UNIVERSITY OF CALIFORNIA, SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Dependence of vesicular stomatitis virus polymerase function on structural flexibility and domain-domain interactions

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Biology

by

John Barragan Ruedas

Committee in charge:

University of California, San Diego

Professor Clodagh O'Shea
Professor Amy Pasquinelli

San Diego State University

Professor Jacques Perrault, Chair
Professor Roland Wolkowicz, Co-Chair
Professor Ralph Feuer

2013

The Dissertation of John Barragan Ruedas is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California, San Diego

San Diego State University

2013

DEDICATION

I dedicate this dissertation to my mother, Angelina Barragan Ruedas, and my father, Johnny Garcia Ruedas, who have always encouraged and supported my aspirations to pursue a love of science and a love of life. Both have instilled in me strength and a desire to never give up on anything I choose to pursue. It is only with their constant love and support that I am able to achieve this degree.

I also dedicate this dissertation to my stepfather, Jose Carmona. Through our mutual love of music, he has always inspired me to aim for the highest-quality in anything I do.

I also dedicate this dissertation to my mentor and dear friend, Professor Jacques Perrault. He recognized in me an ability I didn't know I had. Guided by his wisdom and support, I have been able to achieve a goal I once thought was too high for me to reach.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Jacques Perrault for his support as the chair of my committee. His guidance, encouragement, friendship, and insight have with certainty proved to be invaluable.

I would also like to acknowledge all student and post-doctoral researchers of the Perrault lab whom I have been associated with for their assistance and friendship. We have shared many good times working together in the Perrault lab.

We thank Derek Ostertag for the construction of the VSVwt-EGFP virus and Domenico Spadafora for plasmid constructs used as starting material for the ones described in Chapter 1 of this dissertation.

We also thank the San Diego Chapter of Achievement Rewards for the College Scientist Foundation for scholarship support to John B. Ruedas for work done in Chapter 1 of this dissertation. The initial phases of work presented in Chapter 1 of this dissertation were supported by a grant from NIAID (RO1-A121572) to Jacques Perrault.

Chapter 1, in full, is a reprint of the material as it appears in Journal of Virology, Dec. 2009. Ruedas, John B.; Perrault, Jacques, American Society for Microbiology, 2009. The dissertation author was the primary investigator and author of this paper.

We thank C. Yong Kang for kindly providing plasmids used as starting material for the ones described in Chapter 2 of this dissertation. We also thank Mourad Bendjennat for kindly providing consultation for studies done in Chapter 2 of this dissertation.

We also thank the San Diego Chapter of Achievement Rewards for College Scientists Foundation for scholarship support to John B. Ruedas for work done in Chapter 2 of this dissertation.

Chapter 2, in full, is being prepared for submission for publication. The dissertation author was the primary investigator and author of this manuscript.

VITA

1995	Bachelor of Science, University of Texas, El Paso
2008-2013	Teaching Associate, San Diego State University
2004-2013	Graduate Research Assistant, San Diego State University
2013	Doctor of Philosophy, San Diego State University

PUBLICATIONS

Ruedas, J. B., and J. Perrault. 2009. Insertion of enhanced green fluorescent protein in a hinge region of vesicular stomatitis virus L polymerase protein creates a temperature-sensitive virus that displays no virion-associated polymerase activity in vitro. J Virol **83:**12241-52.

FIELDS OF STUDY

Major Field: Biology (Cell and Molecular Biology)

Studies in Molecular Virology
Professor Jacques Perrault

ABSTRACT OF THE DISSERTATION

Dependence of vesicular stomatitis virus polymerase function on structural flexibility and domain-domain interactions

by

John Barragan Ruedas

Doctor of Philosophy in Biology

University of California, San Diego, 2013
San Diego State University, 2013

Professor Jacques Perrault, Chair

Professor Roland Wolkowicz, Co-Chair

Gene expression and replication of nonsegmented negative strand (NNS) RNA viruses is performed by an RNA-dependent RNA polymerase. The polymerase large (L) protein catalyzes all functions including addition of a cap 1 (m^7GpppNm) structure on nascent transcripts. Alignment studies show L proteins are modular consistent with a multi-domain protein. Reports of L functions being coupled suggest orchestrated interactions among its domains occur however specific domain interactions have not been experimentally confirmed. At the start of this dissertation virtually no structural information was known for NNS L proteins other than an EGFP-tolerant hinge region just

upstream of a methyltransferase domain was discovered in the measles L protein. This dissertation shows a homologous hinge region is also present in the NNS prototype vesicular stomatitis virus (VSV) L protein and aims to gain fundamental insight into putative interactions involving domains near this hinge. Insertion of EGFP in VSV L produced a temperature sensitive polymerase in minigenome assays with minimal effects at 33°C but severe attenuation of RNA synthesis at 37°C. This phenotype was also observed in engineered recombinant virus encoding this modified polymerase. Moreover, this virus exhibited an inability to synthesize transcripts in vitro. The specific effect on RNA synthesis led to the proposal here that the insert restricts hinge movement thus impeding critical domain interactions. Four domains adjacent to the hinge were exchanged with VSV serotype counterparts to determine the effect on polymerase function. Exchanging the methyltransferase domain had no impact on L functions in minigenome assays and recombinant virus indicating it enjoyed a degree of autonomy. Exchange of the other domains however abolished RNA synthesis showing polymerase activity depends heavily on peptide sequences in these domains, supporting a model of interaction between these domains and other polymerase components. Combination exchanges of all four domains also yielded polymerase inactivity with one exception. Co-exchange of the methyltransferase and adjacent C-terminal domain partially rescued a defect that occurred by exchanging the C-terminal domain alone. This combination rescued replication activity

substantially more than transcription (avg. 67% and 13%, respectively) providing evidence of an essential domain-domain interaction for RNA synthesis that is more favorable for replication.

DISSERTATION INTRODUCTION

Vesicular stomatitis virus (VSV) is a well characterized livestock pathogen and prototype of the nonsegmented negative-strand (NNS) RNA viruses. This large group of viruses comprises four families (Rhabdoviridae, Paramyxoviridae, Filoviridae, and Bornaviridae) and includes the lethal pathogens rabies, Marburg, and Ebola. Although they have a diverse host-range these viruses share a common genetic arrangement and replication strategy. Each has a negative-sense RNA genome that is fully encapsidated by the virus nucleoprotein (N) forming a helical ribonucleoprotein (RNP) core (5). Gene expression and replication of the virus genome is carried out by an RNA-dependent RNA polymerase complex comprised of the virus large (L) protein and the essential phosphoprotein (P) cofactor. Two host proteins, elongation factor 1 (EF-1) and guanynyltransferase (GT), have been shown to bind to the polymerase complex (2, 8) but it is unclear whether their association is linked to function. The L protein (~250 kDa) serves as the exclusive catalytic protein of the complex and carries out multiple functions involved in RNA synthesis and transcript modification (5). Functions of the L protein differ biochemically from analogous cellular reactions making the L protein a promising target for researchers looking to develop antiviral therapeutics against these viruses. Protein alignments indicate the NNS L protein is a modular protein and genetic and biochemical studies suggest

important interactions occur internally within the protein. The L protein however has proven quite challenging to study for a number of reasons leaving many unanswered fundamental questions regarding its structure, function, and dynamics. This dissertation work was initiated with the main objective of gaining fundamental insight into the global structure of the VSV L protein and its internal dynamics.

Upon virus entry into the cell, the virion-associated polymerase complex initiates transcription near the 3'-end of the RNP template as it transiently separates small stretches of genomic RNA from its protective N protein sheath to generate monocistronic transcripts of each viral gene. The polymerase L-P complex binds to the RNP template though direct binding of the P protein with the N protein. Interestingly, the L-P polymerase complex fails to utilize naked RNA templates for effective RNA synthesis making it quite distinct from cellular polymerases. The VSV genome represents the most simple gene arrangement of the NNS viruses encoding five genes flanked by 47-nt leader and trailer promoter regions in the order 3'-*le*-N-P-M-G-L-*tr*-5'. All NNS viruses have a similar gene order of these core genes along with additional virus-specific genes positioned between the core genes. During transcription the polymerase dissociates from the RNP template at each gene junction approximately 30% of the time and reinitiates near the genome 3'-end resulting in a gradient where transcripts of genes proximal to the 3'-end accumulate in sequentially greater amounts than those of genes positioned

further away from the 3'-end. The multifunctional L protein also adds a 5' guanosine cap to nascent transcripts during their synthesis and catalyzes two distinct methyltransferase reactions producing the cap 1 (m^7GpppNm) structure (5). Although the virus cap structure is identical to cellular messenger caps, the catalytic reactions carried out by L are biochemically different from the analogous cellular processes. See references for a detailed review of the VSV L protein capping and methyltransferase processes (10, 12, 13) . The L protein also generates 3'-polyA tails on messengers through reiterative transcription of a short polyU tract at the end of each gene. As N protein accumulation increases during cell infection, a shift in the polymerase synthesis mode occurs inducing the polymerase to initiate exactly at the 3'end of the genome and read thru all gene junctions resulting in a complement (+) sense antigenome copy. The nascent antigenome strand is immediately encapsidated by N protein during its synthesis. The resulting antigenome RNP serves as a template for the polymerase to generate replicate copies of genomic (-) sense RNP that are subsequently packaged into newly-assembled virion particles. While the switch in RNA synthesis correlates with the accumulation of N protein, the underlying mechanism(s) that induces the polymerase from a transcription to replication mode is not well understood, although it has been proposed that a tripartite L-P-N complex may serve as a replication mode polymerase (9).

The discovery that L has several catalytic functions suggested early on that it is a modular protein with multiple active sites. Alignments of various NNS L proteins were in agreement with this notion as six conserved regions (CRI-VI) were reported (14). It is believed these conserved regions act as protein domains where the assorted functions of L are carried out. At the start of this dissertation it was known that a polymerase signature motif lies in CRIII which has been confirmed to contain the polymerase active site (16-18). Bioinformatic studies also reported that CRVI shared homology to a class of RNA 2'-*O* methyltransferases (1, 4). This prediction has since been confirmed as CRVI functions as a dual-specificity methyltransferase catalyzing both methylation reactions of the cap structure (7, 10). More recently, capping activity has been associated with conserved residues in CRV which functions as a polyribonucleotidyltransferase, a capping enzyme unique to the NSS viruses, that transfers the 5'-monophosphorylated end of the nascent transcript onto a GDP-acceptor (13). While CRV and CRVI are generally regarded as the capping and methyltransferase domains, respectively, the conserved regions of L do not fully fit the ambiguous definition of a protein domain. Even so, as 'domain' generally describes a distinct structural, functional, or binding unit of a protein, it may be used in place of 'region' in this dissertation. Functions of the other conserved regions of L have not been definitively established. In addition to the conserved regions, three large blocks of nonconserved sequence, each ~300 aa in length, are also present in

all NNS L proteins. The nonconserved regions are located at each of the terminal ends of the protein and between CRV and CRVI. The role of these nonconserved regions is not known but it is possible they also act as domains harboring functions more specific to individual viruses or virus subgroups.

Analyses of VSV L indicate that some of its functions appear coupled. For example, transcripts synthesized by VSV mutants and lacking 5' caps are prematurely terminated once they reach ~50 nt in length (19). This finding implies that an interaction between the polymerase domain and capping domain occurs to prevent synthesis of transcripts lacking the necessary 5'-cap. Also, studies show specific mutations mapped to the methyltransferase domain result in significant attenuation of RNA synthesis and in some cases induce hyperpolyadenlyation of messenger polyA tails (6, 10, 11). Although these correlations are not understood, these findings collectively show that changes in the capping and methyltransferase domain sequences influence activity of the polymerase domain and hence imply domain interactions are taking place. To date however no experimental evidence confirming the existence of and/or the functional significance of specific domain-domain interactions has been reported.

The large modular structure of L with possible multi-domain interactions suggests its overall structure is dynamic with multiple conformations. It is not surprising that vigorous attempts at crystallizing the full L protein, as well as subsections of the protein, have been unsuccessful. Consequently, at the

onset of this study, virtually no structural information was known for any NNS L protein (note that recently reported electron microscopy studies by others and discussed in Chapter 2 of this dissertation now provide a low-resolution structure of VSV L). A study with measles, a member of the morbilliviruses in the family Paramyxoviridae, provided the only experimentally-based insight into the structure of an NNS L protein. An alignment of morbillivirus L proteins identified a short variable region which was subsequently shown to tolerate insertion of EGFP without abolishing polymerase activity, although a recombinant measles virus encoding this modified L protein did exhibit moderate growth attenuation (3). This part of L just upstream of CRVI clearly showed remarkable flexibility leading the authors to propose that it functioned as a hinge. Soon after, studies on VSV L determined the nearby CRVI region was as a dual-specificity methyltransferase (7, 10). Morbilliviruses such as measles are not favorable systems for in-depth analysis of L functions, including methyltransferase activity. Thus the basis of the mutant virus growth attenuation and whether methyltransferase activity was affected by the insert was not determined. Moreover it remained unclear whether this proposed hinge was a structural feature unique only to the morbillivirus L protein.

The first part of this dissertation focused on determining if a homologous region of flexibility (i.e. hinge) exists in the NNS prototype VSV (Rhabdoviridae) L protein. The significance of this finding would show that a hinge just upstream the methyltransferase domain is found in L proteins of two

different NNS families, implicating it as an important structural feature preserved in the evolution of L. Moreover, determining any effect the EGFP insert has on L functions would be more feasible in VSV L and could provide new insight into the structure and function of this protein. I show in this study that a homologous region in the VSV L protein is indeed capable of tolerating insertion of EGFP thus demonstrating a putative hinge at this position of L is preserved across NNS families. I further show the EGFP insert does not have an effect on the nearby methyltransferase domain but instead causes a temperature-sensitive (ts) effect on RNA synthesis with robust polymerase activity at 33°C but a drastic reduction of activity at 37°C. The basis for the polymerase ts phenotype was not immediately clear but observations in these studies led me to favor a hypothesis proposing that at nonpermissive temperature movement of the hinge is restricted by EGFP thus disrupting domain-domain interaction(s) important for RNA synthesis. The lack of a fundamental understanding of domain interactions in L however only permitted this hypothesis to be supported by speculation at that time.

To establish a framework of knowledge that could support this hypothesis, the second half of this dissertation aimed to gain a basic understanding into putative interactions that involved domains positioned near the hinge. The primary objective was to determine which of these domains function autonomously and which function in a manner consistent with an interactive domain. Exploring the behavior of the domains relied on utilizing an

approach that was feasible within the constraints of L. Methods used to search for domain-domain interactions are generally the same techniques used in identifying protein-protein interactions. These include yeast or mammalian hybrid systems, co-immunoprecipitation, crosslinking, fluorescent resonance energy transfer (FRET), and bimolecular fluorescence complementation (BiFC). These approaches however provide a low probability of success or feasibility with regard to the intricacies of the L protein. Generally, these approaches require individual domains be stably expressed in their native state, fused or tagged with other peptides, and have a relatively high-affinity for their binding cognate. While there have been recent advances in stable expression of large fragments of L, expression of smaller fragments representing individual domains of L figures to be more technically challenging in the absence of a high-resolution structure. Moreover, L fragments fail to bind or bind poorly with the essential P cofactor (15). It is very possible domain interactions in L could be facilitated by P as recent EM analysis indicates P is involved not only with repositioning certain domains but in restructuring the shape of the domains themselves (15). Any involvement of P makes it virtually impossible to recreate the native conformations of the L-P complex if studying separately expressed L fragments fused to other proteins (e.g. two-hybrid systems). Moreover, a dynamic L protein with multiple conformations makes it virtually impossible to distinguish between orchestrated function-driven co-localization of domains versus domain co-localization that occurs by chance

amid the fluidity of the protein structure (e.g. crosslinking, FRET, and fluorescence techniques). Most important however, the likelihood of a dynamic L protein suggests its internal interactions occur transiently with low-level affinities that may fall under the detection limit of these techniques. Therefore, I felt it necessary to devise an approach compatible with L and its P cofactor that could provide data telling of domain autonomy or interactivity.

Many protein domains function independently from the rest of the attached protein and often can be swapped into or fused with other multi-domain proteins with no effects on the swapped domain or recipient protein. I reasoned that swapping domains of L with homologous counterparts from related L proteins could provide insights into which domains, if any, function autonomously. This approach is feasible in that it allows study of fully intact L proteins in their native state. Furthermore, in conjunction with established minigenome assays this approach is capable of assessing a functional output of L in response to any change brought about by the heterotypic domain, whether by an effect on domain interactions or by other effects on L. I therefore made VSV-serotype exchanges of four distinct domains near the hinge, either alone or in combinations, and assed these chimeric-L proteins for polymerase activity. I show here that L polymerase activity requires these domains (except the methyltransferase) be of the same serotype as the rest of the protein for L to be functional, indicating these domains are not autonomous from polymerase activity. I also provide experimental evidence of

an essential interaction between the methyltransferase domain and a C-terminal domain of unknown function. Restoration of this interaction by co-exchange of these two domains (homotypic to each other but heterotypic to the rest of the protein) allowed a partial recovery of RNA synthesis with a substantially greater rescue of replication activity over transcription activity. This is to my knowledge the first tangible evidence of a specific domain-domain interaction reported in the NNS L protein. As these two interacting domains are located downstream of the hinge region, their effect on RNA synthesis strongly implies they also interact with the remotely upstream polymerase domain and/or other components of the polymerase complex. These findings provide the foundations for a model of structural fluidity and orchestrated domain-domain interactions in the VSV L protein.

REFERENCES

1. **Bujnicki, J. M., and L. Rychlewski.** 2002. In silico identification, structure prediction and phylogenetic analysis of the 2'-O-ribose (cap 1) methyltransferase domain in the large structural protein of ssRNA negative-strand viruses. Protein Eng **15:**101-8.

2. **Das, T., M. Mathur, A. K. Gupta, G. M. Janssen, and A. K. Banerjee.** 1998. RNA polymerase of vesicular stomatitis virus specifically associates with translation elongation factor-1 alphabetagamma for its activity. Proc Natl Acad Sci U S A **95:**1449-54.

3. **Duprex, W. P., F. M. Collins, and B. K. Rima.** 2002. Modulating the function of the measles virus RNA-dependent RNA polymerase by insertion of green fluorescent protein into the open reading frame. J Virol **76:**7322-8.

4. **Ferron, F., S. Longhi, B. Henrissat, and B. Canard.** 2002. Viral RNA-polymerases -- a predicted 2'-O-ribose methyltransferase domain shared by all Mononegavirales. Trends Biochem Sci **27:**222-4.

5. **Fields, B. N., D. M. Knipe, and P. M. Howley.** 2007. Fields' virology, 5th ed. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia.

6. **Galloway, S. E., and G. W. Wertz.** 2008. S-adenosyl homocysteine-induced hyperpolyadenylation of vesicular stomatitis virus mRNA requires the methyltransferase activity of L protein. J Virol **82:**12280-90.

7. **Grdzelishvili, V. Z., S. Smallwood, D. Tower, R. L. Hall, D. M. Hunt, and S. A. Moyer.** 2005. A single amino acid change in the L-polymerase protein of vesicular stomatitis virus completely abolishes viral mRNA cap methylation. J Virol **79:**7327-37.

8. **Gupta, A. K., M. Mathur, and A. K. Banerjee.** 2002. Unique capping activity of the recombinant RNA polymerase (L) of vesicular stomatitis virus: association of cellular capping enzyme with the L protein. Biochem Biophys Res Commun **293:**264-8.

9. **Gupta, A. K., D. Shaji, and A. K. Banerjee.** 2003. Identification of a novel tripartite complex involved in replication of vesicular stomatitis virus genome RNA. J Virol **77:**732-8.

10. **Li, J., J. T. Wang, and S. P. Whelan.** 2006. A unique strategy for mRNA cap methylation used by vesicular stomatitis virus. Proc Natl Acad Sci U S A **103:**8493-8.

11. **Murphy, A. M., and V. Z. Grdzelishvili.** 2009. Identification of sendai virus L protein amino acid residues affecting viral mRNA cap methylation. J Virol **83:**1669-81.

12. **Ogino, T.** 2013. In vitro capping and transcription of rhabdoviruses. Methods **59:**188-98.

13. **Ogino, T., and A. K. Banerjee.** 2007. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. Mol Cell **25:**85-97.

14. **Poch, O., B. M. Blumberg, L. Bougueleret, and N. Tordo.** 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. J Gen Virol **71 (Pt 5):**1153-62.

15. **Rahmeh, A. A., A. D. Schenk, E. I. Danek, P. J. Kranzusch, B. Liang, T. Walz, and S. P. Whelan.** 2010. Molecular architecture of the vesicular stomatitis virus RNA polymerase. Proc Natl Acad Sci U S A **107:**20075-80.

16. **Schnell, M. J., and K. K. Conzelmann.** 1995. Polymerase activity of in vitro mutated rabies virus L protein. Virology **214:**522-30.

17. **Sleat, D. E., and A. K. Banerjee.** 1993. Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. J Virol **67:**1334-9.

18. **Smallwood, S., T. Hovel, W. J. Neubert, and S. A. Moyer.** 2002. Different substitutions at conserved amino acids in domains II and III in the Sendai L RNA polymerase protein inactivate viral RNA synthesis. Virology **304:**135-45.

19. **Stillman, E. A., and M. A. Whitt.** 1999. Transcript initiation and 5'-end modifications are separable events during vesicular stomatitis virus transcription. J Virol **73:**7199-209.

Chapter 1

Insertion of Enhanced Green Fluorescent Protein in a Hinge Region of Vesicular Stomatitis Virus L Polymerase Protein Creates a Temperature-Sensitive Virus That Displays No Virion-Associated Polymerase Activity In Vitro

John B. Ruedas and Jacques Perrault

Journal of Virology, Dec. 2009

JOURNAL OF VIROLOGY, Dec. 2009, p. 12241–12252 Vol. 83, No. 23
0022-538X/09/$12.00 doi:10.1128/JVI.01273-09

Insertion of Enhanced Green Fluorescent Protein in a Hinge Region of Vesicular Stomatitis Virus L Polymerase Protein Creates a Temperature-Sensitive Virus That Displays No Virion-Associated Polymerase Activity In Vitro[V]

John B. Ruedas and Jacques Perrault*

Department of Biology and Center for Microbial Sciences, San Diego State University, San Diego, California 92182

Received 20 June 2009/Accepted 19 September 2009

The RNA-dependent RNA polymerase of viruses belonging to the order *Mononegavirales* is part of a large multifunctional L protein that also catalyzes viral mRNA capping and cap methylation. The L protein of this diverse group of agents displays six blocks of conserved sequences. The precise relationship between these conserved regions and individual functions is largely unknown, except for "domain" VI that clearly encodes a viral mRNA cap methylase. The L protein of morbilliviruses (family *Paramyxoviridae*) was reported to tolerate insertion of the enhanced green fluorescent protein (EGFP) in a region just upstream of domain VI. Recombinant viruses with this insertion grow well in cell culture but are highly attenuated in animal hosts. We show here that the L protein of vesicular stomatitis virus (VSV), the prototype of the *Rhabdoviridae* family, also tolerates insertion of EGFP at a similar site. The modified protein (L_{EGFP}) and the resultant recombinant virus both demonstrated a sharp temperature-sensitive phenotype for polymerase activity, with reduced activity at 37°C and no activity at 37.5°C. Neither translation nor methylation of mutant virus transcripts was affected at 37°C. Curiously, mutant virus grown at permissive temperature contained about threefold-less L protein than the wild-type virus did and displayed no virion-associated polymerase activity in vitro. These findings support the notion that a flexible "hinge" region separates the cap methylase domain of L proteins from upstream functions and open up a number of avenues for studies of L-protein function in the more-tractable VSV model system.

Viruses whose genomes consist of nonsegmented negative-sense RNA are grouped together as four families within the order *Mononegavirales* (MNV) and include many important human disease agents, such as rabies virus, measles virus (MV), respiratory syncytial virus, Marburg virus, and Ebola virus. The RNA-dependent RNA polymerase encoded by these viruses lies within a larger protein (L) that is also responsible for addition of cap structures to viral mRNAs and their modification by methylation. How this large multifunctional protein orchestrates its various activities during transcription and replication of the viral genome is largely unknown. Vesicular stomatitis virus (VSV), the prototype of the *Rhabdoviridae* family, is a livestock pathogen in the Americas and has long served as a research model for MNV viruses (19). The VSV genome encodes five essential genes in the order 3′-N-P-M-G-L-5′ and exemplifies the simplest coding arrangement within MNV viruses. This report sheds light on VSV L-protein structure and function by demonstrating that a region immediately upstream of the domain involved in methylating viral mRNA caps tolerates insertion of enhanced green fluorescent protein (EGFP).

Six blocks of conserved sequences are shared among MNV virus L proteins (28), and these presumably reflect different functions. A polymerase signature motif lies in conserved region III, and mutagenesis studies have confirmed its essential role for viral polymerase activity (31, 34, 35). Recent studies suggest that conserved block V plays an essential role in capping of viral mRNAs (17, 23). Block VI, however, is the best-characterized region. It encodes the viral mRNA cap methylase (MTase), first recognized through sequence homology to other RNA methylases (2, 10). This has since been confirmed by mutagenesis studies of domain VI motifs predicted to be required for methylase activity for both VSV L protein (13, 18) and the Sendai paramyxovirus L protein (22). The methylase domain is so far the only L-protein region with discernible homology to proteins from other organisms and for which we can be fairly confident of its boundaries. The functions of conserved regions I, II, and IV remain uncertain. The active form of the L protein requires association with the P protein, and the binding site for P protein has been mapped to the first 350 to 380 N-terminal residues of L, at least in the case of paramyxoviruses (6, 14). An oligomerization site for paramyxovirus L proteins has also been mapped to this N-terminal region (5), although the significance of L oligomerization has recently been disputed (7).

Capping and polyadenylation of viral transcripts occur cotranscriptionally (37), which suggests that the separate catalytic activities of the L protein work in concert, conceivably involving dynamic changes in conformation to allow translocation of the growing RNA chain from one protein domain to another. Replication of the viral genome involves a process very different from transcription and is coupled to assembly of the N

* Corresponding author. Mailing address: Department of Biology, NLS 401, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182. Phone: (619) 594-5150. Fax: (619) 594-5676. E-mail: jperrault@sunstroke.sdsu.edu.

[V] Published ahead of print on 30 September 2009.

protein to growing nascent chains initiated at the very 3′ end of the template. Only encapsidated genomes and antigenomes can serve as template for the viral polymerase. What causes the L protein to change from a transcriptive mode to a replicative mode of synthesis and ignore gene junctions is still unclear. The stoichiometry of the L/P complex remains uncertain. Some studies suggest that the transcriptase enzyme is associated with the alpha subunit of host cell translation elongation factor 1 (EF-1) while the replicase consists of an L/P/N complex (29). Regardless of polymerase composition, the switch to replication can take place only after primary transcription of input virion templates and synthesis of viral proteins, since N protein is required for assembly of functional encapsidated templates. Perhaps not surprisingly, L proteins, or any part thereof, have so far resisted attempts at crystallization and/or structure determination.

Given its complexity, it is surprising that the L proteins of MV, rinderpest virus (RPV), and canine distemper virus (CDV), all of which are members of the *Morbillivirus* genus within the *Paramyxoviridae* family, retain function following in-frame insertion of the entire EGFP protein at a specific site within its polypeptide sequence (1, 9, 33). The insertion site was chosen on the basis of L-protein sequence comparisons of this very closely related group of viruses, which revealed two small nonconserved regions, H1 and H2, separating three much larger conserved stretches (20). The H2 region proved to be tolerant to EGFP insertion. Notably, the recombinant viruses encoding the modified L protein were reported to grow nearly as well as wild-type (wt) virus in cell culture, while RPV and CDV recombinants showed strong growth attenuation in infected animals (1, 33). These findings led Brown et al. and Silin et al. (1, 33) to suggest that modifying morbillivirus L proteins by insertions in the H2 "hinge" region offers a novel rational approach for vaccine development. Moreover, the EGFP-modified L proteins provide a means of tracking the virus polymerase in infected cells.

Unbeknownst at the time of the morbillivirus reports, the EGFP insertion site turns out to be immediately upstream of the MTase (domain VI). It remains unclear from these studies whether the EGFP insert alters proper functioning of the viral methylase or other L-protein activities and whether this might explain the attenuating effect on virus growth. We initiated the present study to determine whether the L protein of a virus belonging to a different family within MNV, namely, VSV, would also tolerate EGFP insertion upstream of the MTase domain. We show that this is indeed the case, but the resulting recombinant virus displays unusual properties, including a remarkable temperature-sensitive growth profile, a defect in L-protein packaging, and a surprising lack of virion-associated polymerase activity in vitro.

MATERIALS AND METHODS

Cell culture and viruses. All studies were conducted using baby hamster kidney cells (BHK-21) grown as monolayers in minimal essential medium (MEM) containing 7% newborn calf serum. The wild-type recombinant virus used in this report (VSVwt-EGFP) was described previously (25). This virus was derived from the original recombinant wt virus described by Lawson et al. (16) and encodes the EGFP protein from an additional transcription unit between the M and G genes. The growth properties of VSVwt-EGFP in BHK cells are identical to those of wt virus. The recombinant VSV-L_{EGFP} was also derived from the original wt recombinant virus as described below.

Plasmid constructions. To enable insertion of the EGFP coding sequence in the L gene, we first engineered a unique BstEII site upstream of the methylase domain. The pVSVFL(+) plasmid encoding the entire VSV genome (16) was digested with BamHI and self-ligated to generate a plasmid encoding the C-terminal 971 amino acids of the L protein. This plasmid then served as the template to introduce the BstEII site by site-directed mutagenesis (QuikChange PCR; Stratagene) using a plus-sense primer, 5′-GCTAAGGATAATAATAAggtaccCATGAGCTATCCCCCTTG (BstEII site underlined and mutated nucleotides in lowercase type), and its complement. The NgoMIV/SalI fragment containing the BstEII site was then used to replace the homologous fragment in the pGEM-L(His/HA) expression plasmid to generate pGEM-L_{D139VT}. pGEM-L(His/HA) was derived from the pGEM-LS1 expression vector used in our previous studies (3) by adding a 21-amino-acid-long sequence that encodes a histidine (His) tag followed by a hemagglutinin (HA) tag ($EH_6EVYPYDVP$ DYA) immediately following the L-gene start codon. pGEM-L(His/HA) showed the same activity as pGEM-LS1 in minigenome assays (not shown).

To construct pGEM-L_{EGFP}, the EGFP coding sequence was initially amplified by PCR from the pBI-EGFP plasmid (Clontech) using a mutagenic primer to change the start codon to an NcoI site (5′-TATTTTATccatggTGAGCAAGG-3′) and a primer 95 bases downstream of the stop codon (5′-CTGCATTCTAG TTGTGGTTTGTC-3′). The NcoI/AccI-cut PCR product was then cloned into the corresponding sites of the pTM1 plasmid (21). The pTM1-EGFP plasmid was then employed as a template to generate a PCR fragment encoding the EGFP open reading frame, minus the stop codon, and immediately flanked by BstEII restriction sites using primers 5′-GTTGCTAAGGATAATAATTTggtta ccATGGTGAGCAAG-3′ and 5′-ATTAAATAGCTTTCggtaccTTGTACAGC TCGTC-3′ (lowercase and underlined nucleotides indicate the BstEII site and EGFP coding sequences, respectively). The BstEII-cut PCR product was then cloned into the BstEII site of pGEM-L_{D139VT} to yield pGEM-L_{EGFP}.

To generate the pVSV-L_{EGFP} plasmid for recovery of infectious virus, we first removed the NgoMIV site in the plasmid backbone of pVSVFL(+) to generate pVSVFL*(+). The NgoMIV/SalI fragment from pGEM-L_{EGFP} containing the EGFP insert was then used to replace the corresponding fragment in pVSVFL*(+). Expected sequence changes in the final constructs were verified by sequencing (San Diego State University Microchemical Core Facility).

Minigenome assays. BHK cells grown to ~60% confluence in 5-cm plates were first infected with the vaccinia-T7 recombinant virus, vTF7.3 (11), at a multiplicity of infection (MOI) of 10 in the presence of 40 μg/ml cytosine arabinoside (Ara C) and 10 μg/ml DEAE-dextran. Plasmid transfection was carried as described previously using cationic liposomes (36), except for using 6 μg of pBS-N (16) in lieu of pGEM-N, 3 μg of either pGEM-L(His/HA), pGEM-L_{D139VT}, or pGEM-L_{EGFP}, and various amounts of pGEM-P. After 45 min of virus adsorption at room temperature, the inoculum was replaced with a 0.5-ml volume of transfection mix containing the above plasmids and cells were incubated for 30 min before the addition of 0.5 ml of fresh MEM containing 80 μg/ml Ara C. Following 4 h of incubation at 37°C or 33°C, 60 μl of a cell lysate containing a previously determined optimal amount of minigenome particles encoding the M and G genes (25) was added to the transfection mix and incubated for an additional 45 min before the addition of 2 ml of fresh MEM containing 40 μg/ml Ara C and continued incubation at 37°C or 30°C until 20 h posttransfection. Cytoplasmic extracts were then prepared, and either treated with micrococcal nuclease or left untreated to analyze for the presence of virus replication and transcription products, respectively, followed by RNA extraction and Northern blot analysis using glyoxal gels and [α-^{32}P]GTP-labeled T7 riboprobes corresponding to negative-sense and positive-sense M-gene sequences as described before (24). Labeled bands were quantified using a Storm phosphorimager (Molecular Dynamics).

Recovery of recombinant VSV-L_{EGFP} virus. BHK cells were first infected with vaccinia-T7 recombinant virus as described above for minigenome assays followed by transfection with support plasmids pBS-N (4 μg), pBS-P (10 μg), and pBS-L (6 μg) (16) and 8 μg of the VSV-L_{EGFP} template plasmid. Ara C (40 μg/ml) was added throughout infection and transfection to limit production of infectious vaccinia-T7 virus. At 48 h posttransfection, the supernatants were collected and clarified by centrifugation, and 1 ml was employed to infect fresh BHK monolayers followed by incubation at ~35°C. Successful virus recovery as judged by cytopathic effect (CPE) and green fluorescence was evident by 24 h and confirmed by plaque assay on BHK cells. Mutant virus produced blurred plaques with indistinct margins at 33°C and showed no plaques at 37°C. Clear plaques, however, were obtained by first incubating the plates at 33°C for 24 h and then changing the incubation temperature to 37°C for an additional 20 h. High-titer working stocks of VSV-L_{EGFP} virus (~5 × 10^8 PFU/ml) were

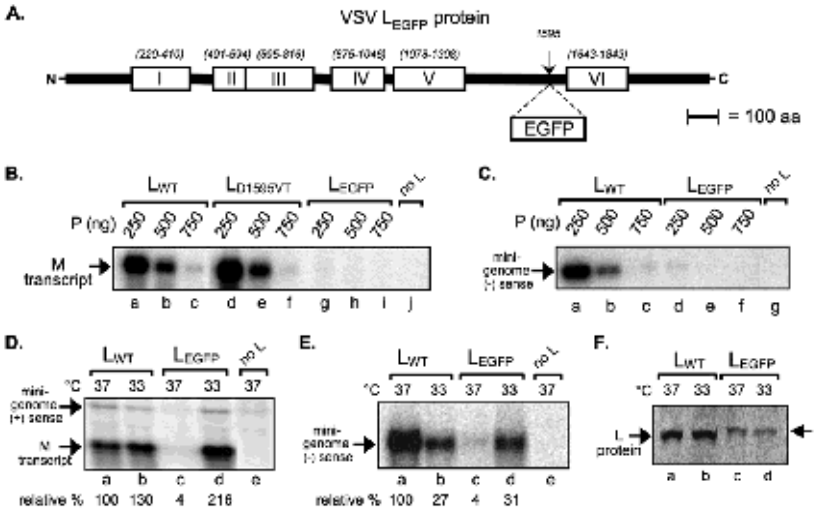

FIG. 1. Minigenome assays comparing viral RNA synthesis of mutant L_{EGFP} protein to the L_{WT} protein. (A) Conserved domains of the VSV L protein and the site of EGFP insertion. Domain positions are indicated in parentheses as reported by Poch et al. (28), except for domain VI, which is shown as reported by Bujnicki and Rychlewski (2). aa, amino acids. (B to E) Minigenome transcription (B) and minigenome replication (C) were measured at 37°C as a function of input P plasmid. The same activities were measured at 33°C versus 37°C using optimum amounts of P plasmid (250 ng) in panels D (transcription) and E (replication). Relative amounts, amounts compared to that of L_{WT} set at 100%, are shown below the lanes. (-) sense, minus sense; (+) sense, plus sense. (F) Accumulation of HA-tagged wt and mutant L proteins in the minigenome assays. M transcript and minigenome accumulation were assayed by Northern blotting, and viral protein accumulation was assayed by Western blotting, all at 18 h posttransfection.

obtained following two consecutive plaque-to-plaque clonings and amplification in BHK cells at 33°C.

Growth curve and infected-cell virus RNA analyses. All growth curves were carried out using BHK cell monolayers grown in sealed 25-cm^2 tissue culture flasks. Identical sets of flasks seeded in parallel were used for each time point and each temperature. Virus adsorption was carried out for 45 min at room temperature (MOI of 10) before the addition of 4 ml of MEM prewarmed to the indicated temperature followed by 1 min of equilibration with 5% CO_2 gas. The flasks were then sealed and immediately submerged in a circulating water bath adjusted to the desired temperature using a digital thermometer (Fisher) accurate to ±0.3°C. One flask from each temperature set was removed at each time point. The medium was clarified by centrifugation (600 × g, 5 min) and immediately frozen for subsequent determination of infectious virus titers by plaque assay or quantitation of released virus particles. Cytoplasmic extracts were then prepared from attached cell monolayers, combined with cells recovered from the medium by centrifugation (later time points only) and analyzed for virus RNA products as described above for minigenome assays, except for using a positive-sense L-gene riboprobe to detect replication products (24). Primary transcription analyses were carried out at an MOI of 50 and with addition of 100 μg/ml cycloheximide to the medium starting 30 min before virus infection and maintaining the same concentration throughout. Infected-cell RNAs were recovered using Tri-Reagent (Molecular Research Center, Inc.) and probed for M transcripts as for minigenome assays. Released particle genomes were analyzed by Northern blotting as described previously (24).

Western blotting. Immunoblots were conducted according to standard sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) procedures using 10% polyacrylamide gels and transfer to polyvinylidene difluoride membranes (GE Healthcare). Membranes were blocked with 1% bovine serum albumin. Anti-HA and VSV anti-G monoclonal antibodies (Sigma) were used at a dilution of 1:10,000, while anti-EGFP antibody (Roche) was used at a dilution of 1:1,000. VSV anti-L antibody prepared against gel-purified protein was a gift from Donald Summers and was used at a dilution of 1:1,000. Horseradish peroxidase-conjugated anti-mouse rabbit and anti-rabbit goat secondary antibodies (Jackson ImmunoResearch Laboratories) were used at a dilution of 1:10,000. Antibody detection was carried using ECL+ chemiluminescent substrate (GE Healthcare), and blots were quantified using the Storm phosphorimager (Molecular Dynamics).

Virus purification and in vitro transcription. Large amounts of VSVwt-EGFP and VSV-L_{EGFP} viruses were grown in BHK cells using 225-cm^2 flasks and purified using sucrose and tartrate gradients as described before (26). Standard in vitro transcription reactions, product purifications, and analysis on glyoxal gels were carried out as described previously (8), except for using 2 μCi [α-^{32}P]GTP per reaction mixture, transferring the gel-separated products to nylon membranes, and quantification using the Storm phosphorimager. Input genomes in the reaction mixtures were quantified using a positive-sense L riboprobe as described above for virus-infected replication products.

Fluorescence microscopy. Virus-infected cells in 5-cm dishes or 25-cm^2 flasks were examined in phase-contrast and fluorescence modes using an Olympus model IX70 inverted tissue culture microscope.

RESULTS

In-frame insertion of EGFP within the VSV L protein. Prior studies established that insertion of EGFP in the variable 23-residue-long H2 hinge region of morbillivirus L proteins has minimal effects on recombinant virus growth in cell culture (1, 9, 33). The VSV L protein, however, displays no obvious H2-like variable region. To engineer a similar insertion, we relied instead on the distance of the morbillivirus insert relative to the first lysine in the K-D-K-E catalytic tetrad of the downstream viral RNA methylase domain (2, 10). We thus inserted the EGFP domain 55 residues upstream of K1651 in the VSV L protein (Fig. 1A), compared to 56 residues upstream of

K1766 in morbilliviruses. To do so, we first introduced a unique BstEII restriction site that resulted in substitution of D1595 with valine and threonine residues ($L_{D1595VT}$). The EGFP coding sequence was then amplified by PCR along with flanking BstEII restriction sites, and inserted into $L_{D1595VT}$. The methionine start codon of the EGFP reading frame immediately followed T1596 of the $L_{D1595VT}$ construct, while the last EGFP codon was followed by valine and threonine residues before continuing into the L-protein reading frame. The predicted size of the L_{EGFP} protein is thus 241 amino acids larger than the wild-type L protein (L_{WT}) (2,350 compared to 2,109 amino acids).

The L_{EGFP} protein is functional but temperature sensitive in minigenome assays. We first determined whether the insertion affected VSV polymerase activity using the well-established vaccinia-T7 expression virus system to drive transcription and replication of a virus minigenome template that encodes VSV M and G genes flanked by genome leader and trailer sequences. For this purpose, we used L expression plasmids encoding His/HA tagged wt and mutant L proteins. Our previous studies showed that the presence of the His/HA tag at the N terminus of L_{WT} has no significant effect on VSV polymerase activities expressed in the vaccinia-T7 virus system (not shown). Plasmids encoding the VSV N, P, and His/HA-tagged wt or mutant L proteins were cotransfected into vaccinia-T7 virus-infected BHK cells and infected 4 h later with VSV minigenome virus particles (see Materials and Methods). RNA products were analyzed by Northern blotting 18 h posttransfection using ^{32}P-labeled negative- and positive-sense M-gene probes to detect minigenome transcription and replication products, respectively. At the temperature normally used for these assays (37°C), L_{EGFP} mutant protein showed only trace activity compared to L_{WT} for transcription (Fig. 1B, lanes g to i versus lanes a to c) as well as replication (Fig. 1C, lanes d to f versus lanes a to c). Peak activity in these assays critically depends on P-protein/L-protein ratio, presumably reflecting proper L/P polymerase complex formation. Separate experiments showed that overall activity decreased at lower P plasmid levels and optimal activity was reproducibly obtained using 250 ng of input P plasmid and 3 μg of L plasmid (data not shown). Less than optimal amounts of these plasmids or minigenome particles displayed no differential effect on transcription versus replication (e.g., Fig. 1B and C). Increasing P plasmid amounts did not compensate for the poor transcription and replication activity of the L_{EGFP} protein (Fig. 1B and C), suggesting no change in affinity for P. The $L_{D1595VT}$ protein, on the other hand, was slightly more active than the L_{WT} protein was for transcription (Fig. 1B, lanes d to f) as well as replication (not shown). The latter result indicates that a minor change in amino acid sequence at the EGFP insertion site does not adversely affect polymerase activity, and no further studies were carried out with this mutant protein.

Several repetitions of the above experiments revealed that while L_{EGFP} activity was severely compromised, it nonetheless functioned at a low but variable level. We speculated that L_{EGFP} might not fold correctly or function optimally at 37°C but perhaps show more activity at lower temperature. Indeed, incubation at 33°C dramatically increased L_{EGFP} RNA synthesis compared to incubation at 37°C (Fig. 1D and E, lane c versus lane d). In the experiment shown, L_{EGFP} transcription activity was 50-fold higher at the lower temperature and reached a level nearly twofold higher than L_{WT} at the same temperature (Fig. 1D, lane b versus lane d). L_{EGFP} replication activity was likewise much higher at 33°C and also reached a level comparable to the level of L_{WT} at the same temperature (Fig. 1E). Note that L_{WT} replication activity was substantially lower at 33°C than at 37°C in this experiment, but the relative activity at the two temperatures varied as much as fivefold in other experiments. Replication activity appears to be particularly sensitive to small fluctuations in temperature unavoidable in minigenome assays, as these entail multiple steps (vaccinia virus infection, transfection, minigenome particle infection). Most importantly, however, the transcription and replication activities of mutant L were always :::4% those of L_{WT} at 37°C and mutant transcription was consistently about twofold higher than wt transcription at 33°C in all individual experiments.

We have previously shown that VSV L protein is unstable with a half-life of 3 to 6 h in the vaccinia-T7 virus system unless coexpressed with the P protein (4). We therefore tested whether enhanced L_{EGFP} degradation might be responsible for its lack of activity at 37°C. Western blot analysis of L-protein accumulation in the minigenome assays (*N*-terminal HA tag) showed about two- to threefold-less L_{EGFP} than L_{WT} but no effect of temperature (Fig. 1F, lanes c and d versus lanes a and b). The levels of $L_{D1595VT}$ protein, on the other hand, were similar to those of L_{WT} (not shown). These results were obtained reproducibly in several different experiments all of which showed no evidence of L_{EGFP}-specific degradation products. We thus exclude degradation of L_{EGFP} protein as the cause for failed RNA synthesis at 37°C. Interestingly, although L_{EGFP} accumulated to levels lower than those of L_{WT}, it synthesized comparable or greater amounts of viral RNA than L_{WT} at 33°C. Conceivably, L-protein levels reach saturation in this assay or, alternatively, L_{EGFP} displays a higher specific activity than L_{WT}, especially with regard to transcription. Regardless, it is clear from these results that L_{EGFP} demonstrates a temperature-sensitive phenotype. The temperature defect clearly applies to replication, which is independent of minigenome transcript accumulation in this assay. However, a block in minigenome replication is also expected to block transcription, since transcription is proportional to the number of available minigenome templates. The results described below, however, strongly support the notion that L_{EGFP} is temperature sensitive for both transcription and replication.

Recombinant VSV-L_{EGFP} virus growth in BHK cells is temperature sensitive. We next attempted to recover a recombinant virus encoding L_{EGFP} protein as reported previously for morbilliviruses. We succeeded in doing so using standard methodology. The recombinant virus, denoted VSV-L_{EGFP}, was plaque purified twice serially and amplified to a working stock at 33°C (see Materials and Methods). VSV-L_{EGFP} plaques on BHK cells displayed a blurred appearance with indistinct margins when incubated at 33°C but showed no readily visible plaques at 37°C. The mutant virus, however, generated clear plaques by first incubating the plates at 33°C for 24 h and then incubating them at 37°C for an additional 20 h. We therefore used this latter procedure for determining the titers of both mutant and wt virus yields in all subsequent analyses. The wt virus employed for these analyses, VSVwt-

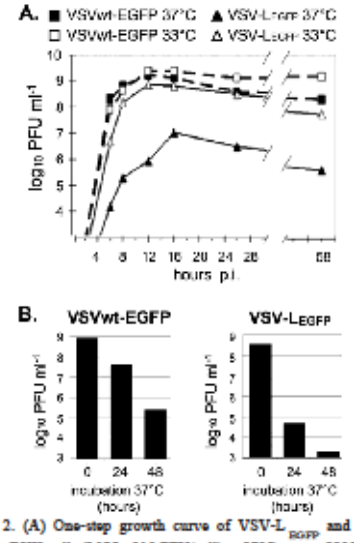

FIG. 2. (A) One-step growth curve of VSV-L_{EGFP} and VSVwt-EGFP in BHK cells (MOI of 10 PFU/cell) at 37°C versus 33°C. (B) Infectious virus titers of VSV-L_{EGFP} and VSVwt-EGFP following 1- and 2-day incubations of infected -ell lysates at 37°C. High-titer lysates were collected from cells previously infected at an MOI of 10 for 24 h at 33°C and clarified by centrifugation before initiating incubation at 37°C. All titers were determined by plaque assay in BHK cells (see Materials and Methods).

EGFP, is identical to the VSV-L_{EGFP} parent virus, except for an additional transcription unit encoding EGFP between the M and G genes. The growth properties of VSVwt-EGFP are identical to those of VSVwt as reported in our previous studies (25). Note that the L protein encoded by both VSVwt-EGFP and VSV-L_{EGFP} lack the His/HA tag present in the expression constructs employed in our minigenome assays. However, as noted above, the tagged form of wild-type L protein showed the same activity as untagged L in the vaccinia-T7 virus system.

We carried out a single-cycle growth analysis of VSV-L_{EGFP} versus VSVwt-EGFP in BHK cells (MOI of 10) at 33°C versus 37°C (Fig. 2A). The mutant grew at nearly the same rate as wt virus at 33°C, but maximum yields obtained at 12 to 16 h p.i. were three- to fourfold lower than those for the wt virus. As expected, the wt virus grew at essentially the same rate at both 37°C and 33°C and generated equivalent titers at 16 h postinfection (p.i.). In stark contrast, VSV-L_{EGFP} grew more slowly than VSVwt-EGFP did at 37°C, and the yields of the mutant virus were 2 log units lower at 16 h pi. The apparent leveling off of mutant virus titers 12 h p.i. at 37°C was not observed in an independent experiment and is likely due to sampling error. These results thus clearly establish that VSV-L_{EGFP} growth is temperature sensitive, as one might expect from the temperature sensitivity of the L_{EGFP} polymerase.

Recombinant VSV-L_{EGFP} virus is temperature labile. We noted in several different experiments that VSV-L_{EGFP} virus titers appeared to decline more than wt virus titers with prolonged incubation of infected-cell lysates, especially at 37°C. The data in Fig. 2B show that VSV-L_{EGFP} infectivity is substantially more temperature labile than VSVwt-EGFP at 37°C. Incubation of mutant virus-infected cell lysates for 24 h at 37°C resulted in a loss of infectivity of ≥4 log units compared to a loss of ≤1.5 log units for wt virus. VSV-L_{EGFP} also appeared slightly more temperature labile than the wt virus at 33°C (Fig. 2A). These results suggest that L_{EGFP} protein activity is somewhat less stable than L_{WT}.

Virus polymerase protein displays autofluorescence in VSV-L_{EGFP}-infected cells. Earlier findings with morbilliviruses showed that inserting EGFP into the L protein enabled localization of the virus polymerase protein in infected cells (1, 9, 33). This also proved to be the case for VSV-L_{EGFP}-infected BHK cells (Fig. 3). At 33°C, cells infected with mutant virus displayed progressive rounding of cells with accompanying fluorescence reaching a maximum at 12 to 16 h p.i. The signal appeared restricted to the cytoplasm, the expected location of the L_{EGFP} protein. Much less fluorescence was evident for VSV-L_{EGFP}-infected cells at 37°C, especially at 8 h p.i., in keeping with its temperature-sensitive phenotype. In contrast, VSVwt-EGFP-infected cells showed essentially equivalent fluorescence at both temperatures, and the signal in this case was evident throughout the cell as expected from passive diffusion of the small EGFP protein (27 kDa) through the nuclear pore. The stronger fluorescence signal observed for VSVwt-EGFP-infected cells presumably reflects, at least in part, the relative abundance of EGFP mRNA compared to the more promoter-distal L_{EGFP} mRNA.

Curiously, some VSV-L_{EGFP}-infected cells displayed an unusual type of cell-to-cell fusion at 12 and 16 h p.i. consisting of large syncytia with many closely packed nuclei surrounded by a halo of adherent and weakly fluorescent cytoplasm (Fig. 3, inserts). The majority of mutant virus-infected cells, however, showed rounded cell CPE very similar to VSVwt-EGFP-infected cells, except for the absence of fluorescence in nuclei. Some adherent cytoplasm also appeared to be present in wt virus-infected cell monolayers, but none showed the characteristically large structures observed in mutant virus-infected cells. The basis for the change in mutant virus-induced cell fusion characteristics is unknown. Clearly, more-detailed studies using higher-resolution microscopy will be necessary to better define changes in CPE as well as trafficking of the mutant virus polymerase in infected cells.

VSV-L_{EGFP} is genetically stable. The mutant virus stock used in the above studies was derived following two sequential plaque isolations (plaque to plaque) on BHK cells before amplification at 33°C. Three additional virus stocks amplified from independent plaques originating from the initial mutant virus recovery also displayed the same phenotype, i.e., temperature sensitivity and EGFP fluorescence in infected cells as described above. To more rigorously test the genetic stability of VSV-L_{EGFP}, we carried out four serial low-MOI passages of the mutant virus at the nonpermissive temperature of 37°C. The fourth passage yielded almost no virus (≤40 PFU/ml) in contrast to parallel passage at 33°C, which yielded ≥10^8 PFU/ml. The mutant virus passaged under nonpermissive conditions showed no change in phenotype compared to the phenotype of the original

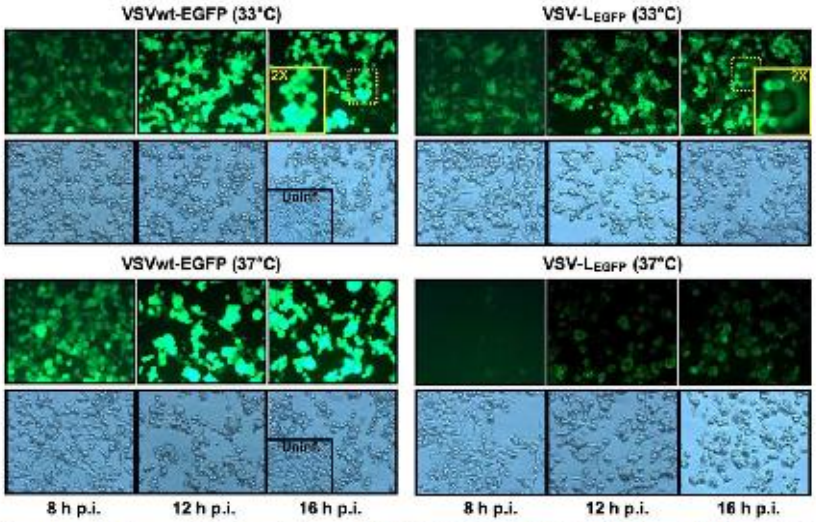

FIG. 3. Fluorescence and phase-contrast microscopic images of BHK cells infected with VSVwt-EGFP versus VSV-L_{EGFP} at 33°C and 37°C (MOI of 10). Identical sets of sealed tissue culture flasks were incubated at each temperature by submersion in circulating water baths. A single flask was used for observation at each time point. Phase-contrast images of uninfected cells (Uninf.) at 16 h are shown as inserts. Enlarged fluorescence images (enlarged 2x) contrasting CPE differences between VSVwt-EGFP and VSV-L_{EGFP} infections are also shown as inserts.

mutant stock. We therefore conclude that VSV-L_{EGFP} is genetically stable even when repeatedly subjected to nonpermissive growth conditions. Previous studies with the analogous morbillivirus recombinant viruses also indicate genetic stability both in cell culture and in infected animals (1, 9, 33).

VSV-L_{EGFP} RNA synthesis is delayed but not abolished at 37°C. Given the properties of the mutant L_{EGFP} protein in minigenome assays, we anticipated that viral RNA synthesis would also show temperature sensitivity in virus-infected cells. Infected-cell samples from the growth curve experiment in Fig. 2A were analyzed for M transcript and genome RNA accumulation by Northern blotting (Fig. 4). The accumulation of mutant M transcript at 37°C was clearly delayed compared to the accumulation of wt M transcript, but, surprisingly, mutant M transcripts reached the same level as that for wt transcript at 16 p.i. In contrast, mutant M transcript accumulation at 33°C was not delayed compared to wt M transcript accumulation and even showed somewhat more rapid kinetics (Fig. 4A). Replication products behaved a bit differently (Fig. 4B). Mutant genome RNA accumulation was about half that of wt genome RNA accumulation at 33°C with similar kinetics but was severely reduced at 37°C except again at 16 h p.i. where it reached ~25% of wt level. Similar results were obtained in several additional experiments, one of which included the same broad range of time points, while others including sampling at either one or two time points. The unexpected recovery of mutant viral transcription and partial recovery of replication at late times p.i. was reproducibly observed. We therefore conclude that the mutant viral polymerase can still function in infected cells at 37°C but with delayed kinetics.

Viral protein synthesis in VSV-L_{EGFP}-infected cells is proportional to transcript accumulation. Since inserting EGFP upstream of the viral L-protein methylase domain might adversely affect cap methylation and thus translation of viral mRNA transcripts, we also examined viral protein accumulation in the infected-cell samples analyzed above for viral RNA. Western blots of VSV G protein showed little difference in accumulation for wt virus-infected cells at 33°C versus 37°C, as expected (Fig. 4C, lower panel). For mutant virus samples, G-protein accumulation at 33°C was similar to that for the wt virus samples but was clearly delayed at 37°C. This delay correlated well with that seen for M transcript accumulation in the same infected-cell samples (Fig. 4A). The accumulation of L protein showed the same pattern as for G protein for all samples (Fig. 4C, top panel), except for the presence of a second lower-molecular-mass band (~180 kDa) for the mutant L_{EGFP} protein. This second band was absent in other experiments, including minigenome assays (Fig. 1F), and presumably originated from enhanced susceptibility of L_{EGFP} protein to cleavage during lysis of samples. Interestingly, the lower L_{EGFP} band was absent when the same blot was probed with anti-EGFP (Fig. 4C, middle panel), but a faint EGFP-reactive band of ~80 kDa was detectable (not shown), suggesting that cleavage takes place close to the EGFP insertion site. These results strongly suggest that viral protein translation, and by inference methylation of viral mRNA caps, is not affected by EGFP insertion in the polymerase protein at 37°C. However, these

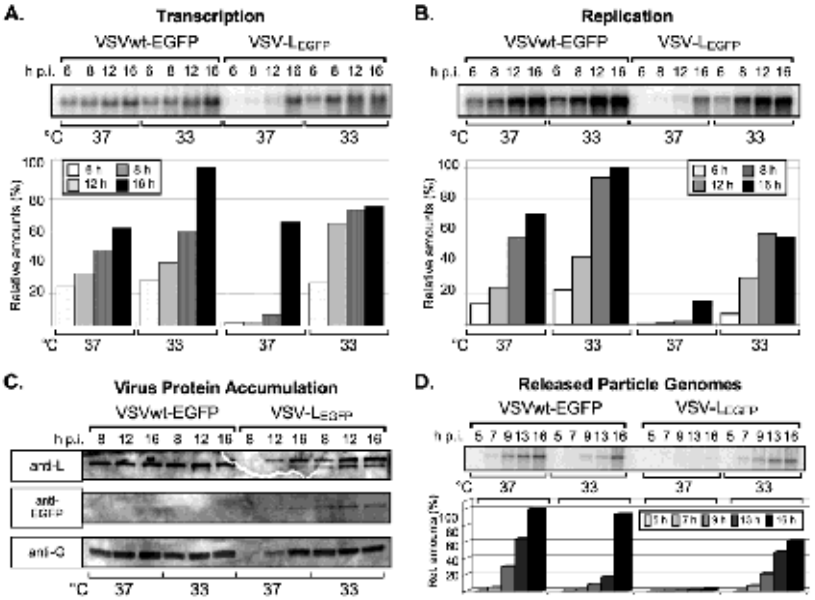

FIG. 4. Kinetics of virus RNA and protein accumulation in VSVwt-EGFP- versus VSV-L$_{EGFP}$-infected BHK cells (MOI of 10) at 33°C and 37°C. (A and B) M transcripts (A) and full-size virus genomes (B) were measured by Northern blotting. (C) Accumulation of L$_{WT}$, L$_{EGFP}$, and G proteins was assayed by Western blotting. A single blot was used in the order anti-G, anti-EGFP, and anti-L with no intervening antibody stripping. A band corresponding to EGFP was observed in VSVwt-EFGP-infected lysates only, as expected (not shown). (D) Virus particle release was assessed by Northern blotting of virus genomes present in clarified infected-cell supernatants. The data in panels A, B, and C were obtained from the samples used for the growth curve shown in Fig. 2A. Released particle data were derived from an independent growth curve experiment. Rel. amounts, relative amounts.

data do not rule out the possibility that the EGFP insertion does in fact compromise viral methylase activity and that a host cell enzyme compensates for this defect. The results of additional experiments not shown here argue strongly against this possibility (see Discussion).

Infectivity of VSV-L$_{EGFP}$ particles is comparable to wt virus. The above findings showed that infectious yields of VSV-L$_{EGFP}$ at 37°C were 2 log units lower than wt virus at 16 h p.i., but intracellular viral protein accumulation at this time was nonetheless comparable to that of wt virus, and virus genome accumulation was only fourfold lower. This discrepancy suggests either a block in assembly or release of mutant virus particles at 37°C and/or release of noninfectious particles. We tested the latter possibility by quantifying released virus particles from a growth curve experiment similar to that shown in Fig. 2A. Northern blot analysis of released virus particle RNA demonstrated a very good correlation with PFU titers for both mutant and wt viruses at either 33°C or 37°C (Fig. 4D). Very few mutant virus particles were released at 37°C, in agreement with a 2-log-unit decrease in released virus infectivity. These results indicate no change in the particle-to-PFU ratio for the mutant virus compared to wt virus. We therefore conclude that the block in infectious virus production for the mutant virus at nonpermissive temperature reflects, at least in part, a block in virus particle assembly and/or budding.

VSV-L$_{EGFP}$ growth and RNA synthesis display a very sharp temperature sensitivity profile. The above findings suggested that the VSV-L$_{EGFP}$ protein still functioned in mutant virus-infected cells at 37°C, albeit in a delayed or less-efficient mode. To shed light on what might cause the temperature-sensitive defect, we examined the effects of different temperatures on mutant virus infectious yields at 18 h p.i. (Fig. 5A). VSVwt-EGFP titers remained essentially the same from 31°C to 37.5°C, as expected, and decreased about 4 log units at 39°C. In contrast, VSV-L$_{EGFP}$ titers declined about 10-fold from 31°C to 35°C and then precipitously dropped about 5 log units at 37.5°C. No mutant virus was detectable at 39°C. Interestingly, VSV-L$_{EGFP}$ yields at 37.5°C were substantially lower than at 37°C (Fig. 2A), suggesting an unusually sharp decline over a very small temperature range. This acute temperature sensitivity was confirmed in an independent experiment measuring intracellular viral RNA accumulation in infected cells at 18 h p.i. over a narrow range of temperatures (Fig. 5B). VSV-L$_{EGFP}$ M transcript accumulation was essentially the same at 33°C and 35.5°C and decreased about twofold at 36.5°C. Remarkably, no transcripts were detectable at 37.5°C. The same pat-

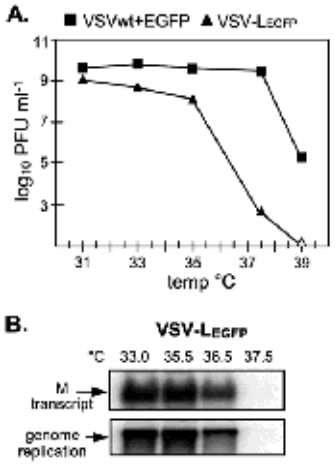

FIG. 5. (A) Temperature sensitivity profile of VSVwt-EGFP versus VSV-L_{EGFP} growth in BHK cells (MOI of 10). Titers at 18 h p.i. were obtained from identical sets of infected tissue culture flasks submerged in circulating water baths at the indicated temperature. The open triangle at 39°C denotes that no PFU was detectable at a 10 dilution of the infected-cell lysate. (B) An independent Northern blotting experiment carried out similarly and measuring accumulation of viral RNA M transcript (top panel) and full-size genome (bottom panel) at 18 h p.i. as a function of temperature.

tern was evident for virus genome RNA accumulation. The mutant L_{EGFP} protein thus displays a unique temperature sensitivity profile for viral RNA synthesis, with a cutoff very close to the physiological temperature. At 37°C, however, as documented above, substantial viral RNA synthesis takes place but is delayed.

VSV-L_{EGFP} virions are transcriptionally inactive in vitro. The above findings clearly showed that the L_{EGFP} protein displays temperature-sensitive polymerase activity both in minigenome assays and virus-infected cells. We thus expected that the L_{EGFP} protein packaged in virus particles would display temperature sensitivity for in vitro transcription. Surprisingly, detergent-activated VSV-L_{EGFP} virus particles consistently showed no polymerase activity in vitro at 30°C, the optimal temperature typically used for such assays (Fig. 6A). No significant incorporation of ^{32}P-labeled precursor above the background level was observed even after 18 h of incubation, and no RNA products were detectable by gel analysis. When transcription was carried out over a range of temperatures (Fig. 6B), VSVwt-EGFP virions showed optimal activity at 30°C, as expected, and ~80% less at 22°C. The mutant virus, however, was inactive at all temperatures tested. Similar results were obtained with an independent preparation of VSV-L_{EGFP} virus (originating from a separate plaque isolate) purified using extra precautions to keep the temperature as low as possible at all times and carrying out transcription assays immediately following partial purification (sucrose gradient only) or full purification (sucrose and tartrate gradients). Likewise, variations in assay conditions (concentrations of nucleotide, magnesium, salt, reducing agent, and detergent) or addition of uninfected BHK or HeLa cell extracts failed to uncover any polymerase activity. Curiously, despite this deficit, purified VSV-L_{EGFP} particles nonetheless productively infected BHK cells at 33°C (not shown). These results suggest that the mutant polymerase protein is very unstable under in vitro transcription conditions. Alternatively, some factor(s) normally required for polymerase activity may be present in wt virions but absent in mutant virions, or L_{EGFP} protein, but not L_{WT} protein, may require a factor(s) present only in host cells.

VSV-L_{EGFP} virions package reduced amounts of L polymerase protein. Given the absence of polymerase activity in VSV-L_{EGFP} particles, we wondered whether this defect might be due to abnormal packaging of viral proteins. SDS-PAGE analysis of purified virus revealed no significant difference in the content or ratio of G, N, P, and M proteins between wt and mutant virion, but surprisingly, the latter showed about threefold-less L protein (Fig. 6C). The L_{EGFP} band migrated slightly slower than the wild-type L band, as expected, with no evidence of cleavage products. These results were confirmed by Western blotting using anti-L and anti-P antibodies (not shown). To our knowledge, this is the first reported instance of a defect in packaging of the viral L polymerase protein in MNV. What causes reduced packaging of L_{EGFP} protein, however, remains unclear. This deficit could conceivably explain the lack of transcription in vitro if such activity requires a full complement of L.

Primary transcription in VSV-L_{EGFP}-infected cells is also temperature sensitive. Primary transcription of the VSV genome by the virion-associated polymerase is required to initiate virus replication in vivo. The mutant virus should thus be capable of carrying out primary transcription in infected BHK cells despite the lack of demonstrable virion-associated polymerase activity in vitro. Figure 6D confirms that VSV-L_{EGFP} does in fact synthesize M transcripts in infected cells when protein synthesis is blocked. For wt virus, transcript amounts increased from 5 to 10 p.i., as expected, and reached levels approximately twofold higher at 37°C than at 33°C. VSV-L_{EGFP} M transcript accumulation at 33°C was slightly delayed compared to that of wt virus at 5 h, perhaps a result of reduced amounts of L_{EGFP} proteins in virus particles, but nonetheless reached the wt level at 10 p.i. At 37°C, however, VSV-L_{EGFP} primary transcription was clearly reduced compared to 33°C, accumulating no more than 20% of the wt level even at 14 h p.i. This reduced and delayed accumulation of primary transcripts at 37°C is consistent with the mutant L behavior in standard virus infections, which also displayed delayed accumulation of transcripts at this temperature (Fig. 4A). The recovery in mutant virus transcription at 16 h p.i. in standard infections is consistent with delayed accumulation of newly synthesized L protein. We conclude that the polymerase packaged in VSV-L_{EGFP} particles is capable of catalyzing efficient primary transcription in vivo at the permissive temperature of 33°C despite the lack of virion-associated activity in vitro. Moreover, the substantially reduced primary transcription activity at 37°C indicates that the transcriptase activity of the L_{EGFP} protein, not just its replicase activity, is temperature sensitive in vivo.

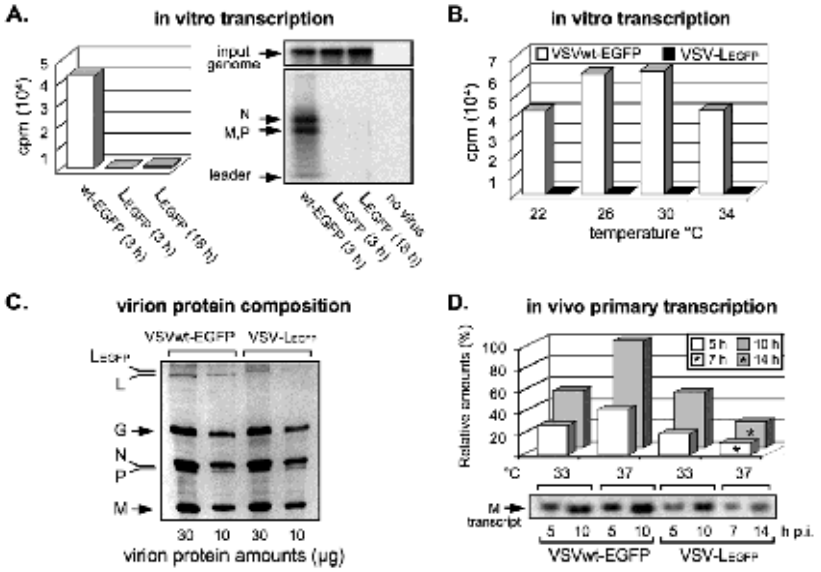

FIG. 6. Virion-associated polymerase activity and protein composition of VSVwt-EGFP versus VSV-L_{EGFP}. (A) In vitro transcription activity of detergent-disrupted purified virions was assayed by incorporation of [a-^{32}P]GTP using standard reaction conditions at 30°C. Transcription reaction products were separated on gels, and the presence of input template in the reactions was assessed by Northern blotting (see Materials and Methods). (B) Transcription activity was measured in parallel at different temperatures. (C) The protein composition of purified virions was assessed by standard SDS-PAGE and Coomassie blue staining. The primary transcription activity of the virion-associated polymerase in vivo (MOI of 50) in the presence of 100 µg/ml cycloheximide is shown in the Northern blot in panel D.

DISCUSSION

We initiated the above work to shed light on the domain structure of the VSV L polymerase protein. Absent structural information, it remains unclear whether the six blocks of conserved sequences shared by all MNV L proteins correspond to independently functioning "domains" or whether the conserved regions must interact with each other to orchestrate genome transcription and replication. On the basis of L-protein sequence alignments within *Morbillivirus*, a genus within the *Paramyxoviridae* family, McIlhatton et al. (20) identified two small hypervariable regions (H1 and H2), which they postulated function as "hinges" between conserved domains. The first of these lies between conserved blocks II and III, and the second lies just upstream of domain VI, which recent studies have clearly identified as the virus methylase responsible for modifying cap structures on viral mRNAs. Remarkably, recombinant morbilliviruses with an EGFP insertion in the H2 region (but not H1) grow nearly as well as wt virus in cultured cells (1, 9, 33). Although *Rhabdoviridae* L-protein sequences show no obvious hypervariable regions corresponding to H1 and H2, our findings here show that EGFP insertion just upstream of the VSV L-protein domain VI is also well tolerated. These findings provide strong support to the notion that all MNV L proteins harbor a flexible "hinge" region that separates the methylase domain from upstream functions.

Given the absence of an H2 region in VSV, we chose to insert EGFP 55 residues from the first lysine of the K-D-K-E catalytic tetrad in the methylase domain, which is very similar to the insert position used for morbillivirus L proteins (56 residues upstream). The resulting L_{EGFP} protein displayed a temperature-sensitive polymerase activity. This was evident in minigenome assays where transcription and replication activities of the modified polymerase were comparable to those of L_{WT} at 33°C but were reduced substantially at 37°C. Similarly, the yields of recombinant VSV-L_{EGFP} virus in BHK cells at 33°C were only slightly reduced compared to wt virus (~3-fold) but were much lower at 37°C (~100-fold). Mutant virus-infected cells accumulated viral RNA and protein to levels similar to those of wt virus-infected cells at the permissive temperature but far more slowly at 37°C and reaching wt levels only very late in infection (16 h p.i.). Primary transcription in VSV-L_{EGFP}-infected cells was also delayed and, in this case, reached only 20% of wt levels. Viral mRNA transcripts synthesized by the mutant L protein, including L mRNA, were translated as efficiently as those by the wt L protein at both 33°C and 37°C. The L_{EGFP} protein showed no evidence of degradation in infected cells. These results lead us to conclude that the mutant L_{EGFP} polymerase functions in infected cells at 37°C but does so with a considerable delay compared to the wt polymerase. Moreover, since mutant virus particle release and

infectious yields were both reduced ~100-fold at this temperature, we infer that delayed viral RNA and protein accumulation leads to a block in virus assembly and/or budding.

Efficient translation of VSV mRNAs, like that of most cellular mRNAs, requires a cap structure methylated at the guanine-*N7* position. To date, all available evidence suggests that the VSV methylase encoded in domain VI catalyzes both the guanine-*N7* methylation and 2′-*O*-adenosine methylation to produce a cap 1 structure (13, 18). Efficient translation of virus transcripts in VSV-L_{EGFP}-infected cells suggests that the viral methylase is not compromised by the EGFP insert. However, BHK cells also harbor an activity that can methylate VSV mRNA caps at the guanine-*N7* position, as shown long ago by Horikami and colleagues using the methylase-deficient VSV *hr1* and *hr8* mutants (15). Notably, the host cell activity does not catalyze 2′-O methylation of viral mRNA caps. We found no difference in the extent of cap 2′-O methylation of mutant virus transcripts versus wt virus transcripts isolated from infected BHK cells and no evidence for a compensatory host cell activity for methylation at this site (unpublished), in agreement with earlier findings (15). Our findings therefore strongly suggest that VSV-L_{EGFP} displays no methylation defect in vivo.

The EGFP insertion lies at the C-terminal end of the longest nonconserved region (~330 residues) evident in MNV L-protein alignments (between conserved blocks V and VI). Other nonconserved regions between conserved blocks are less than 100 residues in length and in some cases, much less. The large size of this nonconserved stretch suggests that it may nonetheless encode a function that differs substantially among MNV families. Interestingly, a recent study demonstrated that amino acid substitutions in a small region of this nonconserved region in the VSV L protein (spanning residues 1450 to 1481) adversely affect both guanine-*N7* and 2′-*O*-adenosine methylation of viral mRNA caps (12). Although the effect on methylase activity could reflect indirect consequences on protein folding, it is nonetheless possible that interaction between some part of the nonconserved region and domain VI is necessary for virus methylase activity. If so, our results also imply that this interaction is not disrupted by the EGFP insertion.

The temperature sensitivity of the VSV L_{EGFP} protein is remarkable in that even a slight increase in temperature above 37°C led to undetectable polymerase activity in infected cells and a drastic reduction in virus yields (~4 log units). This could well be due to misfolding and/or denaturation of the mutant protein above a critical temperature, but we favor an alternative explanation. We postulate that the hinge mediates a crucial interaction between the methylase domain and some upstream domain(s) of the L protein and that EGFP insertion weakens this interaction(s), making it more sensitive to temperature. It should prove instructive in this regard to determine whether the recombinant morbilliviruses with analogous EGFP-modified L proteins also show temperature sensitivity but with a higher threshold. The nature and size of the hinge region may also play a role in the outcome of EGFP insertion. For morbilliviruses, the hinge region is perhaps equivalent to the hypervariable H2 region of 23 residues, but the size of the hinge could vary substantially among different families and/or subgroups within MNV. The nature of the insert could likewise play a role in the outcome of insertions in the postulated hinge region. The compact β-can structure of EGFP, with flexible N- and C-terminal extensions protruding out of the same end of the barrel, is somewhat unique (27), and the effects of other types of inserts of different lengths and structure and at various positions upstream of domain VI should shed light on the nature and function of the hinge region.

Perhaps the most unexpected finding we report here is the apparent complete absence of mutant virion-associated polymerase activity. This is all the more puzzling, since the primary transcription of VSV-L_{EGFP} in vivo was similar to that of wt virus at 33°C and only partially reduced at 37°C. Primary transcription necessarily reflects the activity of the virion-bound polymerase after cell entry, since it is carried out in the presence of a protein synthesis inhibitor. Even so, all of our attempts to uncover polymerase activity in purified mutant virus particles, which included several variations in reaction mixture components, a broad range of temperatures (22°C to 34°C), and addition of uninfected-cell extracts, proved unsuccessful. The purified mutant virus preparations nonetheless retained infectivity. These findings suggest that the virion-associated L_{EGFP} protein is very unstable under transcription conditions in vitro but much less so in the context of infected cells. This is consistent with our finding that mutant virus infectivity was more labile than wt virus infectivity when infected-cell lysates were incubated for prolonged periods at 37°C. It may prove instructive to examine whether the mutant L protein produced in cells infected by the recombinant vaccinia-T7 virus retains activity when reconstituted with purified N-RNA templates in vitro (3).

Equally surprising and perhaps somehow related to the absence of virion-associated polymerase activity, L_{EGFP} protein was about threefold less abundant in virions than L_{WT} was, while other viral proteins showed no difference. To our knowledge, this is the first reported instance of a L polymerase packaging defect in MNV. Our minigenome assays are consistent with the notion that L/P complex formation was unaffected by the EGFP insert, since optimum polymerase activity required the same amount of P protein, but more-direct studies are needed to resolve this question. One possible explanation is that the mutant L/P complex is more easily disrupted when the viral M protein promotes condensation of the virus core at the time of packaging. Even so, the mutant L-protein packaging deficiency did not significantly reduce primary transcription at the permissive temperature in vivo, suggesting that virions package more L protein than required to initiate infection efficiently, at least in highly permissive cells. It is tempting to speculate that this packaging defect is more acute at higher temperatures and could thus explain a block in assembly and release of infectious virus from infected cells at 37°C. If so, the presence of properly folded L protein in virion cores may be a requirement for virus budding. It would be interesting to know whether similar virion-associated defects occur in the analogous recombinant morbilliviruses, but this may be technically challenging, since the latter infections yield far less virus than VSV infections and much lower levels of virion-associated polymerase activity.

As in the morbillivirus case, the recombinant VSV-L_{EGFP} protein proved to be autofluorescent in infected cells, which provides an ideal tool for tracking L polymerase protein dis-

tribution. The low-resolution images offered here are clearly only a beginning but do show that, at permissive temperatures, the L_{EGFP} protein is distributed more or less uniformly in the cytoplasm before maximum rounding of infected cells takes place (8 h p.i.). The same pattern is evident at 37°C but considerably delayed, in line with the Western blot results. Interestingly, VSV-L_{EGFP}-infected-cells also produced an unusual CPE. In contrast to the wt virus used throughout our studies reported here, which also encoded EGFP from an additional transcription unit, VSV-L_{EGFP} infections accumulated a significant number of large syncytia containing several nuclei surrounded by a halo of fluorescent cytoplasm. These were most evident by 12 h p.i. at 33°C. What causes this unusual CPE is unclear, since viral RNA accumulation and protein accumulation were very similar for both mutant and wt virus infections at this temperature. Brown et al. and Duprex et al. reported that the EGFP-tagged L protein in MV- and RPV-infected cells is present in cytoplasmic inclusions (1, 9) and also detectable on the inner leaflet of cellular membranes in RPV-infected cells (1). This was not evident in our VSV-L_{EGFP} infections, but more-detailed studies at higher resolution are needed to clarify the location and distribution of the tagged polymerase protein.

Although the EGFP-modified morbilliviruses show very modest growth attenuation in cell culture, the opposite is true in animal models of infection. The modified RPV shows strongly attenuated pathogenicity for cattle, although not quite as much as a current vaccine strain (1). The modified CDV provides significant protection against challenge virus in the ferret model but appears overattenuated for full protection (33). The more dramatic effects on mutant virus growth in vivo were deemed to result from inefficiencies in the replication machinery allowing more time for an adequate immune response. On the basis of our findings with VSV-L_{EGFP}, we suggest that temperature sensitivity of the virus polymerase may well be responsible for the postulated delay in replication of the modified morbilliviruses in vivo.

As pointed out by Brown et al. and Duprex et al. (1, 9), their studies provide the basis for a novel rational approach to making live attenuated morbillivirus vaccines using a single-step mutagenesis procedure. Importantly, the modified morbilliviruses proved to be genetically stable showing no reversion following passages in cell culture or, in the case of the modified RPV, following 9 days of infection in cattle. Likewise, we found no evidence of reversion for VSV-L_{EGFP} following sequential plaque isolation and passages in cell culture under nonpermissive conditions. This likely results from the improbability of mutational events that would restore wild-type polymerase activity by deleting most or all of the EGFP insert while keeping the correct reading frame. Our findings suggest that this vaccine approach may be feasible for MNV in general. VSV recombinants have also been given serious consideration in recent years as vaccine platforms to induce effective immune responses against a large variety of antigens, including influenza virus and human immunodeficiency virus (30, 32). Given its apparent growth attenuation but efficient protein expression at physiological temperatures, VSV-L_{EGFP} may serve as a desirable platform in these applications because of its ease-of-production advantages at permissive temperatures and very probable strong attenuation in humans. Characterization of VSV-L_{EGFP} pathology in animal models may shed light on its potential use in vaccine applications.

ACKNOWLEDGMENTS

We thank Derek Ostertag for the construction of the VSVwt-EGFP virus and Domenico Spadafora for plasmid constructs used as starting material for the ones described in this paper.

We also thank the San Diego Chapter of Achievement Rewards for College Scientists Foundation for scholarship support to J.B.R. The initial phases of this work were supported by a grant from NIAID (RO1-AI21572) to J.P.

REFERENCES

1. **Brown, D. D., B. K. Rima, I. V. Allen, M. D. Baron, A. C. Banyard, T. Barrett, and W. P. Duprex.** 2005. Rational attenuation of a morbillivirus by modulating the activity of the RNA-dependent RNA polymerase. J. Virol. 79: 14330–14338.
2. **Bujnicki, J. M., and L. Rychlewski.** 2002. In silico identification, structure prediction and phylogenetic analysis of the 2′-O-ribose (cap 1) methyltransferase domain in the large structural protein of ssRNA negative-strand viruses. Protein Eng. 15:101–108.
3. **Canter, D. M., R. L. Jackson, and J. Perrault.** 1993. Faithful and efficient in vitro reconstitution of vesicular stomatitis virus transcription using plasmid-encoded L and P proteins. Virology 194:518–529.
4. **Canter, D. M., and J. Perrault.** 1996. Stabilization of vesicular stomatitis virus L polymerase protein by P protein binding: a small deletion in the C-terminal domain of L abrogates binding. Virology 219:376–386.
5. **Cevik, B., S. Smallwood, and S. A. Moyer.** 2007. Two N-terminal regions of the Sendai virus L RNA polymerase protein participate in oligomerization. Virology 363:189–197.
6. **Chattopadhyay, A., and M. S. Shaila.** 2004. Rinderpest virus RNA polymerase subunits: mapping of mutual interacting domains on the large protein L and phosphoprotein P. Virus Genes 28:169–178.
7. **Chattopadhyay, S., and A. K. Banerjee.** 2009. Phosphoprotein, P of human parainfluenza virus type 3 prevents self-association of RNA-dependent RNA polymerase, L. Virology 383:226–236.
8. **Chuang, J. L., and J. Perrault.** 1997. Initiation of vesicular stomatitis virus mutant polR1 transcription internally at the N gene in vitro. J. Virol. 71: 1466–1475.
9. **Duprex, W. P., F. M. Collins, and B. K. Rima.** 2002. Modulating the function of the measles virus RNA-dependent RNA polymerase by insertion of green fluorescent protein into the open reading frame. J. Virol. 76:7322–7328.
10. **Ferron, F., S. Longhi, B. Henrissat, and B. Canard.** 2002. Viral RNA-polymerases – a predicted 2′-O-ribose methyltransferase domain shared by all *Mononegavirales*. Trends Biochem. Sci. 27:222–224.
11. **Fuerst, T. R., E. G. Niles, F. W. Studier, and B. Moss.** 1986. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. Proc. Natl. Acad. Sci. USA 83:8122–8126.
12. **Grdzelishvili, V. Z., S. Smallwood, D. Tower, R. L. Hall, D. M. Hunt, and S. A. Moyer.** 2006. Identification of a new region in the vesicular stomatitis virus L polymerase protein which is essential for mRNA cap methylation. Virology 350:394–405.
13. **Grdzelishvili, V. Z., S. Smallwood, D. Tower, R. L. Hall, D. M. Hunt, and S. A. Moyer.** 2005. A single amino acid change in the L-polymerase protein of vesicular stomatitis virus completely abolishes viral mRNA cap methylation. J. Virol. 79:7327–7337.
14. **Holmes, D. E., and S. A. Moyer.** 2002. The phosphoprotein (P) binding site resides in the N terminus of the L polymerase subunit of Sendai virus. J. Virol. 76:3078–3083.
15. **Horikami, S. M., F. De Ferra, and S. A. Moyer.** 1984. Characterization of the infections of permissive and nonpermissive cells by host range mutants of vesicular stomatitis virus defective in RNA methylation. Virology 138:1–15.
16. **Lawson, N. D., E. A. Stillman, M. A. Whitt, and J. K. Rose.** 1995. Recombinant vesicular stomatitis viruses from DNA. Proc. Natl. Acad. Sci. USA 92:4477–4481.
17. **Li, J., A. Rahmeh, M. Morelli, and S. P. Whelan.** 2008. A conserved motif in region V of the large polymerase proteins of nonsegmented negative-sense RNA viruses that is essential for mRNA capping. J. Virol. 82:775–784.
18. **Li, J., J. T. Wang, and S. P. Whelan.** 2006. A unique strategy for mRNA cap methylation used by vesicular stomatitis virus. Proc. Natl. Acad. Sci. USA 103:8493–8498.
19. **Lyles, D. S., and C. E. Rupprecht.** 2007. Rhabdoviridae, p. 1363–1408. *In* D. M. Knipe and P. M. Howley (ed.), Fields virology, 5th ed., vol. 1. Lippincott Williams & Wilkins, Philadelphia, PA.
20. **McIlhatton, M. A., M. D. Curran, and B. K. Rima.** 1997. Nucleotide sequence analysis of the large (L) genes of phocine distemper virus and canine distemper virus (corrected sequence). J. Gen. Virol. 78:571–576.

21. **Moss, B., O. Elroy-Stein, T. Mizukami, W. A. Alexander, and T. R. Fuerst.** 1990. Product review. New mammalian expression vectors. Nature **348**:91–92.
22. **Murphy, A. M., and V. Z. Grdzelishvili.** 2009. Identification of Sendai virus L protein amino acid residues affecting viral mRNA cap methylation. J. Virol. **83**:1669–1681.
23. **Ogino, T., and A. K. Banerjee.** 2007. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. Mol. Cell **25**:85–97.
24. **Ostertag, D., T. M. Hoblitzell-Ostertag, and J. Perrault.** 2007. Cell-type-specific growth restriction of vesicular stomatitis virus polR mutants is linked to defective viral polymerase function. J. Virol. **81**:492–502.
25. **Ostertag, D., T. M. Hoblitzell-Ostertag, and J. Perrault.** 2007. Overproduction of double-stranded RNA in vesicular stomatitis virus-infected cells activates a constitutive cell-type-specific antiviral response. J. Virol. **81**:503–513.
26. **Perrault, J., G. M. Clinton, and M. A. McClure.** 1983. RNP template of vesicular stomatitis virus regulates transcription and replication functions. Cell **35**:175–185.
27. **Phillips, G. N., Jr.** 1997. Structure and dynamics of green fluorescent protein. Curr. Opin. Struct. Biol. **7**:821–827.
28. **Poch, O., B. M. Blumberg, L. Bougueleret, and N. Tordo.** 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. J. Gen. Virol. **71**:1153–1162.
29. **Qanungo, K. R., D. Shaji, M. Mathur, and A. K. Banerjee.** 2004. Two RNA polymerase complexes from vesicular stomatitis virus-infected cells that carry out transcription and replication of genome RNA. Proc. Natl. Acad. Sci. USA **101**:5952–5957.
30. **Schell, J., N. F. Rose, N. Fazo, P. A. Marx, M. Hunter, E. Ramsburg, D. Monteflori, P. Earl, B. Moss, and J. K. Rose.** 2009. Long-term vaccine protection from AIDS and clearance of viral DNA following SHIV89.6P challenge. Vaccine **27**:979–986.
31. **Schnell, M. J., and K. K. Conzelmann.** 1995. Polymerase activity of in vitro mutated rabies virus L protein. Virology **214**:522–530.
32. **Schwartz, J. A., L. Buonocore, A. Roberts, A. Suguitan, Jr., D. Kobasa, G. Kobinger, H. Feldmann, K. Subbarao, and J. K. Rose.** 2007. Vesicular stomatitis virus vectors expressing avian influenza H5 HA induce cross-neutralizing antibodies and long-term protection. Virology **366**:166–173.
33. **Sillin, D., O. Lyubomska, M. Ludlow, W. P. Duprex, and B. K. Rima.** 2007. Development of a challenge-protective vaccine concept by modification of the viral RNA-dependent RNA polymerase of canine distemper virus. J. Virol. **81**:13649–13658.
34. **Sleat, D. E., and A. K. Banerjee.** 1993. Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. J. Virol. **67**:1334–1339.
35. **Smallwood, S., T. Hovel, W. J. Neubert, and S. A. Moyer.** 2002. Different substitutions at conserved amino acids in domains II and III in the Sendai L RNA polymerase protein inactivate viral RNA synthesis. Virology **304**:135–145.
36. **Spadafora, D., D. M. Canter, R. L. Jackson, and J. Perrault.** 1996. Constitutive phosphorylation of the vesicular stomatitis virus P protein modulates polymerase complex formation but is not essential for transcription or replication. J. Virol. **70**:4538–4548.
37. **Whelan, S. P., J. N. Barr, and G. W. Wertz.** 2004. Transcription and replication of nonsegmented negative-strand RNA viruses. Curr. Top. Microbiol. Immunol. **283**:61–119.

ACKNOWLEDGMENTS

We thank Derek Ostertag for the construction of the VSVwt-EGFP virus and Domenico Spadafora for plasmid constructs used as starting material for the ones described in Chapter 1 of this dissertation.

We also thank the San Diego Chapter of Achievement Rewards for the College Scientist Foundation for scholarship support to John B. Ruedas for work done in Chapter 1 of this dissertation. The initial phases of work presented in Chapter 1 of this dissertation were supported by a grant from NIAID (RO1-A121572) to Jacques Perrault.

Chapter 1, in full, is a reprint of the material as it appears in Journal of Virology, Dec. 2009. Ruedas, John B.; Perrault, Jacques, American Society for Microbiology, 2009. The dissertation author was the primary investigator and author of this paper.

Chapter 2

Appendage domain interactions in the vesicular stomatitis virus L polymerase protein

ABSTRACT

The large polymerase protein (L) of nonsegmented negative strand (NNS) RNA viruses catalyzes all RNA synthesis and transcript modifications needed for virus multiplication. The protein consists of a core polymerase domain with an attached multi-domain appendage encoding co-transcriptional viral mRNA capping and methylation functions. To gain insight into putative transient interactions between different L domains required for viral RNA synthesis, we determined the effects of exchanging each of the four appendage domains within the L protein of vesicular stomatitis virus (VSV) Indiana serotype with those of the VSV New Jersey serotype. The methyltransferase domain exchange yielded a fully functional polymerase protein in minigenome assays and a fully competent recombinant virus. However single replacement of other domains, including non-conserved regions of unknown function, abolished polymerase activity. Co-exchange of two or more domains also yielded inactive proteins, with the notable exception of the methyltransferase and C-terminal domain combination that generated a protein with substantially more replication than transcription activity. We also show that the large non-conserved region between the capping enzyme and methyltransferase domains does not appear to function purely as a hinge, as

previously proposed, since in-frame EGFP insertions upstream of the previously identified permissive site just upstream of the methyltransferase domain also abrogated polymerase activity. Our findings identify one essential and specific domain-domain interaction within the appendage of the L protein and suggest that other dynamic interactions involving the appendage structure play a role both in transcription and replication.

INTRODUCTION

The large protein (L) encoded by nonsegmented negative-strand (NNS) RNA viruses is the chief protein responsible for virus-directed RNA synthesis. Together with a multimer of the virus phosphoprotein (P), the L protein forms an RNA-dependent RNA polymerase complex and carries out all catalytic functions of the multifunctional P-L complex. These include synthesis and co-transcriptional modifications of virus mRNAs and replication of the virus genome. The L protein (~250 kDa) consists of discrete domains that catalyze the various functions performed during RNA synthesis. The domains are believed to interact based on evidence of coupling between distinct processes (9, 16, 32). The complexities of the L protein however make it challenging to identify specific domain-domain interactions using standard approaches, and little is known about the significance or specifics of these interactions. We initiated this study to gain basic insight into putative L protein domain-domain interactions required for viral RNA synthesis. By analyzing various chimeric L proteins containing exchanges of domains between related virus serotypes, we show that polymerase activity is highly dependent on homotypic integrity of some domains and provide evidence of a specific and hitherto unknown domain-domain interaction.

NNS viruses constitute a large group that includes a number of significant pathogens, most notably the lethal rabies, Marburg, and Ebola viruses. Diverse in their host range, NNS viruses share a similar strategy for

gene expression and replication. Much of what is known of their RNA synthesis machinery is derived from studies of the prototype vesicular stomatitis virus (VSV), an important livestock pathogen. The VSV negative-sense RNA genome encodes five genes in the order 3'-N-P-M-G-L-5' and is completely encapsidated by the virus nucleoprotein (N). The polymerase utilizes the N-encapsidated genome as a template and initiates transcription near the 3' end to generate monocistronic transcripts of each gene. Modifications of these transcripts, catalyzed by the L protein, take place during synthesis. These include a unique GDP:polyribonucleotidyltransferase reaction producing a 5'cap (21), two biochemically different methyltransferase reactions resulting in a cap 1 (m^7GpppNm) structure (10, 18, 25), and generation of a poly-A tail at the 3'end of viral mRNAs. Two host proteins, elongation factor-1 and guanynyltransferase, have been reported to bind specifically to the L-P polymerase complex (6, 11), but the possible functional implication of these associations remains unclear. As N protein accumulates in infected cells following synthesis of virus mRNA transcripts by the virion-associated L-P complex, the polymerase switches to a replication mode by initiating at the 3'end and reading through gene junctions to produce an N-encapsidated positive-sense antigenome strand. The antigenome then serves as a template for generating N-encapsidated negative-sense genome copies. The underlying factor(s) leading to the shift in polymerase synthesis is not well

understood although it has been proposed that a tripartite L-P-N complex serves as a possible replication-mode polymerase (12).

Biochemically different reactions catalyzed by L indicate it has multiple active sites, presumably functioning as a network of interacting protein domains. High-resolution crystal structures are lacking for NNS L proteins, however compelling evidence from other sources supports the concept that L is indeed structured as a multi-domain protein. Comparison of evolutionary divergent NNS L proteins shows partitioning of conserved sequences into six distinct blocks (CRI-VI) postulated to be domains (22). Functions have been firmly linked to three of these regions. CRIII contains a signature polymerase motif and has been confirmed as the polymerase domain (27, 29, 30). Polyribonucleotidyltransferase activity, a novel capping enzyme unique to NNS viruses, has been linked to CRV (21), and the dual-specificity methyltransferase has been mapped to CRVI (10, 18). Functions of the other conserved regions have not been definitively established. Three large blocks of nonconserved sequence (>250 aa) are also present in NNS L proteins positioned at both ends of the protein and between the capping (CRV) and methyltransferase (CRVI) domains. The role of these non-conserved regions is not yet clear but they could conceivably also serve as domains with virus-specific functions.

Consistent with the notion of a multi-domain protein, low resolution electron microscopy (EM) of the VSV L protein (minus P protein) show the N-

terminal ~1100 aa, which include the polymerase domain, forming a ring structure and the C-terminal ~1000 aa forming an attached appendage consisting of three globular structures (24). Notably, the ring and appendage structures must be joined together for L to be functional. The first appendage globule proximal to the ring structure was determined by deletion analysis to be the capping enzyme domain and the second globule was identified as the methyltransferase domain. The third globule mapped to the C-terminal non-conserved region of unknown function. For paramyxovirus L proteins this terminal region was proposed to serve as a GTP- or GDP-binding site involved in the capping function (20) but no evidence supporting this claim for L proteins of other NNS virus families has been reported. Interestingly, as viewed by EM analysis, addition of P protein causes a significant change in the appendage structure, making the globules indistinct from each other and converting the structure to a swirl attached to the ring structure (24). This suggests that P binding induces a rearrangement of the L appendage domains, although the location of the P protein multimer bound to L was not discernible.

The flexibility of the L protein appendage structure suggests the presence of hinge regions between the various domains. One such region, capable of accepting an in-frame EGFP insertion while retaining polymerase activity, was identified earlier and is located near the end of the non-conserved region between the capping enzyme and methyltransferase domains (2, 8,

28). This insertion-tolerant site was first described for paramyxovirus L proteins and subsequently by us for the VSV L protein (26). Interestingly, the non-conserved region as a whole did not form a visible structure in EM images suggesting it has a low level of structure consistent with its functioning as a hinge (24). The span of the hinge has not been established leaving it unclear if the entire non-conserved region, or just a part of it, acts as a hinge. Nonetheless, the hinge is postulated to be an important feature allowing for rearrangement of the appendage domains as illustrated by addition of the P protein in EM analysis.

All of the above supports the view that L operates as a dynamic protein with its domains interacting in orchestrated fashion and asserting influence over one another. Modifications such as capping and methylation occurring during transcript synthesis suggest that the domains carrying out these functions interact with the core polymerase domain. Consistent with this notion are reports showing premature termination of elongating transcripts lacking 5’caps (32) and point mutations in the methyltransferase domain that significantly affect polymerase activity (18, 19). A recent study shows that two fragments of L separated at the hinge and rejoined by dimerization tags require that they be derived from the same virus L protein to reconstitute RNA synthesis (7), indicating a virus-specific interaction between the two fragments. Moreover, these fragments failed to join without dimerization tags indicative of low affinity binding as expected for transient interactions. Identification and

characterization of domain-domain interactions in L proteins could therefore provide insight into the inner workings and dynamics of the polymerase complex and possibly reveal novel targets or strategies for antiviral therapeutics against this class of viruses. It is not clear however which specific domains or regions of L operate autonomously and which participate in interactions that impact function.

Techniques used in identifying domain-domain interactions traditionally involve molecular-based approaches such as two-hybrid systems, co-immunoprecipitation, crosslinking, and fluorescence resonance energy transfer (FRET). Such approaches typically require stable expression of protein domains as well as relatively high-affinity, high-specificity interactions. Only recently have large fragments of L been stably expressed (24) however expression of smaller fragments corresponding to individual domains has proven more difficult. Moreover, fragments of L often fail to bind or bind poorly to the essential P protein (24) and the latter could conceivably be involved in repositioning domains, particularly in the appendage region, to facilitate interactions. Therefore, at present traditional approaches are not likely to succeed in identifying domain-domain interactions in NNS L proteins.

We set out to determine which domains or regions in the L appendage structure operated autonomously and which exhibited a dependence on other regions of L by generating an array of chimeric L proteins derived from two well-characterized VSV serotype viruses, Indiana-serotype (IND) and New

Jersey serotype (NJ). By swapping NJ-based donor fragments into an IND-L recipient we show here that polymerase activity is, for the most part, highly dependent on homotypic combination of the appendage domains in conjunction with a homotypic ring (polymerase) structure. The methyltransferase domain is the only unit of the appendage structure that shows a degree of autonomy, as L remains fully active with a heterotypic VSV methylase in place. Furthermore, we show that a hitherto unknown interaction between the methyltransferase domain and the C-terminal domain is required for polymerase activity, and that the large non-conserved region upstream of the known hinge region likely plays some other role in polymerase function.

MATERIALS AND METHODS

Cell culture and viruses. All studies were conducted using baby hamster kidney cells (BHK-21) grown as monolayers in minimal essential medium (MEM) containing 7% newborn calf serum. Recombinant wt virus described by Lawson et al. (15) was used to compare growth kinetics of the recombinant IND VSV-L($_{NJ}$M) virus reported here. This virus was derived from the recombinant wt virus as described below.

Plasmid constructions. Plasmids encoding chimeric L proteins were engineered using a recipient pGEM-L plasmid (Indiana serotype, Mudd-Summers strain) that encodes L containing a his-HA tag at the N-terminal end as we reported in our previous studies (26). New Jersey L donor fragments derived by PCR or restriction enzyme digest originated from a pBS-L(NJ) plasmid (Ogden strain) kindly provided to us by Dr. C. Yong Kang. Many of the constructs were cloned using a unique BstEII site in IND-L (pGEM-L$_{BstEII}$) that we previously engineered (26). All sequences generated by PCR were sequenced as part of the final plasmid construct for the single region exchanges of C, H, M, and T (SeqXcel, Inc.).

C_{NJ}

To swap the NJ-cap domain (C) into IND L, we used the Circular Polymerase Extension Cloning (CPEC) method described by Quan and Tian (23). PCR products were generated with Phusion High Fidelity polymerase (New England Biolabs). Briefly, we generated two PCR fragments with homologous ends to each other. The first PCR fragment encoded NJ-C amino acids 1070-1311 immediately flanked by IND-L sequences (aa 1062-1069 and 1312-1320) using the primers *forward* 5'-cctctttgacacatttagggaaactt**CACTATCGTTTAGGGGATAATCAGATATGGTCC** and *reverse* 5'-tgggggcgtgtagtccatacttgagtc**CAGTTCTACTTCTTCAATCTCTCTAATGCATCC**; lower case represents IND nucleotides and bold caps represents NJ nucleotides. An NsiI-NsiI fragment was cut from pGEM-L (IND) and the truncated vector was self-ligated to yield a pGEM-LΔNsiI construct that was then used as a template to produce the second PCR fragment. The primers 5'-GAAGTTTCCCTAAATGTGTCAAAGAGGATACC and 5'-ATAGAAGAGATCACCCTGGACTCAAGTATGG (underlined nucleotides represent the complement of the IND sequences in the above forward and reverse primers, respectively) were used to generate the second PCR product yielding a linear copy of pGEM-LΔNsiI where nucleotides encoding IND C (aa 1070-1305) were omitted by primer design. The two PCR fragments were gel-purified and mixed at a 1:1 molar ratio in a PCR reaction where the fragments

were heat denatured, annealed to each other at both homologous ends, and extended to yield double nicked-circular pGEM-LΔNsiI($_{NJ}$C). An aliquot of this reaction was electroporated into XL-1 Blue *E.coli* (Invitrogen) to obtain covalently closed-circular plasmid copies. A BstBI-BstEII fragment containing NJ-C was cut from LΔNsiI($_{NJ}$C) and shuttled into pGEM-L$_{BstEII}$ cut with the same enzymes to yield pGEM-L(C_{NJ}).

U_{NJ}

The NJ-unstructured region (U) was swapped into IND L by the CPEC technique described above. A PCR fragment encoding the NJ (U) region that included additional residues extending into the C- and M-regions (aa 1293-1646) and immediately flanked by IND-L sequences (aa 1282-1292 and 1647-1659) was amplified using primers *forward* 5'-ctgttgcaagagacggatggatcaccagttgtaca**GATCATTACCATATCAGATGTAAAGG** and *reverse* 5'-ccattccatgtaatatactccgaattttataatgagcacc**AGTAGGCAGTTGTCCAAGC**; lower case represents IND sequences and bold caps represents NJ sequences. The pGEM-LΔNsiI template described above and primers 5'-CTGTACAACTGGTGATCCATCC

and 5'-CCTCCAAGAATCCAAAATCCC were used to generate a second PCR fragment corresponding to a linear copy of pGEM-LΔNsiI where nucleotide

sequences encoding the U-region (aa 1293-1627) were omitted by primer design. The underlined sequence in the first primer represents the complement of the IND sequence in the above forward primer; note the second primer initiates synthesis on the template at a position upstream of the IND sequence in the above reverse primer. The two PCR products were processed as described above to generate pGEM-LΔNsiI($_{NJ}$U). A BstBI-BspEI fragment containing NJ-U was cut from pGEM-LΔNsiI($_{NJ}$U) and shuttled into pGEM-L cut with the same enzymes to yield pGEM-L(U$_{NJ}$).

M$_{NJ}$ and VSV- L(M$_{NJ}$)

Cloning of the NJ methylase (M) domain into IND-L was done by standard restriction digest and ligation using the previously reported BstEII site and an NsiI site we engineered in this study by silent mutation of the codon encoding Y1863. The pBS-VSVFL(+) plasmid encoding the VSV (IND) genome (ref) and containing a BstEII site in the L gene was digested with BamHI and self-ligated to generate a mini-L plasmid encoding the C-terminal 971 amino acids of the L protein. This plasmid then served as the template to introduce the NsiI site by site-directed mutagenesis (QuickChange PCR; Stratagene) using a plus-sense primer 5'-GGAAAAACCTGTAtGCATTCCAGTCATCAGAACAGG (NsiI site underlined and mutated nucleotide in lowercase type) and its complement. The resulting

construct was a mini-L (BstEII/NsiI) plasmid. A PCR fragment was generated encoding NJ L amino acids 1596-1863 and immediately flanked by BstEII and NsiI restriction sites using primers 5'-CGATATAATGGTTACCTTTTATCCGAATTGGGGAACAGAGTACATTGG (BstEII site underlined) and 5'-CTTTGAAGCTTCTGAATGCATACTGTTTCTCC (NsiI site underlined). The BstEII-NsiI cut PCR product was then cloned into the respective sites in the mini-L (BstEII/NsiI) plasmid to yield a mini-L($_{NJ}$M) plasmid. A BstEII-AgeI fragment containing the NJ methylase was cut from mini-L($_{NJ}$M) and ligated into pGEM-L$_{BstEII}$ cut with the same enzymes to yield pGEM-L(M_{NJ}). To generate the pVSV-L(M_{NJ}) plasmid used for recovery of infectious virus, a BstEII-HindIII fragment containing the NJ methylase was shuttled from pGEM-L(M_{NJ}) into pBS-VSVFL(+).

T1$_{NJ}$

The NJ-terminal 1 (T1) coding sequence (aa 1863-1958) with an NsiI site placed immediately upstream was PCR amplified from the pBS-L(NJ) template using primers 5'-CAACTGGGAGAAACAGTATGCATTCAGAAGC (NsiI site underlined) and 5'-CACATTTTGTTGACATCCCCATCTGATGG which annealed downstream of a PflMI site common in IND and NJ that we designated as a separation point between T1 and T2. The NsiI-PflMI cut PCR

fragment was ligated into a mini-L(BstEII/NsiI) construct described above cut with the same enzymes to yield a mini-L($_{NJ}$T1) intermediate. A BstEII-AgeI fragment containing NJ-T1 was cut from mini-L($_{NJ}$T1) and ligated into a pGEM-L_{BstEII} recipient cut with the same enzymes to yield pGEM-L($T1_{NJ}$). Note, AgeI is present in the IND-T2 coding sequence.

$T2_{NJ}$

A PflMI-SpeI fragment containing the NJ-T2 coding sequence was cut from pBS-L (NJ) and ligated into the mini-L (BstEII/NsiI) plasmid cut with the same enzymes to yield mini-L($_{NJ}$T2). A BstEII-SpeI fragment containing the NJ-T2 was cut from mini-L($_{NJ}$T2) and ligated into a pGEM-L_{BstEII} recipient cut with BstEII and NheI (compatible ends with SpeI) to yield pGEM-L($T2_{NJ}$).

T_{NJ}

The NsiI-PflMI cut PCR fragment encoding the NJ-T1 fragment described above was ligated into the mini-L($_{NJ}$T2) recipient cut with the same enzymes to yield an intermediate mini-L($_{NJ}$T) construct containing the complete NJ T-region. A BstEII-SpeI fragment containing NJ-T was cut from mini-L($_{NJ}$T) and ligated into a pGEM-L_{BstEII} recipient cut with BstEII and NheI to yield pGEM-L(T_{NJ}).

$(MT2)_{NJ}$

A PflMI-SpeI fragment containing the NJ-T2 coding sequence described above was ligated into the mini-L($_{NJ}$M) plasmid cut with the same enzymes to yield mini-L($_{NJ}$M**T2**). A BstEII-SpeI fragment containing NJ-M and -T2 was removed from mini-L($_{NJ}$M**T2**) and ligated into a pGEM-L_{BstEII} recipient cut with BstEII and NheI to yield pGEM-L$(MT2)_{NJ}$

$(CMT)_{NJ}$

A PCR fragment encoding NJ-M and -T domains (aa 1596-2109) flanked by an upstream BstEII site and an SpeI site downstream of the stop codon was generated using the pBS-L(NJ) plasmid template and the primers 5'-CGATATAAT<u>GGTTACC</u>**TTTTATCCGAATTGGGGAACAGAGTACATTGG** (underlined and bolded nucleotides indicate the BstEII site and NJ coding sequence, respectively) and 5'-GGCTCGTATGTTGTGTGGAATTGTGAGCG (complementary to plasmid backbone). The BstEII-SpeI cut PCR product was then cloned into a pBS-L_{BstEII} recipient vector cut with the same enzymes to yield an intermediate pBS-L($_{NJ}$MT) construct. A BstBI-BstEII fragment containing the NJ-C domain was cut from pGEM-L($_{NJ}$C) and shuttled into pBS-L($_{NJ}$MT) to yield a second intermediate pBS-L($_{NJ}$CMT) construct. A BstBI-NotI fragment was cut from pBS-L($_{NJ}$CMT) and shuttled into pGEM-L($_{NJ}$T) cut with

the same enzymes to yield the pGEM-L(CMT)$_{NJ}$ plasmid. Note that NotI is specific to NJ-T coding sequence.

(MT)$_{NJ}$

A BstBI-BstEII fragment was cut from IND pGEM-L$_{BstEII}$ and shuttled into the pGEM-L($_{NJ}$CMT) plasmid described above cut with the same enzymes to yield pGEM-L(MT)$_{NJ}$, thus reintroducing IND-C back into the construct.

(CM)$_{NJ}$

A BstBI-BstEII fragment containing NJ-C was cut from pGEM-L($_{NJ}$C) and ligated into pGEM-L($_{NJ}$M) cut with the same enzymes yielding pGEM-L(CM)$_{NJ}$.

(CT)$_{NJ}$

A BstBI-BstEII fragment containing NJ-C was cut from pGEM-L($_{NJ}$C) and ligated into pGEM-L($_{NJ}$T) cut with the same enzymes yielding pGEM-L(CT)$_{NJ}$.

(CUMT)$_{NJ}$

An NgoMIV-NotI fragment was cut from pBS-L(NJ) and ligated into a pGEM-L($_{NJ}$CT) recipient cut with the same enzymes yielding pGEM-L(CUMT)$_{NJ}$. Note, NgoMIV site is located in NJ-C and NotI site is located in NJ-T coding sequences.

(CU)$_{NJ}$

A BstBI-SpeI fragment containing NJ-C and part of NJ-U was cut from pGEM-L(CUMT)$_{NJ}$ and ligated into pGEM-L($_{NJ}$U) cut with the same enzymes to yield pGEM-L(CU)$_{NJ}$.

(UT)$_{NJ}$

A BstBI-BspEI fragment containing NJ-U was cut from the pGEM-LΔNsiI($_{NJ}$U) construct described above and ligated into pGEM-L($_{NJ}$T) cut with the same enzymes to yield pGEM-L(UT)$_{NJ}$.

(UMT)$_{NJ}$

A SalI-BlpI fragment was cut from pBS-L(NJ) and ligated into pGEM-$L(_{NJ}UT)$ cut with the same enzymes yielding pGEM-$L(UMT)_{NJ}$. Note, SalI site is located in NJ-U and BlpI site is located in NJ-T.

L_{EGFP} variants

The five newly developed L_{EGFP} variants reported in this study were engineered using the CPEC technique described above. Briefly, each variant was made by first generating two PCR fragments with homologous ends to each other. The first fragment contained the EGFP coding sequence immediately flanked on both ends by in frame L sequences (~25 nt) corresponding to the indicated position of insertion. For example, the L_{EGFP}(1318) variant has the EGFP reading frame placed directly between codons for aa 1318(T) and 1319(P) of VSV-L. The second PCR fragment was a linear copy of pGEM-LΔNsiI described above where primers of opposite polarity were designed to generate a replicate copy of the template linearized at or near the indicated position of insertion. The two PCR products were processed as described above to generate a pGEM-LΔNsiI version with EGFP inserted at the indicated position. A fragment containing the full EGFP reading frame was then subcloned into pGEM-L (full-length) to produce the final L_{EGFP} encoding plasmid. All sequences generated by PCR and subcloned were

sequenced as part of the final plasmid construct (SeqXcel, Inc.). Primer and sequencing information for all L_{EGFP} constructs is available upon request.

Minigenome assays. Minigenome assays were done as we previously described (ref) except that transfection was carried out with polyethyleneimine (PEI). Briefly, cell monolayers were grown to 75% confluency in 6 cm dishes and transfected 18 hours prior to vaccinia-T7 infection. Eight microliters of a 7.5 mM solution of PEI (Polysciences, Inc. PEI “MAX” CAT. No. 24765) was mixed per µg of DNA in 0.4 ml of MEM without serum for 20 min. before adding to the cell monolayer in 4 ml of fresh MEM plus serum. Plasmid amounts were identical to those we reported previously unless otherwise noted (26). A constant PEI:DNA ratio was maintained when plasmid amounts were varied. Transfection supernatants were aspirated and cell monolayers washed twice with saline immediately prior to vaccinia-T7 infection. All experiments were incubated at 37°C for 5 hours post vaccinia-infection before minigenome particles were added and incubation temperature lowered to 33°C until harvest of cytoplasmic extracts 20 h post-infection. Northern blot labeled bands were quantified using a Typhoon phosphorimager (GE Healthcare).

Western blotting. Immunoblots were conducted according to standard sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) procedures using 10% polyacrylamide gels and transfered to polyvinylidene difluoride membranes (GE Healthcare). Membranes were blocked with 1% bovine serum albumin. Anti-HA antibody (Sigma) and horseradish peroxidase-conjugated anti-mouse antibody (Santa Cruz Biotechnology) were used at a dilution of 1:10,000. Antibody detection was carried out by autoradiography using ECL chemiluminescent substrate (Thermo Scientific Pierce) and film (GE Healthcare).

Recovery of recombinant IND VSV-L($_{NJ}$M) virus. Recovery of infectious virus was done as we previously described (26) except that transfection was carried out using PEI as described above. Note that support plasmids pBS-N (0.6 µg), pBS-P (1.4 µg), pBS-L (0.9 µg), and the full-genome template plasmid (1.1 µg) were transfected immediately after infection with vaccinia-T7 virus in this case. Successful virus recovery was judged by cytopathic effect (CPE) and confirmed by plaque assay on BHK cells. RT-PCR using AccuScript reverse transcriptase (Agilent Technologies) was carried out on infected cell extract to confirm the presence of the NJ-methylase coding sequence in the viral genome.

RESULTS

Effect of swapping L protein segments between VSV serotypes on virus polymerase activity. The amino acid sequences of the VSV IND L protein capping enzyme domain (C), “unstructured” domain (U), methyltransferase domain (M), and C-terminal domain (T) are 71%, 67%, 66%, and 42% identical to the homologous segments of the VSV NJ L protein, respectively. We surmised that swapping these segments between the two serotypes might shed light on sequence-specific interactions between appendage domains and/or with other components of the polymerase complex. Accordingly, we constructed a series of T7 polymerase-driven L protein expression plasmids each containing one of the above NJ-specific segments within the backbone of the IND L protein (see Figures 1 and 2). We then used the minigenome system described in our previous study (26) to monitor effects of these modifications on virus polymerase activity. This assay relies on co-transfection of T7 promoter-driven plasmids encoding VSV L, P, and N proteins into vaccinia-T7 virus-infected BHK cells, followed by infection with a virus particle containing a minigenome template. The latter contains all the RNA synthesis control elements found in the standard virus genome but encodes only the VSV M and G genes. Replication and transcription of the minigenome thus relies entirely on plasmid-expressed viral proteins and is assessed by Northern blotting using plus sense and minus sense M gene

probes, respectively. Note that transcription in this assay is detectable only if replication occurs first to provide the needed templates.

We previously found that insertion of the EGFP coding sequence in the VSV IND L protein upstream of the methyltransferase domain resulted in temperature-sensitive polymerase activity (26). To maximize detection of polymerase activity in the experiments reported here, we carried out all assays at the permissive temperature of 33°C. Figure 3 displays the effect of increasing amounts of wild-type L plasmid on accumulation of transcription and replication products under these assay conditions. Transcription reached a maximum with 7 ug input L plasmid while replication, as well as L protein accumulation, peaked at 3 ug. All experiments described in this paper, unless otherwise noted, used 3 ug L plasmid input and optimum amounts of P and N plasmid as determined previously (26).

When the C, U, and T appendage domains from VSV IND L protein were individually replaced with their VSV NJ counterparts, polymerase activity was abolished (Fig. 4A). In contrast, the M domain swap reproducibly yielded transcription and replication activities ranging from slightly lower to somewhat higher than wild-type IND L protein, as shown in Figure 4A. Lack of polymerase activity was clearly not due to defective NJ donor segments since the parent NJ L protein was fully active with VSV IND minigenome templates in the presence of NJ P and NJ N proteins (Fig. 4A, top panel). All chimeric L proteins accumulated about twofold less than wild-type L regardless of activity,

ruling out insufficient L protein expression for lack of polymerase activity (Fig. 4A, lower panel). Attempts to rescue activity for the C, U, and T swaps by increasing L plasmid input or varying P plasmid input proved unsuccessful, as illustrated in Fig. 4B for the U swap.

Although the active methyltransferase domain swap generated a competent polymerase, we could not exclude the possibility that the swapped domain itself was catalytically inactive since transcription of viral genes from the minigenome template or replication of the template do not require co-transcriptional methylation. We addressed this question by recovering a recombinant virus harboring the methyltransferase swap. This virus displayed no significant growth differences or defects compared to wild-type rVSV (not shown). The NJ methyltransferase domain within the context of a VSV IND L protein therefore fully functioned in all aspects of the virus replication cycle.

Since exchange the methyltransferase domain, in contrast to other domains of the appendage, could tolerate sequence differences between the two serotypes, we explored whether this relative degree of autonomy allowed for replacement with the more distant Ebola filovirus methyltransferase domain (22% identity). The resulting chimeric protein accumulated to the same extent as wild-type IND L but showed no activity in the minigenome assay (not shown). It remains to be determined if homologous methyltransferase domains from viruses more closely related to VSV can function in this context.

Loss of polymerase activity for the three other domain swaps suggested that these might participate in serotype-specific contacts either with each other or with some other component of the polymerase complex, and that such interactions are required at least for replicase activity. This was particularly surprising to us with regard to the C-terminal domain of the VSV L protein, whose function is unclear but shows very little sequence conservation between the two serotypes (42% identity). To determine if loss of activity mapped to a specific portion of the T domain, we swapped the first and second halves of the region separately. Neither the T1 nor the T2 sub-region exchanges yielded an active polymerase (Fig. 4C). These results in turn suggest that the T domain likely functions as a single unit, in line with the recent EM study showing its correspondence to a globule at the C-terminal end of the appendage region (24).

Effect of co-swapping multiple domains of the L protein appendage on virus polymerase activity. Hypothesizing that inactivity of the C, U, and T domain swaps might be due to loss of domain-domain interactions specifically within the appendage, we next tested various combinations of exchanged domains for polymerase activity. Compound swaps that included the entire appendage (CUMT), two triple-domain exchanges (CMT and UMT), and four double-domain exchanges (CU, CM, CT, and UT) all failed to rescue activity (Fig. 5). The one telling exception was the MT double-domain

exchange (Fig. 5). Intriguingly, this chimeric MT L protein rescued replication to a greater extent than transcription (67% ± 8.7% and 13% ± 5.3% of wild-type, respectively). Furthermore, no rescue was observed when only the T2 half of the NJ terminal domain was co-exchanged with the NJ M domain (Fig. 5). These results clearly show that the M and T domains must interact with each other in a serotype-specific manner to enable replication and transcription, and that the site(s) of interaction is either present in the T1 region or spans both T1 and T2 regions. Interestingly, functional interaction between these two appendage domains is retained when only the NJ methyltransferase domain is exchanged, as noted above. A summary of activity for all single and combination domain exchanges is shown in Figure 6.

Lack of activity resulting from most compound swaps is also revealing. For example, one would expect that exchange of the entire appendage would retain functional domain-domain interactions within the appendage. If so, loss of activity implies that the capping domain and/or other domains in the appendage requires serotype-specific interactions with the core polymerase and/or some other protein in the polymerase complex to sustain polymerase activity. Loss of activity for the UMT exchanges moreover suggests the large unstructured U domain plays a role in polymerase activity, either as an independent unit interacting with the core polymerase or in fashioning an active configuration of the appendage structure independent of the capping domain.

In-frame EGFP insertions in the "unstructured" domain block polymerase activity. We have so far referred to this region of the appendage as an unstructured domain based on previous EM studies, where this large non-conserved region (~330 residues) between the capping and methyltransferase domains could not be visualized and deemed to function as a large unstructured linker region (24). This assignment finds some support from studies providing evidence for a linker site near the C-terminal end of this region (just upstream of the methyltransferase domain) in VSV (26) as well as several paramyxoviruses (2, 7, 8, 28). Serotype-specificity for this unstructured domain in our swapping experiments was thus unexpected. To further explore the possibility of a linker function for this entire region, we extended our previous study showing that polymerase function is retained when EGFP is inserted at position 1595 of the VSV L protein. Accordingly, we constructed L proteins with EGFP at five different positions (1318, 1374, 1472, 1522, and 1577) spanning the unstructured region (see Fig. 7).

Surprisingly, none of the newly constructed L-EGFP insert variants showed any polymerase activity compared to wild-type L or the original 1595 insert construct (Fig. 7). As noted previously for the 1595 EGFP insert, the additional EGFP insert L proteins examined accumulated to levels 2- to 4-fold lower than wild-type L. The 1522 insert variant however accumulated about tenfold less protein than wild-type (Fig. 7), an amount nonetheless sufficient

for concluding it lacked significant polymerase activity. Reduced L protein accumulation for the EGFP insert mutants was observed repeatedly in independent experiments and correlated with EGFP fluorescence intensity in the transfected cells used in the minigenome assay (not shown). Moreover, increasing input L plasmid amounts from 3 µg to 7 µg failed to significantly increase L protein accumulation for the 1522 insert mutant (not shown). The basis for reduced L protein accumulation remains to be explored but could reflect mutant L protein instability due to defective P binding since the latter protein has been shown to stabilize L from degradation (3).

Notwithstanding, these results show that EGFP insertions throughout the “unstructured” region between the capping and methyltransferase domains is incompatible with replicase function and possibly transcriptase funciton as well. These findings along with our domain exchange results, suggest that the bulk of this “unstructured” domain does not function as a hinge. The latter appears to be restricted to a relatively short span of residues upstream of the methyltransferase domain. Likewise, lack of activity for an insertion only 13 residues downstream of the inferred boundary for the capping domain suggests either the absence of a linker in this region, or the presence of a more restrictive one. Interestingly, the “non-structured” region shows only moderately less sequence conservation than the methyltransferase domain within the *Rhabdoviridae* family, in contrast to the much more variable terminal domain region (not shown). Our findings also find support from an earlier study

where small 10-residue linker insertions spanning the corresponding domain of the measles virus L protein also generated inactive polymerases (7).

DISCUSSION

NNS virus L proteins encode all the enzymatic activities necessary to produce capped and methylated viral mRNA transcripts. These modifications take place on short nascent transcripts necessitating coupling to the polymerization process. EM studies of purified VSV L protein have shown that capping and cap methylation functions map to a flexible C-terminal appendage attached to a core polymerase ring structure (24). This overall architecture suggests that transient interactions between individual appendage domains and other components of the polymerase complex are necessary to orchestrate individual steps of transcription and replication. We sought to gain insight into these putative interactions by exchanging appendage domain components between VSV serotypes and determining effects on polymerase activity in a minigenome system.

The appendage region of the VSV IND L protein consists of four well-defined domains: cap addition enzyme, “unstructured” region, methyltransferase domain, and C-terminal domain. Each of these domains was swapped for the homologous counterpart from VSV NJ L, either one at a time or in various combinations. Our findings shed light on several different aspects of appendage functions. Perhaps the most significant of these is with regard to the methyltransferase domain and the C-terminal domain whose precise function remains unclear. While the methyltransferase could be

exchanged on its own without affecting polymerase function, all other single domain swaps generated inactive polymerases. However, co-exchange of both the methyltransferase and C-terminal domain produced a functional polymerase. This intramolecular complementation provides compelling evidence that a functional linkage exists between two distinct structural units of the L appendage and offers fresh insight into the dynamics of NNS virus L proteins.

VSV IND and VSV NJ serotype L proteins show 78% overall similarity (1) and undoubtedly catalyze essentially identical functions. Previous studies have addressed whether the VSV NJ L protein can function in the context of VSV IND P and N proteins. VSV NJ L was shown to substitute for IND L for in vitro transcription of IND N-RNA templates but only if both IND P and NJ P proteins are present (31). In contrast, VSV NJ L could not replicate IND VSV copy-back DI RNA templates in the presence of IND P and IND N proteins but could do so in the presence of IND P and NJ N (14). These earlier studies provide strong evidence that transcription requires a serotype-specific interaction between NJ L and NJ P proteins and that replication requires a serotype-specific interaction between NJ L and NJ N proteins. Loss of polymerase activity in the context of our studies here could therefore involve disruption of domain-domain interactions within the chimeric L proteins, or disruption of interactions with IND P or IND N proteins. These two possibilities are, of course, not mutually exclusive. Failure to recognize RNA template

sequence features, however, is clearly not involved since the VSV NJ L protein proved to be fully functional with the IND minigenome template when assayed with NJ P and NJ N proteins (Fig. 2).

We documented here that swapping the NJ methyltransferase domain into IND L protein had no apparent effect on L polymerase functions in the minigenome system, or in the context of a recombinant virus expressing the chimeric L protein. While this clearly shows the methyltransferase domain displays a relative degree of autonomy compared to other appendage domains, our findings also indicate that it must interact with the downstream C-terminal domain since swapping the latter on its own produces an inactive polymerase and swapping both rescues activity. Moreover swapping with the more diverged Ebola methyltransferase domain generated an inactive chimeric L protein in the minigenome system where catalytic activity of the swapped domain is not required for RNA synthesis. We presume that, in this case, the Ebola domain blocks essential appendage domain interactions either within L, or with some other component of the polymerase complex. A non-catalytic role for the methyltransferase domain is consistent with reports that mutations within the domain affect polymerase even though the latter does not depend on methylation activity (18). We believe these findings are in better agreement with a model where the methyltransferase domain must actively interact with the C-terminal domain of the appendage, and perhaps additional domains of L, to allow proper execution of transcription and replication.

Despite the capping domains of IND and NJ being more conserved than the methyltransferase domains (71% versus 67 % identity), the IND L protein failed to function as a polymerase with a NJ capping domain in place. The lack of a high-resolution structure for any NNS L protein makes it difficult to precisely map the various domain boundaries. In exchanging the capping domain we relied on reported L protein alignments (17, 22) as well as secondary structure predictions (JPRED 3 server) to select junctions just outside of the domain boundaries and anticipated to have minimal effect (Fig. 2). Even so, we cannot rule out that one or both junctions we chose for the capping domain were part of a critical structure destroyed by the exchange and thus causing the polymerase defect. Conceivably, swapping the capping domain at different junction sites could result in a functional polymerase. Notably, exchange of the complete appendage region (using the same N-terminal junction) also generated an inactive polymerase. In this latter case, interactions purely between appendage domains were presumably preserved. It is also important to keep in mind that loss of polymerase activity in the minigenome assay used here equates to a loss of replicase activity but not necessarily a loss of transcriptase activity, since the latter depends on accumulation of replicated templates. We therefore consider it more likely that the polymerase defect observed with our capping domain exchange is due to sequence differences that negatively affect an interaction between this domain

and the core polymerase required for replication, and perhaps also transcription.

All domain exchanges, except for the aforementioned methyltransferase domain swap and co-exchange of this domain with the C-terminal domain, generated completely inactive polymerases and thus affected replication. This suggests that all domains of the appendage may participate in fashioning a replication-competent form of the polymerase. It has long been speculated that the switch from transcription to replication in NNS viruses requires a conformational change in the polymerase protein, perhaps mediated by recruitment of an N protein subunit (12). It seems reasonable to expect a reconfiguration of the appendage structure when switching from transcription to replication in order to reposition the capping and methyltransferase domains away from the site of polymerization. We in fact showed previously that inhibiting proper functioning of a hinge region just upstream of the methyltransferase domain via EGFP insertion also blocks replication (26). Moreover, EM analysis of the VSV L protein provides direct evidence of reconfiguration of the globular appendage structure to a less well-defined flap over the polymerase ring domain in the presence of the P protein polymerase subunit.

Curiously, the C-terminal domain is the least conserved region of NNS virus L proteins (even VSV IND and VSV NJ show only 42% identity) but it nonetheless maps to a globular structure discernible by EM imaging of the

VSV L protein. No definitive function has yet been assigned to this domain but the equivalent region in paramyxovirus L proteins bears a motif associated with cellular capping enzymes (20). Mutation or deletion of this motif in the human parainfluenza virus type 2 L protein resulted in a transcription-defective but replication-competent polymerase, prompting the authors to propose that this region of L, although more than 500 residues downstream of the capping domain, is a GTP- or GDP-binding motif involved in the capping mechanism. This motif however is not present in NNS virus L proteins other than paramyxoviruses, and evidence for a transcription-specific function for the C-terminal domain of other NNS virus L proteins is lacking (but see below). Our swap of the C-terminal domain generated an inactive polymerase indicating loss of replicase activity but not necessarily transcriptase activity. The same result was obtained when either half of the C-terminal domain was exchanged separately, indicating the defect is not due to a localized stretch of variable residues. Our findings do not contradict the possibility of an accessory capping activity for this domain but do suggest the C-terminal domain has a role in specifying the replicase form of the L protein. Poor sequence conservation moreover suggests that this domain may function differently when comparing more distantly related NNS viruses.

As mentioned above, co-exchange of the C-terminal domain with the upstream methyltransferase domain did produce an active polymerase. Moreover, co-exchange of the NJ methyltransferase with its homotypic T2

sub-region of the C-terminal domain (e.g., separated by IND T1) also showed no polymerase activity, providing further evidence that the NJ methyltransferase was not exclusively responsible for activity in the co-exchange. We interpret these findings as compelling evidence that the NJ methyltransferase and NJ terminal domain in tandem offset the polymerase defect. We find it difficult to envision a scenario where the rescue of activity here is not the result of a functional linkage between the methyltransferase and the C-terminal domain. To our knowledge, this is the first evidence of a specific and essential interaction between two domains of the appendage region of NNS L proteins.

Notably, co-exchange of the above two domains restored polymerase activity asymmetrically: transcription was marginally rescued while replication showed significantly greater rescue (averaging 13% and 67% of wt activity, respectively). This phenotype is consistent with a model where the C-terminal domain plays a significant role in transcript modification such as capping while also being indispensable for replication activity. In each case interaction with the neighboring methyltransferase domain, by direct or indirect means, is likely essential. It is tempting to speculate that additional interactions involving other appendage domains, and perhaps the core polymerase itself, are required for full transcription activity.

An important aspect of our findings also addresses the role of the large non-conserved region (~300 amino acids) between the capping and

methyltransferase domains. We refer to this region as “unstructured” domain on the basis of EM analysis of the VSV L protein showing no discernible structure between the globule forming capping and methyltransferase domains, leading the authors to propose that the entire region consists of an unstructured hinge region (24). For the most part however, this region is not particularly enriched for residues predicted to be disordered in comparison to other domains of L (7). Moreover, for IND and NJ serotypes, this region is as well conserved as the downstream methyltransferase domain (67% versus 66% identity). Nonetheless, the hinge hypothesis is in line with recent studies with several NNS virus L proteins, including our own with VSV L, demonstrating the presence of a hinge near the very end of the region, ~ 50 residues upstream of the methyltransferase domain (2, 8, 26, 28). The span of this hinge region however was not determined in previous reports. As our studies here showed that swapping of the entire “unstructured” region between VSV serotypes inactivated the polymerase, it would seem unlikely that the entire region functions purely as a hinge.

To shed additional light on the function of the “unstructured” domain, we extended our previous approach using in-frame EGFP insertions as a probe for hinge regions. All EGFP insertions spanning the “unstructured” domain blocked polymerase activity, including one just 18 residues upstream of the known hinge site and one 13 residues downstream of the presumed capping domain boundary. These results provide further evidence that the bulk of the

“unstructured” domain does not function as a hinge and that the span of the hinge near the methyltransferase domain is restricted to a relatively small region. These observations are also in agreement with a recent report showing that insertion of small 10-residue linkers at various positions within this region of the measles virus L protein also inactivates polymerase function (7). Given this, it is somewhat puzzling that this region of the VSV L protein remained invisible by EM analysis. However, this conclusion was based on two constructs: one spanning the N-terminal 1114 residues and missing most of the capping domain, and one containing the N-terminal 1593 residues that included the capping domain as well as most of the downstream non-conserved region. The shorter fragment produced a slightly more open ring structure whereas the longer fragment generated the ring structure with the first appendage globule. On the basis of these constructs, it remains formally possible that the first globule closest to the core polymerase domain does in fact correspond to the non-conserved region whereas the capping domain is part of the ring structure. EM analysis of the N-terminal ~1310 residues where the capping domain but not the variable region is included could provide key information to confirm assignment of the first appendage globule to the capping domain.

Interestingly, the inactive EGFP insert constructs accumulated somewhat lower amounts of L protein than those harboring NJ region swaps, especially for the construct with EGFP inserted at position 1522. This

phenomenon could reflect a higher rate of turnover due to a defect in P protein binding, which is known to stabilize L from degradation (3). However, which region(s) of the VSV L protein is responsible for P binding is not entirely clear. Rahme and colleagues reported that the N-terminal portion of VSV L ending just upstream of the methyltransferase domain, and a C-terminal portion containing the remainder of the protein, both show weak P binding, while mixing of the two fragments fails to restore strong binding. In the case of rabies virus, the C-terminal 566 residues of L retain most of the P binding activity (5) while for paramyxovirus L proteins, the binding site(s) map to the N-terminal 400 residues (4, 13). Our findings here suggest that the non-conserved domain between the capping and methyltransferase domains may also play a role in P binding, either directly or perhaps indirectly by facilitating interactions between other domains of L.

In summary, the polymerase complex of NNS viruses is clearly a well-integrated machine with appendage domains that must reposition themselves to orchestrate the various steps involved in transcription and replication. Our study here provides a novel approach for mapping dynamic sites of interactions between individual L protein appendage domains and other components of the VSV polymerase complex. The sequence divergence between VSV IND and VSV NJ domains is sufficiently small to enable relatively straightforward identification of precise sites of interactions by site-directed mutagenesis.

ACKNOWLEDGMENTS

We thank C. Yong Kang for kindly providing plasmids used as starting material in Chapter 2 of this dissertation. We also thank Mourad Bendjennat for kindly providing consultation for studies in Chapter 2 of this dissertation.

We also thank the San Diego Chapter of Achievement Rewards for College Scientists Foundation for scholarship support to John B. Ruedas for work done in Chapter 2 of this dissertation.

Chapter 2, in full, is being prepared for submission for publication. The dissertation author was the primary investigator and author of this manuscript.

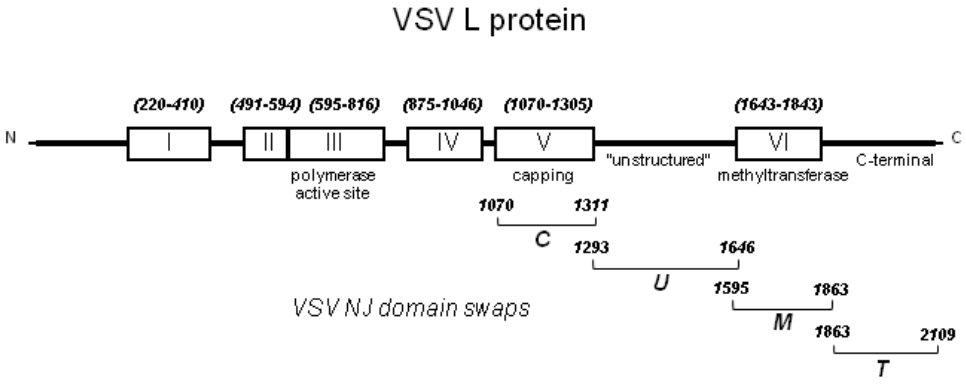

Figure 1. Schematic map showing conserved sequence blocks (I to VI) and assigned functions in the VSV L protein. Amino acid positions of conserved blocks I-IV are as reported by Poch et al., (1990), and blocks V and VI are as indicated by Li et al., (2005) and Bujnicki and Rychlewski (2002), respectively. VSV NJ L protein segments incorporated in place of their VSV IND L counterparts are abbreviated C, U, M, and T. The L protein of the two VSV serotypes is 66% identical and numbering position of conserved and non-conserved regions is identical.

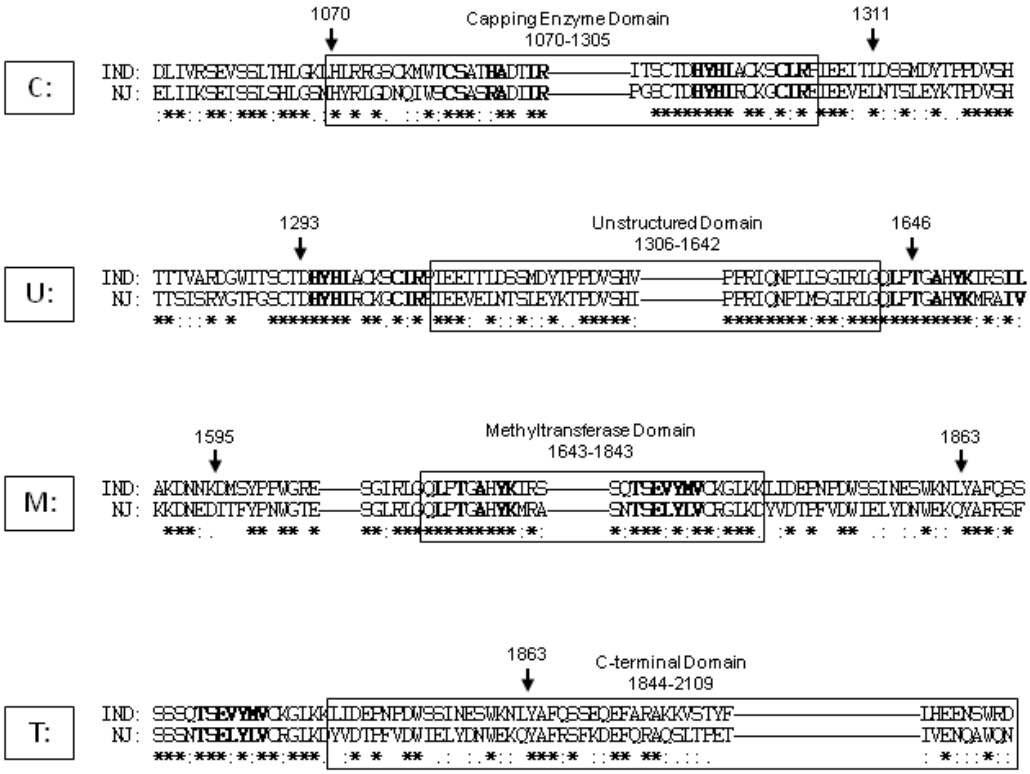

Figure 2. Comparison of VSV IND and VSV NJ amino acid sequences (GenBank accession no. Q98776 and P16379, respectively) showing junction site positions (arrows) for C, U, M, and T domain exchanges. Inferred domain sequences are boxed. The capping enzyme and methyltranferase domain borders were derived from the sequence alignments in Li et al., (2008) and Bujnicki and Rychlewski (2002), respectively. Alignments were based on ClustalW with conventional designation of identical (*), conserved (;) and semi-conserved residues (.). Consensus residues for the capping and methylttransferase enzymes are indicated in bold.

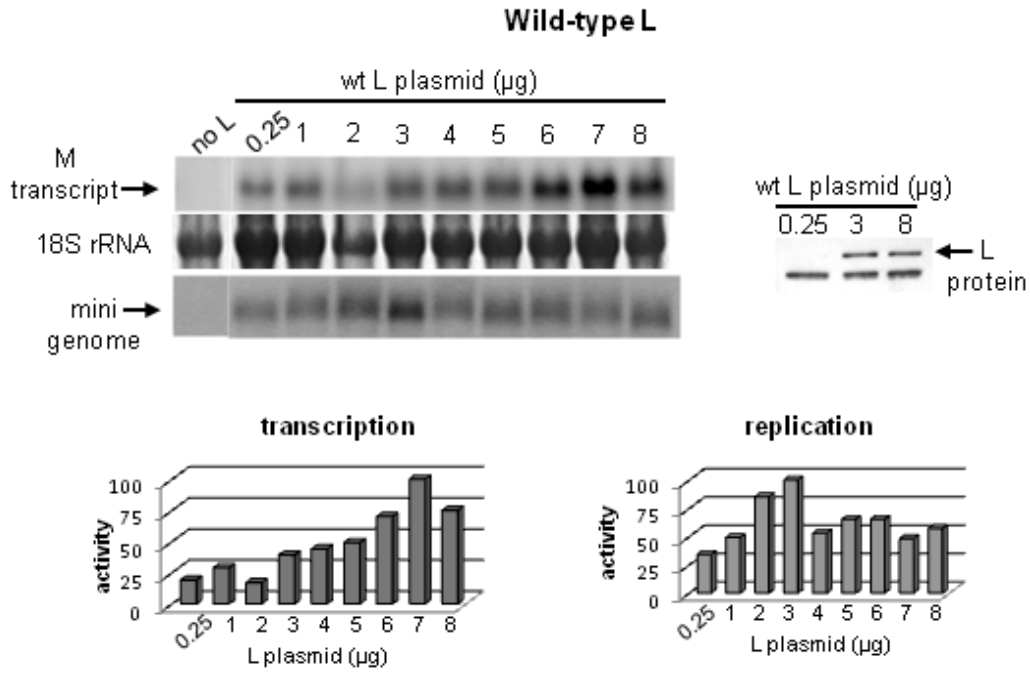

Figure 3. Minigenome assays measuring wild-type L transcription and replication activity as a function of input wt L plasmid. The transcription blot was initially stained with methylene blue to measure 18s RNA (middle panel). Transcription (upper panel) and replication (lower panel) RNA products were determined by Northern blot using ^{32}-P labeled riboprobes. Transcription was measured by probing for the viral M transcript and replication measured by probing for the negative-sense minigenome strand. Transcription values were normalized with the 18s RNA values in this experiment. Activity is expressed as arbitrary values with highest signal set at 100. Wild-type and mutant L proteins are HA-tagged at the N-terminus and their cytoplasmic accumulation measured by Western blot (upper right) and indicated by the arrow; lower band represents an undetermined cellular protein.

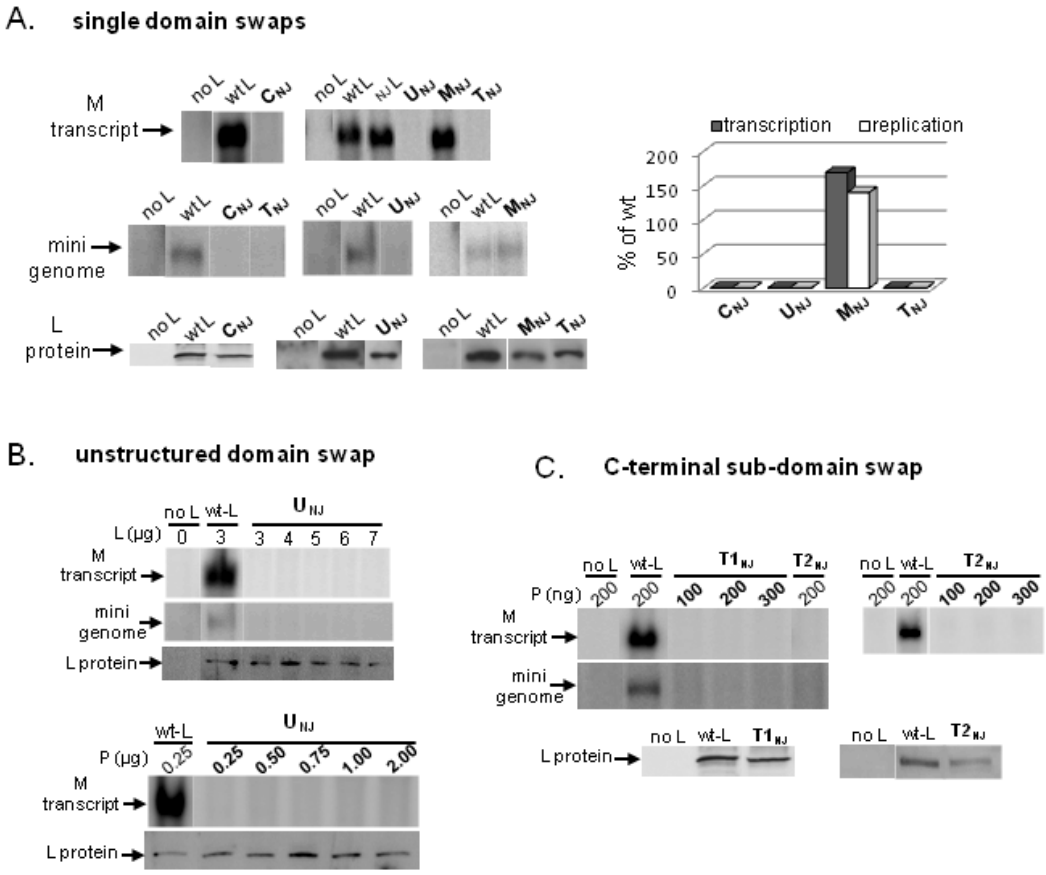

Figure 4. Minigenome assays comparing polymerase activity of single domain swapped L mutants to wild-type IND L protein. (A) Representative experiments showing transcription and replication activity of IND L mutants containing domains swapped with NJ counterparts: C-capping enzyme domain, U-unstructured domain, M-methyltransferase domain, T-terminal domain at C-terminal end. The graph reports mutant activity versus wt L for the representative experiments shown. All activities reported except for M_{NJ} are <1% of wild-type activity. $_{NJ}$-L (A-upper panel) represents the full NJ-L protein assayed using NJ-N and NJ-P encoding plasmids. (B) Minigenome experiments analyzing the unstructured domain (U) mutant as a function of its input L plasmid (top) or of input P plasmid (bottom). (C) Minigenome experiments analyzing sub-region exchanges T1 (aa 1863-1959) and T2 (aa 1960-2109) over a limited range of input P plasmid. Transcription was measured by probing for the viral M transcript and replication measured by probing for the negative-sense minigenome strand. Wild-type and mutant L proteins are HA-tagged at the N-terminus.

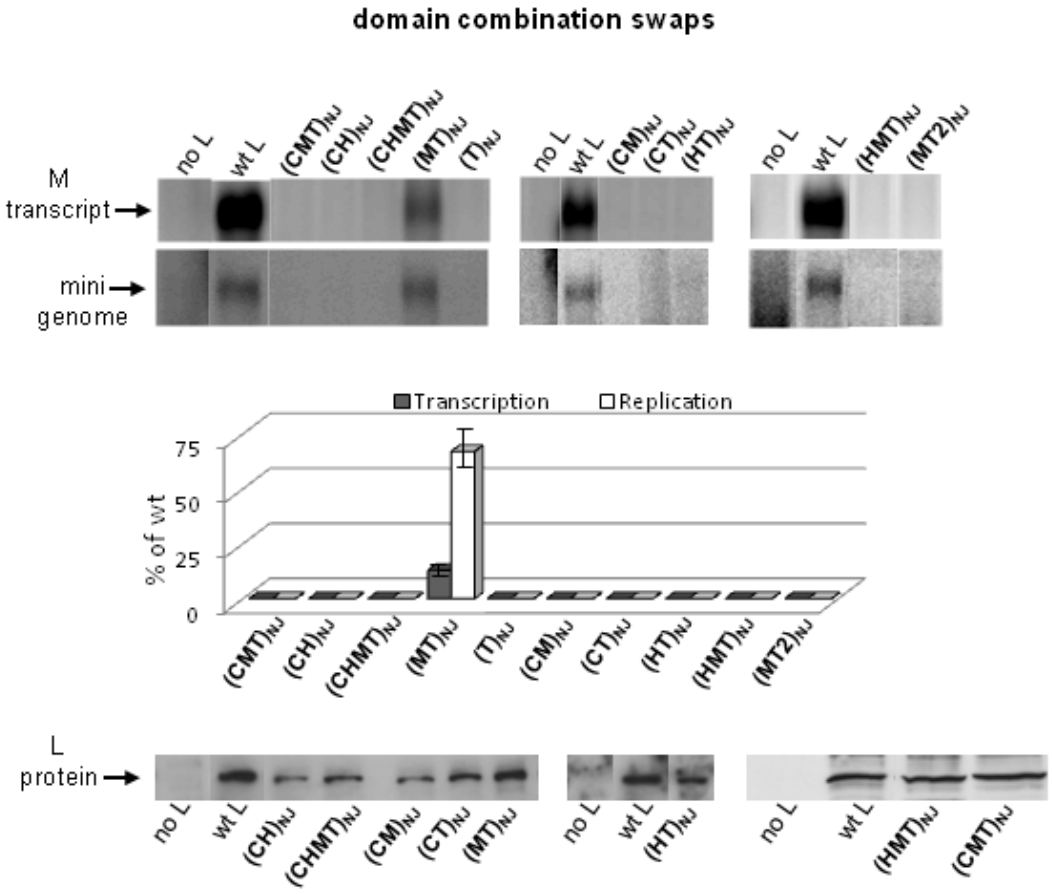

Figure 5. Minigenome assays comparing polymerase activity of domain combination swaps to wild-type IND L protein. Representative experiments show transcription and replication activity (upper sets of panels) and protein accumulations (bottom panels). The graph reports average mutant activity vs. wt L for at least two independent experiments. $(MT)_{NJ}$ mutant transcription is an average of six independent experiments (13% ± SD 5.3%) and replication is an average of four independent experiments (67% ± SD 8.7%) . All other values are <1% of wt L activity. $(MT2)_{NJ}$ mutant has an IND T1 region. All mutant L proteins are HA-tagged at the N-terminus identically as wt L.

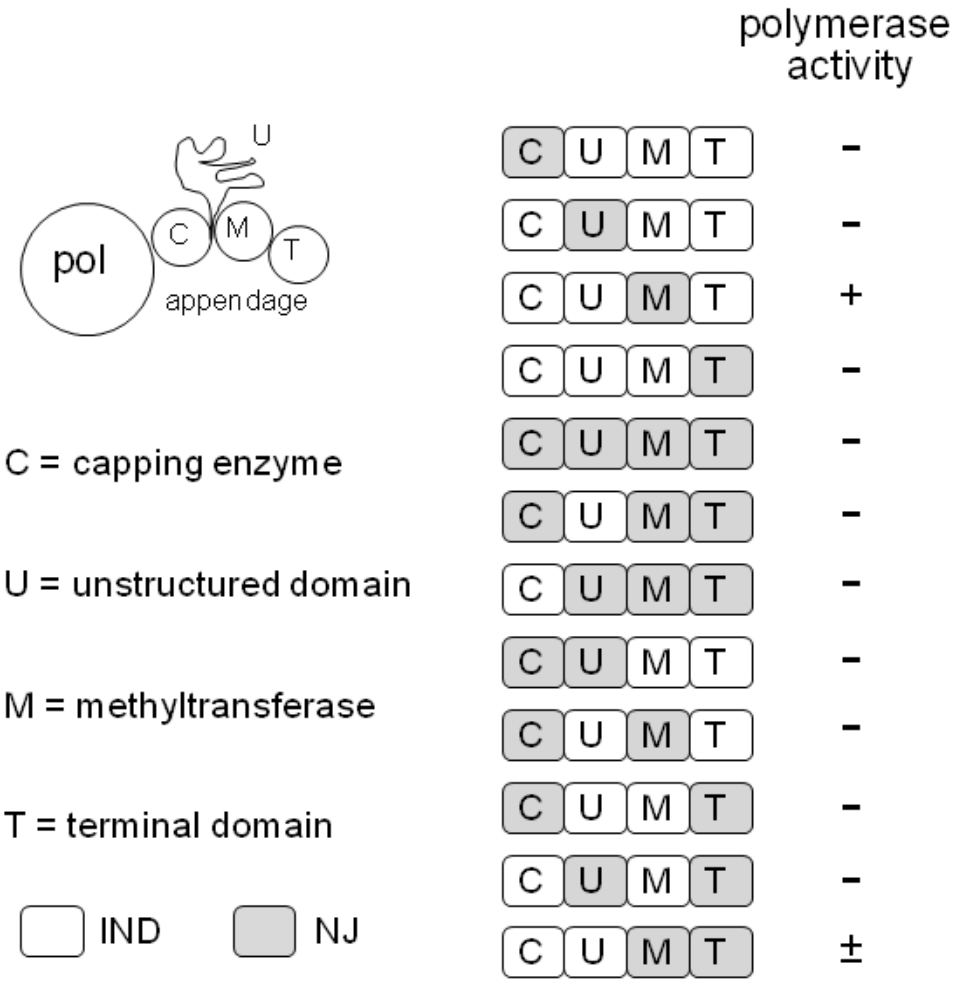

Figure 6. Summary of polymerase activities of the single domain and combination domain swap mutants by minigenome analysis. Schematic is a rendering of the low-resolution structure of L as reported by Rahmeh et al. Boxes represent L appendage domains and shaded boxes are VSV NJ serotype domains exchanged into a full IND L protein. "+" means transcription and replication activity are similar to wild type L; "±" means high replication activity and poor transcription activity.

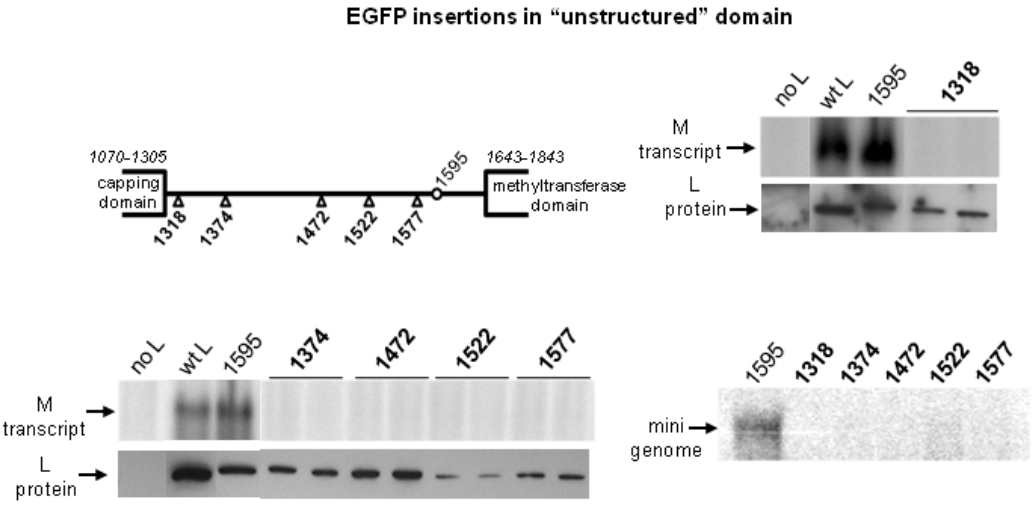

Figure 7. Effects of in-frame EGFP insertions within the "unstructured" domain of VSV IND L protein on polymerase activity in the minigenome assay. Insertion at position 1595 was shown previously to retain polymerase activity. Transcription blots and protein accumulation were determined using the same sample extracts.

REFERENCES

1. **Barik, S., E. W. Rud, D. Luk, A. K. Banerjee, and C. Y. Kang.** 1990. Nucleotide sequence analysis of the L gene of vesicular stomatitis virus (New Jersey serotype): identification of conserved domains in L proteins of nonsegmented negative-strand RNA viruses. Virology **175:**332-7.

2. **Brown, D. D., B. K. Rima, I. V. Allen, M. D. Baron, A. C. Banyard, T. Barrett, and W. P. Duprex.** 2005. Rational attenuation of a morbillivirus by modulating the activity of the RNA-dependent RNA polymerase. J Virol **79:**14330-8.

3. **Canter, D. M., and J. Perrault.** 1996. Stabilization of vesicular stomatitis virus L polymerase protein by P protein binding: a small deletion in the C-terminal domain of L abrogates binding. Virology **219:**376-86.

4. **Chattopadhyay, A., and M. S. Shaila.** 2004. Rinderpest virus RNA polymerase subunits: mapping of mutual interacting domains on the large protein L and phosphoprotein p. Virus Genes **28:**169-78.

5. **Chenik, M., M. Schnell, K. K. Conzelmann, and D. Blondel.** 1998. Mapping the interacting domains between the rabies virus polymerase and phosphoprotein. J Virol **72:**1925-30.

6. **Das, T., M. Mathur, A. K. Gupta, G. M. Janssen, and A. K. Banerjee.** 1998. RNA polymerase of vesicular stomatitis virus specifically associates with translation elongation factor-1 alphabetagamma for its activity. Proc Natl Acad Sci U S A **95:**1449-54.

7. **Dochow, M., S. A. Krumm, J. E. Crowe, Jr., M. L. Moore, and R. K. Plemper.** 2012. Independent structural domains in paramyxovirus polymerase protein. J Biol Chem **287:**6878-91.

8. **Duprex, W. P., F. M. Collins, and B. K. Rima.** 2002. Modulating the function of the measles virus RNA-dependent RNA polymerase by insertion of green fluorescent protein into the open reading frame. J Virol **76:**7322-8.

9. **Galloway, S. E., and G. W. Wertz.** 2008. S-adenosyl homocysteine-induced hyperpolyadenylation of vesicular stomatitis virus mRNA requires the methyltransferase activity of L protein. J Virol **82:**12280-90.

10. **Grdzelishvili, V. Z., S. Smallwood, D. Tower, R. L. Hall, D. M. Hunt, and S. A. Moyer.** 2005. A single amino acid change in the L-polymerase protein of vesicular stomatitis virus completely abolishes viral mRNA cap methylation. J Virol **79:**7327-37.

11. **Gupta, A. K., M. Mathur, and A. K. Banerjee.** 2002. Unique capping activity of the recombinant RNA polymerase (L) of vesicular stomatitis virus: association of cellular capping enzyme with the L protein. Biochem Biophys Res Commun **293:**264-8.

12. **Gupta, A. K., D. Shaji, and A. K. Banerjee.** 2003. Identification of a novel tripartite complex involved in replication of vesicular stomatitis virus genome RNA. J Virol **77:**732-8.

13. **Holmes, D. E., and S. A. Moyer.** 2002. The phosphoprotein (P) binding site resides in the N terminus of the L polymerase subunit of sendai virus. J Virol **76:**3078-83.

14. **Kim, G. N., and C. Y. Kang.** 2005. Utilization of homotypic and heterotypic proteins of vesicular stomatitis virus by defective interfering particle genomes for RNA replication and virion assembly: implications for the mechanism of homologous viral interference. J Virol **79:**9588-96.

15. **Lawson, N. D., E. A. Stillman, M. A. Whitt, and J. K. Rose.** 1995. Recombinant vesicular stomatitis viruses from DNA. Proc Natl Acad Sci U S A **92:**4477-81.

16. **Li, J., A. Rahmeh, V. Brusic, and S. P. Whelan.** 2009. Opposing effects of inhibiting cap addition and cap methylation on polyadenylation during vesicular stomatitis virus mRNA synthesis. J Virol **83:**1930-40.

17. **Li, J., A. Rahmeh, M. Morelli, and S. P. Whelan.** 2008. A conserved motif in region v of the large polymerase proteins of nonsegmented negative-sense RNA viruses that is essential for mRNA capping. J Virol **82:**775-84.

18. **Li, J., J. T. Wang, and S. P. Whelan.** 2006. A unique strategy for mRNA cap methylation used by vesicular stomatitis virus. Proc Natl Acad Sci U S A **103:**8493-8.

19. **Murphy, A. M., and V. Z. Grdzelishvili.** 2009. Identification of sendai virus L protein amino acid residues affecting viral mRNA cap methylation. J Virol **83:**1669-81.

20. **Nishio, M., M. Tsurudome, D. Garcin, H. Komada, M. Ito, P. Le Mercier, T. Nosaka, and D. Kolakofsky.** 2011. Human parainfluenza virus type 2 L protein regions required for interaction with other viral proteins and mRNA capping. J Virol **85:**725-32.

21. **Ogino, T., and A. K. Banerjee.** 2007. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. Mol Cell **25:**85-97.

22. **Poch, O., B. M. Blumberg, L. Bougueleret, and N. Tordo.** 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. J Gen Virol **71 (Pt 5):**1153-62.

23. **Quan, J., and J. Tian.** 2011. Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. Nat Protoc **6:**242-51.

24. **Rahmeh, A. A., A. D. Schenk, E. I. Danek, P. J. Kranzusch, B. Liang, T. Walz, and S. P. Whelan.** 2010. Molecular architecture of the vesicular stomatitis virus RNA polymerase. Proc Natl Acad Sci U S A **107:**20075-80.

25. **Rhodes, D. P., S. A. Moyer, and A. K. Banerjee.** 1974. In vitro synthesis of methylated messenger RNA by the virion-associated RNA polymerase of vesicular stomatitis virus. Cell **3:**327-33.

26. **Ruedas, J. B., and J. Perrault.** 2009. Insertion of enhanced green fluorescent protein in a hinge region of vesicular stomatitis virus L polymerase protein creates a temperature-sensitive virus that displays no virion-associated polymerase activity in vitro. J Virol **83:**12241-52.

27. **Schnell, M. J., and K. K. Conzelmann.** 1995. Polymerase activity of in vitro mutated rabies virus L protein. Virology **214:**522-30.

28. **Silin, D., O. Lyubomska, M. Ludlow, W. P. Duprex, and B. K. Rima.** 2007. Development of a challenge-protective vaccine concept by modification of the viral RNA-dependent RNA polymerase of canine distemper virus. J Virol **81:**13649-58.

29. **Sleat, D. E., and A. K. Banerjee.** 1993. Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. J Virol **67:**1334-9.

30. **Smallwood, S., T. Hovel, W. J. Neubert, and S. A. Moyer.** 2002. Different substitutions at conserved amino acids in domains II and III in the Sendai L RNA polymerase protein inactivate viral RNA synthesis. Virology **304:**135-45.

31. **Smallwood, S., E. Richards-Sumners, and S. A. Moyer.** 1994. Determinants of serotype specificity in transcription of vesicular stomatitis virus synthetic nucleocapsids. Virology **199:**11-9.

32. **Stillman, E. A., and M. A. Whitt.** 1999. Transcript initiation and 5'-end modifications are separable events during vesicular stomatitis virus transcription. J Virol **73:**7199-209.

DISSERTATION DISCUSSION

A highly flexible region just upstream of the methyltransferase domain was discovered by insertion of EGFP in the measles (family Paramyxoviridae) polymerase L protein (2). As a result, this region was proposed to function as a hinge however it was unclear if this structural feature was unique to measles. This dissertation study was initiated with the objective of determining if a homologous region is also present in the L protein of the prototypic NNS virus, VSV (Rhabdoviridae). The measles L protein is 74 amino acids longer than VSV L (2109 aa). As a result sequence gaps produced imprecise region borders between these proteins during alignment making it difficult to predict with accuracy a homologous sequence in VSV L. By aligning equal length L proteins from various VSV serotypes I observed a small variable stretch (~20 aa) positioned approximately the same distance upstream from the methyltransferase domain as where EGFP was inserted in the measles L protein.

My insertion of EGFP in this variable stretch of VSV L resulted in a functional polymerase in minigenome assays indicating that a homologous flexible region is indeed present in VSV L. The more tractable VSV system allowed for a more in-depth analysis of this modified protein. Interestingly, the

nearby dual-specificity methyltransferase domain did not show any negative effects in response to the EGFP insert however RNA synthesis surprisingly did show a temperature-sensitive (ts) effect. Transcription and replication in minigenome assays was slightly less active for this mutant compared to wt L at 33°C however total RNA synthesis was severely attenuated versus wt L at 37°C. In addition, a recombinant virus encoding this modified polymerase (VSV-L_{EGFP}) showed robust growth at 33°C but was growth attenuated at 37°C and also failed to produce transcripts in vitro. The drastic effect on RNA synthesis was unexpected since the polymerase domain is positioned nearly 800 amino acids upstream of the insert, although these structures could have close contact in the protein tertiary structure. In light of recent findings in the literature and additional studies I conducted below, a possible explanation for the ts phenotype is proposed at the end of this section.

Large-scale recombination or deletion events are rarely observed in the RNA genomes of the NNS viruses. Therefore it was not surprising that serial passages of VSV-L_{EGFP} did not produce revertants as autofluorescence and temperature-sensitivity were continuously observed at each passage demonstrating the genetic stability of the EGFP insert. Growth attenuation was also reported for the recombinant measles L_{EGFP} virus and other paramyxoviruses reported thereafter with the same polymerase modification (1, 2, 5). The basis for the growth defect of these viruses has not been determined. In light of my findings, it is possible these viruses also have a ts

phenotype for RNA synthesis. Determining this could be informative as the authors propose that rationale attenuation of morbillivirues and other NNS viruses by EGFP insertion in L can be used to produce live-attenuated genetically-stable vaccines. Interestingly, alignments show Filoviridae L proteins (Marburg and Ebola) may be more closely related to Rhabdoviridae L than the Paramyxoviridae L proteins (unpublished). Thus it can be inferred that a homologous hinge is likely present in the Filoviridae L protein as well.

Recently, electron microscopy studies have provided low-resolution images of VSV L showing it consists of two dominant features, a polymerase ring structure and an attached appendage consisting of three globules (4). The first two globules of the appendage were determined to be, in order, the capping and methyltransferase domains. Interestingly, the long variable region between these domains where EGFP was inserted did not produce any visible structure indicating it could be an unstructured or disordered domain. This visual evidence of a lack of structure in this part of L supports the notion that a flexible hinge is positioned here and further supports my findings and those of the morbilliviruses studies. Curiously, expression of the P cofactor with L in the EM studies resulted in significant restructuring of the appendage with the globules no longer being visibly distinct from each other and the appendage appearing as a loose flap over the polymerase ring structure. This observation led the authors to propose the full length of the unstructured domain (>300 aa) serves as a long hinge that facilitates domain restructuring in the appendage.

Subsequently, when I swapped the unstructured domain between two closely related VSV serotypes with 66% of the residues being identical it was surprising that the chimeric L protein failed to synthesize any detectable RNA (see below). This finding was not in agreement with the notion that the entire unstructured domain serves as a hinge and alternatively suggested that at least some part of it harbors a serotype-specific function critical for polymerase activity.

To gain additional insight into the extent of the hinge in this part of the L protein I inserted EGFP at other positions in the unstructured domain upstream of the original insertion site (aa 1595). In each instance the L protein failed to synthesize RNA including one construct that had EGFP inserted just 18 residues upstream of the original position. Cytoplasmic accumulation of the 1522 mutant was substantially lower than the other L_{EGFP} variants despite increasing its input plasmid suggesting this mutant is structurally instable and prone to higher turnover rates. These experiments showed that the degree of structural tolerance exhibited by the hinge does not extend significantly upstream of aa 1595 in VSV L. This demonstrates the hinge is limited to a small region at the C-terminal end of the unstructured domain and provides additional support that a function may be linked to the bulk of the unstructured domain. Interestingly, closer scrutiny of the constructs used to designate domains in the EM structure of L suggests an alternative possibility where the capping domain is part of the ring structure and the long variable region (e.g.,

“unstructured domain) instead comprises the first appendage globule. The findings I report here invite reexamining the EM structures to confirm the L region that correlates with the first appendage globule. Collectively, my findings above support a structural model of the NNS L protein where a hinge of limited length (~20-50 aa) is positioned just upstream of the methyltransferase domain. These studies also show perturbation of this hinge, despite its distant position from the polymerase domain, can have a major impact on overall RNA synthesis.

I showed attenuation of L_{EGFP} polymerase activity occurred at temperatures approaching 37°C however the basis of this defect was unclear. Cytoplasmic accumulation of this mutant occurred in similar amounts at permissive and nonpermissive temperature indicating protein instability was not likely a contributing factor. It is possible this mutant misfolds at nonpermissive temperature however in the absence of a high-resolution structure of wild-type L this cannot be determined. The insertion of EGFP in a hinge between two domains (e.g. capping and methyltransferase) thought to interact with the polymerase domain makes it tempting to speculate the insert interferes with a domain-domain interaction(s). The literature suggests domain interactions occur in the L protein, at least during transcript synthesis, however detailed information on these interactions and their impact on function remain unknown. To probe if the EGFP insert interferes with L domain interactions I

initially set out to characterize the interactive nature of domains adjacent to the hinge.

The primary objective of the second half of this dissertation was to determine if the four domains of the L appendage region (capping, unstructured, methyltransferase, and terminal) function autonomously or in a manner consistent with an interactive domain. A secondary objective was to gather experimental evidence of any functional domain-domain interaction(s). As discussed in the introduction of this dissertation, the intricacies of the NNS L protein makes it challenging to investigate its domain interactions by standard approaches. I therefore relied on generating chimeric L proteins to gain information on the appendage domains. By individually exchanging the four appendage domains with homologous counterparts from a closely related VSV serotype I showed polymerase activity is critically dependent on the sequence integrity of these regions. Only exchange of the methyltransferase domain allowed the polymerase to remain functional indicating that it could possibly function autonomously. However, swapping the methyltransferase with a homologue counterpart from another NNS virus (Ebola) also resulted in a nonfunctional protein showing that polymerase activity is also dependent, perhaps more loosely, on the methyltransferase sequence.

Cytoplasmic accumulations of these chimeric L mutants were not significantly lower than wt L and preliminary co-immunoprecipitation experiments (not shown) indicated they bind to the P cofactor, although

affinities between the chimeric L mutants and P as well as stochiometric ratios in their complexed form are not known versus wt L. Nonetheless, it is reasonable to conclude the heterotypic domain in these mutants fails to be recognized by another part of L or the polymerase complex thereby abolishing RNA synthesis. This interpretation reveals that, from the point of view of the polymerase domain, these domains are not autonomous as polymerase activity depends heavily on their sequence information and therefore suggests interactions, by direct or indirect means, occur among these L domains. Of note, it is not possible in these assays to determine if the function of the individual domain such as capping or methylation activity also requires interaction or instead acts autonomously since the nascent RNA substrate fails to be synthesized and in NNS systems these functions do not act on exogenous substrates. Therefore, while this study shows each of the appendage domains behave in a manner consistent with an interactive domain, this is reported specifically for RNA synthesis only.

Although substituting the NJ serotype terminal domain into IND L rendered the protein inactive, co-exchange of the NJ terminal domain with the adjacent NJ methyltransferase domain (MT mutant) partially restored RNA synthesis. Co-exchange of a partial NJ terminal region (e.g. IND T1/NJ T2) with the NJ methyltransferase yielded a nonfunctional polymerase indicating activity necessitates the full NJ terminal domain together with the NJ methyltransferase. These experiments show an intriguing case where overall

protein activity is dependent on the homotypic coexistence of two structurally distinct domains that are otherwise heterotypic to the rest of the protein. As a result, it is difficult to envision a scenario where the rescue of activity here is not the result of a functional linkage between the methyltransferase and terminal domain. I interpret these findings as compelling evidence that these two domains interact with each other and that in tandem they also interact with another part L or the polymerase complex in a manner that is critical for polymerase activity. To my knowledge, this is the first tangible evidence of a specific and essential domain-domain interaction reported in the NNS L protein. Based on the similar arrangements of conserved and nonconserved blocks in NNS L proteins (3), it can be inferred that a homologous methyltransferase-terminal domain interaction occurs in other NNS L proteins as well. Presumably, all interactions are preserved in the context of a NJ methyltransferase and an IND terminal domain as is the case in the highly functional methyltransferase-swapped chimera. Moreover, it is possible the Ebola methyltransferase failed to interact with the IND terminal domain explaining, at least in part, the lack of activity for the Ebola methyltransferase-swapped chimera.

Interestingly, the MT mutant showed an asymmetrical rescue of activity as significant replication was recovered (avg. 67% wt activity) but only a low level of transcription was restored (avg. 13% wt activity). It is possible secondary interactions involving the NJ MT domain pair and other parts of the

IND polymerase complex occur with higher efficiency in a replication conformation than in a transcription conformation. It was logical to initially hypothesize that suboptimal interactions between the NJ MT domain pair and the nearby IND capping and/or IND unstructured domain was the primary cause of reduced RNA synthesis in this mutant. However, chimera proteins representing these NJ domain combinations (C/M/T, U/M/T, and C/U/M/T exchanges) showed no polymerase activity indicating this is not the case. Instead, reduced RNA synthesis by the MT mutant appears to be caused by suboptimal interactions between the MT domain pair and other parts of L and/or the polymerase complex. This possibility is intriguing since the hinge is positioned where it could conceivably regulate interaction between the downstream MT domain pair and the upstream polymerase ring structure. In light of this finding, I propose the EGFP insert causes a block of this specific interaction, and possibly other interactions yet to be identified as well.

The use of chimeric proteins in this study has provided telling information about the interactive nature of domains in the L appendage region and moreover identified an essential interaction between the methyltransferase and terminal domain. While this approach cannot confirm a direct physical interaction, it shows a functional linkage exists between the two domains therefore providing compelling evidence that a physical interaction is likely occurring. This approach also shows for the first time a polymerase dependence on a homotypic capping and homotypic unstructured domain in

the L protein. Although the basis of the homotypic requirement is not yet known it may be worthwhile to swap smaller sub-sections of these domains in an effort to gain insight into the homotypic requirement as well as to identify the variable residue(s) or motifs responsible for it. In addition, small sub-section swaps of those domains with unknown functions (e.g. unstructured and terminal domain) may yield partially active polymerases that could then be used to investigate the function of the manipulated domain.

In summary, this dissertation demonstrates that sequence information in domains of the L appendage region has a critical impact on polymerase activity and illustrates the need for a flexible hinge upstream of the methyltransferase to be preserved in the NNS L protein. This dissertation also provides the foundations for a model of structural fluidity and orchestrated domain-domain interaction exhibited by the NNS L protein.

REFERENCES

1. **Brown, D. D., B. K. Rima, I. V. Allen, M. D. Baron, A. C. Banyard, T. Barrett, and W. P. Duprex.** 2005. Rational attenuation of a morbillivirus by modulating the activity of the RNA-dependent RNA polymerase. J Virol **79:**14330-8.

2. **Duprex, W. P., F. M. Collins, and B. K. Rima.** 2002. Modulating the function of the measles virus RNA-dependent RNA polymerase by insertion of green fluorescent protein into the open reading frame. J Virol **76:**7322-8.

3. **Poch, O., B. M. Blumberg, L. Bougueleret, and N. Tordo.** 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. J Gen Virol **71 (Pt 5):**1153-62.

4. **Rahmeh, A. A., A. D. Schenk, E. I. Danek, P. J. Kranzusch, B. Liang, T. Walz, and S. P. Whelan.** 2010. Molecular architecture of the vesicular stomatitis virus RNA polymerase. Proc Natl Acad Sci U S A **107:**20075-80.

5. **Silin, D., O. Lyubomska, M. Ludlow, W. P. Duprex, and B. K. Rima.** 2007. Development of a challenge-protective vaccine concept by modification of the viral RNA-dependent RNA polymerase of canine distemper virus. J Virol **81:**13649-58.