent the portions of the cells shown at higher magnification in the panels on the right. The Pearson's correlation coefficient r values were as follows: $r = 0.81$ for p36 and Myc-Vps23, $r = 0.71$ for Myc-Vps23 and CoxII, and $r = 0.21$ for Myc-Vps23 and CoxII. (E to H) Representative epifluorescence micrographs of BY-2 cells cotransformed (as indicated) with either p36 and either N-terminal HA epitope- or GFP-tagged Vps23 (E); p36 and a Myc-tagged *N. tabacum* Vps23 (F), p36 and Myc-tagged Vps28 or Vps25 (G), or p36 and various organelle marker fusion proteins, including Cherry-PTS1 (peroxisome), GFP-Syp52 (early endosome/trans-Golgi network), GFP-Syp21 (late endosome), TIC40-RFP (plastid), PAP26-Cherry (lytic vacuole) (H). Bar = 10 μ m.

CIRV specifically recruits the ESCRT-I component Vps23 to mitochondria in plant cells and that the p36 replicase protein is the minimal viral component necessary for this recruitment.

The cytosol-facing N-terminal region of p36 recruits Vps23 to mitochondria. On the basis of previously published models (11, 26) and as illustrated in Fig. 4A, p36 is an integral outer mitochondrial membrane protein that is orientated with both the N- and C-terminal regions facing the cytosol. p36 also possesses two TMDs and an intervening loop region that contains the mitochondrial targeting signal and is predicted to face the intermembrane space (11). Keeping in mind this predicted topology, we began to explore regions within p36 that could be responsible for recruitment of Vps23 to mitochondria by examining mutant versions of p36 lacking the cytosol-facing N- and/or C-terminal portions of the protein, since these should be available to interact with cytosolic Vps23.

As shown in Fig. 4B, Myc-Vps23 localized to the cytosol and not to mitochondria when coexpressed with p36⁹¹⁻³³⁰, a mutant of p36 in which the protein's cytosol-facing, N-terminal 90 amino acid residues were deleted (see the legend to Fig. 4B for the *r* value of coexpressed Myc-Vps23 and p36⁹¹⁻³³⁰, based on confocal microscopy). In contrast, p36¹⁻¹⁹⁰-CAT, which is a previously characterized p36 mutant wherein the cytosol-facing C-terminal region of the protein (i.e., residues 191 to 330) was replaced with the passenger protein CAT (11), retained the ability to recruit Myc-Vps23 to mitochondria (Fig. 4B). Taken together, these data indicate that the N terminus, but not the C terminus, of p36 is necessary for the recruitment of Vps23 to mitochondria.

Consistent with this conclusion, Myc-Vps23 was not recruited to mitochondria when coexpressed with p36⁹⁰⁻¹⁹⁰-CAT (Fig. 4B), which lacks the N-terminal 90 residues of p36, and also has its cytosol-facing C terminus replaced with CAT (11). These results further indicate that the intervening loop region of p36 alone is not capable of recruiting Vps23 to mitochondria, as expected from the protein's predicted topology (Fig. 4A), wherein the intervening loop faces the mitochondrial intermembrane space and, thus, is inaccessible to interact with Vps23 in the cytosol. Confirmation that the intervening loop in both full-length p36 and p36⁹⁰⁻¹⁹⁰-CAT is oriented toward the mitochondrial intermembrane space was obtained using BY-2 cell differential detergent permeabilization assays (Fig. 4D). For instance, applied antibodies specific for a peptide sequence within the intervening loop of p36 were able to immunodetect the protein (or p36⁹⁰⁻¹⁹⁰-CAT) in cells only when both the plasma membrane and mitochondrial membranes (and all other organelle membranes) were permeabilized with Triton X-100, but not when only the plasma membrane was permeabilized with digitonin. We also confirmed that all of the p36 mutants mentioned above were properly targeted to mitochondria based on their colocalization with an endogenous mitochondrial marker protein (results are presented for p36⁹¹⁻³³⁰ only [Fig. 4E]).

To further demonstrate the importance of the N-terminal region (i.e., residues 1 to 90) of p36 in the recruitment of Vps23 to mitochondria, we replaced the N terminus of a mitochondrial outer membrane protein (TraB) with that of p36 and tested the fusion protein's ability to recruit coexpressed Vps23 to mitochondria. TraB is a "tail"-anchored membrane protein (Fig. 4A) (60), consisting of (i) an N-terminal region that faces the cytosol and represents the majority of the protein, (ii) a single TMD located near its C terminus, and (iii) a short C-terminal tail region orien-

tated toward the intermembrane space that, along with the TMD, constitutes the protein's mitochondrial outer membrane targeting information (11). As shown in Fig. 4C, Myc-p36¹⁻⁹⁰-TraB, consisting of a Myc epitope tag fused to the N-terminal 90 amino acids of p36 and the C-terminal TMD and tail of TraB (i.e., residues 348 to 371), colocalized with coexpressed HA epitope-tagged Vps23 (HA-Vps23; refer also to Fig. 3E) at mitochondria in BY-2 cells; mitochondrial recruitment was confirmed by colocalization of Myc-p36¹⁻⁹⁰-TraB and endogenous CoxII in the same batch of Myc-p36¹⁻⁹⁰-TraB- and HA-Vps23-cotransformed cells (Fig. 4E). In contrast, HA-Vps23 did not colocalize with full-length Myc-TraB at mitochondria (Fig. 4C and E) (refer also to the Fig. 4C legend for the *r* value for coexpressed Myc-TraB and HA-Vps23, based on confocal microscopy), a result that was not due to an aberrant topology of Myc-TraB, which like Myc-p36¹⁻⁹⁰-TraB is orientated with its N terminus facing the cytosol (Fig. 4D).

Overall, the data presented in Fig. 4 indicate that the cytosol-facing N-terminal portion of p36, which represents the first 90 amino acids of the protein, is both necessary and sufficient for recruitment of Vps23 to mitochondria in plant cells.

The N-terminal region of p36 interacts directly with Vps23. We investigated next whether the N-terminal region of p36 interacts directly with Vps23 by employing several different approaches. For example, yeast two-hybrid assays revealed that coexpression of p36¹⁻⁹⁰ and Vps23 resulted in significant yeast growth on high selection media (Fig. 5A), indicating that the two fusion proteins interact. Similarly, coexpression of Vps23 and ESCRT-I protein Vps28 resulted in growth on high selection media, which has been reported previously using the yeast two-hybrid assay (36, 61, 62). As a negative control and consistent with our *in vivo* data (Fig. 3G), no interaction was observed between p36¹⁻⁹⁰ and the ESCRT-II component Vps25. Likewise, coexpression of p36¹⁻⁹⁰, Vps23, Vps28, or Vps25 with the corresponding empty vectors yielded no yeast growth on high selection media (Fig. 5A).

The interaction between p36 and Vps23 was confirmed using *in vitro* coimmunoprecipitation. Briefly, recombinant S-epitope-tagged Vps23 (S-Vps23) from cleared *E. coli* cell lysate was immobilized on S-protein agarose and then incubated with *in vitro*-translated, Myc-tagged versions of either full-length p36 (Myc-p36), the N-terminal 90 amino acids of p36 (Myc-p36¹⁻⁹⁰), or Vps28 (Myc-Vps28); Myc-Vps28 served as a positive control based on its previously published interaction with Vps23 via coimmunoprecipitation (35, 63). For a negative control, the same three *in vitro*-translated proteins were incubated with S-protein agarose preincubated with cleared cell lysate prepared from *E. coli* transformed with empty vector, but containing an approximately equal amount of total protein (Fig. 5B, right panel). As shown also in Fig. 5B, both Myc-p36 and Myc-p36¹⁻⁹⁰, similar to Myc-Vps28, were coimmunoprecipitated with immobilized S-Vps23, but not by the corresponding empty vector negative controls (marked by solid arrowheads in Fig. 5B, left panel), indicating that both full-length p36 and the N terminus of p36 interact directly with Vps23.

We further demonstrated the interaction between p36 and Vps23 using the *in vivo* BiFC assay. In this experiment, p36 or p36 lacking its N terminus (p36⁹¹⁻³³⁰) fused to the C-terminal half of YFP (p36-cYFP and p36⁹¹⁻³³⁰-cYFP, respectively) were each coexpressed with Vps23 fused to the N-terminal half of YFP (nYFP-Vps23) in BY-2 cells. In addition, cells were transformed with cytosolic RFP, which was used to identify transformed cells and

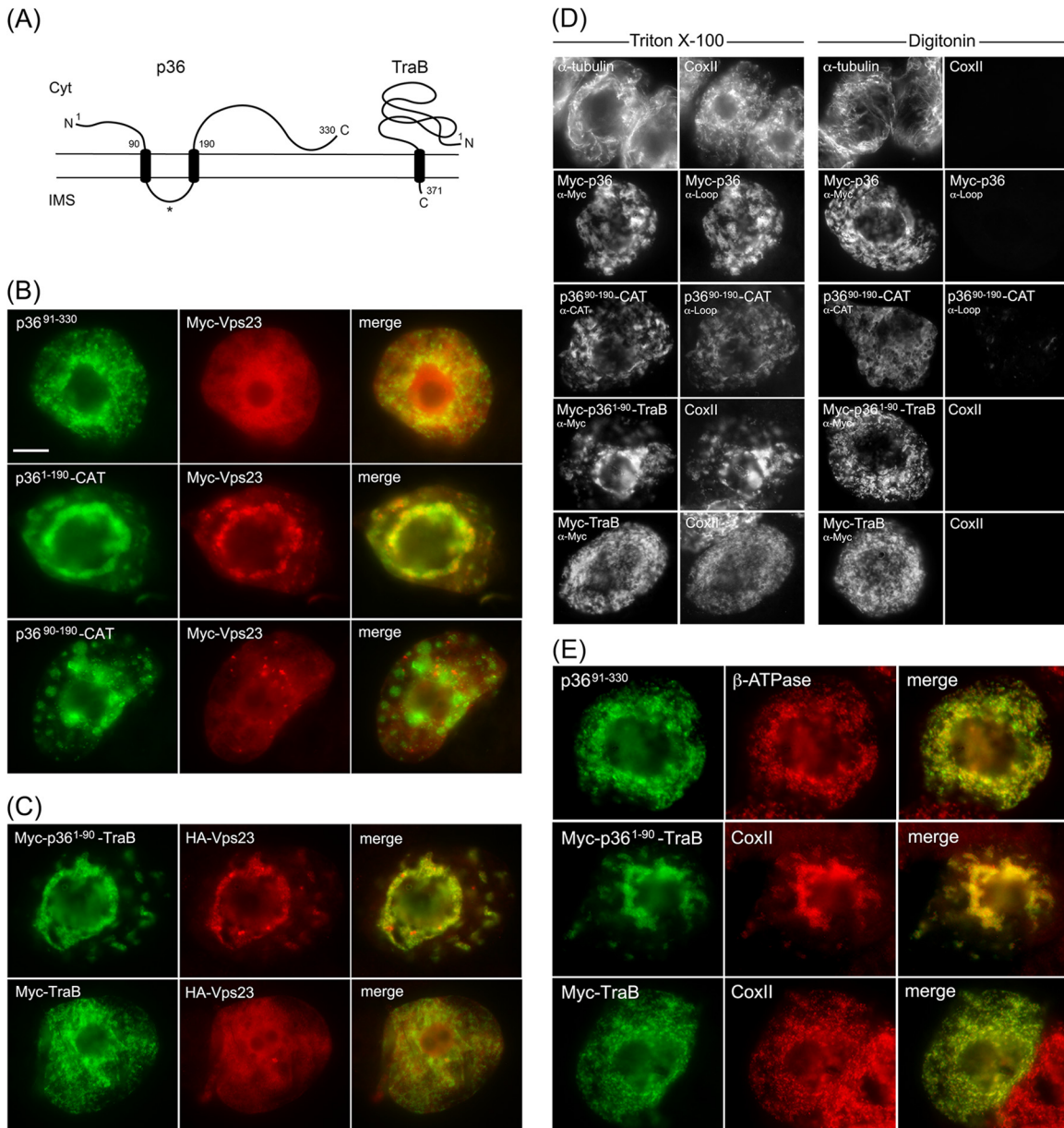


FIG 4 The cytosol-facing N-terminal region of p36 is both necessary and sufficient for relocating Vps23 to mitochondria in BY-2 cells. (A) Topology models of p36 and TraB in the mitochondrial outer membrane based on previously published results (11, 26, 50) and those presented in panel D. The numbers show the numbers of specific amino acid residues in full-length p36 (330 residues) and TraB (371 residues), including those in the names of the modified versions of p36 and p36-CAT (and p36-TraB) shown in panels B to E. The asterisk represents the relative position of the peptide sequence in the intervening loop of p36 used to generate the antiloop antibodies used in panel D. The N and C termini of the proteins are shown. Cyt, cytosol; IMS, intermembrane space. (B and C) Representative epifluorescence micrographs of BY-2 cells cotransformed with proteins as indicated by panel labels. Cells were formaldehyde fixed and processed for immunofluorescence epimicroscopy as described in the legend to Fig. 3. The Pearson's correlation coefficient r values based on confocal microscopy of cells coexpressing p36⁹¹⁻³³⁰ and Myc-Vps23 or Myc-TraB and HA-Vps23 were $r = 0.31$ and $r = 0.25$, respectively. Bar = 10 μm. (D) Representative epifluorescence micrographs of BY-2 cells either nontransformed (top row) or cotransformed with the indicated proteins, fixed with formaldehyde, and then permeabilized with either Triton X-100 (which permeabilizes both the plasma membrane and all organellar membranes) or digitonin (which permeabilizes only the plasma membrane) (16). The cells were then processed for immunofluorescence microscopy using primary antibodies raised against endogenous cytosolic α-tubulin, endogenous mitochondrial matrix protein CoxII (anti-CoxII [α-CoxII]), the Myc epitope tag (anti-Myc [α-Myc]), an amino acid sequence in the intervening loop of p36 (antiloop [α-Loop]), or the bacterial passenger protein CAT (anti-CAT [α-CAT]), as indicated by panel labels. Note that the presence of immunostaining in digitonin-permeabilized cells indicates that endogenous α-tubulin and the expressed protein's appended Myc epitope tag(s) or CAT moiety are exposed to the cytosol. Conversely, endogenous CoxII or the intervening loop sequence of p36 is not immunodetectable in the same corresponding digitonin-permeabilized cells. (E) Representative epifluorescence micrographs of BY-2 cells transformed with p36⁹¹⁻³³⁰, Myc-p36¹⁻⁹⁰-TraB, or Myc-TraB and immunostained for endogenous mitochondrial β-ATPase or CoxII. Note that the mitochondria in the cells expressing p36¹⁻⁹⁰-TraB, but not Myc-TraB, were aggregated, similar to the mitochondrial aggregation observed in cells expressing full-length p36 (Fig. 3).

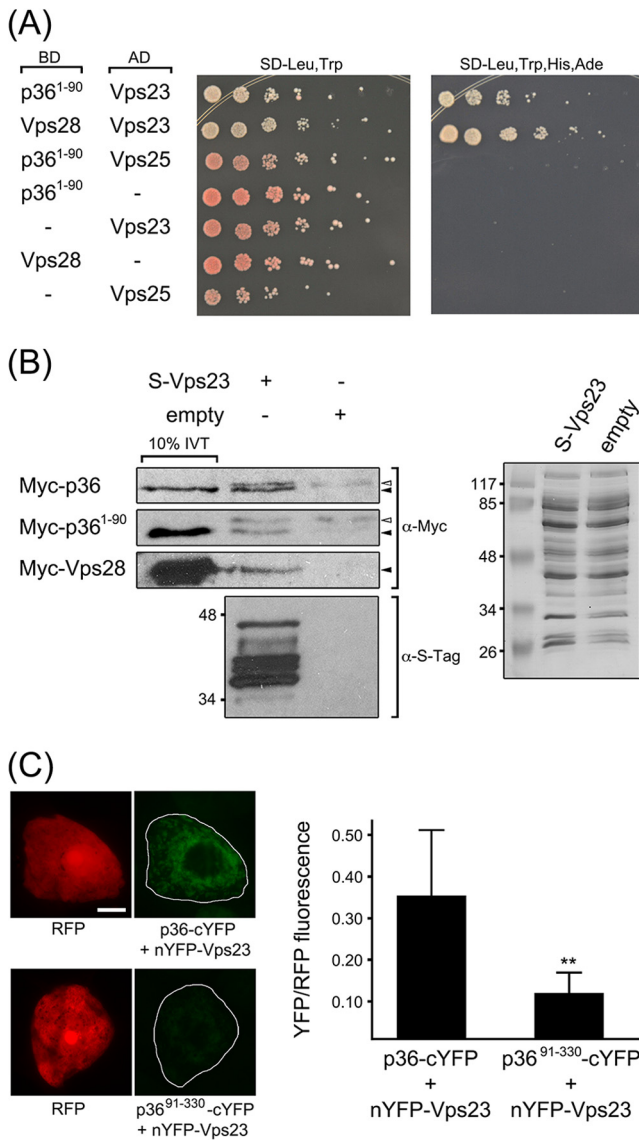


FIG 5 The N-terminal region of p36 interacts with Vps23. (A) p36¹⁻⁹⁰ interacts with Vps23 in the yeast two-hybrid assay. Yeast strains were cotransformed with the indicated pairs of GAL4-binding domain (BD) and GAL4-activating domain (AD) fusion proteins or the corresponding empty BD or AD control plasmids, denoted by a hyphen in the BD or AD column. Serial (1:5) dilutions of cells were spotted onto agar plates containing either low-stringency media (SD-Leu, Trp) or high-stringency media (SD-Leu, Trp, His, Ade), where cell growth is dependent on two-hybrid protein interactions. (B) *In vitro* coimmunoprecipitation of p36 or p36¹⁻⁹⁰ and Vps23. Whole-cell lysates of *E. coli* transformed with S-Vps23 or the corresponding empty vector were incubated with S-protein agarose and *in vitro*-translated (IVT) full-length Myc-p36, Myc-p36¹⁻⁹⁰, or Myc-Vps28, and washed. Eluted proteins (and 10% of the total IVT reaction mixture) were subjected to SDS-PAGE and protein blotting and then probed with either anti-Myc or anti-S-tag antibodies. The solid (black) arrowheads indicate the relative positions of specific immunodetected proteins; the open (white) arrowheads indicate nonspecific, immunodetected proteins. Note that the immunodetected proteins indicated by the solid arrowheads are not present in the coimmunoprecipitation reactions with lysates of *E. coli* containing empty vector. Shown to the right is a Coomassie blue-stained gel of total soluble protein from lysates containing S-Vps23 or empty vector prior to incubation with S-protein agarose, confirming equivalent input. The numbers to the left of the gels are the molecular masses (in kilodaltons) of protein standards. (C) Interaction of p36 and Vps23 in the BiFC assay. Representative epifluorescence micrographs of BY-2 cells

also served as an internal normalization for reconstituted YFP fluorescence, minimizing differences due to cell-to-cell variability in expression levels. p36⁹¹⁻³³⁰, which does not recruit Vps23 to mitochondria in plant cells (Fig. 4B), served as a negative control. As shown in Fig. 5C, BY-2 cells cotransformed with p36-cYFP and nYFP-Vps23 displayed significantly more YFP/RFP fluorescence than cells cotransformed with p36⁹¹⁻³³⁰-cYFP and nYFP-Vps23. While we cannot rule out the possibility that the relatively small amount of YFP/RFP fluorescence still observed upon coexpression of p36⁹¹⁻³³⁰-cYFP and nYFP-Vps23 may be a result of weaker or, as discussed below, “secondary” interactions between the C terminus of p36 and Vps23, these results indicate that, just as in yeast two-hybrid and *in vitro* coimmunoprecipitation assays (Fig. 5A and B), p36 and Vps23 interact in plant cells in a manner that is dependent upon the N terminus of p36.

A unique polypeptide sequence near the N terminus of p36 is both necessary and sufficient for recruiting Vps23 to mitochondria. Since the N-terminal 90-amino-acid sequence of p36 contains no sequence motifs that match the peptide late-domain motifs (e.g., P[S/T]AP) responsible for the recruitment of ESCRT by enveloped retroviral proteins (20) or by TBSV p33 (23, 25), we carried out a mutational analysis of this region to define any potentially novel Vps23 recruitment motifs. The first mutants examined were those in which either the N-terminal 67 (p36⁶⁸⁻³³⁰) or 22 (p36²³⁻³³⁰) amino acid residues were deleted from the protein. As shown in Fig. 6A, both p36⁶⁸⁻³³⁰ and p36²³⁻³³⁰, which were properly targeted to mitochondria (results are presented for p36²³⁻³³⁰ only [Fig. 6A]), did not recruit Myc-Vps23. Instead, Myc-Vps23 localized in these cells primarily to the cytosol and/or late endosomes, just as it does when it is expressed on its own (Fig. 3A) or coexpressed with the p36 lacking the entire 90-amino-acid-long N-terminal region (p36⁹¹⁻³³⁰) (Fig. 4B). These data suggest that at least the first 22 amino acids of the p36 N terminus are essential for recruiting Vps23 to mitochondria.

Comparison of the N-terminal sequences of p36 with its counterparts from various other tombusviruses (Fig. 6B) revealed that only p36 and the replicase protein of *Pelargonium necrotic spot virus* (PeNSV), which like CIRV, replicates at host cell membranes derived specifically from mitochondria (64), possess a unique and identical stretch of 16 amino acids (residues 7 to 22) that is not present in the replicase proteins of tombusviruses that replicate at peroxisome-derived membranes (i.e., TBSV p33, *Cucumber necrosis virus* [CNV] p33, *Cymbidium ringspot virus* [CyRSV] p33). Notably, this unique sequence in CIRV p36 is due to the initiation codon of the p36 open reading frame being positioned further upstream than the corresponding codon in TBSV or CyRSV (5, 8) and while it contains a negative determinant for satellite RNA replication (65), it is not suspected to play a significant role in virus symptom development (66). Nonetheless, we

triple transformed (as indicated) with RFP, nYFP-Vps23, and either p36-cYFP or p36⁹¹⁻³³⁰-cYFP. Relative RFP and YFP fluorescence values of transformed cells, which are delineated by the solid lines in the panels on the right were quantified using ImageJ (refer to Materials and Methods for details). Note the relatively low YFP fluorescence in the representative cell coexpressing the negative control, p36⁹¹⁻³³⁰-cYFP. At least 25 transformed cells were analyzed from at least three independent experiments, and the mean YFP-to-RFP ratios (plus standard deviation [SD]) are plotted in the bar graph on the right. The two asterisks denote a statistically significant difference between the two samples ($P \leq 0.001$).

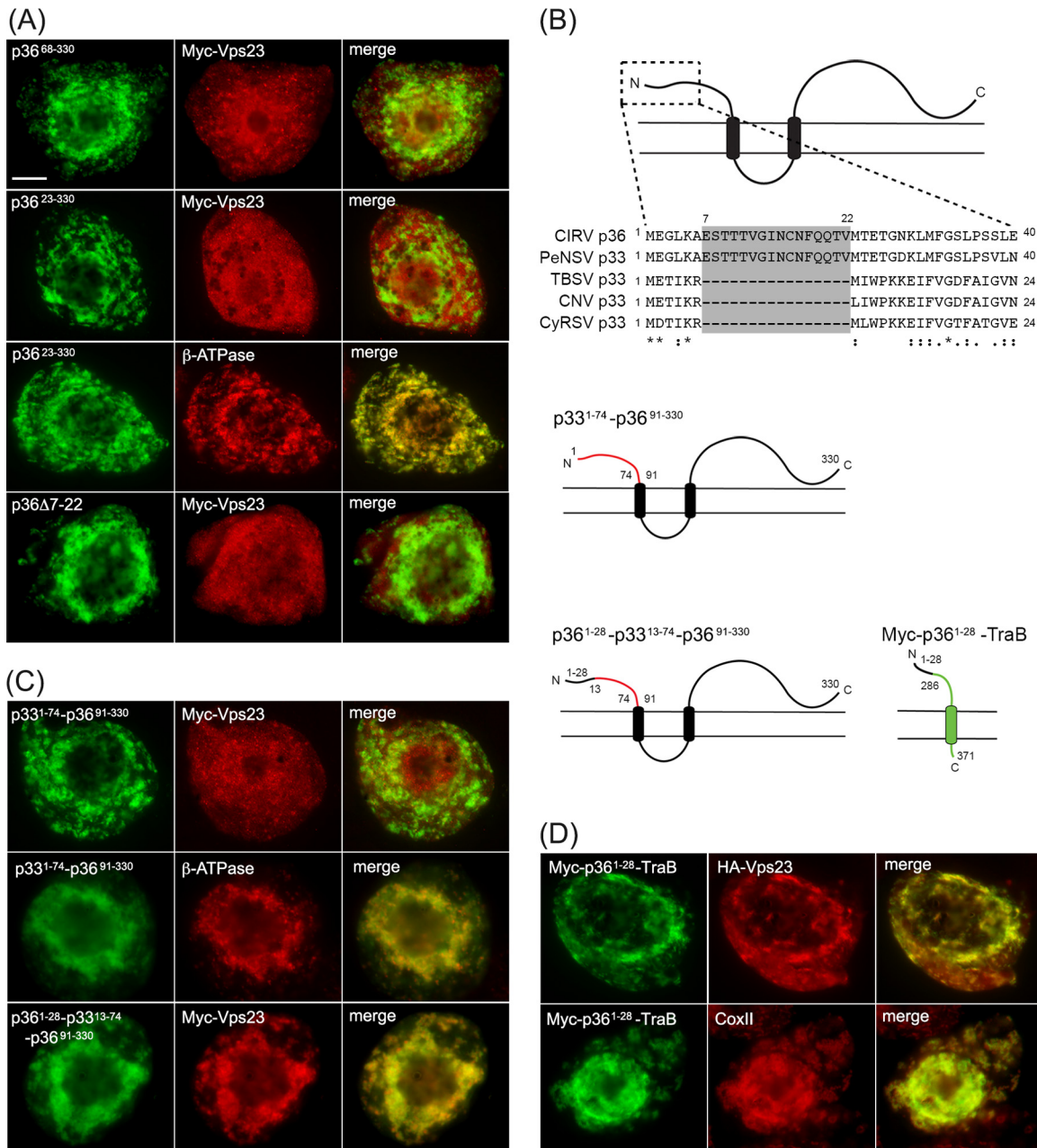


FIG 6 A unique 16-amino-acid-long sequence at the N terminus of p36 is both necessary and sufficient for relocating Vps23 to mitochondria in BY-2 cells. (A) Representative epifluorescence micrographs of BY-2 cells cotransformed with proteins as indicated by panel labels and immunostained (as indicated) for endogenous mitochondrial β -ATPase. Numbers in the name of the construct denote the specific amino acid residues derived from full-length p36 (330 residues) or specific residues deleted from p36. Bar = 10 μ m. (B) Amino acid sequence alignment of the N termini of various tombusvirus p36 and p33 proteins and cartoon illustrations of p33-p36 and p36-TraB hybrid proteins. Sequences were obtained from GenBank and aligned using ClustalW. Identical and similar amino acids in each protein are indicated with asterisks and colons or periods, respectively. The unique amino acid sequence present in CIRV p36 and PeNSV p33 (residues 7 to 22) but absent in TBSV, CNV, and CyRSV p33 are shaded gray. Cartoons depict the structure and topology of p33-p36 and p36-TraB hybrid proteins in the mitochondrial outer membrane. Lines representing amino acid sequences from p33, p36, and TraB are colored red, black, and green, respectively. Numbers represent specific amino acid residues derived from either full-length p33 (296 residues), p36 (330 residues), or TraB (371 residues) and correspond to the numbers in the names of the p33-p36 and p36-TraB hybrid proteins described in panels C and D. (C and D) Representative epifluorescence micrographs of BY-2 cells cotransformed with proteins as indicated by panel labels and immunostained (as indicated) for endogenous mitochondrial β -ATPase. Numbers in the name of the construct denote the specific amino acid residues derived from full-length p36 or p33 and are as illustrated in panel B.

chose to focus on whether this unique region in p36 (and PeNSV p33) is involved in the recruitment of Vps23 based on the inability of p36²³⁻³³⁰ to recruit Vps23 to mitochondria. As shown in Fig. 6A, Myc-Vps23 did not localize to mitochondria when coex-

pressed with a p36 mutant lacking residues 7 through 22 (p36 Δ 7-22), supporting the premise that this region in p36 is specifically involved in Vps23 recruitment. As a complementary approach, we also replaced the N-terminal 90 residues of p36 with the equiva-

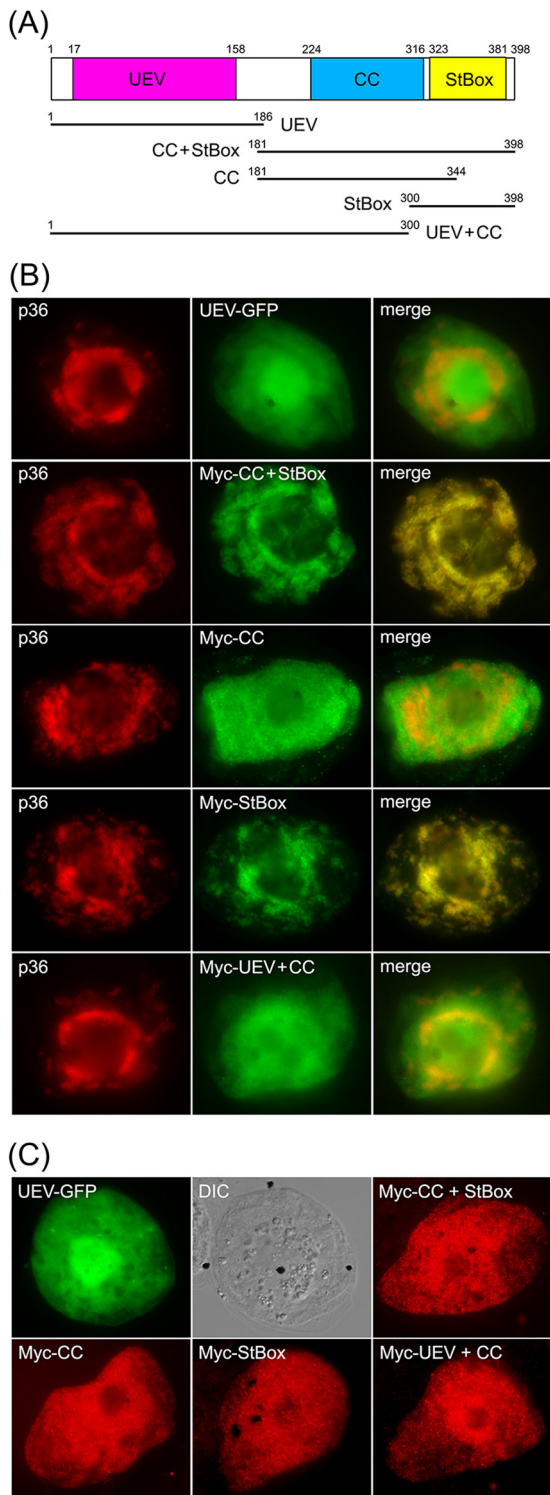


FIG 7 The C-terminal StBox of Vps23 is necessary and sufficient for its recruitment to mitochondria by p36 in BY-2 cells. (A) Schematic diagram illustrating the domain organization of *Arabidopsis* Vps23 based on Spitzer et al. (63). Numbers denote the specific amino acid residues derived from full-length Vps23 (398 residues) that delineate the protein's UEV, CC, and StBox domains, which are colored red, blue, and yellow, respectively, and which are depicted below as lines representing the various Vps23 truncation mutants described in panels B and C. Numbers denote the specific amino acid residues at the N and C termini of each Vps23 mutant. (B and C) Representative epifluorescence micrographs of BY-2 cells (co)transformed with proteins as indicated by panel labels.

lent N-terminal soluble region of TBSV p33 (residues 1 to 74), which is similarly orientated toward the cytosol (4, 26), but does not contain the unique sequence found within the N terminus of p36 (Fig. 6B). Moreover, this N-terminal region of p33 is not considered on its own to be sufficient for interacting with Vps23 (23). As shown in Fig. 6C, the resulting fusion protein, p33¹⁻⁷⁴-p36⁹¹⁻³³⁰, did not colocalize with Myc-Vps23 at mitochondria (refer to Fig. 6C, which shows colocalization between p33¹⁻⁷⁴-p36⁹¹⁻³³⁰ and endogenous β -ATPase for evidence that this fusion protein was properly targeted to mitochondria). In contrast, when the first 28 residues of p36 were reintroduced into p33¹⁻⁷⁴-p36⁹¹⁻³³⁰, yielding p36¹⁻²⁸-p33¹³⁻⁷⁴-p36⁹¹⁻³³⁰ (Fig. 6B), the recruitment of Myc-Vps23 to mitochondria was restored (Fig. 6C).

We also showed that the N-terminal 28 amino acid residues of p36 were capable of recruiting Vps23 to mitochondria when this region was fused to the N terminus of TraB. That is, similar to Myc-p36¹⁻⁹⁰-TraB (Fig. 4C), Myc-p36¹⁻²⁸-TraB, which as illustrated in Fig. 6B, consists of the N-terminal 28 residues of p36 fused to a portion of the cytosol-facing domain of TraB (in order to preserve the overall length of the soluble N-terminal region), as well as its C terminus, also recruits HA-Vps23 to mitochondria in coexpressing cells (Fig. 6D).

Taken together, the data presented in Fig. 6 indicate that the unique N-terminal 16-amino-acid-long sequence present in CIRV p36, but absent in the replicase proteins from tombusviruses that rely on peroxisomal membranes for their replication, contains a novel Vps23 recruitment motif.

Recruitment of Vps23 to mitochondria by p36 requires its StBox domain rather than the UEV domain. As illustrated in Fig. 7A, *Arabidopsis* Vps23, similar to Vps23 in all species examined, consists of three unique structural/functional domains, including (i) an N-terminal UEV domain, (ii) a central coiled-coil (CC) region, which is involved in ESCRT-I protein-protein interactions, and (iii) a C-terminal steadiness box (StBox), which is also involved in protein interactions with other ESCRT-I components, as well as regulating Vps23 protein turnover (63). Given the importance of the UEV domain in facilitating the interaction between Vps23 and p33 during recruitment to peroxisomes (23) and the interactions between Tsg101 and HIV Gag p6 at the plasma membrane (20) and between certain ESCRT-0 proteins and Vps23/Tsg101 during MVB biogenesis in yeasts and mammals (13), we postulated that the UEV domain might also be involved in the recruitment of Vps23 to mitochondria by p36.

To investigate this possibility, the UEV domain of Vps23 fused to green fluorescent protein (UEV-GFP) was coexpressed with full-length p36 in BY-2 cells. As shown in Fig. 7B, UEV-GFP was not localized to mitochondria in cells coexpressing p36, but instead localized throughout the cytosol, just as it does when expressed on its own, as do all other Vps23 mutants described in this study [see below] (Fig. 7C). The lack of recruitment of the UEV domain by p36 was somewhat unexpected and led us to investigate next whether the CC and/or StBox domains play a role in its recruitment to mitochondria by p36. As shown in Fig. 7B, a Vps23 mutant consisting of both the CC and StBox domains fused to an

Names of mutants represent the specific domain(s) derived from Vps23, as illustrated in panel A. The corresponding merged images (B) and the corresponding DIC images of the UEV-GFP-transformed BY-2 cells (C) are also shown.

N-terminal Myc epitope tag (i.e., Myc-CC+StBox; refer to schematic diagram in Fig. 7A), colocalized with coexpressed p36 at mitochondria in BY-2 cells, reinforcing the notion that the recruitment of Vps23 by p36 is mediated by a region(s) other than the UEV domain. As shown also in Fig. 7B, a Vps23 mutant consisting of the CC domain alone (Myc-CC) localized to the cytosol in cells coexpressing p36, while the StBox domain (Myc-StBox) alone was recruited to p36-containing mitochondria. These results suggest that Vps23 recruitment is mediated by the StBox domain. Indeed, the Vps23 mutant consisting of the UEV and CC domains (Myc-UEV+CC) did not colocalize with coexpressed p36 but instead remained in the cytosol (Fig. 7B), just as it does when coexpressed without full-length p36 (Fig. 7C). Together, these results indicate that the C-terminal StBox domain is critical for the recruitment of Vps23 to mitochondria by p36 and that the UEV domain either does not play a role or plays a minor role that is undetectable using this method.

Consistent with this conclusion, yeast two-hybrid assays performed with p36¹⁻⁹⁰ and the equivalent set of Vps23 mutants as those described above revealed that only those proteins containing the StBox (i.e., CC+StBox and StBox) conferred yeast growth on high selection media (Fig. 8A). A notable exception was the UEV+CC mutant, which when coexpressed with p36¹⁻⁹⁰, also resulted in growth on high selection media, albeit relatively less than that observed for yeast coexpressing CC+StBox, StBox, or full-length Vps23 (Fig. 8A). However, as mentioned above, the equivalent UEV-CC mutant is not recruited to mitochondria by full-length p36 in plant cells (Fig. 7B), suggesting that if the CC domain does indeed participate in the recruitment of Vps23 by p36, it does so to a lesser extent than the StBox domain does. While we confirmed that UEV+CC (or any of the other Vps23 mutants) does not autoactivate the yeast two-hybrid reporter genes (Fig. 8B), we cannot rule out the possibility that the observed p36¹⁻⁹⁰/UEV+CC interaction may be the result of a “bridge” by an endogenous yeast ESCRT component(s). Moreover, and as discussed in more detail below, we also confirmed that all of the Vps23 mutants (and full-length Vps23) interacted in yeast two-hybrid assays with the *Arabidopsis* ESCRT-I proteins Vps28 and Vps37 (Fig. 8C and D) in ways that are consistent with how these domains are involved in the corresponding ESCRT protein-protein interactions in yeasts and mammals (27–30). For instance, like their yeast and mammalian counterparts, the interaction of *Arabidopsis* Vps28 and Vps23 is mediated by the Vps23 StBox domain (Fig. 8C), while the interaction of *Arabidopsis* Vps37 and Vps23 is mediated by the Vps23 CC domain, although the Vps23 UEV domain also appears to be involved in Vps23-Vps37 binding (Fig. 8D).

DISCUSSION

Understanding how viruses appropriate host cell organelles in order to carry out their replication is not only an important step in the development of strategies to prevent infection, it often provides novel insights into the molecular mechanisms underlying organelle biogenesis and dynamics in uninfected cells. For example, recent studies examining how the *Bunyavirus Tomato spotted wilt virus* buds from the endoplasmic reticulum (ER) in plant cells or how *Turnip mosaic virus* infection leads to the fusion of ER-derived vesicles with chloroplasts has shed new light on the formation of ER exit sites (67) and the role(s) of the previously uncharacterized SNARE protein Syp71 (68). Likewise, the charac-

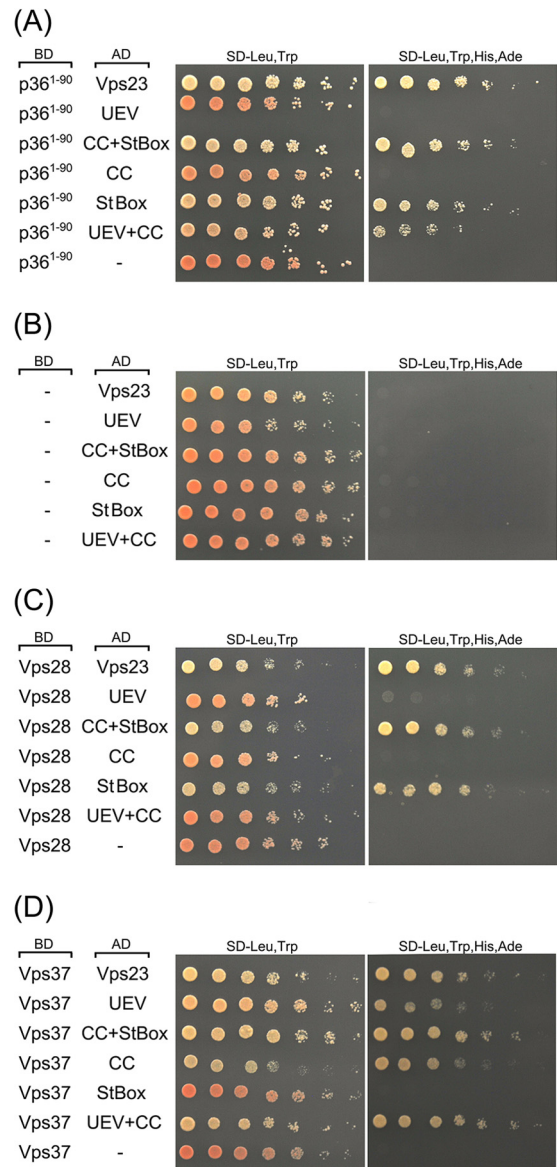


FIG 8 The N-terminal region of p36, like Vps28, interacts with the C-terminal StBox of Vps23 in the yeast two-hybrid assay. (A to D) Yeast strains were cotransformed with the indicated pairs of GAL4-binding domain (BD) and GAL4-activating domain (AD) fusion proteins or the corresponding empty BD or AD control plasmids (indicated by a hyphen in the BD or AD column). Serial (1:5) dilutions of cells were spotted onto agar plates containing either low-stringency medium (SD-Leu,Trp) or high-stringency medium (SD-Leu,Trp,His,Ade), where cell growth is dependent on two-hybrid protein interactions. Note that neither full-length Vps23, nor any of the truncation mutants of Vps23, Vps28, and Vps37 autoactivated the two-hybrid reporter gene system on their own.

terization of the late-domain motifs in certain retroviral structural proteins and their interaction with the host cell ESCRT-I protein Tsg101 (69) was a key step in the subsequent discovery that similar peptide motifs exist in certain ESCRT-0 proteins. As such, it is now known that, in uninfected cells, the sequential recruitment of the ESCRT subcomplexes (i.e., ESCRT-0, -I, -II, and -III) during MVB biogenesis begins with ESCRT-0 binding to ESCRT-I and that retroviral proteins such as the HIV Gag p6 mimic ESCRT-0 in order to hijack ESCRT-I (Tsg101) to the sites of viral budding at the plasma membrane (20).

Presently, our understanding of ESCRT in plants is relatively poor, and thus, working models based on studies of ESCRT in yeasts and mammals are often employed to infer the mechanisms of ESCRT assembly and function in plants. While this approach has generally proven to be valid because of the overall conserved nature of ESCRT function among evolutionarily diverse organisms, particularly during MVB biogenesis (35, 61, 62), there are also several features of ESCRT in plants that are distinct, including the ability of TOL proteins to compensate for the lack of ESCRT-0 proteins in terms of ubiquitinated cargo recognition and sorting (16, 17). Thus, understanding how different tombusviruses appropriate the plant ESCRT machinery, apparently through different forms of molecular mimicry, as discussed below, not only gives important insight into how tombusviruses manipulate membranes to facilitate their replication but may also provide valuable clues to previously unexplored aspects of normal ESCRT assembly and function in plants.

The recruitment of ESCRT to mitochondria by CIRV is mediated by a unique sequence located near the N terminus of p36. Plant tombusviruses are a well-established model for studying (+)RNA virus replication and virus-host interactions (70) and are also emerging as an especially useful tool for understanding how (+)RNA viruses recruit and manipulate specific host cell organelle membranes in order to create unique compartments that house the viral replication machinery. For instance, based on studies that focused on TBSV and employed yeast as a model organism for tombusvirus replication (71), Nagy and coworkers recently extended an earlier working model (72) whereby ESCRT is hijacked by TBSV in order to facilitate the progressive invagination of the peroxisomal boundary membrane, which according to the model, leads to the concentration and assembly of viral replicase complexes within nascent spherules (23). While several aspects of this model remain to be confirmed experimentally, one of the most notable tenets is that the TBSV replicase protein p33 binds to the ESCRT-I protein Vps23 via a peptide sequence (i.e., PSVP) in p33 that resembles the proline-rich late-domain motif found in HIV Gag p6. Consequently, p33, analogous to Gag p6, is considered to “mimic” ESCRT-0 in terms of its ability to bind and recruit Vps23 (ESCRT-I) (21, 22). However, as mentioned above, the lack of plant homologs of ESCRT-0 proteins implies that the precise nature by which p33 interacts with Vps23 is not completely analogous to the interaction between HIV Gag p6 and Tsg101 and, instead, may reflect a unique mechanism for ESCRT recruitment in plants, which may or may not involve the TOL proteins, and which is currently an open question.

Here we investigated whether an ESCRT recruitment process is utilized by CIRV, another member of the tombusvirus family that also causes pronounced changes to a specific organellar membrane in infected plant cells, but unlike TBSV, does so at the mitochondrial outer membrane (73). Our initial interest in investigating CIRV stemmed from the observation that the late-domain-like motif (i.e., PSVP) identified in TBSV p33 (23) is not conserved in CIRV p36 (Fig. 1), suggesting that, if CIRV does rely on the ESCRT machinery, it does so in a distinct manner compared to TBSV. To begin to test this hypothesis, we showed that the replication efficiency of CIRV does indeed depend on ESCRT. That is, coexpression of a mutant version of Vps4 (Vps4^{E232Q}), which is one of the best-characterized ESCRT components in plants (33, 35, 55, 63), inhibited CIRV replication in *N. benthamiana* plants (Fig. 2), presumably by disrupting the ESCRT-depen-

dent assembly of the replicase complex, a conclusion that was reinforced by the observation that TBSV, but not TRV, replication was also inhibited by Vps4^{E232Q}, as expected (23). We also showed that the ESCRT-I component Vps23 is redirected to mitochondria in plant cells expressing either full-length CIRV or the p36 replicase-associated protein alone (Fig. 3), indicating that p36, like TBSV p33 (23), is the minimal viral component necessary for Vps23 recruitment. Moreover, the recruitment of Vps23 to mitochondria by p36 was shown to be mediated by the viral protein's N-terminal 90 amino acids (Fig. 4 and 5), a region that precedes the first of the protein's two TMDs and is orientated toward the cytosol (Fig. 4A), which we considered a prerequisite for interaction with cytosolic Vps23. In contrast, neither the cytosol-facing C-terminal portion, nor the intervening loop sequence of p36, which contains the mitochondrial targeting signal and is orientated toward the mitochondrial intermembrane space (11) (Fig. 4), were directly involved in Vps23 recruitment (Fig. 4). The latter observation is particularly notable because the intervening loop region of TBSV p33, which contains the PSVP late-domain-like motif implicated in recruitment of Vps23 to peroxisomes (23), is also considered to be orientated toward the peroxisomal matrix (11, 26). Thus, it remains to be determined how the PSVP motif in p33 is accessible to bind to Vps23 in the cytosol, although one possibility is that the association of nascent p33 with cytosolic chaperones prior to its targeting to and insertion into the peroxisomal membranes (74) may render this motif accessible to Vps23. While we did not detect *in vivo* recruitment activity for the intervening loop region and C-terminal region of p36, we also cannot rule out the possibility that additional motifs within these regions of p36 interact, albeit weakly, with Vps23 and that these secondary interactions might help CIRV to more efficiently compete with the cellular ESCRT machinery, for example, Vps28 (as discussed below) for Vps23. Indeed, while the BiFC results presented in Fig. 5C support the premise that p36 and Vps23 interact in plant cells and do so in a manner that is dependent upon the N terminus of p36, the observation that coexpression of the p36⁹¹⁻³³⁰ mutant and Vps23 yielded a BiFC signal, although much weaker than that from full-length p36 and Vps23, implies that such secondary interactions might exist.

To gain an understanding of the mechanism by which the N terminus of p36 recruits Vps23 to mitochondria, we carried out a mutational analysis of this region, the results of which highlighted a 16-amino-acid-long sequence (i.e., residues 7 to 22) that is present in tombusviruses that replicate at mitochondria (i.e., CIRV and PeNSV) and absent in tombusviruses that replicate at peroxisomes (i.e., TBSV and *Cucumber necrosis virus* [CNV]) (Fig. 6B). We showed that when this unique polypeptide sequence in p36 was specifically deleted, Vps23 recruitment to mitochondria was abolished (Fig. 6A). Conversely, the inclusion of this sequence from p36 within the context of various hybrid fusion proteins consisting of other portions of p36 and either p33 or TraB conferred the ability to recruit Vps23 to mitochondria (Fig. 6C and D). This implies that this region in p36 can function as a Vps23-binding domain outside of the context of the full-length viral protein. Despite its involvement in Vps23 recruitment, results from previous studies indicate that this unique N-terminal sequence in p36 is not essential for virus replication, since a mutant version of CIRV with this region deleted induced the same necrotic symptoms in infected *N. benthamiana* plants as that observed for the wild-type virus, although the appearance of these symptoms was

conspicuously delayed (66). These results suggest that the N-terminal motif may mediate a more robust infection. Indeed, the retention of this motif in CIRV indicates that it confers a competitive advantage to the virus, which we suggest is related to its demonstrated role in Vps23 recruitment. Nonetheless, the virus viability observed in the absence of the motif implies that CIRV might rely on other secondary interactions between p36 and Vps23, as mentioned above, and/or additional ESCRT factors (e.g., Bro1), analogous to how other viruses, such as HIV and TBSV, interact with multiple ESCRT proteins (20, 23). This study, in combination with the rapidly growing interest in the plant ESCRT machinery in general, should now facilitate a more detailed analysis of the other ESCRT factors involved in CIRV replication.

It is also worth mentioning that while the N terminus of p36, including the sequence corresponding to amino acid residues 7 through 22, is devoid of any obvious functional domains that resemble those used by other viruses to recruit Vps23/Tsg101 (e.g., proline- or leucine-rich late-domain motifs) (75), a portion of this region is predicted to be relatively disordered (i.e., based on IUPred [76]). Disordered (unstructured) regions are commonly found in viral proteins, where they are known to mediate interactions with different host cell components in dynamic ways (77). Indeed, intrinsically disordered regions in TBSV p33 have been implicated in its RNA chaperone activity (78), as well as potentially mediating interactions with host proteins and viral factors (79). Currently, we are assessing the structural attributes (or lack thereof) of the p36 N terminus, as well as employing peptide “mimetics” (80) to determine the possible structure-function relationship of this region of p36 and its binding to Vps23.

p36, similar to the ESCRT-I protein Vps28, interacts with the C-terminal StBox domain of Vps23. Additional evidence that CIRV utilizes a distinct mechanism in order to hijack ESCRT was our finding that the binding and recruitment of Vps23 by p36 do not depend on the Vps23 N-terminal UEV domain but instead are mediated by an interaction between p36 and the C-terminal StBox domain of Vps23 (Fig. 7 and 8). In mammalian cells, the StBox in Tsg101 is involved in coordinating the regulation of the protein's stability and its interaction with Vps28 during ESCRT-I assembly (81). Similarly, we showed that *Arabidopsis* Vps28 and Vps23 interact via the Vps23 StBox (Fig. 7 and 8). As such, it is possible that p36 “mimics” Vps28 by binding to the StBox in Vps23, thereby facilitating its recruitment to mitochondria. It now remains to be determined whether p36 and Vps28 compete for an interaction with Vps23 in plant cells, based on our assumption that the molecular mechanisms underlying ESCRT-I assembly in plants and mammals, including the interaction of Vps28 with the Vps23 StBox, are similar. This appears to be plausible, since the core components of *Arabidopsis* ESCRT-I (i.e., Vps23, Vps28, and Vps37) are already known to interact in a manner similar to their mammalian counterparts (33, 35, 55, 63), and as we showed in this study (Fig. 8), *Arabidopsis* Vps37 and Vps28 interact with the UEV/CC domains and StBox of *Arabidopsis* Vps23, respectively, which is consistent with models of ESCRT-I assembly in mammals and yeasts (13). Presently, we are comparing the nature of the interaction between p36 or Vps28 and the Vps23 StBox in plant cells and comparing these with the well-characterized interactions between Vps28 and Vps23 in yeasts and mammals (27–30).

Our model that the binding of p36 and Vps28 to Vps23 is mutually exclusive also prompts the question of how additional

downstream ESCRT components would be recruited and assembled at the mitochondrial membranes in CIRV-infected cells, since Vps28 is considered to act as a bridge between ESCRT-I and ESCRT-II during normal MVB biogenesis (82). However, there is also a growing appreciation that not all ESCRT components participate in the cellular processes involving ESCRT-dependent membrane deformation. For instance, during HIV infection in mammalian cells, the binding of Tsg101 by HIV Gag p6 does not lead to the recruitment of Vps28 and subsequently ESCRT-II. Instead, Tsg101 and Gag p6 interact with the ESCRT accessory protein Alix, which recruits ESCRT-III, which ultimately facilitates the viral budding process (20). Interestingly, a mutant of Bro1, the Alix ortholog in plants, was shown to inhibit TBSV replication and TBSV p33 interacts with yeast Bro1p (23). This suggests a role for Bro1 in ESCRT assembly at peroxisomes in TBSV-infected plant cells that may be analogous to the role for Alix during HIV budding. Likewise, it is possible that while CIRV and TBSV seem to rely on distinct mechanisms for the initial recruitment of Vps23, the recruitment of additional ESCRT components, either in parallel with the initial interaction(s) of p36 and Vps23 or further downstream, may be mechanistically similar, including the recruitment of Bro1. Future experiments aimed at dissecting the recruitment of additional ESCRT components, including Bro1 and ESCRT-III, during CIRV replication, will be important for obtaining a clear picture of the molecular events that underlie spherule formation, which is critical for understanding the spatial dynamics of tombusvirus replication at host cell membranes.

ACKNOWLEDGMENTS

We thank M. Otegui, S. Gelvin, T. Elthon, and S. Subramani for their generous gifts of plasmids or antibodies used in this study and T. Nguyen for maintaining tobacco plants and BY-2 cell cultures. We are also indebted to all the members of our laboratories and J. Dyer for their helpful discussions during the course of this work. We also thank the anonymous referees for their constructive comments that led to an improved version of the article.

This study was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to R. T. Mullen and K. A. White, as well as an NSERC CGS-D Postgraduate Scholarship to L. G. L. Richardson. R. T. Mullen holds a University of Guelph Research Chair, and K. A. White holds a Canada Research Chair in plant biotechnology and structural biology.

REFERENCES

1. Russo M, Burgyan J, Martelli GP. 1994. Molecular biology of *Tombusviridae*. *Adv. Virus Res.* 44:381–428. [http://dx.doi.org/10.1016/S0065-3527\(08\)60334-6](http://dx.doi.org/10.1016/S0065-3527(08)60334-6).
2. den Boon JA, Ahlquist P. 2010. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annu. Rev. Microbiol.* 64:241–256. <http://dx.doi.org/10.1146/annurev.micro.112408.134012>.
3. Laliberté JF, Sanfaçon H. 2010. Cellular remodeling during plant virus infection. *Annu. Rev. Phytopathol.* 48:69–91. <http://dx.doi.org/10.1146/annurev-phyto-073009-114239>.
4. McCartney AW, Greenwood JS, Fabian MR, White KA, Mullen RT. 2005. Localization of the tomato bushy stunt virus replication protein p33 reveals a peroxisome-to-endoplasmic reticulum sorting pathway. *Plant Cell* 17:3513–3531. <http://dx.doi.org/10.1105/tpc.105.036350>.
5. Burgyan J, Rubino L, Russo M. 1996. The 5'-terminal region of a tombusvirus genome determines the origin of multivesicular bodies. *J. Gen. Virol.* 77:1967–1974. <http://dx.doi.org/10.1099/0022-1317-77-8-1967>.
6. Scholthof KB, Scholthof HB, Jackson AO. 1995. The tomato bushy stunt virus replicase proteins are coordinately expressed and membrane associated. *Virology* 208:365–369. <http://dx.doi.org/10.1006/viro.1995.1162>.
7. Oster SK, Wu B, White KA. 1998. Uncoupled expression of p33 and p92

- permits amplification of tomato bushy stunt virus RNAs. *J. Virol.* 72: 5845–5851.
8. Rubino L, Burgyan J, Russo M. 1995. Molecular cloning and complete nucleotide sequence of *carnation Italian ringspot tobravirus* genomic and defective interfering RNAs. *Arch. Virol.* 140:2027–2039. <http://dx.doi.org/10.1007/BF01322690>.
 9. Cimino PA, Nicholson BL, Wu B, Xu W, White KA. 2011. Multifaceted regulation of translational readthrough by RNA replication elements in a tobravirus. *PLoS Pathog.* 7:e1002423. <http://dx.doi.org/10.1371/journal.ppat.1002423>.
 10. Navarro B, Rubino L, Russo M. 2004. Expression of the *Cymbidium ringspot virus* 33-kilodalton protein in *Saccharomyces cerevisiae* and molecular dissection of the peroxisomal targeting signal. *J. Virol.* 78:4744–4752. <http://dx.doi.org/10.1128/JVI.78.9.4744-4752.2004>.
 11. Hwang YT, McCartney AW, Gidda SK, Mullen RT. 2008. Localization of the *Carnation Italian ringspot virus* replication protein p36 to the mitochondrial outer membrane is mediated by an internal targeting signal and the TOM complex. *BMC Cell Biol.* 9:54. <http://dx.doi.org/10.1186/1471-2121-9-54>.
 12. Nagy PD, Barajas D, Pogany J. 2012. Host factors with regulatory roles in tobravirus replication. *Curr. Opin. Virol.* 2:691–698. <http://dx.doi.org/10.1016/j.coviro.2012.10.004>.
 13. Hanson PI, Cashikar A. 2012. Multivesicular body morphogenesis. *Annu. Rev. Cell Dev. Biol.* 28:337–362. <http://dx.doi.org/10.1146/annurev-cellbio-092910-154152>.
 14. Otegui MS, Spitzer C. 2008. Endosomal functions in plants. *Traffic* 9:1589–1598. <http://dx.doi.org/10.1111/j.1600-0854.2008.00787.x>.
 15. Schellmann S, Pimpl P. 2009. Coats of endosomal protein sorting: retromer and ESCRT. *Curr. Opin. Plant Biol.* 12:670–676. <http://dx.doi.org/10.1016/j.pbi.2009.09.005>.
 16. Herman EK, Walker G, van der Giezen M, Dacks JB. 2011. Multivesicular bodies in the enigmatic amoeboflagellate *Breviata anathema* and the evolution of ESCRT 0. *J. Cell Sci.* 124:613–621. <http://dx.doi.org/10.1242/jcs.078436>.
 17. Korbei B, Moulinier-Anzola J, De-Araujo L, Lucyshyn D, Retzer K, Khan MA, Luschign C. 2013. *Arabidopsis* TOL proteins acts as gatekeepers for vacuolar sorting of PIN2 plasma membrane protein. *Curr. Biol.* 16:2500–2505.
 18. Morita E. 2012. Differential requirements of mammalian ESCRTs in multivesicular body formation, virus budding and cell division. *FEBS J.* 279: 1399–1406. <http://dx.doi.org/10.1111/j.1742-4658.2012.08534.x>.
 19. Jouvenet N, Zhadina M, Bieniasz PD, Simon SM. 2011. Dynamics of ESCRT protein recruitment during retroviral assembly. *Nat. Cell Biol.* 13:394–401. <http://dx.doi.org/10.1038/ncb2207>.
 20. Martin-Serrano J, Neil SJ. 2011. Host factors involved in retroviral budding and release. *Nat. Rev. Microbiol.* 9:519–531. <http://dx.doi.org/10.1038/nrmicro2596>.
 21. Pornillo O, Higginson DS, Stray KM, Fisher RD, Garrus JE, Payne M, He GP, Wang HE, Morham SG, Sundquist WI. 2003. HIV Gag mimics the Tsg101-recruiting activity of the human Hrs protein. *J. Cell Biol.* 162: 425–434. <http://dx.doi.org/10.1083/jcb.200302138>.
 22. Im YJ, Kuo L, Ren X, Burgos PV, Zhao XZ, Liu F, Burke TR, Jr, Bonifacino JS, Freed EO, Hurley JH. 2010. Crystallographic and functional analysis of the ESCRT-I/HIV-1 Gag PTAP interaction. *Structure* 18:1536–1547. <http://dx.doi.org/10.1016/j.str.2010.08.010>.
 23. Barajas D, Jiang Y, Nagy PD. 2009. A unique role for the host ESCRT proteins in replication of *Tomato bushy stunt virus*. *PLoS Pathog.* 5:e1000705. <http://dx.doi.org/10.1371/journal.ppat.1000705>.
 24. Serviene E, Shapka N, Cheng CP, Panavas T, Phuangrat B, Baker J, Nagy PD. 2005. Genome-wide screen identifies host genes affecting viral RNA recombination. *Proc. Natl. Acad. Sci. U. S. A.* 102:10545–10550. <http://dx.doi.org/10.1073/pnas.0504844102>.
 25. Barajas D, Nagy PD. 2010. Ubiquitination of tobravirus p33 replication protein plays a role in virus replication and binding to the host Vps23p ESCRT protein. *Virology* 397:358–368. <http://dx.doi.org/10.1016/j.virol.2009.11.010>.
 26. Rubino L, Russo M. 1998. Membrane targeting sequences in tobravirus infections. *Virology* 252:431–437. <http://dx.doi.org/10.1006/viro.1998.9490>.
 27. Stuchell MD, Garrus JE, Müller B, Stray KM, Ghaffarian S, McKinnon R, Kräusslich HG, Morham SG, Sundquist WI. 2004. The human endosomal sorting complex required for transport (ESCRT-I) and its role in HIV-1 budding. *J. Biol. Chem.* 279:36059–36071. <http://dx.doi.org/10.1074/jbc.M405226200>.
 28. Eastman SW, Martin-Serrano J, Chung W, Zhang T, Bieniasz PD. 2005. Identification of human VPS37C, a component of endosomal sorting complex required for transport-I important for viral budding. *J. Biol. Chem.* 280:628–636.
 29. Kostelansky MS, Sun J, Lee S, Kim J, Ghirlando R, Hierro A, Emr SD, Hurley JH. 2006. Structural and functional organization of the ESCRT-I trafficking complex. *Cell* 125:113–126. <http://dx.doi.org/10.1016/j.cell.2006.01.049>.
 30. Teo H, Gill DJ, Sun J, Perisic O, Vepintsev DB, Vallis Y, Emr SD, Williams RL. 2006. ESCRT-I core and ESCRT-II GLUE domain structures reveal role for GLUE in linking to ESCRT-I and membranes. *Cell* 125:99–111. <http://dx.doi.org/10.1016/j.cell.2006.01.047>.
 31. Hearne PQ, Knorr DA, Hillman BI, Morris TJ. 1990. The complete genome structure and synthesis of infectious RNA from clones of tomato bushy stunt virus. *Virology* 177:141–151. [http://dx.doi.org/10.1016/0042-6822\(90\)90468-7](http://dx.doi.org/10.1016/0042-6822(90)90468-7).
 32. Liu Y, Schiff M, Dinesh-Kumar SP. 2002. Virus-induced gene silencing in tomato. *Plant J.* 31:777–786. <http://dx.doi.org/10.1046/j.1365-313X.2002.01394.x>.
 33. Haas TJ, Sliwinski MK, Martinez DE, Preuss M, Ebine K, Ueda T, Nielsen E, Odorizzi G, Otegui MS. 2007. The *Arabidopsis* AAA ATPase SKD1 is involved in multivesicular endosome function and interacts with its positive regulator LYST-INTERACTING PROTEIN5. *Plant Cell* 19: 1295–1312. <http://dx.doi.org/10.1105/tpc.106.049346>.
 34. Chung SM, Frankman EL, Tzfira T. 2005. A versatile vector system for multiple gene expression in plants. *Trends Plant Sci.* 10:357–361. <http://dx.doi.org/10.1016/j.tplants.2005.06.001>.
 35. Richardson LG, Howard AS, Khuu N, Gidda SK, McCartney A, Morphy BJ, Mullen RT. 2011. Protein-protein interaction network and subcellular localization of the *Arabidopsis thaliana* ESCRT machinery. *Front. Plant Sci.* 2:20. <http://dx.doi.org/10.3389/fpls.2011.00020>.
 36. Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM. 2006. Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell* 18:2294–2313. <http://dx.doi.org/10.1105/tpc.106.043695>.
 37. Uemura T, Ueda T, Ohniwa RL, Nakano A, Takeyasu K, Sato MH. 2004. Systematic analysis of SNARE molecules in *Arabidopsis*: dissection of the post-Golgi network in plant cells. *Cell Struct. Funct.* 29:49–65. <http://dx.doi.org/10.1247/csf.29.49>.
 38. Ching SL, Gidda SK, Rochon A, van Cauwenberghe OR, Shelp BJ, Mullen RT. 2012. Glyoxylate reductase isoform 1 is localized in the cytosol and not peroxisomes in plant cells. *J. Integr. Plant Biol.* 54:152–168. <http://dx.doi.org/10.1111/j.1744-7909.2012.01103.x>.
 39. Dhanoa PK, Richardson LG, Smith MD, Gidda SK, Henderson MP, Andrews DW, Mullen RT. 2010. Distinct pathways mediate the sorting of tail-anchored proteins to the plastid outer envelope. *PLoS One* 5:e10098. <http://dx.doi.org/10.1371/journal.pone.0010098>.
 40. Hurley BA, Tran HT, Marty NJ, Park J, Snedden WA, Mullen RT, Plaxton WC. 2010. The dual-targeted purple acid phosphatase isozyme AtPAP26 is essential for efficient acclimation of *Arabidopsis* to nutritional phosphate deprivation. *Plant Physiol.* 153:1112–1122. <http://dx.doi.org/10.1104/pp.110.153270>.
 41. Citovsky V, Lee LY, Vyas S, Glick E, Chen MH, Vainstein A, Gafni Y, Gelvin SB, Tzfira T. 2006. Subcellular localization of interacting proteins by bimolecular fluorescence complementation in planta. *J. Mol. Biol.* 362: 1120–1131. <http://dx.doi.org/10.1016/j.jmb.2006.08.017>.
 42. Restrepo MA, Freed DD, Carrington JC. 1990. Nuclear transport of plant potyviral proteins. *Plant Cell* 2:987–998.
 43. Clark SM, Di Leo R, Dhanoa PK, Van Cauwenberghe OR, Mullen RT, Shelp BJ. 2009. Biochemical characterization, mitochondrial localization, expression, and potential functions for an *Arabidopsis* gamma-aminobutyrate transaminase that utilizes both pyruvate and glyoxylate. *J. Exp. Bot.* 60:1743–1757. <http://dx.doi.org/10.1093/jxb/erp044>.
 44. Scholthof HB. 1999. Rapid delivery of foreign genes into plants by direct rub-inoculation with intact plasmid DNA of a tomato bushy stunt virus gene vector. *J. Virol.* 73:7823–7829.
 45. White KA, Morris TJ. 1994. Nonhomologous RNA recombination in tobamoviruses: generation and evolution of defective interfering RNAs by stepwise deletions. *J. Virol.* 68:14–24.
 46. MacFarlane SA, Vassilakos N, Brown DJ. 1999. Similarities in the genome organization of *tobacco rattle virus* and *pea early-browning virus* isolates that are transmitted by the same vector nematode. *J. Gen. Virol.* 80:273–276.

47. Lingard MJ, Gidda SK, Bingham S, Rothstein SJ, Mullen RT, Trelease RN. 2008. *Arabidopsis* PEROXIN11c-e, FISSION1b, and DYNAMIN-RELATED PROTEIN3A cooperate in cell cycle-associated replication of peroxisomes. *Plant Cell* 20:1567–1585. <http://dx.doi.org/10.1105/tpc.107.057679>.
48. Lee MS, Mullen RT, Trelease RN. 1997. Oilseed isocitrate lyases lacking their essential type 1 peroxisomal targeting signal are piggybacked to glyoxysomes. *Plant Cell* 9:185–197.
49. Luethy MH, Horak A, Elthon TE. 1993. Monoclonal antibodies to the α - and β -subunits of the plant mitochondrial F1-ATPase. *Plant Physiol.* 101: 931–937.
50. Frelin O, Agrimi G, Laera VL, Castegna A, Richardson LG, Mullen RT, Lerma-Ortiz C, Palmieri F, Hanson AD. 2012. Identification of mitochondrial thiamin diphosphate carriers from *Arabidopsis* and maize. *Funct. Integr. Genomics* 12:317–326. <http://dx.doi.org/10.1007/s10142-012-0273-4>.
51. Gidda SK, Shockey JM, Falcone M, Kim PK, Rothstein SJ, Andrews DW, Dyer JM, Mullen RT. 2011. Hydrophobic-domain-dependent protein-protein interactions mediate the localization of GPAT enzymes to ER subdomains. *Traffic* 12:452–472. <http://dx.doi.org/10.1111/j.1600-0854.2011.01160.x>.
52. Lee LY, Wu FH, Hsu CT, Shen SC, Yeh HY, Liao DC, Fang MJ, Liu NT, Yen YC, Dokladal L, Sykorova E, Gelvin SB, Lin CS. 2012. Screening a cDNA library for protein-protein interactions directly in planta. *Plant Cell* 24:1746–1759. <http://dx.doi.org/10.1105/tpc.112.097998>.
53. Clontech Laboratories. 2007. Matchmaker™ GAL4 two-hybrid system 3 & libraries user manual. PT3247-1 (PR742219). Clontech Laboratories, Inc., Mountain View, CA.
54. Park S, Gidda SK, James CN, Horn PJ, Khuu N, Seay DC, Keereetaweep J, Chapman KD, Mullen RT, Dyer JM. 2013. The α/β hydrolase CGI-58 and peroxisomal transport protein PXA1 coregulate lipid homeostasis and signaling in *Arabidopsis*. *Plant Cell* 25:1726–1739. <http://dx.doi.org/10.1105/tpc.113.111898>.
55. Shahriari M, Keshavaiah C, Scheuring D, Sabovljevic A, Pimpl P, Hausler RE, Hulskamp M, Schellmann S. 2010. The AAA-type ATPase AtSKD1 contributes to vacuolar maintenance of *Arabidopsis thaliana*. *Plant J.* 64:71–85.
56. Babst M, Wendland B, Estepa EJ, Emr SD. 1998. The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function. *EMBO J.* 17:2982–2993. <http://dx.doi.org/10.1093/emboj/17.11.2982>.
57. Brandizzi F, Irons S, Kearns A, Hawes C. 2003. BY-2 cells: culture and transformation for live cell imaging. *Curr. Protoc. Cell Biol.* Chapter 1:Unit 1.7. <http://dx.doi.org/10.1002/0471143030.cb0107s19>.
58. Laporte C, Vetter G, Loudes AM, Robinson DG, Hillmer S, Stussi-Garaud C, Ritzenthaler C. 2003. Involvement of the secretory pathway and the cytoskeleton in intracellular targeting and tubule assembly of *Grapevine fanleaf virus* movement protein in tobacco BY-2 cells. *Plant Cell* 15:2058–2075. <http://dx.doi.org/10.1105/tpc.013896>.
59. Wei T, Zhang C, Hong J, Xiong R, Kasschau KD, Zhou X, Carrington JC, Wang A. 2010. Formation of complexes at plasmodesmata for potyvirus intercellular movement is mediated by the viral protein P3N-PIPO. *PLoS Pathog.* 6:e1000962. <http://dx.doi.org/10.1371/journal.ppat.1000962>.
60. Duncan O, Taylor NL, Carrie C, Eubel H, Kubiszewski-Jakubiak S, Zhang B, Narsai R, Millar AH, Whelan J. 2011. Multiple lines of evidence localize signaling, morphology, and lipid biosynthesis machinery to the mitochondrial outer membrane of *Arabidopsis*. *Plant Physiol.* 157:1093–1113. <http://dx.doi.org/10.1104/pp.111.183160>.
61. Shahriari M, Richter K, Keshavaiah C, Sabovljevic A, Huelskamp M, Schellmann S. 2011. The *Arabidopsis* ESCRT protein-protein interaction network. *Plant Mol. Biol.* 76:85–96. <http://dx.doi.org/10.1007/s11103-011-9770-4>.
62. Ibl V, Csaszar E, Schlager N, Neubert S, Spitzer C, Hauser MT. 2012. Interactome of the plant-specific ESCRT-III component AtVPS2.2 in *Arabidopsis thaliana*. *J. Proteome Res.* 11:397–411. <http://dx.doi.org/10.1021/pr200845n>.
63. Spitzer C, Schellmann S, Sabovljevic A, Shahriari M, Keshavaiah C, Bechtold N, Herzog M, Muller S, Hanisch FG, Hulskamp M. 2006. The *Arabidopsis elch* mutant reveals functions of an ESCRT component in cytokinesis. *Development* 133:4679–4689. <http://dx.doi.org/10.1242/dev.02654>.
64. Heinze C, Wobbe V, Lesemann DE, Zhang DY, Willingmann P, Adam G. 2004. Pelargonium necrotic spot virus: a new member of the genus *Tombusvirus*. *Arch. Virol.* 149:1527–1539.
65. Céliz A, Burgyan J, Rodriguez-Cerezo E. 1999. Interactions between tombusviruses and satellite RNAs of tomato bushy stunt virus: a defect in sat RNA B1 replication maps to ORF1 of a helper virus. *Virology* 262:129–138. <http://dx.doi.org/10.1006/viro.1999.9865>.
66. Burgyan J, Hornyik C, Szittyta G, Silhavy D, Bisztray G. 2000. The ORF1 products of tombusviruses play a crucial role in lethal necrosis of virus-infected plants. *J. Virol.* 74:10873–10881. <http://dx.doi.org/10.1128/JVI.74.23.10873-10881.2000>.
67. Ribeiro D, Jung M, Moling S, Borst JW, Goldbach R, Kormelink R. 2013. The cytosolic nucleoprotein of the plant-infecting bunyavirus tomato spotted wilt recruits endoplasmic reticulum-resident proteins to endoplasmic reticulum export sites. *Plant Cell* 25:3602–3614. <http://dx.doi.org/10.1105/tpc.113.114298>.
68. Wei T, Zhang C, Hou X, Sanfacon H, Wang A. 2013. The SNARE protein Syp71 is essential for turnip mosaic virus infection by mediating fusion of virus-induced vesicles with chloroplasts. *PLoS Pathog.* 9:e1003378. <http://dx.doi.org/10.1371/journal.ppat.1003378>.
69. Pornillos OP, Garrus JE, Sundquist WI. 2002. Mechanisms of enveloped RNA virus budding. *Trends Cell Biol.* 12:569–579. [http://dx.doi.org/10.1016/S0962-8924\(02\)02402-9](http://dx.doi.org/10.1016/S0962-8924(02)02402-9).
70. White KA, Nagy PD. 2004. Advances in the molecular biology of tombusviruses: gene expression, genome replication, and recombination. *Prog. Nucleic Acid Res. Mol. Biol.* 78:187–226. [http://dx.doi.org/10.1016/S0079-6603\(04\)78005-8](http://dx.doi.org/10.1016/S0079-6603(04)78005-8).
71. Nagy PD, Pogany J. 2006. Yeast as a model host to dissect functions of viral and host factors in tombusvirus replication. *Virology* 344:211–220. <http://dx.doi.org/10.1016/j.virol.2005.09.017>.
72. Mullen RT, McCartney AW, Flynn CR, Smith GST. 2006. Peroxisome biogenesis and the formation of multivesicular peroxisomes during tombusvirus infection: a role for ESCRT? *Can. J. Bot.* 84:551–564. <http://dx.doi.org/10.1139/b06-005>.
73. Di Franco A, Russo M, Martelli GP. 1984. Ultrastructure and origin of cytoplasmic multivesicular bodies induced by *Carnation Italian ringspot virus*. *J. Gen. Virol.* 65:1233–1237. <http://dx.doi.org/10.1099/0022-1317-65-7-1233>.
74. Serva S, Nagy PD. 2006. Proteomics analysis of the tombusvirus replicase: Hsp70 molecular chaperone is associated with the replicase and enhances viral RNA replication. *J. Virol.* 80:2162–2169. <http://dx.doi.org/10.1128/JVI.80.5.2162-2169.2006>.
75. Ren X, Hurley JH. 2011. Proline-rich regions and motifs in trafficking: from ESCRT interaction to viral exploitation. *Traffic* 12:1282–1290. <http://dx.doi.org/10.1111/j.1600-0854.2011.01208.x>.
76. Dosztányi Z, Csizmok V, Tompa P, Simon I. 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21:3433–3434. <http://dx.doi.org/10.1093/bioinformatics/bt1541>.
77. Xue B, Williams RW, Oldfield CJ, Goh GK, Dunker AK, Uversky VN. 2010. Viral disorder or disordered viruses: do viral proteins possess unique features? *Protein Pept. Lett.* 17:932–951. <http://dx.doi.org/10.2174/092986610791498984>.
78. Stork J, Kovalev N, Sasvari Z, Nagy PD. 2011. RNA chaperone activity of the tombusviral p33 replication protein facilitates initiation of RNA synthesis by the viral RdRp *in vitro*. *Virology* 409:338–347. <http://dx.doi.org/10.1016/j.virol.2010.10.015>.
79. Nagy PD, Richardson CD. 2012. Viral replication—in search of the perfect host. *Curr. Opin. Virol.* 2:663–668. <http://dx.doi.org/10.1016/j.coviro.2012.11.001>.
80. Ross NT, Katt WP, Hamilton AD. 2010. Synthetic mimetics of protein secondary structure domains. *Philos. Trans. A Math Phys. Eng. Sci.* 368: 989–1008. <http://dx.doi.org/10.1098/rsta.2009.0210>.
81. McDonald B, Martin-Serrano J. 2008. Regulation of Tsg101 expression by the steadiness box: a role of Tsg101-associated ligase. *Mol. Biol. Cell* 19:754–763.
82. Gill DJ, Teo H, Sun J, Perisic O, Veprintsev DB, Emr SD, Williams RL. 2007. Structural insight into the ESCRT-I/II link and its role in MVB trafficking. *EMBO J.* 26:600–612. <http://dx.doi.org/10.1038/sj.emboj.7601501>.