

The nucleocapsid of vesicular stomatitis virus

LUO Ming

Department of Microbiology, the University of Alabama at Birmingham, Birmingham 35294, USA

Received December 22, 2011; accepted March 2, 2012

The nucleocapsid of vesicular stomatitis virus serves as the genomic template for transcription and replication. The viral genomic RNA is sequestered in the nucleocapsid in every step of the virus replication cycle. The structure of the nucleocapsid and the entire virion revealed how the viral genomic RNA is encapsidated and packaged in the virus. A unique mechanism for viral RNA synthesis is derived from the structure of the nucleocapsid and its interactions with the viral RNA-dependent RNA polymerase.

RNA cavity, cross-molecular interactions, RNA access, phosphoprotein binding to the nucleocapsid, bullet-shaped virion

Citation: Luo M. The nucleocapsid of vesicular stomatitis virus. *Sci China Life Sci*, 2012, 55: 291–300, doi: 10.1007/s11427-012-4307-x

Negative strand RNA viruses (NSV) have a single strand RNA genome that is encapsidated by a virally encoded protein, called the nucleocapsid protein (abbreviated as NP or N). The complex structure formed by the viral RNA genome and the nucleocapsid protein is termed as nucleocapsid. Most NSVs have a genome of one single RNA strand, such as *Rhabdoviridae* and *Paramyxoviridae*, and others may have multiple fragments of single RNA strands, such as *Orthomyxoviridae* and *Bunyaviridae* (see ICTV; <http://ictvonline.org>). The RNA genome of NSV has a negative sense (or antisense) because the genes in a NSV genome must be first transcribed into mRNAs by the viral RNA-dependent RNA polymerase (RdRp). As a result, the virion of a NSV contains not only the nucleocapsid, but also the viral RdRp. During infection, the viral RdRp enters the host cell together with the nucleocapsid and immediately starts transcription of viral genes. The key characteristic that separates NSV from all other viruses is that the template used by the viral RdRp for synthesis of virus-specific RNA, both in transcription and replication, is not the free form of

the viral RNA genome, but the nucleocapsid. If a copy of free viral genomic RNA is exposed to the viral RdRp, the viral RdRp cannot recognize it as a template for RNA synthesis. The viral RdRp binds only the nucleocapsid and is capable of using the genomic RNA sequestered inside the nucleocapsid as the template to carry out viral RNA synthesis. It is therefore essential to understand the molecular mechanism of the assembly of NSV nucleocapsid and its implications in virus replication. The study of the vesicular stomatitis virus (VSV) nucleocapsid in my laboratory may be an example to illustrate this mechanism since VSV has been used as the prototype NSV for many studies.

1 Replication cycle of VSV

VSV has a single strand RNA genome that comprises approximately 11K nucleotides. The VSV encodes five viral proteins: the nucleocapsid protein (N), the phosphoprotein (P), the matrix protein (M), the glycoprotein (G), and the large protein of the viral RdRp (L) (Figure 1A). There is a leader sequence (Le) at the 3' end and a trailer sequence (Tr)

email: mingluo@uab.edu

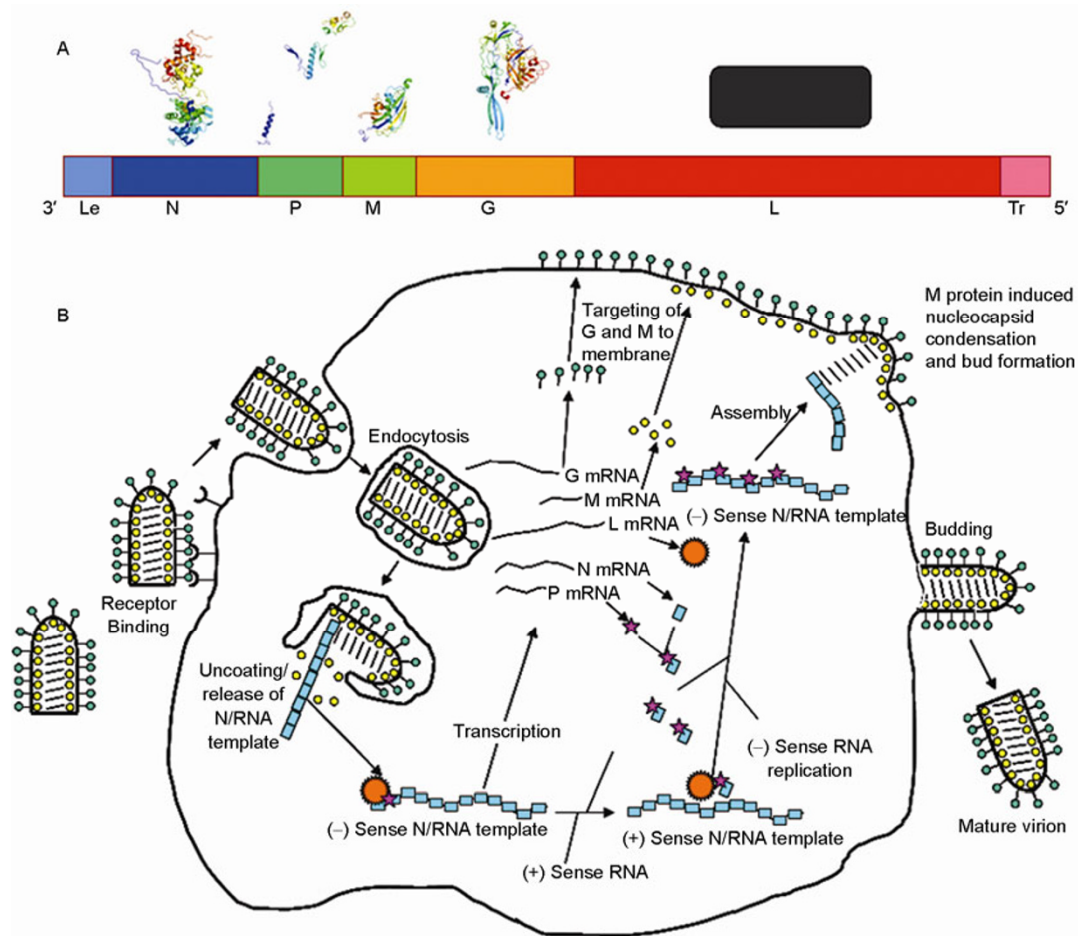


Figure 1 A, The genomic structure of vesicular stomatitis virus. Starting from the 3' end of the viral RNA genome, five viral genes are encoded: nucleocapsid (N), phosphoprotein (P), matrix (M), glycoprotein (G) and the large protein of the viral RNA-dependent RNA polymerase (L). There is a leader sequence at the 3' end and a trailer sequence at the 5' end. The crystal structure of N, M and G has been solved as shown above the genome sketch. The crystal structure of P has been solved as fragments and is shown as a possible composition of the fragments. The crystal structure of L is unknown (black box). B, A schematic illustration of the replication cycle of vesicular stomatitis virus.

at the 5' end. There is a significant degree of sequence complementarity between Le and Tr that is required for efficient virus replication and assembly [1]. All five viral proteins are packaged in the virion, which suggests that all of them play a role in virus assembly. VSV initiates its infection by attachment to a host receptor on the cell surface (Figure 1B). A specific molecule on cell surface has not yet been identified as the host receptor for VSV [2]. Since VSV can infect a large number of cell types, the VSV host receptor is either a common molecule among different cell types or VSV may be able to use a number of different surface molecules as a receptor. Attachment of the host receptor by VSV triggers endocytosis that leads to VSV uptake in partially clathrin-coated vesicles [3]. When the interior of virus-containing vesicles becomes acidic after entry, the conformation of the glycoprotein G on the virus membrane envelope undergoes a dramatic change, which induces fusion of viral membrane with the cellular membrane of the vesicle. The membrane fusion resulted in the release of the

nucleocapsid packaged in the virion, together with the viral RdRp, and viral transcription is then initiated. During the entry process, the viral RNA genome is never unprotected from the nucleocapsid. The structurally stable nucleocapsid is, however, a functional template for viral transcription.

2 Structure of the nucleocapsid

To find out how the viral RdRp gains access to the RNA genome encapsidated in the nucleocapsid, the first step is to unveil the structure of the nucleocapsid. When the nucleocapsid is released from the virion, it looks like a linear thread, with a somewhat coiled appearance (Figure 2A) [4,5]. The nucleocapsid is composed of subunits of a single viral protein, N, and the genomic RNA. The integrity of the nucleocapsid could not be destroyed unless 3 mol L⁻¹ guanidine was used to treat the purified nucleocapsid [6]. The capsid formed with subunits of N is thus very stable

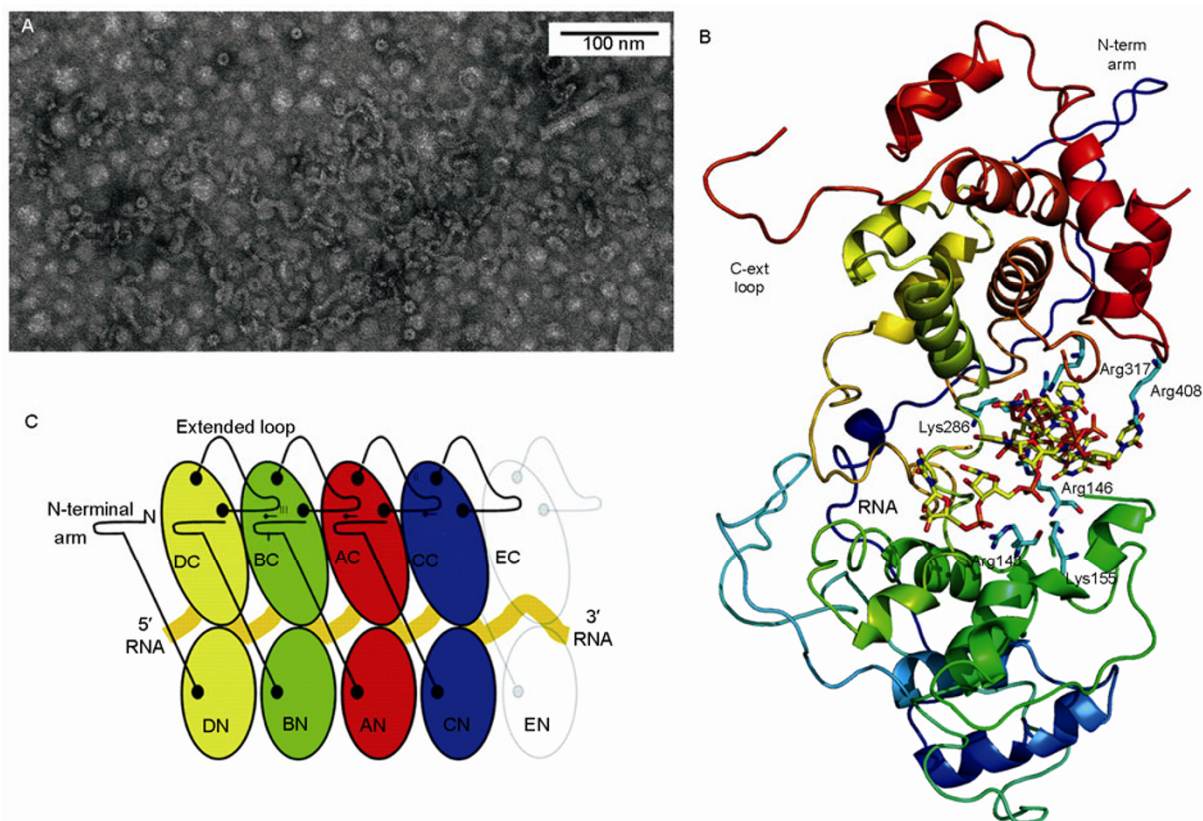


Figure 2 A, An electron micrograph of purified nucleocapsid of VSV. B, A ribbon drawing of the crystal structure of an N subunit in complex with a piece of 9 nucleotide RNA. The amino acid sequence is traced by color from the N-terminus in blue to the C-terminus in red. The six residues that coordinate with the backbone phosphate groups of RNA are drawn as stick models with labels. The N-terminal arm and the C-terminal lobe extended loop are also marked. The RNA is shown as stick models. C, A schematic illustration of the interactions between neighboring N subunits in the nucleocapsid. The N-terminal arm of the red N subunit makes contacts with the back of the C-terminal lobe of the green N subunit (Contact I). The C-terminal lobe extended loop makes contacts with the back of the C-terminal lobe of the blue N subunit (Contact II). The N-terminal arm of the red N subunit also makes contacts with the C-terminal lobe extended loop of the yellow N subunit (Contact III). The small gap between the N-terminal lobes indicates lesser side-by-side interactions compared to those between the C-terminal lobes. The encapsidated RNA is shown as an orange ribbon in the center of the N subunits.

and can protect the RNA genome from digestion by RNA nuclease [7,8]. To study the atomic structure, we set out to crystallize a nucleocapsid-like particle (NLP) of VSV. We first designed a unique vector to express recombinant proteins of N and P simultaneously at a 1:1 stoichiometry in *E. coli* [9]. The coding regions for the two genes were cloned in one transcription unit with a separate ribosomal binding site in front of the P gene, following the end of the N gene. This design is based on the observation that the 1:1 ratio of N and P expression was optimal for supporting VSV replication [10,11]. A his₆ tag was fused at the N-terminus of P to assist in purification. When the recombinant protein was purified with a Nickel-affinity column, a large complex of N and P proteins was found to correspond to 10 subunits of N and P, respectively. A single strand of RNA is present in the complex and its length is 90-bases. It was shown later that the sequence of the 90-base RNA is random (Luo M, unpublished data). The P protein could be dissociated from the complex when the complex was incubated at pH 4.0, yielding an N-RNA complex. However, the RNA remains

to be protected from RNA nuclease in this N-RNA complex and the complex is hard to disrupt. We therefore named this N-RNA complex a NLP. When the NLP was examined under electron microscope, its appearance is like a donut. The 3D reconstructed structure using negatively stained electron micrographs revealed that the NLP is built with the N subunit parallelly aligned side-by-side in a circular manner, with one end of the N subunits tightly associated together, and the RNA strand in the middle of the ring formed by the N subunits [12]. The purified NLP can be easily crystallized, but the quality of X-ray diffraction was poor initially. When a deliberated procedure of soaking the crystals of NLP in a uranyl acetate solution was developed, high resolution X-ray diffraction data could be collected from the treated crystals [13]. The atomic structure of NLP revealed all the details of the remarkable arrangement of the N subunits and its mechanism of encapsidating the RNA [14]. The ring structure formed by the N subunits has the C-terminal part tightly associated together, similar to what was observed in the EM structure [12]. The N-terminal part still involves

subunit interactions within the ring, but the area of contact is one third of that between the C-terminal ends. The N protein can be divided into two lobes, the N- and C-terminal lobe, with each primarily composed of α -helices. The two lobes are linked together by one helix from each lobe angled like a letter V (Figure 2B). There is an extended arm in the N-terminal lobe comprising the first 22 amino acids, and an extended loop in the C-terminal lobe comprising residues 340 to 375. These extra structural features play important roles in the construction of the nucleocapsid. The NLP contains only 10 subunits of N arranged in a ring. To extrapolate the structure of the NLP to that of the nucleocapsid, one can imagine that the NLP is cut open and straightened as a linear structure, like a tube of protein capsid with a single strand RNA inside of it. A schematic illustration is presented in Figure 2C. The N subunits are associated with each other in a parallel orientation, with the C-terminal lobe on top and the N-terminal lobe at the bottom. The back of the protein capsid faces the reader and the opening to the genomic RNA is on the opposite side. If the red N subunit is used as the reference, its N-terminal arm interacts with the back of the C-terminal lobe of the green N subunit on its left. We term this patch of interactions Contact I. Meanwhile, the C-terminal extended loop of the red N subunit interacts with the C-terminal lobe of the blue N subunit on its right (Contact II). Finally, the N-terminal arm of the red N subunit interacts with the C-terminal extended loop of the yellow N subunit two units on its left (Contact III). The unique contacts across N subunits are repeated in group of four subunits throughout the entire nucleocapsid. The symmetry of the NSV nucleocapsid is therefore linear, as compared to spherical (icosahedral) and helical symmetry of other virus capsids. These cross-molecular interactions, together with the side-by-side interactions of the parallel N subunits, maintain the integrity of the viral capsid.

This pattern of virus capsid architecture is universal in all virus capsids, including spherical viruses and helical viruses. The virus capsid is built with repeated protein subunits arranged in symmetry. The protein subunits contact each other in the capsid to put together a shell in order to protect the nucleotide viral genome inside. In addition, it has been observed in many virus structures that the capsid protein subunit always has extended structural features (such as long termini and large loops) that make cross-molecular interactions to stabilize the virus capsid. The VSV capsid is no exception [15]. Experiments by our lab and others showed that if these interactions are interrupted by mutation, such as deletion of the N-terminal arm or portion of the C-terminal loop, and mutations to alanines at the side-by-side contact between neighboring C-terminal lobes, no stable nucleocapsid can be formed and the mutant N protein no longer forms a complex with RNA [16,17]. This confirms that the N protein of NSV assembles a capsid to encapsidate the viral genomic nucleotide following the principle that is the same as other virus capsid proteins.

Inside the tube of VSV capsid, the encapsidated RNA is situated in a cavity that is mainly framed by five α -helices from the N-terminal lobe and three α -helices from the C-terminal lobe, with the V-shaped two helix linker in the center (Figure 2B). Each N subunit covers 9 nucleotides in the VSV nucleocapsid. If the first nucleotide at the 5' is numbered as 1, the bases of the first four nucleotides in the 9 nucleotide stretch are stacked together as half of the A form double stranded RNA. On the other hand, bases 5, 7, 8 and 9 are stacked together and more importantly face the interior of the RNA cavity. The structure of the RNA in the nucleocapsid strongly suggests that the cavity must be opened and RNA must be released temporarily during viral RNA synthesis in order to allow the viral RdRp to access the sequence of the genomic RNA sequestered in the nucleocapsid. This relationship also provides an explanation why the viral RdRp must use the nucleocapsid as the template for RNA synthesis. There must be conformational changes induced by interactions between the N protein and the viral RdRp when RdRp binds the nucleocapsid. There are no specific interactions between the residues in the cavity and the RNA, except the sidechains of six positively charged residues, Arg143, Arg146, and Lys155, Lys286, Arg317, and Arg408, that coordinate with the phosphate groups of the RNA backbone. However, these residues are not conserved in the N protein of even closely related rabies virus [14,15,18]. There are in general a number of positively charged residues in the RNA cavity, but the pattern of these residues interacting with RNA is rather variable. The interactions with these residues are not required for RNA encapsidation either [19]. The viral genomic RNA is largely a resident in the nucleocapsid, rather a molecule tightly bound by the N protein. This observation is further confirmed by two experiments. First, an N protein mutant, Ser290→Trp, was made in our laboratory. The sidechain of Ser290 does not have specific interactions with the RNA, but its position is close to the RNA backbone. Substitution of a serine residue with tryptophan introduces a bulky sidechain that imposes steric hindrance to RNA. As the result, a ring structure similar as the NLP was produced when the mutant N protein was coexpressed with P, but it was devoid of RNA [16]. Second, the RNA encapsidated in the NLP or authentic nucleocapsid may be completely removed by extensive treatment of RNase A, especially at high temperature [20]. The empty NLP generated by RNase A treatment and the mutant N (Ser290→Trp) ring have the same structure as that of the NLP, indicating the assembly and the structural integrity of the protein capsid do not necessarily require encapsidation of RNA. This notion does not, however, preclude the role of encapsidated RNA in stabilization of the nucleocapsid. The fact that RNA is encapsidated in the NLP when wt N protein is coexpressed with the P protein implies that the RNA contributes positively to the assembly of a stable NLP. In the case of a N-terminus truncated N protein fused with

maltose-binding-protein, a ring structure similar to the NLP was only stable when the encapsidated RNA was replaced by the N-terminal fragment of P that restored to a large extent Contact I and III in the mutant capsid [17].

Currently, the structure of a NLP has been solved for three NSVs, including VSV [14], rabies virus [18], and respiratory syncytial virus [21]. The amino acid sequence homology is fairly low among the three N proteins. In addition, the N subunit in the NLP of respiratory syncytial virus covers only 6 nucleotides instead of 9. Despite these differences, the RNA cavity and the RNA stacking are highly homologous in their NLP structures. It is possible that the N proteins in NSV have a conserved structural motif for the assembly of the nucleocapsid that effectively encapsidates viral genomic RNA. We name this structural motif (5H+3H) since the RNA cavity is supported by the V-shaped two helices that are grouped with 4 helices in the N-terminal lobe (1+4=5H) and 2 helices in the C-terminal lobe (1+2=3H), respectively [22]. This motif is readily visible in the structure of Borna disease virus N protein [23]. The motif is likely to be present in the N protein of other NSV when the structure of N in complex with RNA becomes available.

3 Nucleocapsid as the template

As shown by the crystal structure, the viral genomic RNA is sequestered in the nucleocapsid. How the viral RdRp recognizes the nucleocapsid as the template and gains access to the genomic RNA is a puzzling question. The recognition of the nucleocapsid by viral RdRp requires the cofactor, P. The P protein (265 amino acids in length) can be associated with the L protein to form a stable complex. The P protein in the complex allows it to recognize the nucleocapsid [24–28]. The pattern of the P protein interactions with the L protein, the RNA-free N protein and the nucleocapsid is complicated. The P protein is present in multiple phosphorylation forms. When the P protein is phosphorylated near its N-terminus at sites Ser60, Thr62 and Ser64, it is involved in viral transcription [29,30], whereas phosphorylation at sites Ser226 and Ser227 is required for viral replication [31]. The P protein is therefore divided into three domains, the N-terminal domain (P_{NTD}), the central domain (P_{CD}), and the C-terminal domain (P_{CTD}). Each domain can function independently. For instance, P_{CTD} was shown to bind the nucleocapsid [32,33]. On the other hand, P_{NTD} could bind probably the L protein to support transcription [33–35]. However, the P protein always functions in a dimer form as shown by complementary experiments by coexpression of two P mutants [36]. Structural analyses also support the domain division of the P protein and its functional form as a dimer [17,37–40] (Figure 3). P_{CD} is the key domain for the formation of a P dimer. The crystal structure of a P fragment corresponding to residues 107–177 revealed that the dimer is constructed

by domain swapping of a β -hairpin in each subunit that participates in collating a β -sheet of four antiparallel strands on each side of the dimer [38] (Figure 3A). The core of the P dimer is consisted of two parallel α -helices brought together by many hydrophobic interactions. Interruption of the dimer association is detrimental to P functions [41,42].

To initiate transcription or replication, the RdRp of the L/P complex binds the nucleocapsid and reads the 3' end for replication or the promoter for transcription. However, the two different types of RNA synthesis may require the viral RdRp to interact with the nucleocapsid template in a different manner. What factors influence the choice of transcription over replication by the viral RdRp has been studied extensively. It has been postulated that the viral RdRp is first loaded at the 3' end and initiates RNA synthesis. This corresponds to an event of replication initiation. If there is present a sufficient amount of encapsidation competent N protein, called N^0 that is kept RNA free by association with the P dimer (N^0 - P_2) [16,17,43], elongation of replication continues [44]. Otherwise, the RNA synthesis terminates and a leader sequence is produced [45]. This termination may completely release the RdRp from the nucleocapsid, or allow it to move toward the first promoter to initiate transcription of the N gene [46]. In fact, transcription initiation of each gene in the VSV genome is always restarted when the transcription of the prior gene is terminated [47–49]. The reinitiation results in that each transcript is capped by a unique capping mechanism carried within the L protein [50] and polyadenylated at the 3' end. At the same time, the amount of each mRNA is reduced by approximately 30% sequentially in the order of N, P, M, G and L because the initiation of any transcription run always starts at the promoter for the N gene. However, it has been discovered later that the viral RdRp in the transcriptase form can enter the promoter for the N gene directly [51,52]. In contrast, the replicase is actually composed of L/P/N that initiates replication at the 3' end [51]. The tripartite replicase is produced in sufficient amounts when the N protein is fully expressed, consistent with the requirement of *de novo* synthesis of N to sustain replication [44]. In any event, the viral RdRp must specifically recognize the nucleocapsid template. This unique recognition relies on the association of P_{CTD} with the nucleocapsid. The crystal structure of P_{CTD} in complex with the NLP presented a clear picture how this specific association is achieved [40] (Figure 3B). P_{CTD} consists of four α -helices and two antiparallel β -strands, and is sandwiched between two neighboring C-terminal lobes of N when P is associated with the nucleocapsid. The main contact areas between P_{CTD} and the nucleocapsid are the C-terminal extended loop and the last α -helix in one C-terminal extended lobe, and the C-terminal extended loop in another neighboring C-terminal lobe. The last α -helix of the C-terminal lobe builds the bottom of the P_{CTD} binding site and the two C-terminal extended loops erect the walls to accommodate P_{CTD} . If the N protein is monomeric as a RNA-free protein,

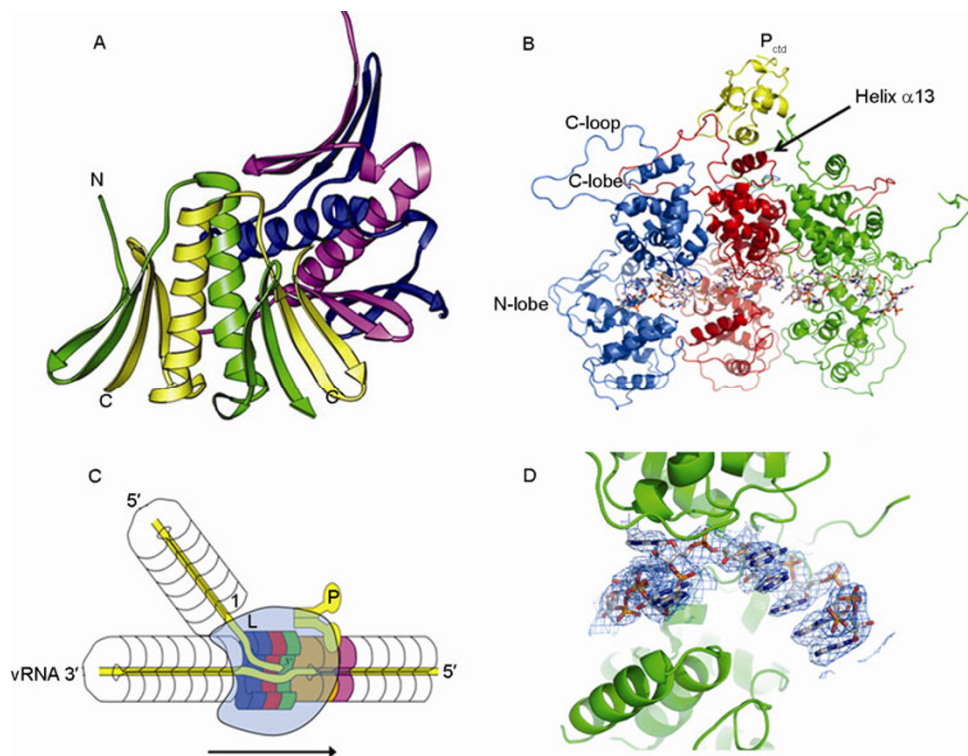


Figure 3 A, A ribbon drawing of the structure of the P_{CD} dimer. The N terminus of the green subunit and the C-termini of the green and yellow subunits are labeled for the P_{CD} dimer. The domain swab of the β -hairpin can be seen in the green and yellow P_{CD} dimer. Two parallel α -helices are in the center of the P_{CD} dimer. A second P_{CD} dimer (blue and purple) is also shown indicating the formation of the β -sheet in the dimer of P_{CD} dimers. B, The ribbon drawing of the structure showing the interaction of P_{CTD} (yellow) with the C-terminal lobe extended loops of the red and green N subunits. The helix $\alpha 13$ of the red N subunit is at the bottom of the P_{CTD} binding site. The access to the RNA cavity is facing the reader. C, A schematic illustration for the replication model. The N subunit is shown as a horse shoe. The viral RNA is in the center of the nucleocapsid. A P dimer lodges the L protein on the nucleocapsid, which leads to unveiling the viral RNA inside the nucleocapsid. The template viral RNA is temporarily released from up to three N subunits to allow copying of the viral genome. The newly synthesized viral RNA is concomitantly encapsidated by the N subunits delivered to the replication site by other P dimers (in the form of N^0 - P_2) while the viral RdRp proceeds along the viral genome. D, The electron density (wire cage) for the tightly stacked adenine bases of a poly(rA) in NLP. The orientation of the N subunit and RNA is the same as in (B) with the C-terminal lobe on top and the N-terminal lobe at the bottom.

two third of this P_{CTD} binding site is still present, which could be part of the P-N interactions in the N^0 - P_2 complex, but the interactions may be weak. It is therefore required that the N-terminal domain of P participates in the interactions with the N groove corresponding to the RNA cavity in order to keep N free of RNA [17]. The P_{CTD} binding site walled by the two neighboring C-terminal lobes in the nucleocapsid must have a higher affinity in order to render the recognition of the nucleocapsid by the L/P RdRp complex. Since this site can only be created by N subunits that are in the nucleocapsid, this unique feature of the nucleocapsid establishes the mechanism for the viral RdRp to bind only the nucleocapsid. In the crystal structure of the P_{CTD} -NLP complex, there are significant conformational changes in P and N when compared to that of the uncomplexed structures [14,39,40], which further confirmed that the P_{CTD} binding site is a high affinity site because of induced fit in the complex structure. In addition, the two phosphorylation sites in P_{CTD} , Ser226 and Ser227, are located closely to the opening of the RNA cavity, a position consistent with its role of bringing the viral RdRp to bind the nucleocapsid template

and gain access to the genomic RNA.

It has been shown that the assembly of the nucleocapsid is concomitant with the replication of the viral RNA genome [44]. The crystal structure of the NLP showed that the N protein has no interactions with specific bases in the RNA, and thus there could be no encapsidation signal in the viral RNA genome, which is different from what was previously thought. It is conceivable that the specific encapsidation of the genomic RNA is the result of nucleocapsid assembly at the site of replication, as suggested in a model [40] (Figure 3C). The P dimer may play a critical role in this process. First of all, the newly synthesized N protein is kept RNA-free and encapsidation competent by the P dimer. Otherwise, the N protein will likely encapsidate random RNA sequences as shown by expression of the N protein alone [53]. Most importantly, we propose that the P dimer actually shuttles the N protein subunits to the site of replication in order to encapsidate newly synthesized genomic RNA. The implication of this P protein function can be derived from the observation that the oligomerization domain (P_{CD}) dimer may be further associated to generate a tetramer

of two P dimers through the extension of the β -sheets [38]. Since a P dimer is already present in the viral RdRp as a cofactor to bind the nucleocapsid template, the additional interactions between the P dimers, one from the RdRp, and the other from the N⁰-P₂ complex, may serve as a docking platform for the newly synthesized N subunits to be delivered to the viral RdRp for the nucleocapsid assembly. It is likely that other interactions also participate in implementing the nucleocapsid assembly site, but the association between the two P dimers is a key for correct N orientation. Once the N subunit is added to the nucleocapsid under construction, the P dimer may be released and recycled for delivery of more N subunits. This active process makes the encapsidation of the viral genomic RNA much more efficient and rapid so that the majority of N subunits will be utilized before encapsidation of random RNA sequences may occur. There is no requirement for an encapsidation signal in the viral genome during this process. Some of the P dimers may remain associated with the fully assembled nucleocapsid, in addition to being present in the nucleocapsid associated viral RdRp, to facilitate its transportation to the budding site at the plasma membrane [54].

For transcription, de novo synthesis of the N protein or the P protein is not required [28,44]. The P dimer present in the viral RdRp renders RdRp binding to the nucleocapsid template, but the viral RdRp must be able to distinguish the promoter sequence in the nucleocapsid. Recent studies showed that mutations in the C-terminal extended loop of N that interacts with the P protein, or in the N protein near the RNA cavity have a profound effect on the level of transcription, but no effects on replication [55,56]. This indicated that the N protein plays a critical role in transcription initiation. There must be a unique feature near the promoter sequence in the nucleocapsid that allows a productive transcription initiation complex to be organized. We developed a novel experiment to map the structural details near the promoter sequence [20]. By extensive RNase A treatment, the random RNA molecules could be completely removed from the NLP. A NLP with a specific RNA sequence can be reconstituted by incubating the empty NLP with RNA molecules of a specific sequence. Our initial structures included those reconstituted with poly(rA), poly(rG), poly(rC), and poly(rU). Again, no specific interactions between the RNA bases and the N protein were observed in these structures. However, the stability of the RNA motifs has significant differences for different RNA sequences. For instance, the base stacking of poly(rA) in the NLP is most stable among the four types of ribonucleotides, contrary to poly(rA) in solution [20] (Figure 3D). The environment in the RNA cavity, including, for instance, the shape of the space in the cavity, hydrophobic patches, or neutralization of the backbone phosphate groups by the positively charged residues, could together favor the formation of a more stable RNA structure near the promoter sequence. Such a stable struc-

ture can induce the viral RdRp that is scanning the nucleocapsid to pause and change its structure so that the promoter sequence is used for initiation of transcription. On the other hand, when the viral RdRp reaches the end of a viral gene which has a poly(rU) track prior to the gene junction site, the structure of poly(rU) stacking motifs is less stable, preventing RNA synthesis from preceding. This model is consistent with the observation that mutations that dampened RNA transcription were associated with decreases in the stability of the nucleocapsid RNA cavity or the complex formed between the viral RdRp and the nucleocapsid [55,56]. Studies of the structure for the promoter and gene junction sequences are in progress (Luo M, unpublished data).

4 Packaging the nucleocapsid in the virion

To spread infection, the newly formed nucleocapsid needs to be packaged in the virion. The first step is to transport the nucleocapsid to the site of virion assembly at the interior of plasma membrane. This transportation involves other viral proteins, probably M, P dimer and/or the viral RdRp containing L and P, and is microtubule-mediated [56,57]. The viral glycoprotein G is separately transported to the plasma membrane via the ER/Golgi system, and is clustered in the membrane raft with the ectodomain on the outside and the cytoplasmic tail on the inside [58]. The M protein plays the role of driving virion assembly. The M protein is first clustered at microdomains of the plasma membrane, whereas the nucleocapsid is associated with the microdomains where the G protein is clustered [59]. The M protein is then recruited to the budding site and the virion assembly is initiated. The virion of VSV has a unique bullet shape, with a pointed head budding from the plasma membrane first and ending with a blunt bottom. Based on the cryoEM structure of VSV virion, it is concluded that the M protein drives this budding process to create the bullet shape [60]. At the center of the budding site starting with the 3' end of the viral genome, the nucleocapsid is likely to initiate budding with a circular structure with 10 N subunits which is similar to that of the NLP found in the crystal structure. In order to continue to package the nucleocapsid, a second turn needs to be added surrounding the initial circle leading to a spiral structure. The second turn must have a larger radius and its interactions with the first turn must be sufficient to stabilize the structure (Figure 4A). The cryoEM structure showed that a turn of M subunits is situated between the two turns of the nucleocapsid, with the N-terminal lobes of N facing the M protein and the C-terminal lobes facing the helical interior (Figure 4B). The M protein bridges the two turns of the nucleocapsid, thus stabilizing the spiral structure. As each turn is added to package the nucleocapsid into the virion, a turn of M subunits is continuously added between the consecutive turns to stabilize the spiral structure, but the

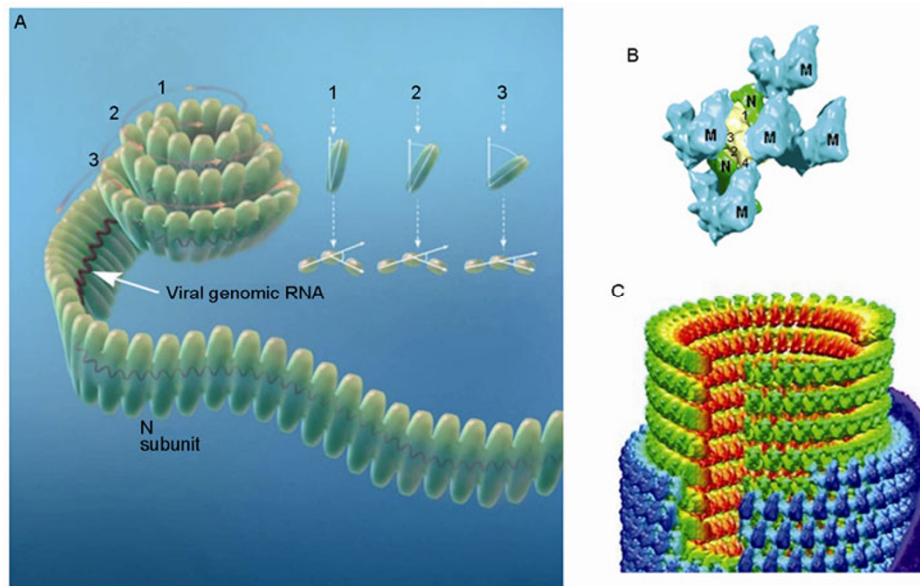


Figure 4 A, A schematic illustration of the assembly process of the nucleocapsid during budding. The N subunit is represented by a spherically ended cylinder. The viral genomic RNA is represented by a red string. The orientation of the N subunits in the 1st, 2nd, and 3rd turns, and the tangent angle between them to indicate the radius are illustrated at the upper right corner. B, Five M subunits from three turns are shown as cyan surface rendered objects (labeled M), and two N subunits underneath the M subunits are shown as green surface rendered objects (labeled N). The yellow portion of the central M subunit represents its N-terminal hob. Numbers 1 and 2 indicate interactions of M with the top and bottom N subunits, respectively, whereas numbers 3 and 4 indicate interactions of M with the M subunits on the left and at below, respectively. B, Surface rendered illustration of the cylindrical portion of the VSV virion. The nucleocapsid is in the center and the mesh formed by the M subunits (blue) surrounds the helix of the nucleocapsid. The purple layer indicates the inner leaflet of the viral membrane.

radius of each consecutive turn is larger than the prior turn up to the 7th turn. Following the 7th turn, the nucleocapsid is wound up as cylindrical turns of the same radius finishing up at the 5' end of the viral genome, with each turn containing 37.5 N subunits, or with 75 N subunits to complete two turns (Figure 4C). The N subunit in each turn is interdigitated with the N subunits in the turn above and below, with neighboring turns stabilized by the M subunits.

The M subunits follow the same symmetry as the N subunits, but the close contacts between turns are provided by the M subunits only. The bullet shaped structure of the nucleocapsid packaged in the virion must be stabilized by the M protein. Otherwise, the nucleocapsid would not be able to shape like a bullet [5,61]. The M subunits establish a complete two-dimensional mesh covering over the nucleocapsid. In the 2D mesh, M subunits maintain four points of close contacts with each other laterally and vertically [60,62]. The rigidity, the shape and the separation between turns of the bullet shaped nucleocapsid are entirely dictated by the M subunits. Meanwhile, the back of the M subunit (the C-terminal domain of M [63]) interacts with the cytoplasmic tail of the trimeric G protein, thus completing the assembly of an infectious virion. The match between the 3-fold symmetry of the G protein trimers and the helical symmetry of the M protein mesh is coincidental due to the dimension of the M-nucleocapsid structure in bullet shape, resulting in the most stable virion.

5 Conclusion

The nucleocapsid of VSV is formed by the N subunits lined-up in a parallel fashion. The side-by-side and cross-molecular interactions among the N subunits maintain the integrity of the capsid to sequester the viral RNA genome. The nucleocapsid is packaged in the virion in a bullet shape and is released together with the viral RdRp into the cytoplasm during virus entry. The linear nucleocapsid in the cytoplasm serves as the template for viral transcription and replication. With the help of the polymerase cofactor, the P protein, the viral RdRp can recognize the nucleocapsid and access the viral RNA genome encapsidated in the nucleocapsid. The nucleocapsid of VSV, and also other NSVs, is functionally the same as other viral nucleocapsids, with a protein capsid that encapsidates the viral RNA genome inside. There are no fundamental differences between the nucleocapsid of NSVs and that of other viruses, except for symmetry. The VSV nucleocapsid has a linear symmetry, whereas other viral nucleocapsids may have a spherical (icosahedral) symmetry, e.g., poliovirus, or a helical symmetry, e.g., tobacco mosaic virus. The general properties of VSV nucleocapsid are common among the nucleocapsid of all NSVs. For further reading, please see [64]. The ribbon drawing for the protein structure in all figures in this publication was done with PyMol [65].

I thank my colleagues, who are coauthors of the original publications, for their contributions to all the VSV work from my laboratory, supported by the National Institutes of Health (Grant No. AI050066).

- 1 Whelan S P, Wertz G W. Regulation of RNA synthesis by the genomic termini of vesicular stomatitis virus: identification of distinct sequences essential for transcription but not replication. *J Virol* 1999, 73: 297–306
- 2 Coil D A, Miller A D. Phosphatidylserine is not the cell surface receptor for vesicular stomatitis virus. *J Virol*, 2004, 78: 10920–10926
- 3 Cureton D K, Massol R H, Saffarian S, *et al.* Vesicular stomatitis virus enters cells through vesicles incompletely coated with clathrin that depend upon actin for internalization. *PLoS Pathog*, 2009, 5: e1000394
- 4 Heggeness M H, Scheid A, Choppin P W. Conformation of the helical nucleocapsids of paramyxoviruses and vesicular stomatitis virus: reversible coiling and uncoiling induced by changes in salt concentration. *Proc Natl Acad Sci USA*, 1980, 77: 2631–2635
- 5 Newcomb W W, Brown J C. Role of the vesicular stomatitis virus matrix protein in maintaining the viral nucleocapsid in the condensed form found in native virions. *J Virol*, 1981, 39: 295–299
- 6 Blumberg B M, Giorgi C, Rose K, *et al.* Preparation and analysis of the nucleocapsid proteins of vesicular stomatitis virus and sendai virus, and analysis of the sendai virus leader-NP gene region. *J Gen Virol*, 1984, 65: 769–779
- 7 Chanda P K, Banerjee A K. Two distinct populations of vesicular stomatitis virus ribonucleoprotein cores with differential sensitivities to micrococcal nuclease. *Biochem Biophys Res Commun*, 1979, 91: 1337–1345
- 8 Keene J D, Thornton B J, Emerson S U. Sequence-specific contacts between the RNA polymerase of vesicular stomatitis virus and the leader RNA gene. *Proc Natl Acad Sci USA*, 1981, 78: 6191–6195
- 9 Green T J, Macpherson S, Qiu S, *et al.* Study of the assembly of vesicular stomatitis virus N protein: role of the P protein. *J Virol*, 2000, 74: 9515–9524
- 10 Peluso R W, Moyer S A. Viral proteins required for the *in vitro* replication of vesicular stomatitis virus defective interfering particle genome RNA. *Virology*, 1988, 162: 369–376
- 11 Howard M, Wertz G. Vesicular stomatitis virus RNA replication: a role for the NS protein. *J Gen Virol*, 1989, 70: 2683–2694
- 12 Chen Z, Green T J, Luo M, *et al.* Visualizing the RNA molecule in the bacterially expressed vesicular stomatitis virus nucleoprotein-RNA complex. *Structure*, 2004, 12: 227–235
- 13 Green T J, Luo M. Resolution improvement of X-ray diffraction data of crystals of a vesicular stomatitis virus nucleocapsid protein oligomer complexed with RNA. *Acta Crystallogr D Biol Crystallogr*, 2006, 62: 498–504
- 14 Green T J, Zhang X, Wertz G W, *et al.* Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science*, 2006, 313: 357–360
- 15 Luo M, Green T J, Zhang X, *et al.* Conserved characteristics of the rhabdovirus nucleoprotein. *Virus Res*, 2007, 129: 246–251
- 16 Zhang X, Green T J, Tsao J, *et al.* Role of intermolecular interactions of vesicular stomatitis virus nucleoprotein in RNA encapsidation. *J Virol*, 2008, 82: 674–682
- 17 Leyrat C, Yabukarski F, Tarbouriech N, *et al.* Structure of the vesicular stomatitis virus N-P complex. *PLoS Pathog*, 2011, 7: e1002248
- 18 Albertini A A, Wernimont A K, Muziol T, *et al.* Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science*, 2006, 313: 360–363
- 19 Rainsford E W, Harouaka D, Wertz G W. Importance of hydrogen bond contacts between the N protein and RNA genome of vesicular stomatitis virus in encapsidation and RNA synthesis. *J Virol*, 2010, 84: 1741–1751
- 20 Green T J, Rowse M, Tsao J, *et al.* Access to RNA encapsidated in the nucleocapsid of vesicular stomatitis virus. *J Virol*, 2011, 85: 2714–2722
- 21 Tawar R G, Duquerroy S, Vonrhein C, *et al.* Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science*, 2009, 326: 1279–1283
- 22 Luo M, Green T J, Zhang X, *et al.* Structural comparisons of the nucleoprotein from three negative strand RNA virus families. *Virology*, 2007, 4: 72
- 23 Rudolph M G, Kraus I, Dickmanns A, *et al.* Crystal structure of the borna disease virus nucleoprotein. *Structure*, 2003, 11: 1219–1226
- 24 Imblum R L, Wagner R R. Inhibition of viral transcriptase by immunoglobulin directed against the nucleocapsid NS protein of vesicular stomatitis virus. *J Virol*, 1975, 15: 1357–1366
- 25 Clinton G M, Burge B W, Huang A S. Effects of phosphorylation and pH on the association of NS protein with vesicular stomatitis virus cores. *J Virol*, 1978, 27: 340–346
- 26 Thornton G B, De B P, Banerjee A K. Interaction of L and NS proteins of vesicular stomatitis virus with its template ribonucleoprotein during RNA synthesis *in vitro*. *J Gen Virol*, 1984, 65: 663–668
- 27 De B P, Banerjee A K. Requirements and functions of vesicular stomatitis virus L and NS proteins in the transcription process *in vitro*. *Biochem Biophys Res Commun*, 1985, 126: 40–49
- 28 Rahmeh A A, Schenk A D, Danek E I, *et al.* Molecular architecture of the vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci USA*, 2010
- 29 Gao Y, Lenard J. Multimerization and transcriptional activation of the phosphoprotein (P) of vesicular stomatitis virus by casein kinase-II. *EMBO J*, 1995, 14: 1240–1247
- 30 Pattnaik A K, Hwang L, Li T, *et al.* Phosphorylation within the amino-terminal acidic domain I of the phosphoprotein of vesicular stomatitis virus is required for transcription but not for replication. *J Virol*, 1997, 71: 8167–8175
- 31 Hwang L N, Englund N, Das T, *et al.* Optimal replication activity of vesicular stomatitis virus RNA polymerase requires phosphorylation of a residue(s) at carboxy-terminal domain II of its accessory subunit, phosphoprotein P. *J Virol*, 1999, 73: 5613–5620
- 32 Emerson S U, Schubert M. Location of the binding domains for the RNA polymerase L and the ribonucleocapsid template within different halves of the NS phosphoprotein of vesicular stomatitis virus. *Proc Natl Acad Sci USA*, 1987, 84: 5655–5659
- 33 Paul P R, Chattopadhyay D, Banerjee A K. The functional domains of the phosphoprotein (NS) of vesicular stomatitis virus (Indiana serotype). *Virology*, 1988, 166: 350–357
- 34 Chattopadhyay D, Banerjee A K. Phosphorylation within a specific domain of the phosphoprotein of vesicular stomatitis virus regulates transcription *in vitro*. *Cell*, 1987, 49: 407–414
- 35 Chattopadhyay D, Banerjee A K. Two separate domains within vesicular stomatitis virus phosphoprotein support transcription when added in trans. *Proc Natl Acad Sci USA*, 1987, 84: 8932–8936
- 36 Chen M, Ogino T, Banerjee A K. Interaction of vesicular stomatitis virus P and N proteins: identification of two overlapping domains at the N terminus of P that are involved in N0-P complex formation and encapsidation of viral genome RNA. *J Virol*, 2007, 81: 13478–13485
- 37 Ding H, Green T J, Luo M. Crystallization and preliminary X-ray analysis of a proteinase-K-resistant domain within the phosphoprotein of vesicular stomatitis virus (Indiana). *Acta Crystallogr D Biol Crystallogr*, 2004, 60: 2087–2090
- 38 Ding H, Green T J, Lu S, *et al.* Crystal structure of the oligomerization domain of the phosphoprotein of vesicular stomatitis virus. *J Virol*, 2006, 80: 2808–2814
- 39 Ribeiro E A Jr., Favier A, Gerard F C, *et al.* Solution structure of the C-terminal nucleoprotein-RNA binding domain of the vesicular stomatitis virus phosphoprotein. *J Mol Biol*, 2008, 382: 525–538
- 40 Green T J, Luo M. Structure of the vesicular stomatitis virus nucleocapsid in complex with the nucleocapsid-binding domain of the small polymerase cofactor, P. *Proc Natl Acad Sci USA*, 2009, 106: 11713–11718
- 41 Das S C, Pattnaik A K. Role of the hypervariable hinge region of phosphoprotein P of vesicular stomatitis virus in viral RNA synthesis and assembly of infectious virus particles. *J Virol*, 2005, 79: 8101–8102
- 42 Chen M, Ogino T, Banerjee A K. Mapping and functional role of the

- self-association domain of vesicular stomatitis virus phosphoprotein. *J Virol*, 2006, 80: 9511–9518
- 43 Mavrakis M, Iseni F, Mazza C, *et al.* Isolation and characterisation of the rabies virus N degrees-P complex produced in insect cells. *Virology*, 2003, 305: 406–414
- 44 Patton J T, Davis N L, Wertz G W. N protein alone satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. *J Virol*, 1984, 49: 303–309
- 45 Colonna R J, Banerjee A K. *In vitro* RNA transcription by the New Jersey serotype of vesicular stomatitis virus. II. Characterization of the leader RNA. *J Virol*, 1978, 26: 188–194
- 46 Testa D, Chanda P K, Banerjee A K. Unique mode of transcription in vitro by Vesicular stomatitis virus. *Cell*, 21: 267–275
- 47 Ball L A, White C N. Order of transcription of genes of Vesicular stomatitis virus. *Proc Natl Acad Sci USA*, 1976, 73: 442–446
- 48 Ball L A. Transcriptional mapping of vesicular stomatitis virus in vivo. *J Virol*, 1977, 21: 411–414
- 49 Abraham G, Banerjee A K. Sequential transcription of the genes of vesicular stomatitis virus. *Proc Natl Acad Sci USA*, 1976, 73: 1504–1508
- 50 Ogino T, Banerjee A K. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell*, 2007, 25: 85–97
- 51 Qanungo K R, Shaji D, Mathur M, *et al.* Two RNA polymerase complexes from vesicular stomatitis virus-infected cells that carry out transcription and replication of genome RNA. *Proc Natl Acad Sci USA*, 2004, 101: 5952–5957
- 52 Whelan S P, Wertz G W. Transcription and replication initiate at separate sites on the vesicular stomatitis virus genome. *Proc Natl Acad Sci USA*, 2002, 99: 9178–9183
- 53 Iseni F, Barge A, Baudin F, *et al.* Characterization of rabies virus nucleocapsids and recombinant nucleocapsid-like structures. *J Gen Virol*, 1998, 79: 2909–2919
- 54 Das S C, Nayak D, Zhou Y, *et al.* Visualization of intracellular transport of vesicular stomatitis virus nucleocapsids in living cells. *J Virol*, 2006, 80: 6368–6377
- 55 Harouaka D, Wertz G W. Mutations in the C-terminal loop of the nucleocapsid protein affect vesicular stomatitis virus RNA replication and transcription differentially. *J Virol*, 2009, 83: 11429–11439
- 56 Nayak D, Panda D, Das S C, *et al.* Single-amino-acid alterations in a highly conserved central region of vesicular stomatitis virus N protein differentially affect the viral nucleocapsid template functions. *J Virol*, 2009, 83: 5525–5534
- 57 Whelan S P, Wertz G W. The 5' terminal trailer region of vesicular stomatitis virus contains a position-dependent cis-acting signal for assembly of RNA into infectious particles. *J Virol*, 1999, 73: 307–315
- 58 Lyles D S, McKenzie M, Parce J W. Subunit interactions of vesicular stomatitis virus envelope glycoprotein stabilized by binding to viral matrix protein. *J Virol*, 1992, 66: 349–358
- 59 Swintek B D, Lyles D S. Plasma membrane microdomains containing vesicular stomatitis virus M protein are separate from microdomains containing G protein and nucleocapsids. *J Virol*, 2008, 82: 5536–5547
- 60 Ge P, Tsao J, Schein S, *et al.* Cryo-EM model of the bullet-shaped vesicular stomatitis virus. *Science*, 2010, 327: 689–693
- 61 Lyles D S, McKenzie M O, Kaptur P E, *et al.* Complementation of M gene mutants of vesicular stomatitis virus by plasmid-derived M protein converts spherical extracellular particles into native bullet shapes. *Virology*, 1996, 217: 76–87
- 62 Graham S C, Assenberg R, Delmas O, *et al.* Rhabdovirus matrix protein structures reveal a novel mode of self-association. *PLoS Pathog*, 2008, 4: e1000251
- 63 Gaudier M, Gaudin Y, Knossow M. Crystal structure of vesicular stomatitis virus matrix protein. *EMBO J*, 2002, 21: 2886–2892
- 64 Luo M, ed. *Negative Strand RNA Virus*. Singapore: World Scientific, 2011
- 65 PyMol. The PyMOL Molecular Graphics System, Version 1.3. Schrödinger, LLC

Biographical Sketch

Dr. Luo Ming received his Bachelor of Science degree from the Department of Chemistry (now College of Chemistry and Molecular Science), Wuhan University, in 1982. Through the China-United States Biochemistry Examination and Application (CUSBEA) program, he went to study biological sciences in Purdue University, Indiana, USA, and received his Doctor of Philosophy degree in 1987. He joined the faculty of the University of Alabama at Birmingham in the Department of Microbiology in 1987. He was appointed Professor in 1996 and is now Associate Director of the Center for Biophysical Sciences and Engineering. His research interest focuses on structure of viruses and viral proteins, including vesicular stomatitis virus, influenza virus, HIV and hepatitis viruses. He also develops innovative methods for X-ray protein crystallography. Design of novel antiviral drugs is another of his research interests.



Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.