Molecular architecture of the vesicular stomatitis virus RNA polymerase

Amal A. Rahmeh[a], Andreas D. Schenk[b], Eric I. Danek[b], Philip J. Kranzusch[a], Bo Liang[a], Thomas Walz[b,c], and Sean P. J. Whelan[a,1]

[a]Department of Microbiology and Molecular Genetics, [b]Department of Cell Biology, and [c]The Howard Hughes Medical Institute, Harvard Medical School, Boston MA 02115

Edited by Robert A. Lamb, Northwestern University, Evanston, IL, and approved October 12, 2010 (received for review September 9, 2010)

Nonsegmented negative-strand (NNS) RNA viruses initiate infection by delivering into the host cell a highly specialized RNA synthesis machine comprising the genomic RNA completely encapsidated by the viral nucleocapsid protein and associated with the viral polymerase. The catalytic core of this protein–RNA complex is a 250-kDa multifunctional large (L) polymerase protein that contains enzymatic activities for nucleotide polymerization as well as for each step of mRNA cap formation. Working with vesicular stomatitis virus (VSV), a prototype of NNS RNA viruses, we used negative stain electron microscopy (EM) to obtain a molecular view of L, alone and in complex with the viral phosphoprotein (P) cofactor. EM analysis, combined with proteolytic digestion and deletion mapping, revealed the organization of L into a ring domain containing the RNA polymerase and an appendage of three globular domains containing the cap-forming activities. The capping enzyme maps to a globular domain, which is juxtaposed to the ring, and the cap methyltransferase maps to a more distal and flexibly connected globule. Upon P binding, L undergoes a significant rearrangement that may reflect an optimal positioning of its functional domains for transcription. The structural map of L provides new insights into the interrelationship of its various domains, and their rearrangement on P binding that is likely important for RNA synthesis. Because the arrangement of conserved regions involved in catalysis is homologous, the structural insights obtained for VSV L likely extend to all NNS RNA viruses.

L protein | structure and function | viral replication

Nonsegmented negative-strand (NNS) RNA viruses initiate infection by delivering into the host cell a highly specialized RNA synthesis machine. This machine consists of a ribonucleoprotein complex (RNP) comprising the genomic RNA completely coated by the viral nucleocapsid (N) protein and associated with the RNA-dependent RNA polymerase (RdRP) complex (1). The catalytic core of the RNP is a single large (L) 250-kDa protein that contains enzymatic activities for nucleotide polymerization, mRNA cap addition, cap methylation, and polyadenylation (2–6). The location of all of the enzymatic activities necessary for transcription within a single polypeptide chain contrasts with the arrangement exhibited by the host cell and many other viruses, in which the different activities reside within separate proteins that assemble into a larger transcription complex (7).

Our understanding of the different activities of NNS RNA virus L proteins has been largely shaped by studies of vesicular stomatitis virus (VSV) because it is the only member of this order of viruses for which robust transcription can be reconstituted in vitro. The enzymatic activities of VSV L have been mapped at the single amino acid level. Within the primary sequence of L are six conserved regions (CRs I–VI) shared among all NNS RNA virus L proteins (8). The RdRP activity maps to CRIII (3), and it is also required for polyadenylation, which occurs through polymerase slippage on a template U tract (5). The capping activities of L differ from those of other viruses and their eukaryotic hosts. Specifically, an RNA:GDP polyribonucleotidyltransferase (PRNTase) activity present within CRV transfers 5′ monophosphate RNA onto a GDP acceptor through a covalent L–pRNA intermediate (4, 6). The resulting mRNA cap is subsequently modified by a dual specificity methyltransferase (MTase) activity within CRVI whereby ribose 2′-O methylation precedes and facilitates subsequent guanine-N-7 (G-N-7) methylation (2, 9).

The location of the nucleotide polymerization, capping, and cap methylation activities within separate regions of L (Fig. 1*A*) has led to the notion that L may be organized as a series of independent structural domains. Consistent with this idea, a C-terminal fragment of Sendai virus L retains the ability to G-N-7 methylate RNA in vitro (10). These domains, however, influence one another functionally, because failure to cap the nascent RNA chain results in the premature termination of transcription, and blocking methylation results in hyperpolyadenylation (11). These latter observations demonstrate that the 5′ mRNA processing activities of L intimately regulate its nucleotide polymerization activity and suggest that the 3D arrangement of the functional domains likely serves a key regulatory role during RNA synthesis.

The template for RNA synthesis in NNS RNA viruses is not naked RNA, but is instead a protein–RNA complex in which an oligomer of N protein completely coats the template. The atomic structure of the VSV N–RNA complex showed that the bases are sequestered between two lobes of each N molecule (12). During RNA synthesis, N must therefore be displaced for the polymerase to copy the RNA. L alone is not capable of copying the N–RNA template, but requires a viral cofactor, the phosphoprotein (P). P is an oligomeric protein that contains three domains: an N-terminal domain (P_{NTD}) that binds to L, a central oligomerization domain, and a C-terminal domain (P_{CTD}) that binds the N–RNA template (13). These domains of P are separated from one another by flexible linkers. The structure of P_{CTD} bound to an N–RNA complex revealed that P binds adjacent N molecules (14). The binding of L to P forms the RdRP complex and is thought to help position L to gain access to the RNA bases of the N–RNA template. Therefore, the active machine for RNA synthesis is a higher order protein–RNA complex within which the L catalytic core plays a central role.

To understand how the different functions of L are arranged within the overall architecture of the protein, we expressed and purified full-length and truncation mutants of L and visualized their structure by negative stain electron microscopy (EM). The results show that the RdRP of L resides within a ring-like domain that is similar to the overall architecture of other RdRPs (15). A set of flexible globular domains, which are appended to this ring, contain the capping machinery. In an L–P complex, the organization of the

Author contributions: A.A.R., T.W., and S.P.J.W. designed research; A.A.R., A.D.S., E.I.D., P.J.K., and B.L. performed research; A.A.R., T.W., and S.P.J.W. analyzed data; and A.A.R. and S.P.J.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: Sean_Whelan@hms.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1013559107/-/DCSupplemental.

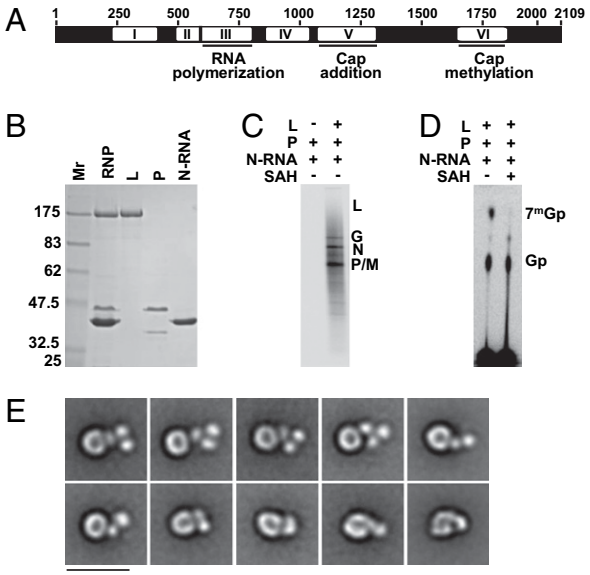

Fig. 1. Functional and structural organization of VSV L. (*A*) A schematic of the linear map of VSV L depicts the six conserved regions (CR I–VI) among NNS virus L proteins as white boxes separated by variable regions shown in black. The RNA polymerization, cap addition, and cap methylation activities have been mapped to CRIII, CRV, and CRVI, respectively. (*B*) Purification of recombinant RNP components. L and P were individually expressed and purified, and N–RNA was purified from recombinant VSV as previously described (4). Proteins were analyzed by 10% SDS/PAGE and stained with Coomassie blue, M_r, molecular weight marker. (*C*) Reconstitution of RNA synthesis from purified N–RNA, P, and L. Transcription reactions were performed in the presence of [α-^{32}P] GTP and the products analyzed by electrophoresis on acid-agarose gels. The identity of the five VSV mRNAs P, M (matrix), N, G (glycoprotein), and L is shown at *Right*. (*D*) Analysis of the mRNA cap structure. Transcription reactions were performed as in *C* in the absence or presence of the MTase inhibitor S-adenosyl homocysteine (SAH). The RNA products were digested with TAP and the products resolved by TLC. The mobilities of Gp and ^{7m}Gp are indicated at *Right*. (*E*) EM characterization of L. The 10 presented class averages show the ring-shaped core domain and illustrate the structural variability of the appendage. (Scale bar: 20 nm.)

domains of L is altered. This work provides structural insights into the L protein of an NNS RNA virus and offers a framework for understanding the coordination of the enzymatic activities of L within structurally distinct but functionally coordinated domains.

Results

Molecular Architecture of VSV L. To determine the overall architecture of L, we expressed and purified the protein as previously described (4). When combined with separately purified P and N–RNA template (Fig. 1*B*), the purified L catalyzed the synthesis of the five full-length viral mRNAs in vitro (Fig. 1*C*). The transcripts contained a methylated cap structure as shown by the release of ^{7m}Gp by tobacco acid pyrophosphatase (TAP) (Fig. 1*D*). These results demonstrate that the purified L is fully active for mRNA synthesis.

Electron microscopy (EM) of negatively stained L revealed monodispersed particles (Fig. S1*A*). We collected 60°/0° tilt pairs to calculate 3D reconstructions of L using the random conical tilt approach (16). A total of 8,806 particle pairs were selected, and the particles from the images of the untilted specimen were classified into 50 classes (Fig S1*B*). The 10 representative class averages (Fig. 1*E*) illustrate the overall appearance of L, which comprises a core ring-like domain decorated by an appendage. The ring is 90–100 Å in diameter with a 23–27 Å stain-filled center. The appendage comprises up to three globules, each 45–50 Å in diameter. The globules occupy variable spatial positions relative to the ring, and in some averages only one or two globules are visible. The structural variability seen in the class averages of L suggests flexible connections between the appendage and the ring, and between the individual globules. Three-dimensional reconstructions calculated for several of the classes using the particle images from the tilted specimen confirmed the features seen in the projection averages but did not provide further insights due to the limited resolution of the density maps (Fig S2*A*).

We noticed that ~10% of the L particles showed two L molecules in contact with each other. To assess whether the two L proteins in these "double particles" make specific interactions, we selected 1,099 particles and classified them into 10 classes (Fig S1*C*). The resulting class averages clearly resolved the two rings but not the appendages, suggesting that L has a tendency to dimerize, but not through highly specific interactions.

Therefore, the EM images of L provide evidence for the presence of four distinct structural domains: a core ring domain, decorated by an appendage that itself comprises three globular domains. Because RdRP molecules often adopt a ring-like conformation (15), we hypothesized that the RdRP activity of L would reside within the ring domain and that the capping enzymes would likely be localized within the appendage.

Mapping the Activities of L to the Domains of the Protein. As a first step toward defining the regions of L that constitute the appendage and ring-like structures, we performed limited proteolysis of L. Trypsin cleavage of L resulted in the release of a major fragment (F1) and two minor fragments (F2 and F3), each of which retained the N terminus (Fig. 2*A*). Estimation of the size of those fragments indicated that trypsin cleaved at the end of the long variable region between CRV and CRVI (F1), between CRIV and CRV (F2), and between CRIII and CRIV (F3). The resulting fragments separated the CRs of L that contain the polymerase, capping, and cap methylase activities, supporting the idea that the CRs map to distinct structural domains. Guided by these results, as well as by

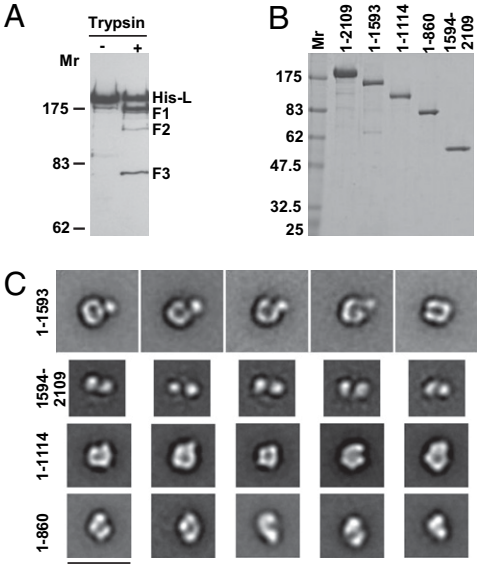

Fig. 2. The structural domains of VSV L. (*A*) Limited proteolysis of L. Purified L was digested with trypsin, before separation by 6% SDS/PAGE. The cleavage products were detected using an anti-His antibody that recognizes the N-terminal tag. Fragments with intact N termini that were selected for further analysis are identified as F1, F2, and F3. (*B*) Fragments corresponding to those released by trypsin digestion (1–1,593 for F1, 1–1,114 for F2, and 1–860 for F3) and a complementary fragment to F1 (1,594–2,109) were designed with an N-terminal 6× His tag, expressed and purified. The purified proteins were separated by 10% SDS/PAGE and visualized by Coomassie blue. (*C*) Electron microscopic characterization of the L fragments. The class averages show the structural features of the various L fragments. (Scale bar: 20 nm.)

sequence conservation, hydrophobicity profile and secondary structure predictions, we separately expressed and purified amino acids 1–1,593, 1–1,114, and 1–860 as the approximate equivalents to F1, F2, and F3, respectively. We also expressed amino acids 1,594–2,109 as the complementary fragment to F1. The mobility of those purified recombinant proteins on SDS/PAGE was consistent with that of the corresponding tryptic fragments (Fig. 2*B*). Additional fragments of L were also designed and tested for expression, but yielded no or low levels of soluble protein and were not pursued further (Table S1).

To determine which parts of L comprise its different structural features, we examined the above fragments by EM and calculated 50 class averages for each (Fig. S3, representative class averages are shown in Fig. 2*C*). EM analysis revealed that amino acids 1–1,593 comprise the ring and one of the globules of the appendage, and 1,594–2,109 form the complementary two globules. The appearance of the two fragments combined resembles that of full-length L, suggesting that these regions can fold independently of one another. Consistent with the hypothesis that the ring domain contains the RdRP and the capping enzymes reside within the appendage, residues 1–1,593 contain the ring and one globule of the appendage. Further support for this idea was provided by inspection of residues 1–1,114, which adopted a similar 90–100 Å ring. In some classes, however, the ring formed by 1–1,114 adopted a slightly more open or less compact conformation (averages 4 and 5), suggesting that the C-terminal regions of L may contribute to the stability of the ring and provide important connections with the appendage. Further truncation of L to residue 860, which still retains the RdRP motif, resulted in an open C-like structure, demonstrating that CRIV forms a key structural component of the ring domain. The ring is thus demarcated by the N-terminal region of L spanning CRI–CRIV.

As residues 1–1,593 contain the ring as well as a globule from the appendage, and 1–1,114 contains only the ring, it appears that the capping enzyme (CRV) is part of the globular appendage. The size of this globular domain does not appear large enough, however, to account for the entire 480-residue difference between the two fragments. We favor the hypothesis that CRV (residues 1,069–1,305) constitutes the globule and that the variable region between CRV and CRVI (1,306–1,593) constitutes a largely unstructured hinge region that is not readily discernible by negative stain EM. Such a hinge region could also account for the variable orientation of the two distal globules provided by 1,594–2,109 in full-length L (Fig. 1*E*). Residues 1,594–2,109 encompass the methyltransferase CRVI (1,640–1,835), and a more variable C-terminal region (1,836–2,109). It is likely, therefore, that each forms an independent globular domain, which would account for the two globules visible for this fragment.

Taken together, these data define the general boundaries of the structural domains of L as an RdRP-containing ring domain extending from the N terminus through the end of CRIV, and an appendage extending from CRV through the C terminus. The appendage can be further delineated as a set of three globular domains that comprise a capping (CRV), a cap methylating (CRVI), and a C-terminal domain.

Overall Architecture of L Is Critical for Full Function. As residues 1–1,593 and 1,594–2,109 fold independently, and together constitute the full-length protein, we tested whether the two fragments could complement one another functionally. Equimolar amounts of the two fragments did not reconstitute RNA synthesis in vitro (Fig. 3*A*). However, this experiment did not permit us to differentiate between a loss of recruitment to the N–RNA template and a loss of catalytic activity of the fragments. Because template association requires P binding, we next tested whether the fragments retained the ability to bind P. Using a fully functional variant of P, in which the N terminus is tagged with green fluorescent protein (eGFP) (17), we tested its ability to associate with L using a coprecipitation assay. Full-length L efficiently coprecipitated eGFP–P, whereas fragments 1–1,593 and 1,594–2,109 bound substantially less eGFP–P, and efficient binding was not restored by combining the two fragments (Fig. 3*B*). This result suggests that P binding is not simply achieved through a bivalent binding to both termini of L, but requires a tertiary structure that depends on an intact L. Because P binding is an essential requirement for recruitment of L to the N–RNA template, it is not surprising that the two fragments could not complement one another for RNA synthesis. We therefore proceeded to evaluate whether the architecture of L alters upon binding to P.

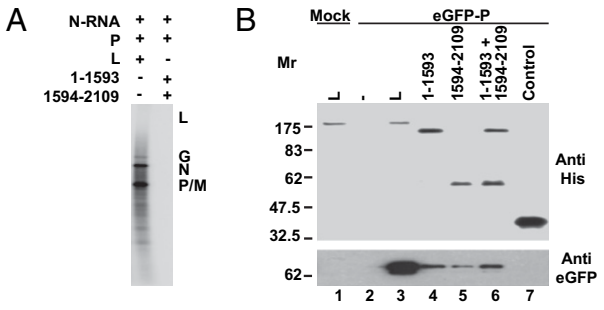

Fig. 3. Functional analysis of the complementary N- and C-terminal fragments of VSV L. (*A*) RNA synthesis. Equimolar quantities of L or 1–1,593 + 1,594–2,109 were used to reconstitute RNA synthesis in vitro and the products analyzed as in Fig 1*C*. (*B*) P protein binding. Full-length L, or fragments 1–1,593, 1,594–2,109, 1–1,593 + 1,594–2,109 were used as bait to pull down eGFP–P expressed in BSRT7 cells. Western blots show the His-tagged proteins (*Upper*) and the corresponding captured eGFP–P (*Lower*). Mock transfected BSRT7 cells (lane 1), no L (lane 2), and His-tagged N-terminal domain of anthrax lethal factor (lane 7) are shown as specificity controls.

Architecture of L Alters upon Binding to P. To examine the architecture of a functional L–P complex, we first purified the complex by size exclusion chromatography (Fig. 4*A*). L [calculated molecular weight (Mwt) = 241 kDa] eluted in a single peak corresponding to ~120 kDa. This elution profile is likely reflective of the compaction of part of L into the ring domain and indicative that L is mostly a monomer in solution. P (calculated Mwt = 29 kDa), eluted in a single peak corresponding to 134 kDa, indicative both of its elongated structure and oligomerization. Mixing L with a molar excess of P resulted in the formation of a complex containing both L and P that eluted in a peak corresponding to ~306 kDa, whereas the uncomplexed excess P eluted at 134 kDa. This L–P complex was fully competent for mRNA synthesis in vitro (Fig. 4*B*).

Like L alone, EM images revealed that the L–P complex consisted of a mixture of single and double particles, but the double particles were much more prevalent for the L–P complex, accounting for approximately one-third of the particle population (Fig S4*A*). We collected 60°/0° tilt pairs of the L–P complex and selected 5,885 particle pairs for the single particles and 2,780 particle pairs for the double particles. The particles from the images of the untilted specimen were classified into 50 classes for the single particles (Fig S4*B*) and into 20 classes for the double particles (Fig S4*C*). The approximate dimensions of the single and double species are 130 × 95 Å and 240 × 180 Å, respectively. Three-dimensional reconstructions calculated using particles from the images of the tilted specimen were consistent with the features seen in the projection averages but again did not provide further insights (Fig. S2*B*).

The single species of L–P adopts a general conformation resembling a "6," where a clamp-like region extends into a curved arm. The rearrangement of L upon binding P has important implications for its function, because it is the L–P complex that is active for RNA synthesis. Although the limited resolution prevented us from de-

MICROBIOLOGY

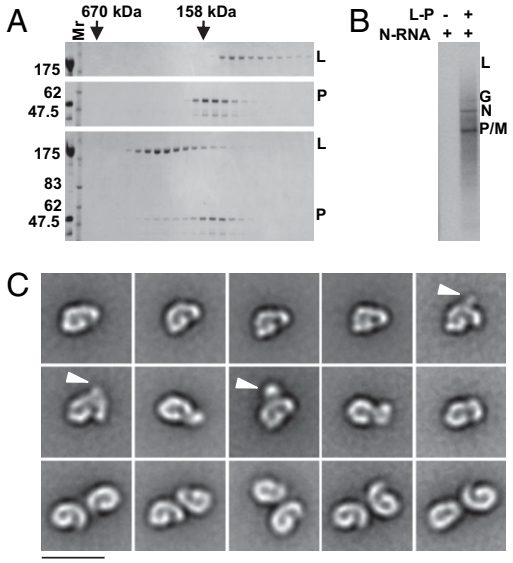

Fig. 4. Functional and structural characterization of the L–P complex. (*A*) Size exclusion chromatography of L (*Top*), P (*Middle*), and L–P (*Bottom*). The fractions eluted from a Superdex 200 column were analyzed by 10% SDS/PAGE and the proteins visualized by Coomassie blue. The elution positions of two Mwt standards are indicated by arrows. (*B*) RNA synthesis in vitro. Transcription reactions were reconstituted using N–RNA and the L–P complex peak fraction and the products analyzed as in Fig 1*C*. (*C*) EM characterization of L–P. The *Top* two rows show representative averages of the single L–P complexes and the *Bottom* row shows representative averages of the double species. The arrowheads indicate the additional globular density occasionally seen in averages of the single L–P complexes. (Scale bar: 20 nm.)

finitively assigning regions of L or P to the various structural features of the L–P complex, by comparing the general features of L alone with the L–P complex, we posit the following model. The single species of L–P resembles a subset of averages of uncomplexed L in which the three globules are not discernible (compare averages in the two *Top* rows in Fig. 4*C* with averages 8–10 in Fig. 1*E*). This similarity suggests that P binding may orient the domains of the L appendage in a specific conformation and that the visible features of the L–P complex are mostly those of L. Because P has a relatively small mass that is distributed throughout an elongated shape comprising highly flexible regions, it is not surprising that distinctive features of P cannot be discerned. The clamp-like region of the L–P complex (90–100 Å) and its stain-accessible center (~25 Å) are similar in dimensions to those of the L ring domain. We suggest that the two structures are related and that P binding facilitates changes in the ring domain. In this model, the curved arm would represent a rearrangement of the globular domains of the appendage.

Unlike for the double particles of L alone, class averages of the double particles of the L–P complex resolve the entire molecule (*Bottom* row in Fig. 4*C*), indicating that the dimerization is more specific. The architecture of the individual members of the double species is similar to the single species, suggesting that they represent a higher order organization containing two L–P complexes. The orientation of the particles in each pair is variable but the connection is often close to the position where the appendage connects to the clamp. This is the same position, where the averages of the single species occasionally show an additional globular density (indicated by an arrowhead in Fig. 4*C*). The association between the two L–P complexes thus appears to be mediated by a molecular bridge. Because P itself is an oligomeric protein, it likely provides the organizing center for the two L molecules. The variable arrangement of the two L molecules is compatible with the presence of known flexible linkers in P connecting the L-binding P_{NTD} to the oligomerization domain.

Discussion

This study provides a view of the molecular architecture of the VSV L protein, maps its different catalytic activities to distinct structural domains, and shows rearrangements in L following complex formation with its cofactor P. The molecular anatomy of L provides insights into the functions of the protein during RNA synthesis, including how the polymerase might engage an encapsidated template and how the various enzymatic activities of L can be coordinated within one multifunctional protein. Furthermore, in combination with the atomic level structures of the N–RNA template and portions of P, the EM images of VSV L now provide an overview of each component of the RNP, offering a framework to understand the functional coordination of the components of this RNA synthesis machine. Because the L proteins of all NNS RNA viruses display a homologous arrangement of the various conserved regions, the structural insights obtained here with VSV will likely apply to those viruses as well. The structural arrangement of L also provides some insights into the evolution of this multifunctional catalytic protein as well as a basis for further structural and functional studies.

Domain Organization of L and Rearrangements on Complex Formation with P. Negative-stain EM of L shows that it always forms a core ring structure with an appendage (Fig. 5*A*). Deletion analysis showed that, in addition to containing CRIII, which includes the active site of the RdRP, the ring comprises CRI, -II, and -IV. Although functions have not yet been assigned to CRI and -II of VSV L, experiments with Sendai virus implicated CRI in binding P (18) and CRII in binding the RNA template (19). The demonstration that CRI and CRII are components of the RdRP-containing ring domain is consistent with their suggested roles in facilitating RNA synthesis. These CRs, as well as CRIV, seem to play an additional structural role in maintaining the integrity of the ring domain.

Whereas many polymerases appear ring-like (15), the ring domain of VSV L shares particular parallels with the structures of the RdRPs of double strand (ds) RNA viruses. Specifically, for dsRNA viruses, the RdRP is enclosed within a large cage-like structure of ~80 Å built from a 1,100–1,300 amino acid protein (20). The ring-like domain of VSV of ~90–100 Å is approximately built from the N-terminal 1,114 amino acids of L. Functionally, the cage-like structure provides multiple channels that allow spatial coordination for template entry and RNA product exit in the crowded interior of a dsRNA virus core (20). For VSV L, the ring structure might help the RdRP negotiate the specific constraint of copying an encapsidated template. The stain-admitting center of the L ring is similar in diameter (~25 Å) to the template entry channel of other RdRPs (15), suggesting that it serves an analogous function. However, the VSV template is markedly different because it is completely coated by N, rendering the RNA bases inaccessible (12) and mandating a transient displacement of

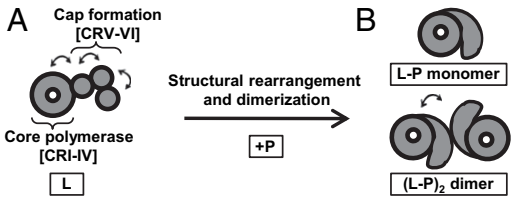

Fig. 5. Molecular architecture of VSV L. (*A*) L is organized into a core ring structure harboring the RdRP domain and a flexible appendage containing the activities necessary for cap formation (capping + methylation). The arrows depict putative flexible linkers. (*B*) Upon binding to P, L undergoes a structural rearrangement. The L–P complex exists as a monomer or a dimer in which the L pairs are likely bridged by interaction with an oligomer of P. The arrow represents the variable orientation of L proteins in the dimers.

N during RNA synthesis. In a putative model in which the RNA chain could thread through the center of the L ring (or the structurally related clamp-like region of the L–P complex) and N molecules are sequentially dissociated from the RNA, the thickness of the L ring (~50 Å) suffices to accommodate a stretch of RNA released from one or two molecules of N (width of N is ~27 Å). The mechanism of N displacement remains uncertain, but likely involves P, which would interact with the elongating L through its P_{NTD} and modulate/displace N through its P_{CTD}.

The appendage spans the C terminus of L starting at CRV. Finer mapping of the appendage indicated that CRV containing the capping enzyme likely occupies the globular domain juxtaposed to the ring, and the two distal globules correspond to the MTase provided by CRVI and a structured C-terminal domain of L. A specific function has not yet been assigned to this latter region. Our results also suggest that the variable region between CRV and CRVI provides a flexible hinge between the proximal globule and the two distal globules. This notion is consistent with genetic experiments, in which the function of L was retained upon in-frame insertion of eGFP within this variable region, suggesting that it constitutes a hinge region in L (21).

The L–P complex adopts an overall conformation resembling a "6," in which a clamp-like region extends into a curved arm (Fig. 5*B*). In contrast to the organization of L into distinct structural domains, the L–P complex exhibits less separate features. Comparative analysis of the structural organization of L and L–P suggests a putative model, in which the clamp-like region corresponds to the ring domain of L and the curved arm reflects a rearrangement in the appendage. Distinctive features of P are not discerned due to its relatively small mass and elongated shape comprising highly disordered regions. In the dimeric species, two rearranged molecules of L are bridged by an oligomer of P. A higher resolution structure is needed to confirm this model.

Insights into the Regulation of the Enzymatic Activities of the Multifunctional L Protein. The finding that the domains containing the different enzymatic activities of L are rearranged on binding P provides a framework for understanding the regulation of L activities. Detection of the RdRP activity of L in the in vitro reconstituted system is absolutely dependent on the presence of P (4), and attempts to reconstitute RNA synthesis using a nonencapsidated model RNA template have proved unsuccessful (22). Thus, whether P has a direct effect on RdRP activity or whether it is only required to recruit the polymerase to the template is unclear. The results of this study indicate that the role of P is not simply to physically recruit L to the N–RNA template. In the L–P complex, a clearly delineated RdRP ring is no longer discernible. Instead, a clamp-like region of similar dimensions is observed, suggesting that P binding influences the dynamic structural properties of the RdRP-containing region.

Upon binding P, the globular domains visible in L that contain the capping activities are also significantly rearranged. We and others have shown that purified L alone catalyzes the separate reactions of capping and methylation of exogenous synthetic RNA substrates and that P binding does not alter those catalytic properties of L (6, 9). The rearrangement of L on P binding might thus be involved in coordinating the capping activities with ongoing RNA synthesis. For instance, flexibility between the ring and the appendage may allow the repositioning of the capping apparatus such that the nascent RNA chain has access to it during mRNA synthesis, but not during genome replication. Furthermore, because during transcription the capping activities of L influence polymerase elongation (11), the observed rearrangements might facilitate the coordination between the capping apparatus and the RdRP domain.

The observation of dimeric L–P complexes raises the possibility that the functional polymerase may include a higher order L–P complex. Earlier work with Sendai virus has indicated that the polymerase may function as a higher order oligomer, possibly a dimer (23). This conjecture is supported by the complementation of L mutants in a cell-based reconstitution assay for RNA synthesis (24). Therefore, the dimeric L–P complex observed here might be functionally relevant for RNA synthesis. This study, however, cannot unambiguously distinguish which form of the L–P complex is at play during RNA synthesis, because a dynamic equilibrium likely exists between the single and double species.

Evolutionary Implications. The ring domain of VSV L appears similar to other polymerases, which is consistent with the common ancestral origin of all polymerases (15). Appended to that ring domain is a capping machine that contains an unconventional PRNTase domain and a dual specificity MTase domain. The MTase domain shares clear homology with known ribose 2′-O MTases (25), supporting the existence of a common ancestor of CRVI and those proteins. The PRNTase, however, is unique to NNS RNA virus L proteins (6). How the capping activities were coopted within the same protein is uncertain.

The L proteins of NNS RNA viruses are distinguished by the homologous arrangement of CRs in their linear sequence, which suggests that their molecular architecture will likely be similar. In parallel experiments with the L protein of a member of the *Arenaviridae* family of segmented negative-sense RNA viruses, Machupo virus (MACV), we also observe a ring domain similar in dimensions to that of VSV L (26). For MACV, the ring is decorated by a distinct appendage, likely reflective of the cap stealing mechanism used by those viruses to produce capped mRNA (27). Despite the lack of extensive sequence conversation between the L proteins of MACV and NNS RNA viruses, the similarities between the two structures are striking and suggest a likely evolutionary pathway, in which a common ancestral RdRP was framed in a ring structure onto which distinct capping activities were appended.

Materials and Methods

Protein Expression and Purification. Recombinant L and L fragments were expressed in Sf21 cells, and P was expressed in BL21 (DE3) as described in *SI Materials and Methods*. N–RNA template, L and P were prepared as previously described (4). L fragments were purified by Ni-NTA agarose (Qiagen) followed by ion exchange chromatography as described in *SI Materials and Methods*. For gel filtration experiments, 117 μg of L or 60 μg of P were individually passed through a Superdex 200 HR 10/30 (GE Healthcare) or first mixed together for 1 h on ice. The column was run at 0.25 mL/min and 250-μL fractions were collected. Apparent Mwts were extrapolated from a standard curve calculated from the elution volumes of a gel filtration standard (Biorad).

Electron Microscopy and Image Processing. Samples were adsorbed to glow-discharged, carbon-coated EM grids and stained with 0.75% (wt/vol) uranyl formate as described (28). The 60°/0° image pairs for full-length L and the L–P complex or only images of untilted specimens for all of the other samples were collected and processed as described in detail in *SI Materials and Methods*.

Trypsin Digestion. Lyophilized trypsin (Worthington) was resuspended at 0.5 μg/μL in 1 mM HCl. A total of 5 μg of His-tagged L were mixed with 10 ng trypsin (500:1 wt/wt) in 20 μL digestion buffer (50 mM Tris HCl pH 8.8, 200 mM NaCl, 2 mM $CaCl_2$) at 20 °C for 15 min. Digestion was stopped by addition of 2 mM PMSF for 5 min, then boiling in SDS/PAGE loading buffer. Products were separated by 6% SDS/PAGE, transferred to a nitrocellulose membrane, and blotted with an anti His-tag monoclonal antibody (Clontech).

Ni-NTA Pulldown Assay. Total cell lysates expressing eGFP–P were prepared and used for coprecipitation of eGFP–P with His-tagged L or L fragments using Ni-NTA agarose beads as described in *SI Materials and Methods*. The precipitated proteins were separated on duplicate 10% SDS/PAGE gels, transferred to a nitrocellulose membrane, and blotted with anti-His-tag monoclonal antibody or Aequora victoria GFP (A.v. peptide) polyclonal antibody detecting eGFP (BD Biosciences).

RNA Transcription and Cap Analysis. In vitro transcription reactions and cap analysis were essentially performed as previously described (4) with modifications as described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We acknowledge the exceptional support in protein production by Robin Ross and Lauren Perry in Core D of the New England Regional Center of Excellence in Biodefense and Emerging Infectious Diseases (NERCE-BEID). We thank Silvia Piccinotti for critical reading of the manuscript. We thank John Collier and Brad Pentelute (Harvard Medical School, Boston, MA) for providing the N-terminal domain of anthrax lethal factor. This study was supported by National Institutes of Health Grants AI059371 (to S.P.J.W.) and AI057159 (NERCE-BEID). S.P.J.W. is a recipient of a Burroughs Wellcome Investigators in the Pathogenesis of Infectious Disease Award. A.D.S. is supported by a Swiss National Science Foundation fellowship. T.W. is a Howard Hughes Medical Institute investigator.

1. Emerson SU, Wagner RR (1972) Dissociation and reconstitution of the transcriptase and template activities of vesicular stomatitis B and T virions. *J Virol* 10:297–309.
2. Li J, Fontaine-Rodriguez EC, Whelan SP (2005) Amino acid residues within conserved domain VI of the vesicular stomatitis virus large polymerase protein essential for mRNA cap methyltransferase activity. *J Virol* 79:13373–13384.
3. Sleat DE, Banerjee AK (1993) Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. *J Virol* 67:1334–1339.
4. Li J, Rahmeh A, Morelli M, Whelan SP (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.
5. Barr JN, Whelan SP, Wertz GW (1997) cis-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation. *J Virol* 71:8718–8725.
6. Ogino T, Banerjee AK (2007) Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell* 25:85–97.
7. Koonin EV, Moss B (2010) Viruses know more than one way to don a cap. *Proc Natl Acad Sci USA* 107:3283–3284.
8. Poch O, Blumberg BM, Bougueleret L, Tordo N (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.
9. Rahmeh AA, Li J, Kranzusch PJ, Whelan SP (2009) Ribose 2′-O methylation of the vesicular stomatitis virus mRNA cap precedes and facilitates subsequent guanine-N-7 methylation by the large polymerase protein. *J Virol* 83:11043–11050.
10. Ogino T, Kobayashi M, Iwama M, Mizumoto K (2005) Sendai virus RNA-dependent RNA polymerase L protein catalyzes cap methylation of virus-specific mRNA. *J Biol Chem* 280:4429–4435.
11. Li J, Rahmeh A, Brusic V, Whelan SP (2009) Opposing effects of inhibiting cap addition and cap methylation on polyadenylation during vesicular stomatitis virus mRNA synthesis. *J Virol* 83:1930–1940.
12. Green TJ, Zhang X, Wertz GW, Luo M (2006) Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* 313:357–360.
13. Ding H, Green TJ, Lu S, Luo M (2006) Crystal structure of the oligomerization domain of the phosphoprotein of vesicular stomatitis virus. *J Virol* 80:2808–2814.
14. Green TJ, Luo M (2009) 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* 106:11713–11718.
15. Ferrer-Orta C, Arias A, Escarmís C, Verdaguer N (2006) A comparison of viral RNA-dependent RNA polymerases. *Curr Opin Struct Biol* 16:27–34.
16. Radermacher M, Wagenknecht T, Verschoor A, Frank J (1987) Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of Escherichia coli. *J Microsc* 146:113–136.
17. Schott DH, Cureton DK, Whelan SP, Hunter CP (2005) An antiviral role for the RNA interference machinery in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 102:18420–18424.
18. Chandrika R, Horikami SM, Smallwood S, Moyer SA (1995) Mutations in conserved domain I of the Sendai virus L polymerase protein uncouple transcription and replication. *Virology* 213:352–363.
19. Smallwood S, Easson CD, Feller JA, Horikami SM, Moyer SA (1999) Mutations in conserved domain II of the large (L) subunit of the Sendai virus RNA polymerase abolish RNA synthesis. *Virology* 262:375–383.
20. McDonald SM, Tao YJ, Patton JT (2009) The ins and outs of four-tunneled Reoviridae RNA-dependent RNA polymerases. *Curr Opin Struct Biol* 19:775–782.
21. Ruedas JB, Perrault J (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: 12241–12252.
22. Moyer SA, Smallwood-Kentro S, Haddad A, Prevec L (1991) Assembly and transcription of synthetic vesicular stomatitis virus nucleocapsids. *J Virol* 65:2170–2178.
23. Cevik B, Smallwood S, Moyer SA (2007) Two N-terminal regions of the Sendai virus L RNA polymerase protein participate in oligomerization. *Virology* 363:189–197.
24. Smallwood S, Cevik B, Moyer SA (2002) Intragenic complementation and oligomerization of the L subunit of the sendai virus RNA polymerase. *Virology* 304:235–245.
25. Ferron F, Longhi S, Henrissat B, Canard B (2002) Viral RNA-polymerases—a predicted 2′-O-ribose methyltransferase domain shared by all Mononegavirales. *Trends Biochem Sci* 27:222–224.
26. Kranzusch PJ, et al. (2010) Assembly of a functional Machupo virus polymerase complex. *Proc Natl Acad Sci USA* 107:20069–20074.
27. Lelke M, Brunotte L, Busch C, Günther S (2010) An N-terminal region of Lassa virus L protein plays a critical role in transcription but not replication of the virus genome. *J Virol* 84:1934–1944.
28. Ohi M, Li Y, Cheng Y, Walz T (2004) Negative staining and image classification - powerful tools in modern electron microscopy. *Biol Proced Online* 6:23–34.

Supporting Information

Rahmeh et al. 10.1073/pnas.1013559107

SI Materials and Methods

Recombinant Protein Expression. Plasmids carrying functional VSV L and P genes were described previously (1). Fragments of L were amplified from the L gene and confirmed by sequencing. A hexahistidine (6× His) tag was introduced at the N terminus of L and one set of L fragments (Table S1). In a second set, the L fragments were inserted in frame of a 6× His-SUMO* tag described in ref. 2 (Table S1). The genes were inserted into pFASTBAC DUAL vector (Invitrogen) under control of the polyhedrin promoter as described previously (3). Recombinant baculovirus (BV) were recovered following transfection of bacmid DNA into *Spodoptera frugiperda* Sf21 cells using Cellfectin (Invitrogen). The recombinant BVs were amplified and used to infect Sf21 cells growing in spinner flasks at 1.5×10^6 cells/mL at a multiplicity of infection (MOI) of 3–5. The cells were collected at 65–72 h postinfection. Protein expression was evaluated by Western blot with an anti-His monoclonal antibody (Clontech).

P was cloned downstream of the His tag of a pET–16b vector (Novagen). The plasmid was transformed in BL21 (DE3) *Escherichia coli*. The cells were grown in LB containing 100 μg/mL ampicillin and induced at A_{600} of 0.6 with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) at 30 °C for 4 h.

Purification of L Fragments. All fragments were first purified using Ni-NTA agarose chromatography (Qiagen), as described previously for full-length L and eluted in 50 mM NaH_2PO_4 pH 7.4, 300 mM NaCl, 200 mM Imidazole (3). Peak fractions were diluted in an equal volume of buffer A (50 mM Tris-HCl pH 7.0, 10% glycerol, 1 mM DTT). Fragments 1–1,593, 1–1,114, and 1–860 were individually loaded on a Mono S 5/50 GL column (GE Healthcare) equilibrated with buffer A containing 150 mM NaCl. The column was washed with five column volumes (CV) of equilibration buffer and eluted with a 10 CV continuous gradient of 150 mM to 1 M NaCl in buffer A. Fragment 1,594–2,109 was first loaded on a similarly equilibrated Mono Q 5/50 GL column (GE Healthcare) and collected from the flow through, then subsequently purified on the Mono S column as described for the other fragments.

Electron Microscopy Image Collection and Processing. The 60°/0° image pairs for full-length L and the L–P complex or only images of untilted specimens for all of the other samples were collected with a Tecnai T12 electron microscope (FEI) equipped with an LaB_6 filament and operated at an acceleration voltage of 120 kV. Images were recorded on imaging plates at a magnification of 67,000× and a defocus of about −1.5 μm using low-dose procedures. Imaging plates were read out with a scanner (DITABIS) using a step size of 15 μm, a gain setting of 20,000, and a laser power setting of 30%; 2 × 2 pixels were averaged to yield a pixel size of 4.5 Å at the specimen level (4). Class averages and 3D reconstructions were calculated using the SPIDER software package (5). BOXER, part of the EMAN software package (6), was used to interactively select particles from images of the L fragments. For the 1–1,593 fragment, 8,874 particles were selected from 112 images and windowed into 64 × 64-pixel images. For the 1,594–2,109 fragment, 12,888 particles were selected from 56 images and windowed into 44 × 44 pixel images. For the 1–1,114 fragment, 15,954 particles were selected from 40 images and windowed into 50 × 50 pixel images. For the 1–860 fragment, 10,518 particles were selected from 80 images and windowed into 50 × 50 pixel images. The particles in the data sets were classified using the SPIDER software package (5). The particles were rotationally and translationally aligned and subjected to 10 cycles of multireference alignment. Each round of multireference alignment was followed by *K*-means classification into 50 classes. The references used for the first multireference alignment were randomly chosen from the particle images.

For full-length L, 9,905 pairs of particles (8,806 single and 1,099 double particles) were selected from 108 60°/0° tilt pairs using WEB, which is part of the SPIDER software package, windowed into 120 × 120 pixel images, and the particles selected from the untilted specimen were classified into 50 classes (single particles) or 10 classes (double particles) as described above. For the L–P complex, 8,665 pairs of particles (5,885 single and 2,780 double particles) were selected from 217 60°/0° tilt pairs using WEB, windowed into 120 × 120 pixel images, and the particles selected from the untilted specimen were classified into 50 classes (single particles) or 20 classes (double particles) as described above. A total of 44 of the classes for the single species of L and 17 of the classes for the single species of the L–P complex were selected to calculate individual 3D reconstructions using the corresponding particles selected from the tilted specimen. The 3D reconstructions were generated by using the back-projection and angular refinement procedures implemented in SPIDER and visualized using DNG from the OpenStructure framework (7). Selected 3D reconstructions showing the structural features seen in the respective class averages are presented in Fig S2.

In Vitro Transcription. The mRNAs were synthesized in vitro using 5 μg of N–RNA template, 0.5 μg of P, and 1 μg of L or (1 μg of 1–1,593 + 1 μg of 1,594–2,019) or 1.5 μg of the S200 peak fraction for the L–P complex. Reactions were performed in the presence of 1 mM ATP; 0.5 mM CTP, GTP, and UTP; 15 μCi of [α-^{32}P] GTP (PerkinElmer), 30% vol/vol rabbit reticulocyte lysates (Promega), 0.05 μg/μL actinomycin D (Sigma) in transcription buffer [30 mM Tris-HCl pH 8.0, 33 mM NH_4HCl, 50 mM KCl, 4.5 mM $Mg(OAC)_2$, 1 mM DTT, 0.2 mM spermidine, 0.05% Triton X-100]. Reactions included 1 mM of the MTase inhibitor S-adenosyl homocysteine (SAH) (Sigma) where indicated. The products were purified by phenol/chloroform extraction, separated by acid-agarose urea gel electrophoresis as previously described (3). and visualized using a PhosphorImager (GE Healthcare).

Tobacco Acid Pyrophosphatase (TAP) Digestion. A total of 50 ng of purified RNA from in vitro transcription reactions performed in the absence or presence of 1 mM SAH were digested with 1 U of TAP (Epicenter) in 10 μL TAP reaction buffer (50 mM sodium acetate pH 5.0, 1 mM EDTA, 0.1% β-mercaptoethanol and 0.01% Triton X-100) at 37 °C for 2 h, and the products were analyzed by TLC on PEI-F cellulose sheets (EM Biosciences) using 1.2 M LiCl as solvent as described previously (8). Spots were visualized using a PhosphorImager (GE Healthcare). Markers 7^mGp and Gp were visualized by UV shadowing at 254 nm.

Preparation of Total Cell Extracts Expressing eGFP–P. Plasmid peGFP–P was created by cloning the eGFP–P ORF from the previously described recombinant vesicular stomatitis virus (9) into the T7 expression vector pGEM3 (Promega). Confluent 6-cm dishes with 2×10^6 BSR-T7 cells were infected at an MOI of 3 with recombinant vaccinia virus (vTF73) expressing T7 RdRP (10). After 1 h, cells were washed and transfected with 3.8 μg peGFP–P plasmid using Lipofectamine 2000 standard protocol (Invitrogen) and harvested at 24 h. Total cell extracts were prepared by lysing

cells in lysis buffer containing 10 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS and 10 mM imidazole. The lysates were clarified by centrifugation at 10,000 *g* for 20 min at 4 °C.

Ni-NTA Pulldown Assay. A total of 300 μg of total cell lysates expressing eGFP–P were incubated with 2 μg of 6× His-tagged L or 1–1,593, 1,594–2,109, (2 μg of 1–1,593 + 2 μg of 1,594–2,019 preincubated on ice for 1 h) or 2 μg of the N-terminal domain of anthrax lethal factor as control and 15 μL of Ni-NTA agarose beads (Qiagen) at 4 °C for 1 h. The beads were precipitated by centrifugation at 2,000 × *g* for 3 min and washed five times with 500 μL of lysis buffer containing 30 mM imidazole. The beads were boiled in 2× SDS/PAGE loading buffer and the proteins were separated by 10% SDS/PAGE.

1. Pattnaik AK, Wertz GW (1990) Replication and amplification of defective interfering particle RNAs of vesicular stomatitis virus in cells expressing viral proteins from vectors containing cloned cDNAs. *J Virol* 64:2948–2957.
2. Liu L, Spurrier J, Butt TR, Strickler JE (2008) Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. *Protein Expr Purif* 62: 21–28.
3. Li J, Rahmeh A, Morelli M, Whelan SP (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.
4. Li Z, Hite RK, Cheng Y, Walz T (2010) Evaluation of imaging plates as recording medium for images of negatively stained single particles and electron diffraction patterns of two-dimensional crystals. *J Electron Microsc (Tokyo)* 59:53–63.
5. Frank J, et al. (1996) SPIDER and WEB: Processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* 116:190–199.
6. Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128:82–97.
7. Biasini M, et al. (2010) *OpenStructure: A Flexible Software Framework for Computational Structural Biology* (Bioinformatics, Oxford, England).
8. Li J, Fontaine-Rodriguez EC, Whelan SP (2005) Amino acid residues within conserved domain VI of the vesicular stomatitis virus large polymerase protein essential for mRNA cap methyltransferase activity. *J Virol* 79:13373–13384.
9. Schott DH, Cureton DK, Whelan SP, Hunter CP (2005) An antiviral role for the RNA interference machinery in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 102: 18420–18424.
10. Fuerst TR, Niles EG, Studier FW, Moss B (1986) Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci USA* 83:8122–8126.

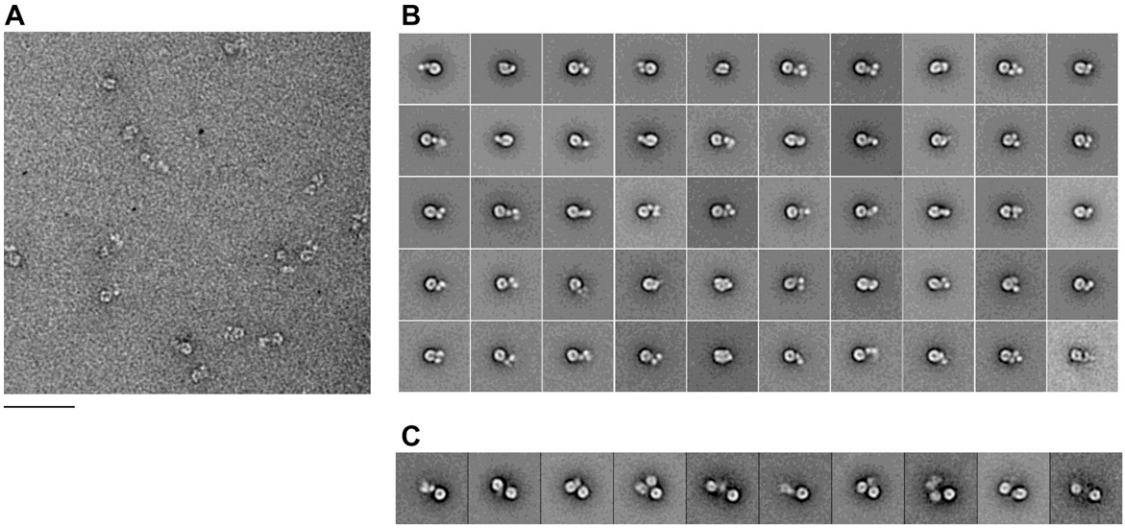

Fig. S1. Single particle analysis of L. (*A*) Representative EM image of L in negative stain. (Scale bar: 50 nm.) (*B*) Class averages of L singles obtained by classification of 8,806 particles into 50 classes. (*C*) Class averages of L doubles obtained by classification of 1,099 particles into 10 classes. The side length of the individual panels in *B* and *C* is 54 nm.

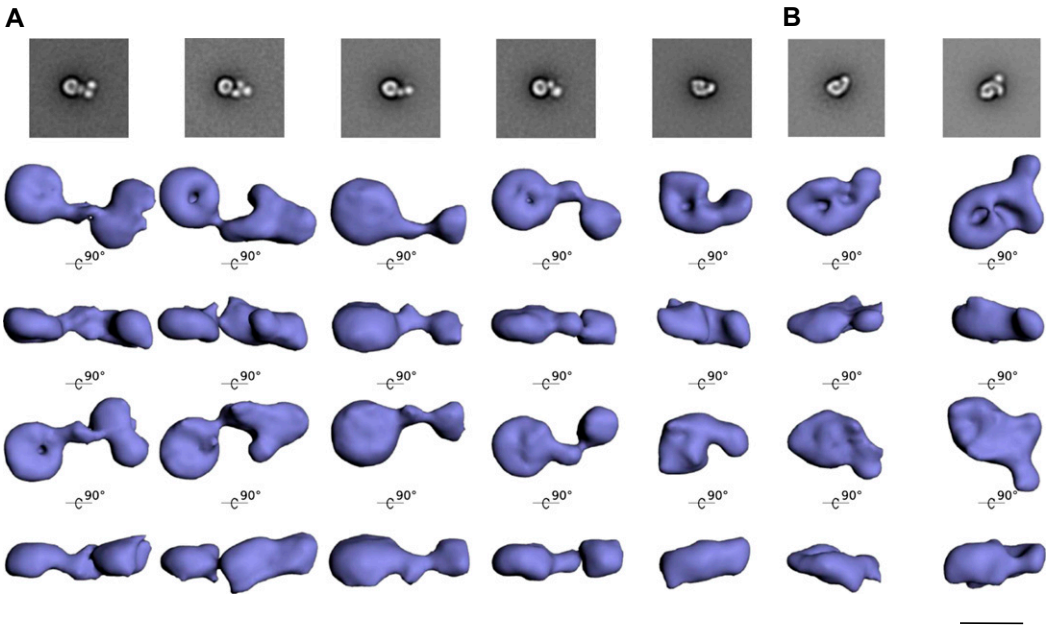

Fig. S2. The 3D reconstructions of L and the L–P complex. Class averages and corresponding 3D reconstructions of (*A*) L and (*B*) the L–P complex in negative stain. Different views of each reconstruction are shown rotated by 90° about the horizontal axis as indicated by the arrows. (Scale bar: 10 nm.) The 3D reconstructions are consistent with the respective class averages, revealing the ring-shaped core domain and the appendage. Due to flattening of the particles introduced by the negative staining procedure and the limited resolution of the density maps, potentially due to remaining structural heterogeneity within individual classes of particles, the 3D maps do not provide more information than the projection averages. In particular, the density maps of L and the L–P complex do not have sufficient structural detail to reveal the conformational change that occurs in the appendage of L upon binding of P.

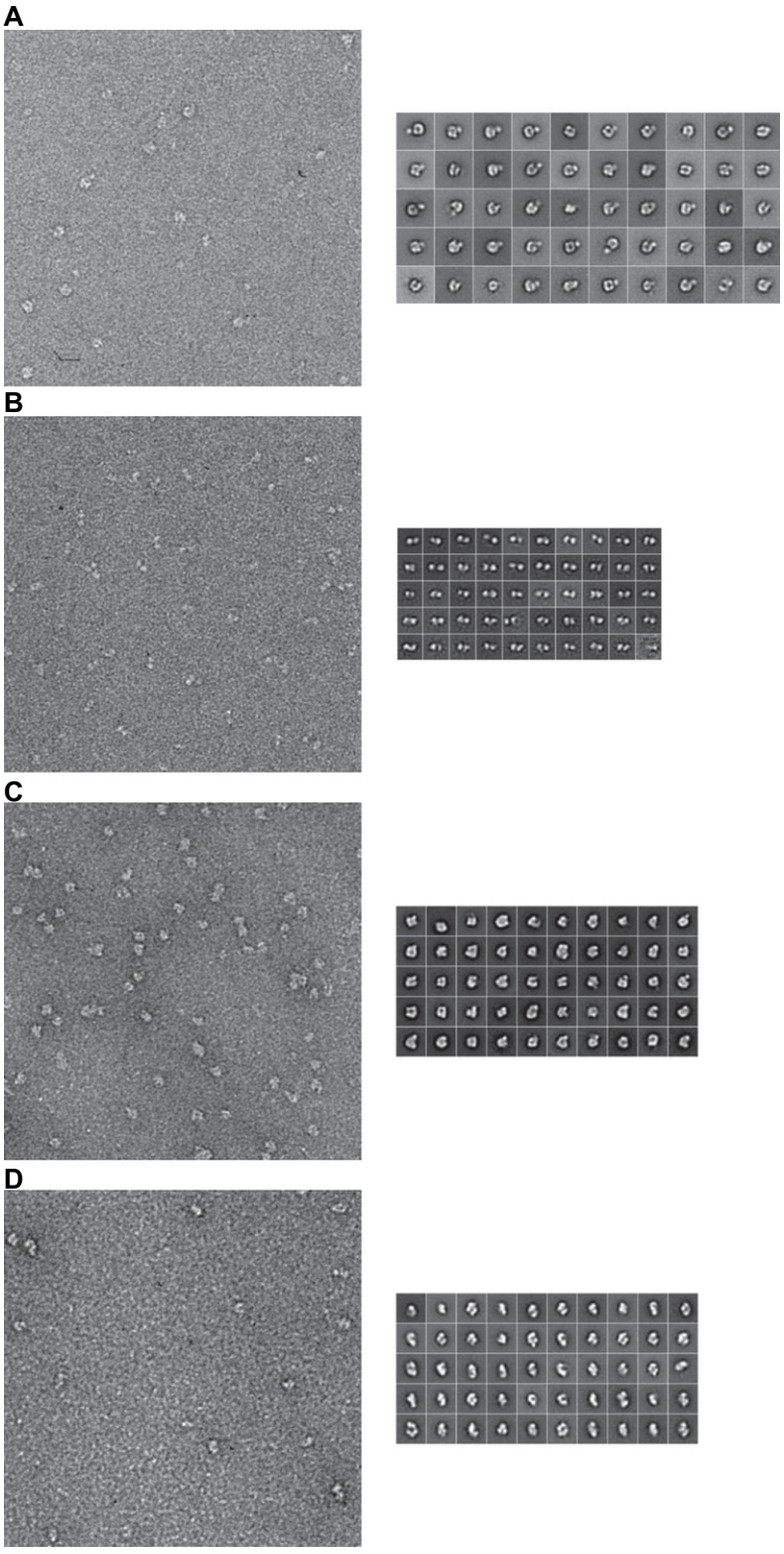

Fig. S3. Single-particle EM analysis of negatively stained L fragments. *Left* shows representative EM images (Scale bar: 50 nm) and *Right* shows the 50 class averages for (*A*) the 1–1,593 fragment, for which the side length of the individual panels is 29 nm, (*B*) the 1,594–2,109 fragment, for which the side length of the individual panels is 20 nm, (*C*) the 1–1,114 fragment, for which the side length of the individual panels is 22 nm, and (*D*) the 1–860 fragment, for which the side length of the individual panels is 22 nm.

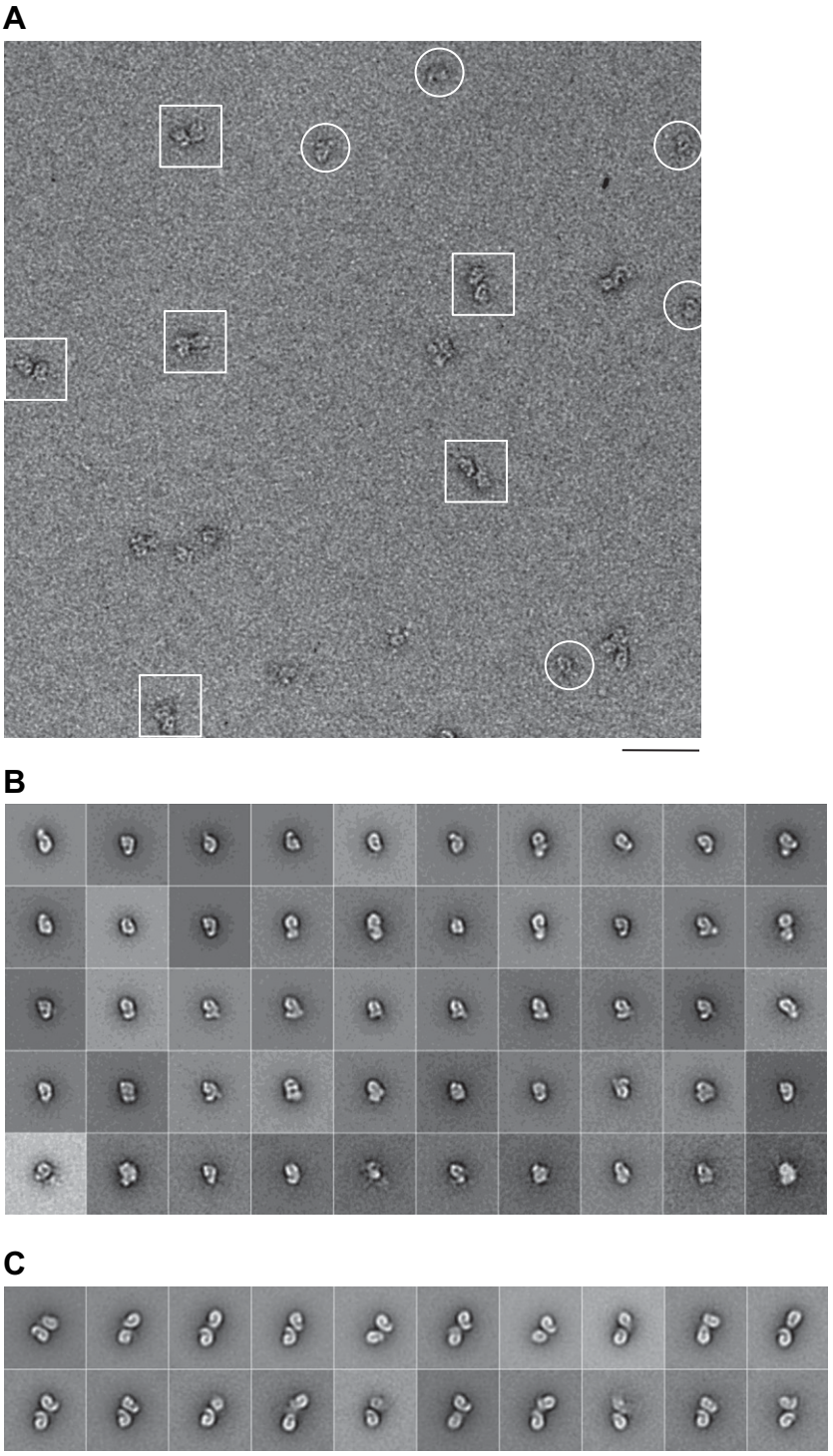

Fig. S4. Single-particle EM analysis of negatively stained L–P complex. (*A*) Representative EM image of the L–P complex with single particles indicated by white circles and double particles by white squares. (Scale bar: 50 nm.) (*B*) Class averages of single L–P particles obtained after classification of 5,885 particles into 50 classes. (*C*) Class averages of double L–P particles obtained after classification of 2,780 particles into 20 classes. The side length of the individual panels in *B* and *C* is 54 nm.

Table S1. Summary of designed VSV L fragments

Construct	Expression
(*i*) N–His	
1–2,109	+++
1–860	+++
1–1,045	+
1–1,080	–
1–1,114	++
1–1,254	+
1–1,378	+
1–1,498	–
1–1,560	–
1–1,593	++
861–2,109	–
1,046–2,109	–
1,594–2,109	+++
1,046–1,593	–
1,594–1,860	+
100–1,593	–
261–1,593	–
1,255–2,109	–
(*ii*) N–His–SUMO*	
194–2,109	–
304–2,109	–
341–2,109	–
413–2,109	–
588–2,109	–
599–2,109	–
861–2,109	–
920–2,109	–
941–2,109	–
1,046–2,109	–
1,070–2,109	–
1,332–2,109	±
1,378–2,109	±
1,476–2,109	–

Full-length L (1–2,109) and fragments were expressed in insect cells as (*i*) N–His-tagged or (*ii*) N–His–SUMO*-tagged. Expression levels are marked from highest (+++) to lowest (–).