Structure of a paramyxovirus polymerase complex reveals a unique methyltransferase-CTD conformation

Ryan Abdella[a,b,1], Megha Aggarwal[a,c,1], Takashi Okura[a,c], Robert A. Lamb[a,c,2], and Yuan He[a,b,d,e,2]

[a]Department of Molecular Biosciences, Northwestern University, Evanston, IL 60208; [b]Interdisciplinary Biological Sciences Program, Northwestern University, Evanston, IL 60208; [c]Howard Hughes Medical Institute, Northwestern University, Evanston, IL 60208; [d]Chemistry of Life Processes Institute, Northwestern University, Evanston, IL 60208; and [e]Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Northwestern University, Chicago, IL 60611

Contributed by Robert A. Lamb, January 10, 2020 (sent for review November 12, 2019; reviewed by Jason McLellan and Felix A. Rey)

Paramyxoviruses are enveloped, nonsegmented, negative-strand RNA viruses that cause a wide spectrum of human and animal diseases. The viral genome, packaged by the nucleoprotein (N), serves as a template for the polymerase complex, composed of the large protein (L) and the homo-tetrameric phosphoprotein (P). The ∼250-kDa L possesses all enzymatic activities necessary for its function but requires P in vivo. Structural information is available for individual P domains from different paramyxoviruses, but how P interacts with L and how that affects the activity of L is largely unknown due to the lack of high-resolution structures of this complex in this viral family. In this study we determined the structure of the L–P complex from parainfluenza virus 5 (PIV5) at 4.3-Å resolution using cryoelectron microscopy, as well as the oligomerization domain (OD) of P at 1.4-Å resolution using X-ray crystallography. P-OD associates with the RNA-dependent RNA polymerase domain of L and protrudes away from it, while the X domain of one chain of P is bound near the L nucleotide entry site. The methyltransferase (MTase) domain and the C-terminal domain (CTD) of L adopt a unique conformation, positioning the MTase active site immediately above the poly-ribonucleotidyltransferase domain and near the likely exit site for the product RNA 5′ end. Our study reveals a potential mechanism that mononegavirus polymerases may employ to switch between transcription and genome replication. This knowledge will assist in the design and development of antivirals against paramyxoviruses.

Cryo-EM | paramyxovirus | polymerase | transcription | replication

The family *Paramyxoviridae* consists of enveloped viruses with nonsegmented, negative strand (NNS) RNA genomes and includes some of the most pathogenic viruses, including measles virus (MeV), mumps virus (MuV), Nipah virus (NiV), Hendra virus (HeV), Newcastle disease virus (NDV), and parainfluenza viruses 1 to 3 (PIV1–3). PIV5 has proven to be an excellent model system for the study of paramyxovirus proteins (1–5). The RNA genome of all paramyxoviruses is packaged by the N protein into a nucleoprotein complex that acts as the template for genome replication and transcription by the polymerase complex consisting of the large (L) protein and the phosphoprotein (P) (6).

L is a remarkable enzyme because it both transcribes and replicates the viral genome (7–10). The L protein encounters the RNA genome at the 3′ end, beginning with the leader sequence, and then transcribes or replicates the entire genome, which requires separating the RNA sequestered by the N protein. During replication, an antigenome copy of the RNA genome is generated, which serves as a new template to synthesize genomic RNA (6). Transcription involves the production of messenger RNA (mRNA) harboring a methylated 5′ cap and poly-A tail. However, the initiation and regulatory mechanisms of genome replication and transcription are poorly understood. There can be either a single pool or two different pools of polymerase within an infected cell, depending on whether a single polymerase can transition between genome replication and transcription (11, 12). The initiation of transcription and replication has been described for respiratory syncytial virus (RSV), a member of the closely related *Pneumoviridae* family, where two different initiation sites in the first 13 nucleotides at the 3′ end of RNA are required (13, 14). However, no evidence for two initiation sites exists for paramyxoviruses. The amount of N protein is implicated in the transition from transcription to genome replication for many NNS virus families, suggesting a common regulatory mechanism (14–16).

Structures of L from other NNS viruses have identified five conserved domains: RNA-dependent RNA polymerase (RdRp), poly-ribonucleotidyltransferase (PRNTase), connecting domain (CD), methyltransferase (MTase), and the C-terminal domain (CTD) (17–19). Early studies on NNS viruses identified six conserved regions in L (CR I–VI), which form crucial motifs involved in its function and interactions with P (20, 21). The functions of these particular regions have been revealed from structural and biochemical studies (10, 22, 23). Studies of the L protein from vesicular stomatitis virus (VSV), a member of the rhabdovirus family, revealed that a fragment of P is necessary to induce stable association of the CD-MTase-CTD module to the RdRp-PRNTase module (17, 24). The CD-MTase-CTD module in both RSV and

Significance

Parainfluenza virus 5 belongs to the family *Paramyxoviridae*, whose members pose significant health burdens to humans and animals. The viral RNA genome is packaged within a nucleoprotein complex that serves as the template for genome replication and transcription. The L protein is responsible for RNA synthesis, capping, and methylation, and requires a cofactor; the P protein, for RNA synthesis in vivo. This study provides a near-atomic resolution structure of a complete paramyxovirus L–P complex, an attractive target for drug design against paramyxoviruses. Comparisons to structures of other mononegavirus polymerases identify a large conformational rearrangement of the methyltransferase and CTD relative to the RdRp domain, a potential mechanism that could allow the complex to switch between genome replication and transcription.

Author contributions: R.A.L. designed research; R.A., M.A., T.O., and Y.H. performed research; R.A., M.A., and Y.H. analyzed data; and R.A., M.A., R.A.L., and Y.H. wrote the paper.

Reviewers: J.M., University of Texas at Austin; and F.A.R., Institut Pasteur.

The authors declare no competing interest.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 6V85, 6V86, and 6VAG) and the Electron Microscopy Databank, https://www.ebi.ac.uk/pdbe/emdb/ (ID codes EMDB-21095 and EMDB-21096).

[1]R.A. and M.A. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: ralamb@northwestern.edu or yuanhe@northwestern.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1919837117/-/DCSupplemental.

First published February 19, 2020.

human metapneumovirus (hMPV) L–P complexes was completely flexible in the final reconstruction, preventing visualization of those domains (18, 19). Crystal structures of substrate-bound MTase and CTD from hMPV indicate that these domains stably interact and define binding sites for the 3′ end of RNA produced by L (23).

P performs several important functions in the viral life cycle and is composed of three domains: N-terminal, central oligomerization (OD), and C-terminal X domain (XD). The N terminus of P is generally unstructured and binds a nascent N monomer (N^0), forming the N^0-P complex to prevent premature N oligomerization and nonspecific RNA binding (2, 25). P-XD is also intrinsically disordered in solution but forms a three-helix bundle under crystallizing conditions and associates with the molecular recognition element (MoRE) of the C terminus of N (N-tail) (26–28). The central region of P self-oligomerizes to form either dimers or tetramers and directly interacts with the L protein, tethering the polymerase to the N-coated RNA template (25). To date, crystal structures of the P-OD from Sendai virus (SeV), MeV, NiV, MuV, and VSV are available (29–33). The paramyxoviruses (SeV, MeV, NiV, and MuV) share a common tetrameric coiled-coil structure arranged in parallel fashion except for the MuV P-OD in which two parallel chains of helices are arranged in an antiparallel orientation to the other pair (29–31, 33). In addition, the structure of the three-helix bundle P-XD has been solved on its own (MuV and MeV) and in complex with the MoRE region of N-tail (MeV) (26–28).

The interaction between L and P is essential for paramyxovirus replication and transcription. Biochemical characterization of the interaction between PIV5 L and P reveals the importance of the N-terminal half of the L protein, while the C-terminal half is dispensable for the interaction with P (34). Similar results have been observed in MeV, SeV, and HPIV3 in which the N-terminal region of L is responsible for L–P interaction (35–38). However, in rabies virus (RaV), the C terminus of L and first 19 residues at the N terminus of P are essential for the interaction between L and P (39). The C terminus of SeV P has been identified as an L-binding region by mutational analysis (40). A bipartite interaction between L and P has been observed in MeV, where the P-OD and P-XD bind at two separate surfaces of L (41). The recently published cryoelectron microscopy (cryo-EM) structures of L–P from RSV and hMPV demonstrate a tentacular arrangement of P with significant interfaces between L and P (18, 19). Each of the four chains of P have distinct conformations in the region C-terminal to the P-OD tetramer and interact with unique locations on L, suggesting structural plasticity in this interaction. Whether this is a common feature of L–P complexes among NNS RNA viruses or limited to viruses within the *Pneumoviridae* family remains to be determined.

Here, we used cryo-EM to solve the structure of the PIV5 L–P complex at 4.3-Å resolution. All five domains (RdRp, PRNTase, CD, MTase, and CTD) of L are well resolved in the density map. The structure reveals the existence of two discrete binding interfaces on the L protein surface for the binding of P-OD and P-XD, respectively. Our 1.4-Å resolution crystal structure of the PIV5 P-OD confirms the presence of a four-helix bundle arranged in a parallel orientation, consistent with other paramyxoviruses, except for MuV. The tetrameric P-OD protrudes away from L, forming minimal interactions with L, while a single copy of the P-XD is bound near the nucleotide entry site of L. The priming loop of L adopts a transcription elongation conformation, and an intrusion loop occupies the active site of the RdRp domain. A significant conformational rearrangement of the domains responsible for nascent mRNA 5′ methylation highlights the dynamic nature of the L protein, revealing a crucial mechanism for the spatial–temporal regulation of RNA synthesis. Comparisons with the structure of the VSV and RSV L–P complexes reveal key differences in both L and P across different NNS viruses, critical for our understanding of the full catalytic cycle of these important enzymes.

Results

Cryo-EM Structure of PIV5 L–P Complex. In order to generate a stable L–P complex, we constructed a dual expression system of baculovirus expressing full-length PIV5 L and P (Fig. 1*A*). We verified that tagged L and P were sufficient for polymerase activity (*SI Appendix*, Fig. S1*A*). A stable complex of L–P was purified using M2 FLAG agarose beads followed by size-exclusion chromatography (*SI Appendix*, Fig. S1*B*). Single-particle cryo-EM of the purified complex gave rise to a three-dimensional (3D) reconstruction at 4.3-Å resolution (Fig. 1*B*, Table 1, and *SI Appendix*, Fig. S2). Local-resolution estimation for the entire complex ranged from 4.0 Å near the core of the RdRp and PRNTase domain to >6.0 Å for the P-OD. All domains of L could be unambiguously assigned to the density and four copies of P form a long four-helix bundle that extends from the RdRp domain with a single P-XD bound near the nucleotide entry tunnel of L (Fig. 1 *B* and *C*). We generated a homology model for PIV5 L based on the VSV L structure and used it as a guide to manually build a complete model from residues Arg5 to His2225.

Structural Architecture of PIV5 L. The 3D arrangement of different domains within L is important for RNA synthesis (42). The N-terminal half of L forms a canonical RdRp domain, whereas the C-terminal half is responsible for the auxiliary functions including PRNTase and MTase activities. The RdRp and PRNTase domains form a stable ring-like structure as seen in other NNS viruses (17, 42, 43). The RdRp domain adopts a similar architecture, composed of fingers, palm, and thumb subdomains, as observed in other viral polymerases (*SI Appendix*, Fig. S3*A*) (17, 19). The fingers subdomain contains CR II, which is involved in RNA binding (21, 44). Flexible loops connect the subdomains together, allowing subtle mobility among them. A conserved Gly840 at the intersection of the palm and thumb subdomains provides additional flexibility. CR III forms the palm subdomain, which folds into a small antiparallel β-sheet. The RdRp active site is formed by the conserved GDN (772–774) motif present between two of the β-strands. This active site motif is surrounded by several loop regions, which form the tunnel that the template RNA travels through toward the catalytic site. The palm subdomain also contains a conserved aromatic residue, Tyr667, which is positioned away from the active site, similar to VSV and RSV L structures, contributing to L stability. The fingers and thumb subdomains are primarily helical except for a few small β-sheets. These two subdomains form most of the interface with the PRNTase domain.

The PRNTase domain utilizes a completely different mechanism to cap the nascent RNA strand than eukaryotes. In NNS viruses, the first transcribed nucleotide forms a covalent bond to a conserved histidine residue located in a conserved histidine-arginine (HR) motif found in CR V (45). Cap addition occurs via nucleophilic attack by GDP on the covalent pRNA-His bond. In eukaryotes a nucleophilic 5′ ppRNA attacks a GTP (46). The PRNTase domain is located above the RdRp, in the same position it occupies in the VSV and RSV structures (Fig. 1 *B* and *C*). A second conserved motif in CR V, GxxT, is involved in binding of the capping guanosine nucleotide (17). The PRNTase domain also contains CR IV and adopts a very similar fold as in the VSV and RSV structures, except for differences in some key areas (*SI Appendix*, Fig. S3*B*). The priming loop is highly flexible in our PIV5 structure, but we were able to trace its path, which adopts a noninitiation conformation similar to the RSV priming loop (Fig. 2*A*). This positions the GxxT (1218–1220) motif much closer to the HR motif, in a conformation that likely could not accommodate the GDP. The HR motif is located within the flexible intrusion loop, which does not impinge on the central cavity in the RSV and VSV structures, but instead projects out into

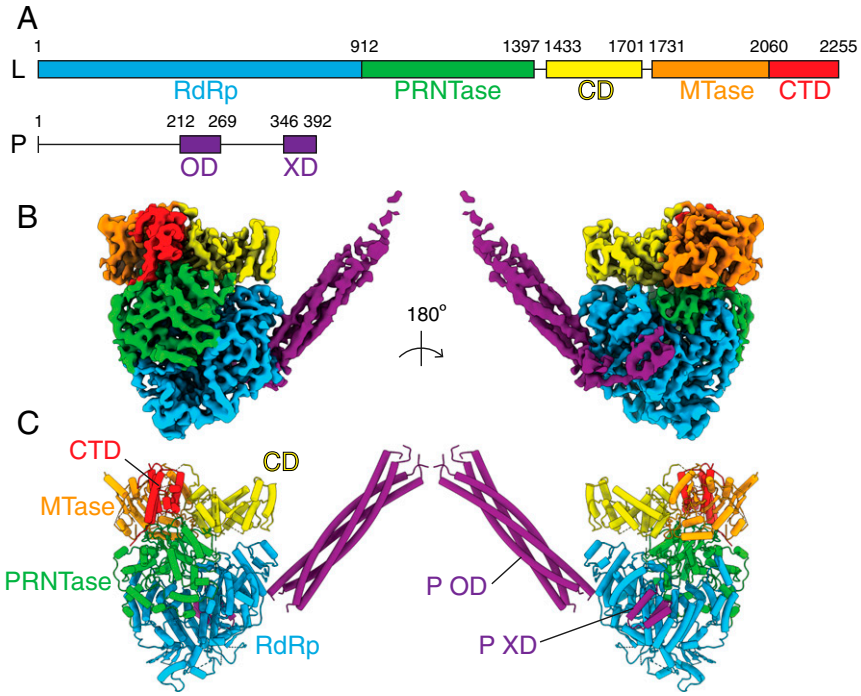

Fig. 1. Architecture of the PIV5 L–P complex. (*A*) Domain diagrams of PIV5 L and P proteins. RdRp, cyan; PRNTase, green; CD, yellow; MTase, orange; CTD, red; P-OD, purple; P-XD, purple. (*B* and *C*) Electron density (*B*) and atomic model (*C*) of the PIV5 L–P complex with domains colored as depicted in *A*.

the central cavity in our structure and would clash with the position of the priming loop in the initiation conformation (Fig. 2 *A* and *B*). Displacement of the intrusion loop is required to accommodate RNA in the active site (44), suggesting a possible tug-of-war between the priming loop and intrusion loop that could regulate transcription initiation.

Although the CD has not been observed to perform any enzymatic activity, it is required for proper polymerase function and cannot tolerate insertions in its sequence (47). The CD is very weakly conserved among NNS viruses, although common secondary structure motifs allowed us to build a model (*SI Appendix*, Figs. S3*C* and S7). A flexible loop (1398–1432) connects the CD to the PRNTase domain and another flexible loop (1702–1730) connects the CD to the MTase domain. These flexible linkers likely help facilitate the domain rearrangements, which will be discussed later.

The MTase performs a key catalytic reaction by first methylating the 2′O of the very first product nucleotide, followed by methylating N7 of the cap (23). Its active site is highly conserved, containing a catalytic K-D-K-E motif common to 2′O methyltransferases (23, 48). In PIV5 those resides are K1786-D1911-K1947-E1984 (*SI Appendix*, Figs. S3*D* and S7). K1786 and an atypical AxGxG motif are located within CR VI. The AxGxG motif forms part of the SAM binding site, despite the fact that the first residue is not a glycine (*SI Appendix*, Fig. S3*D*). Superposition of the hMPV MTase-CTD crystal structure positions all conserved residues in the same location in our PIV5 structure, suggesting a conserved binding mode for substrates SAM and GTP (*SI Appendix*, Fig. S3*D*).

Like the CD, the CTD contains very little sequence conservation and is therefore structurally divergent between PIV5, RSV, and VSV (*SI Appendix*, Fig. S3*E*). PIV5 differs from VSV and RSV in that it does not contain the α1 helix but contains α2, -3, -4, and -6 found in those structures. Instead of an α5 helix, PIV5 contains multiple small helices named α5, α5′, and α5″ that occupy a similar location as a β-sheet in the VSV CTD (*SI Appendix*, Fig. S3*E*). In addition to the conserved HR motif in the PRNTase domain discussed earlier, paramyxoviruses also contain a KxxxKxxG motif located in the CTD, a canonical eukaryotic guanylyl transferase (GTase) motif (23, 46). This motif is located in α6 of the CTD, while it is in a β-sheet in eukaryotic GTases (*SI Appendix*, Fig. S3*E*).

Structure of Tetrameric P-OD. The PIV5 P-OD is located at the central core region of the P protein (Fig. 1*A*). Disrupting the stability of the coiled-coil region or the kink near residues 339–341 in MeV P-OD results in disruption of MeV gene expression (49). In our reconstruction, a tetrameric helical bundle is clearly visible, directly associating with the RdRp domain of L and protruding away from it (Fig. 1 *B* and *C*). However, we could not unambiguously determine the orientation of each chain of P in our density due to the limited resolution in this region (*SI Appendix*, Fig. S2*B*). We were able to crystallize the P-OD from residue 172–278 and the structure was determined at 1.4-Å resolution (Fig. 3*A*, Table 2 and *SI Appendix*, Fig. S4*A*). PIV5 P forms an all-parallel four-helix bundle with two long parallel helices in the asymmetric unit. The all-parallel orientation of P-OD is consistent with the vast majority of NNS P-OD crystal structures that have been determined to date. Residues 198–271 were visible in the crystallographic electron density map and the coiled-coil region comprises residues 203–270. The residues immediately N-terminal to the P-OD from NiV and SeV form a helical cap; however, we were unable to identify a similar feature in either of our crystallographic and cryo-EM density maps. Sequence comparison suggest that those N-terminal helices in NiV and SeV P are likely insertions, which is absent in the PIV5 P protein (*SI Appendix*, Fig. S8).

Bipartite Interface on L for Binding P-OD and P-XD. In order to fit our crystal structure of the P-OD into the cryo-EM density map, we needed to first determine the correct orientation of the P-OD relative to the EM density. We were able to take advantage of the identification of a single P-XD associated with L, which will be described in detail later. Even though the density between the

Table 1. Cryo-EM data collection, refinement, and validation statistics

	PIV5 L–P class1 (EMDB-21095; PDB 6V85)	PIV5 L–P class004 (EMDB-21096; PDB 6V86)
Data collection and processing		
Magnification	30,000	30,000
Voltage (kV)	200	200
Electron exposure (e^{-}/Å^2)	76.5	76.5
Defocus range (μm)	−1.5 to −4.5	−1.5 to −4.5
Pixel size (Å)	1.12	1.12
Symmetry imposed	C1	C1
Initial particle images (*n*)	717,008	717,008
Final particle images (*n*)	102,493	78,547
Map resolution (Å)	4.38	4.63
FSC threshold	0.143	0.143
Map resolution range (Å)		
Refinement		
Initial model used (PDB code)		
Model resolution (Å)		
FSC threshold		
Model resolution range (Å)		
Map sharpening *B* factor (Å^2)	−197	−265
Model composition		
Nonhydrogen atoms	17,895	17,798
Protein residues	2,240	2,229
Ligands	ZN:2	ZN:2
B factors (Å^2)		
Protein	77.11	127.37
Ligand	280.69	280.69
RMSD		
Bond lengths (Å)	0.006	0.004
Bond angles (°)	0.878	0.825
Validation		
MolProbity score	2.47	2.49
Clashscore	21.82	23.19
Poor rotamers (%)	0.40	0.70
Ramachandran plot		
Favored (%)	86.05	86.63
Allowed (%)	13.54	13.14
Disallowed (%)	0.41	0.23

P-OD and P-XD is poor, the P-OD must anchor to L at its C terminus due to the localization of the P-XD, resulting in its N terminus projecting away from L (Fig. 3*B*). The distance between the first residue of P-XD (L342) and the end of P-OD that anchors to L is 43 Å, while the distance to the end of P-OD that is far away from L is 145 Å. With 77 residues between P-OD and P-XD, the more plausible orientation of P-OD is with the C terminus binding to L. This is consistent with the orientation of the P-OD of RSV and hMPV (18, 19).

With the correct orientation determined, our crystal structure could be docked into our EM density as a rigid body without significant changes. The C terminus of chains P_1 and P_4 of the P-OD interacts with a single helix of L comprising residues 392–412 (Fig. 3*B* and *SI Appendix*, Fig. S4*B*). The RSV P-OD, although significantly shorter than PIV5 P-OD, also engages with the homologous helix of L (residues 455–476) (Fig. 4 *A* and *B* and *SI Appendix*, Fig. S4*C*). This helix maps perfectly toward the end of a 408-residue fragment of the closely related MuV L, which was shown to be both necessary and sufficient for interacting with P (36). Interaction between SeV L and P required a larger fragment of L—residues 1 to 1,146—although it is possible that smaller fragments lacked stability of helix 392–412, preventing stable association (22). Replacement of the MeV P-OD up to residue 360 (264 in PIV5) with the GCN4 tetrameric domain did not affect binding of L and P, consistent with the minimal interaction between L and P-OD in our structure (41). Indeed, L only binds to P-OD C-terminal of residue 264 in our structure (Fig. 3*B*).

There is visible density for the flexible linker C-terminal to the P-OD that we were not able to assign to any individual chain of P with confidence or build a model for (Fig. 3*C* and *SI Appendix*, Fig. S4*B*). This flexible linker region clearly associates with loop (384–391) of L, immediately C-terminal of helix 392–412. Replacement of the homologous region of the MeV P with the GCN4

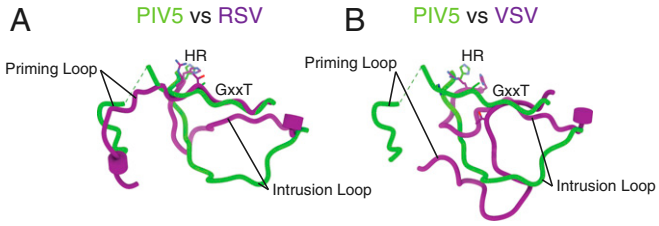

Fig. 2. Comparison of the PIV5, RSV, and VSV priming loop and intrusion loop. (*A*) The PIV5 priming loop adopts the same elongation conformation as in the RSV structure. The intrusion loop projects out into the central cavity between the RdRp and PRNTase domains. (*B*) In the VSV structure the priming loop in the initiation conformation would sterically clash with the position of the PIV5 intrusion loop. The VSV and RSV intrusion loops are in the same position with minor differences.

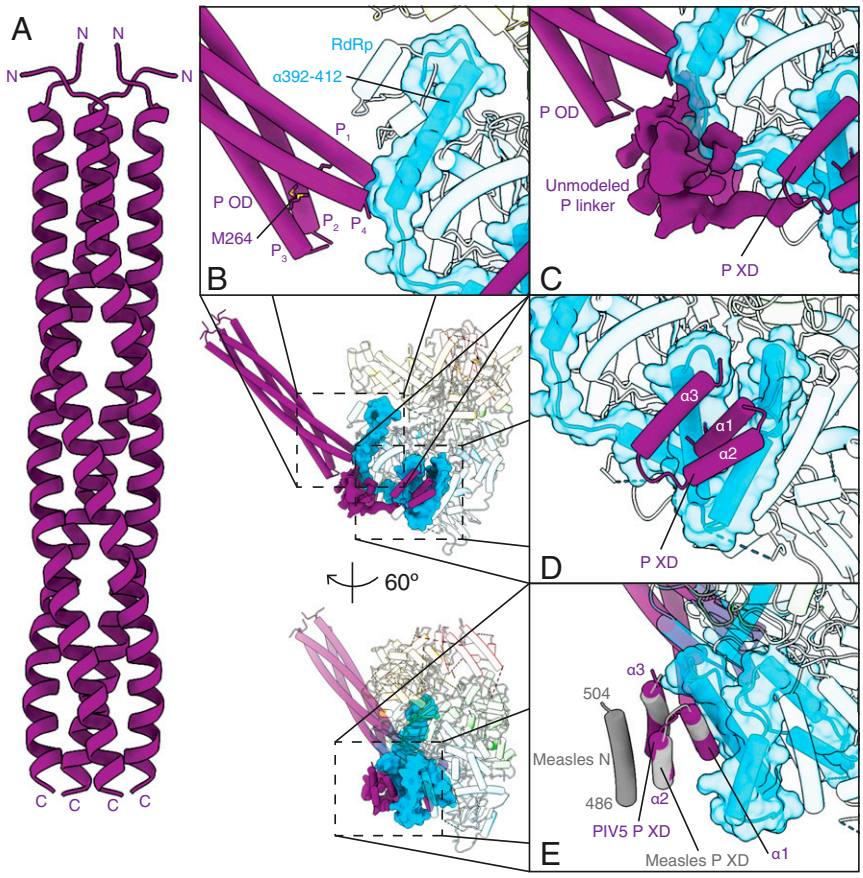

Fig. 3. Interaction interfaces of PIV5 L and P proteins. (*A*) Crystal structure of the OD of the PIV5 P protein. The OD forms an all-parallel four-helix bundle with one helix from each of four chains of P. (*B*) Interfaces between P-OD (purple) and L (cyan). The fragment of L that is necessary and sufficient to interact with P is shown as an opaque surface, the rest of L is shown as transparent cartoon. Helix 392–412 is the only portion of L that interacts with P-OD. (*C*) Interaction between the unmodelled P density and L. This density is not as well resolved as the OD or XD and does not form extensive contacts with L except at the base of P-OD. (*D* and *E*) Interaction of P-XD (residues 346–392) and L. Helices α1 and α3 of P-XD form the interface of the XD with L. The portion of L that interacts with P-XD spans residues 303–350. Superposition of MeV P-XD bound to a C-terminal fragment of N (486–504) with PIV5 P-XD (*E*).

tetrameric domain disrupted the interaction between L and P, highlighting the importance of this interface (41). The flexible linker appears to correspond to P_4 based on homology to RSV, although our density is not resolved enough to be confident in this assignment (*SI Appendix*, Fig. S4 *B* and *C*). The path of the linker is also different from RSV, where it interacts with helix 829–850 of L. In PIV5 the density loops away from L, and the lack of a stabilizing interaction with L is the likely cause of its poor resolution. We see no evidence for additional interactions between the P linker helices of P_1, P_2, and P_3 like those present in RSV (Fig. 4 *A* and *B*).

As mentioned earlier, a single P-XD is bound to L near the nucleotide entry site (Fig. 3*D*). A homology model of PIV5 P-XD based on the MuV P-XD structure fits perfectly into our EM density, with the N terminus of the P-XD naturally connecting to the unmodeled P linker density (Fig. 3*D*). Helices α1 and α3 of P-XD form the interface that interacts with L, consistent with mutational studies of MeV P (41). The part of L that interacts with P-XD spans residues 303–350. Interestingly, the N-terminal 298 residues of MeV L were not sufficient to bind P, defining a fragment of L from 303 to 412 as the region that is sufficient to bind MeV P (36). This region is a subset of CR I, which spans residues 227–419. A separate study on MeV implicated that P-XD residues V463, S466, and H498 are important for the interaction with L (41). The structurally equivalent residues in PIV5 P-XD T349, Q351, and K384 are all in direct proximity to the surface of L (*SI Appendix*, Fig. S4*D*). T349 and Q351 form a hydrogen bond network with L-H315 and L-Q349. These interactions are conserved in the closely related hPIV2 but not in other paramyxoviruses where a T349M mutation in simian virus 41 (SV41) P has a corresponding L-H315Y mutation, creating a π–aromatic interaction that would stabilize the interface in SV41 (*SI Appendix*, Figs. S7 and S8). In MeV, a hydrophobic core is created by P-V463, L-L305, P-S466, and L-I339. These findings are consistent with the biochemical studies on MeV and suggests that this interface may have coevolved and is maintained beyond just the rubulaviruses, of which MeV and PIV5 are members (41). RSV lacks the same P-XD fold as PIV5 and other paramyxoviruses; instead, a single helix in RSV P is found in roughly the same position as α1 of the PIV5 P-XD (Fig. 4*B* and *SI Appendix*, Fig. S4*E*). There are only an additional 13 residues C-terminal to this helix of RSV P, not enough to form an additional two helices. The interaction between this helix of RSV P and L shifts down one turn of helix 350–378 (+3 residues) and forms a hydrophobic core containing residues P-S220, P-L223, L-L361, and L-I398 that stabilizes the interface (*SI Appendix*, Fig. S4 *D* and *E*). Because the interaction between L and P does not involve a large interface, it could allow transient binding of P to L, which would be essential for an association–disassociation mechanism of L–P interaction.

Dual Binding Surfaces on P-XD. Helices α1 and α3 of the P-XD directly interact with L, while α2 is solvent-exposed (Fig. 3*D*). Interestingly, MeV P-XD binds the N-MoRE through an interface formed by α2 and α3 (41). Superposition of the MeV

Table 2. Data collection and refinement statistics for P-OD

	PIV5 P-OD (PDB 6VAG)
Data collection and processing	
Space group	P 2_1 2_1 2
Cell dimensions	
a, *b*, *c* (Å)	29.893, 36.048, 140.016
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)*	1.40 (1.42–1.40)
R_{merge} %*	6.5 (88.2)
I/σ_I*	38.12 (2.08)
Completeness (%)*	90.8 (80.6)
Redundancy*	6.5 (7.1)
Refinement	
Resolution (Å)	1.40
No. relfections	27,970
R_{work}/R_{free} (%)	15.6/18.0
No. of residues	146
No. of solvent molecules	105
RMSD	
Bond lengths (Å)	0.005
Bond angles (°)	0.693
Ramachandran plot	
Favored (%)	100
Allowed (%)	0
Disallowed (%)	0

*Values in parentheses are for highest-resolution shell.

P-XD–N-MoRE structure with our PIV5 P-XD reveals that this association does not interfere with its interaction with L, and there are no significant structural changes in the P-XD as well (Fig. 3*E*). Thus, there are two completely independent binding sites on the P-XD for L and N. This could provide a stable contact between L and the genome while the polymerase scans along the N-encapsidated genome. Hence, this arrangement might be responsible for the proper functioning of the polymerase and aids in preventing the premature release of the polymerase from the nucleocapsid.

The MTase and CTD Domains Adopt a Unique Conformation. The MTase and CTD domains of NNS viruses are implicated in both capping and methyltransferase activities (10, 23, 50). The PIV5 MTase-CTD dimer maintains the same architecture as observed in the VSV and hMPV structures, placing the conserved KxxKxxG motif-containing α6 helix proximal to the MTase active site (*SI Appendix*, Fig. S5 *A–C*). This is consistent with the dual function of this module in both methylation and capping. Exchange of both of these domains from one serotype of VSV (New Jersey) to another (Indiana) confirmed a cooperative action of the MTase-CTD module (47). We observe slight shifts of helices α2 and α3 between PIV5 and hMPV and a flipping of α4 in the CTD (*SI Appendix*, Fig. S5*A*). This flipping is likely prevented in PIV5 due to interactions with the CD (*SI Appendix*, Fig. S5*D*). There are relatively larger shifts of helices in PIV5 compared to VSV (*SI Appendix*, Fig. S5*B*). There is a concerted shift of helices α2, α3, and α4 away from the MTase-CD interface in the VSV structure, suggesting interactions with the CD might drive this rearrangement (*SI Appendix*, Fig. S5*D*).

The position of the MTase-CTD dimer relative to the RdRp and PRNTase domains in our PIV5 structure is significantly different from the VSV structure (Figs. 4 and 5). In our structure, the MTase active site sits directly on top of the PRNTase domain, resulting in a much closer distance between the conserved HR motif and the active site of 25.7 Å, rather than 52.5 Å in the VSV structure. Because of the proximity of the MTase active site to the HR motif, we hypothesize that the conformation we observed in our PIV5 L–P structure represents a transcriptionally competent form of the complex. Indeed, domain exchange experiments with different strains of VSV indicated that swapping the MTase and CTD domains from the New Jersey strain into the Indiana strain caused a preference for replication over transcription (47). It is very surprising to us that only a single conformation is present in both the PIV5 and the VSV structure. All our attempts to look for a conformation of the PIV5 complex with the MTase-CTD in the same position as in the VSV structure failed. While we cannot conclude that there are absolutely no complexes that adopt that configuration, they are likely well below 5% occupancy of the full dataset and therefore too difficult to sort out from the rest of the data.

This rearrangement of the MTase and CTD results in completely different interfaces between them and the CD and is likely facilitated by the flexible linker region between the CD and MTase, as mentioned earlier (*SI Appendix*, Fig. S5*D*). In the recently published structure of the RSV L–P complex, the CD-MTase-CTD module are not visible in the density, despite the fact that the full-length complex was used and MTase activity was shown to be present (19). This suggested that these three domains can separate from the RdRp-PRNTase module where they could rearrange between the two stable conformations before reassociating. Consistent with this, we were able to further separate the particles into two classes that slightly differed in the position of the CD-MTase-CTD module relative to the RdRp-PRNTase module (*SI Appendix*, Fig. S6). The observed movement of the CD-MTase-CTD module indicates that this module largely behaves as a rigid body, implying that the CD does not stay stably associated with the RdRp-PRNTase module while the MTase-CTD undergoes the conformational rearrangement.

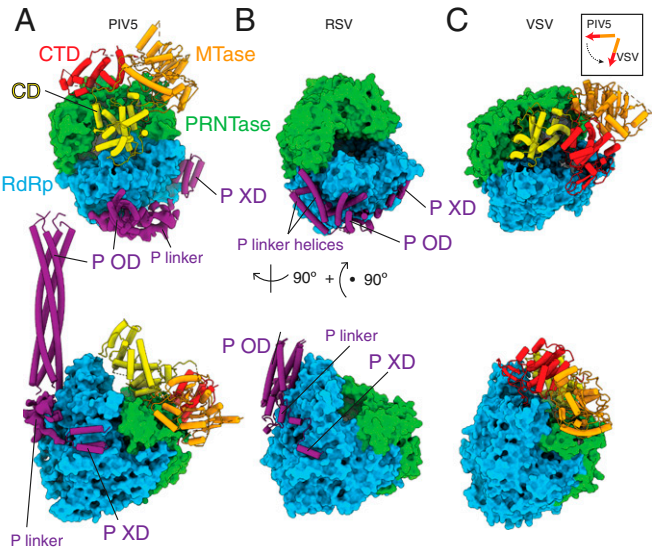

Fig. 4. Comparison of PIV5, VSV, and RSV L–P complexes. (*A–C*) PIV5 (*A*), RSV (*B*), and VSV (*C*) L–P complexes were aligned based on the RdRp domain (cyan). The PRNTase domains (green) are in similar positions relative to the RdRp in all three structures. The P-OD of PIV5 is significantly longer than the P-OD of RSV and protrudes further away from the RdRp. The PIV5 P-XD is in roughly the same position as the single C-terminal helix in the RSV structure. There are no visible P linker helices in the PIV5 structure as there are in the RSV structure. The large rearrangement of the MTase (orange) and (red) is visible between the PIV5 and VSV structures. The relative orientations of the domains are shown in the *Inset* with the arrowhead representing the direction of the protein backbone. CD, connecting domain (yellow); P protein (purple).

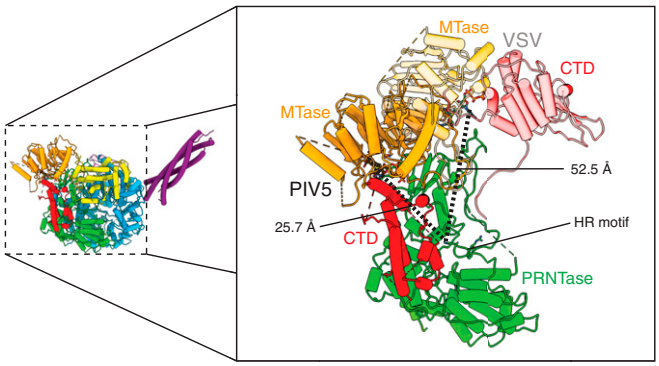

Fig. 5. Comparison of the positions of the conserved HR motif and the MTase active site between the PIV5 and VSV structures. The MTase and CTD are positioned directly above the PRNTase domain in the PIV5 structure (opaque). The distance between the HR motif and the GTP modeled into the MTase active site is 25.7 Å. In the VSV structure the distance increases to 52.5 Å due to the active site projecting away from the RdRp domain.

Discussion

Paramyxoviruses require both L and P to form competent polymerase complexes to function on the genomic RNA wrapped around the N protein. Until recently, only two NNS virus polymerase structures have been determined with key modules missing from these studies, leading to many questions about how L and P cooperate during viral RNA synthesis. Our study presents a structure of a complete viral L protein, with all domains visible, in complex with full-length P. Major differences in the extent of the L–P interface as well as the location of individual domains of L derived from pair-wise structural comparisons help elucidate the mechanistic cycle of these important polymerase complexes.

Mechanism of L–P Procession along N-Coated RNA. The differences in the extent of the L–P interfaces as well as the presence of the P-XD only in paramyxoviruses lead to many questions about how L–P interacts with the viral genome. Recent biochemical studies have shown that one copy of P-XD is essential and sufficient for L–P or N–P interaction, and the MoRE motif-containing N-tail is not essential for the preliminary binding of the polymerase to the RNA template (41, 51). These results are consistent with a model in which repeated association and separation of P-XD with N-tail allows P-XD to cartwheel along the N-coated genome to enable the polymerase to scan along the RNA template (52). We hypothesize that the P-XD might play a role in bridging N to L through two nonoverlapping interfaces (Fig. 3*E*). Anchoring of an N molecule by this single copy of P-XD would limit the movement of the globular domain of N, positioning N to recapture the emerging template genomic RNA. This leads us to further hypothesize a mechanism where once the N–RNA interaction has been disrupted upon the RNA entering the L template entry tunnel, N is captured through the MoRE–P-XD interaction and eventually brought to the anchoring position on L. Capturing N monomers after dissociation from the genomic RNA is likely critical during initial infection when no new N protein has been synthesized in order to recycle all existing N protein to coat the genome after it has been transcribed (Fig. 6*A*). This agrees with existing evidence that transcription of mRNA precedes genome replication, which would require additional N protein to coat the new (anti)-genomic RNA strand as well as the template strand (14, 15). With four strands of P per polymerase, up to four monomers of N could be retained during transcription, allowing up to 24 bases to be displaced from the N-encapsidated genome at a time.

Interactions between P-XD and N or L are thought to be dynamic, leading to two possibilities for how N eventually ends up bound through the P-XD to L. One possibility involves the rotation of the P tetramer relative to L, allowing the different P-XD domains to cycle through interactions with N and L as the

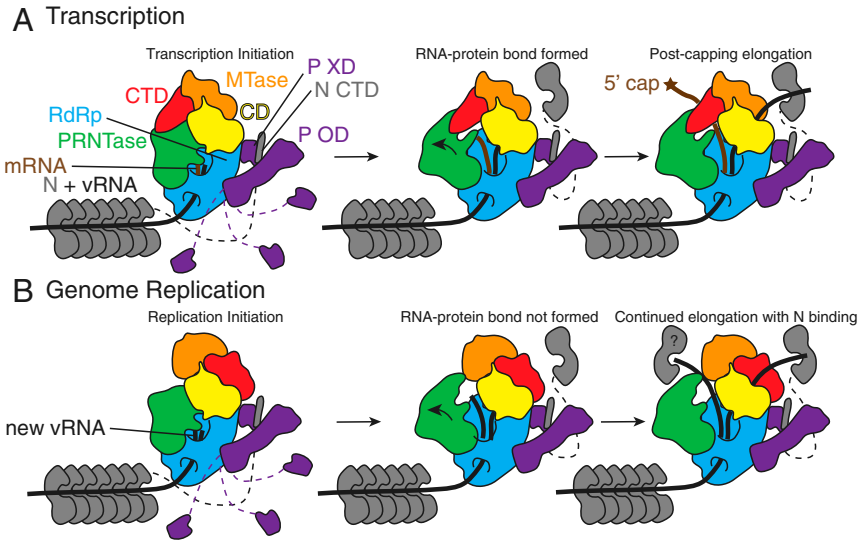

Fig. 6. Model of transcription and genome replication for L–P. The viral RNA (vRNA) genome is separated from N to allow it to enter the template entry tunnel of L. (*A*) In transcription initiation, the MTase-CTD module is positioned directly above the PRNTase domain, as in the PIV5 structure. This positions the active site of the MTase as that the covalently linked RNA-PRNTase is pushed up into it leading to productive capping and methylation. This requires an outward movement of the PRNTase domain to accommodate the growing RNA strand. We hypothesize that a P-XD captures the monomer of N that is no longer bound to the genomic RNA, keeping it in close proximity to be used to recapture the genomic RNA reemerging from the template exit channel. Additional P-XDs are shown in the first panel but removed for clarity from subsequent panels. (*B*) In genome replication, the MTase-CTD module is positioned further away from the PRNTase domain. No covalent linkage is formed between the RNA and conserved histidine, but an outward movement of the PRNTase domain is still required to accommodate the growing RNA strand. The newly released monomer of N is captured in the same way as in *A* and a second copy of N is used to coat the newly synthesized antigenome by an unknown mechanism.

RNA–N complex disassembles and reassembles. The other mechanism occurs without rotation, where an N monomer could be passed from one P-XD to another until it arrives at the P-XD bound to L due to the transient interactions between N, P, and L. In both cases, once N is positioned to coat the genome emerging from the template exit site, it is now furthest away from the flexible P-XD domains and transient dissociation could lead to complete dissociation from the polymerase complex entirely. Despite our structural observation, biochemical studies have suggested that a single copy of P-XD in MeV is unable to interact with both L and N at the same time (41). Further studies are necessary to rationalize these competing observations. VSV lacks the P-XD motif but has been shown to bind N (53). Whether the resulting mechanism is conserved outside of paramyxoviruses or if other virus families adopt a completely different mechanism of interaction with the genome remains to be determined. The highly divergent nature of P both structurally and sequence-wise suggests that differences in regulatory mechanisms likely exist.

Coupling of Conformational Rearrangements and Transcription/Replication Switch. Based on steric clashes and a poorly positioned active site in published structures, it has been hypothesized that the PRNTase domain needs to undergo a substantial conformational change in order to accommodate the growing RNA strand and possibly form the proper active site for capping and methylation (17). This is corroborated by the strict requirement for 31 nucleotides synthesized before capping and methylation (50). Our study provides a structural basis for this hypothesis in that the rearrangement of the MTase-CTD relative to the CD positions the active site for methylation directly adjacent to the PRNTase domain. We propose a mechanism for transcription where the MTase and CTD are in the conformation seen in our PIV5 structure (Fig. 6*A*). The transition of initiation to precapping elongation requires an opening of the PRNTase domain. Continued transcription could push the flexible intrusion loop with RNA bound to the conserved histidine, toward the active site of the MTase-CTD. This would create the active site for capping, followed immediately by methylation.

This model then leads to the conclusion that the conformation of the MTase-CTD observed in the VSV structure is likely utilized during genome duplication (Fig. 6*B*). The transition from initiation to the equivalent of the transcriptional precapping elongation state still requires opening of the PRNTase domain. However, no capping or methylation occurs and RNA synthesis continues into the equivalent of the postcapping elongation phase. Because transcription precedes genome replication in the viral life cycle, questions remain as to: 1) What the mechanism that favors folding L into the conformation seen in PIV5 is; 2) what triggers the change to the state seen in VSV; and 3) whether the same L–P complex transitions between the two states. The CD-MTase-CTD module in the recently solved RSV L–P complex was not visible due to its high flexibility (19). In vivo there are likely to be other factors that could affect the folding pathways of these complexes including the presence of large amounts of N protein (14–16).

Capping Mechanism. It has been well documented that all mononegavirus polymerases contain the conserved HR and GxxT motifs required for a PRNTase capping mechanism (54). Paramyxoviruses and filoviruses contain a conserved KxxxKxxG motif common to eukaryotic GTases (23). In eukaryotic GTases, this motif lies in a β-sheet, while based on the structures of hMPV and now our PIV5 polymerase complex, this motif is found in an α-helix instead. The two lysines are on the same face of either the sheet or the helix, which could result in functional equivalence between the two enzymes. Because of the large rearrangements discussed earlier that are necessary for proper capping and methylation, it still remains unknown if this motif plays a catalytic role in the capping reaction in paramyxoviruses and filoviruses. An abolition of gene expression and production of infectious virions was observed when this motif was deleted from HPIV2 L (45). Mutation of any of the single residue or both lysines in the KxxxKxxG motif did not result in diminished gene expression; however, when all three critical residues were mutated, a reduction in gene expression was observed (45). This suggests the glycine is more functionally important than the lysines. Given the absolute requirement of a lysine in eukaryotic GTases and the high conservation of the HR and GxxT motifs in mononegaviruses, it is likely the case that the KxxxKxxG motif functions to help position the capped RNA in the correct position near the MTase active site, adjacent to the SAM substrate rather than in a catalytic capacity.

Location of the N Terminus of P. The N terminus of P interacts with nascent N^0-P to prevent premature encapsidation of RNA (2). The angle at which the P-OD projects away from L positions its N terminus far away from L. Some NNS viruses including RSV and VSV have much shorter P-ODs, leading to wide variation in the positioning of the N terminus relative to L (19, 32). There are ∼200 residues that are unstructured between the PIV5 P-OD and the N-terminal domain, making it difficult to predict the 3D organization of the N terminus (*SI Appendix*, Fig. S8). The N termini could interact with the transiently tethered N monomers to prevent premature encapsidation of RNA during both transcription and genome replication. Alternatively, it could function only during genome replication as a hub for coating the newly synthesized (anti-)genome using the N monomers recycled from the template.

Outlook. Our structure of the PIV5 L–P complex provides valuable structural information that furthers our understanding of the lifecycle of these polymerases. Differences in the positioning of the MTase and CTD lead to intriguing hypotheses regarding how the polymerase chooses between genome replication and transcription. Designing mutants that favor one conformation or another and testing the effect on capping and methylation are key to interrogating the mechanism. The biggest challenge remains trapping an L–P complex at specific points during transcription in order to dissect these intermediate steps and test the hypotheses described here.

Materials and Methods

Cells and Viruses. Baby hamster kidney (BHK) -21 and BSR T7/5 (BHK cells constitutively expressing T7 RNA polymerase) cells were maintained in DMEM supplemented with 10% FBS, tryptose phosphate broth (TPB), and 1% penicillin-streptomycin. Next, 500 μg/mL G418 was also added every third passage of BSR T7/5 cells. *Spodoptera frugiperda* (Sf9) cells, maintained in SF900 II SFM (Gibco) medium containing 10% FBS and 1% penicillin-streptomycin, were used for generating recombinant baculovirus (rBV) stocks and protein expression.

Plasmid Construction. A FLAG tag was introduced into the C terminus of L, and a histidine tag into the N terminus of P. These L and P genes were separately cloned into the pCAGGS vector for expression in mammalian cells. To generate rBVs coexpressing L–P, the codon-optimized L under the control of the polyhedrin promoter and P under p10 promoter were simultaneously cloned into the pFastBac Dual vector (Invitrogen). Eight different constructs (165–278, 172–278, 183–278, 203–278, 207–278, 214–278, 1–278, 178–392) encoding the PIV5 (strain W3A) P were amplified and subcloned in pET28a expression vector. All of the constructs were in frame to produce the N-terminal 6-His tag followed by thrombin cleavage site.

Expression and Purification of PIV5 L–P and P-OD. The rBV expressing L–P were generated in Sf9 cells by following transfection of bacmid DNAs with Cellfectin (Invitrogen). Following the determination of virus titers by plaque assay, Sf9 cells were infected with the rBVs at a multiplicity of infection of 1. At 96 to 120 h postinfection, the cell lysates were harvested by centrifugation (5,000 rpm, 50 min, 4 °C) and the cell pellets were suspended with

buffer (50 mM Tris·HCl pH 7.4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM DTT, 1 mM EDTA, 1 mM PMSF and EDTA-free complete protease inhibitor [Roche]). The lysates were briefly sonicated and centrifuged (25,000 rpm, 60 min, 4 °C) to obtain the supernatant containing soluble L–P complex. L–P was first purified by affinity chromatography using anti-FLAG M2 affinity gel (Sigma) and eluted with elution buffer (25 mM Hepes pH 7.4, 500 mM NaCl, 10% glycerol, 0.2% Tween 20, 1 mM DTT and 100 μg/mL FLAG peptide [Sigma]). For further purification, the L–P complex was passed through a Superdex 200 10/300 column (GE Healthcare) in buffer (25 mM Hepes pH 7.4, 500 mM NaCl, 5% glycerol, 0.2% Tween 20, 1 mM DTT, and 6 mM $MgSO_4$). The peak fractions were collected for analysis by EM.

The cloned constructs of PIV5 P were transformed into *Escherichia coli* BL21 T7 Express *LysY/I*^{q} cells (New England BioLabs) and the P fragments were expressed by growing the culture and then induced using IPTG (isopropyl β-D-1-thiogalactopyranoside) at 18 °C for 16 h. The cells were harvested and resuspended in 20 mM Tris·HCl pH 7.6, 300 mM NaCl, 5% glycerol, protease inhibitor mixture, DNase, RNase, lysozyme, and 20 mM imidazole. After cell lysis by sonication and clarifying the lysate via centrifugation, the soluble fraction was loaded onto Ni-NTA beads (Pre-equilibrated with buffer 20 mM Tris·HCl pH 8.0, 1 M NaCl, 5% glycerol, and 20 mM imidazole). After extensive washing, the elution was done with buffer containing 20 mM Tris·HCl pH 8.0, 400 mM NaCl, and 250 mM imidazole. The eluted fractions were pooled, dialyzed for removal of imidazole and treated with biotinylated thrombin overnight to cleave the His tag. Thrombin was removed using streptavidin beads and the sample was applied to Ni-NTA beads again. The flow-through containing the His-tag cleaved purified P protein fragment was pooled. Size-exclusion chromatography was carried out in buffer 20 mM Tris·HCl pH 8.0, 400 mM NaCl to further purify each P fragment and check their oligomerization state.

Crystallization, Data Collection, and Structure Determination of PIV5 P-OD. Each of the purified PIV5 P fragments were pooled and concentrated to ∼10 mg/mL. The crystallization screens were set up for each of the P fragments using the hanging-drop vapor diffusion method. The P fragment 172–278 was crystallized in buffer 0.1 M MIB buffer (malonic acid, imidazole, boric acid), pH 6.0 and 25% (wt/vol) PEG (polyethylene glycol) 1500 in 15 d at 20 °C. The crystals were soaked in 20% glycerol and the X-ray diffraction data were collected. We could not find phases by molecular replacement using the known structures of paramyxovirus P-OD. The successful structure solution was done using Ample (55) using module molecular replacement by ideal helices. Also, the ccp4 online server was used to run the program Ample in which the structures were provided by Quark ab initio protein model (56). Both of these led to the solution of one chain of P-OD, which was further used for molecular replacement in Phaser (57). The structure was refined using iterative cycles of Phenix real space refine (58) and manual building in Coot (59). The structure was validated using Molprobity (60) in Phenix real space refine.

Recovery of PIV5 by Reverse Genetics. To examine the functional viral polymerase activity of recombinant L–P complex, reverse genetics was performed as described before (61) using wild-type L and P or epitope-tagged L and P to rescue the virus. Briefly, BSR-T7/5 cells were transfected with the genome plasmid and support plasmids (pCAGGS-N, pCAGGS-P, and pCAGGS-L) using lipofectamine and PLUS reagent according to manufacturer protocol. After 5 d posttransfection, the supernatant containing P1 virus was harvested and used to infect BHK cells. The cells were observed for visualization of syncytia and photographed using an AMG EVOS xl inverted microscope (Fisher Scientific).

Electron Microscopy. Negative-stain samples were prepared using 400-mesh copper grids with a thin layer of continuous carbon that was glow discharged in air for 10 s with 25 W of power. Purified L–P in Buffer A (25 mM Hepes pH 7.4, 500 mM NaCl, 5 mM DTT, 5% glycerol, 6 mM $MgSO_4$, and 0.2% Tween 20) was diluted to 33 nM in Buffer A and incubated for 10 min on a grid in a homemade humidity chamber at 4 °C. The grid was sequentially incubated on 4, 50-μL drops of 2% uranyl formate solution for 5, 10, 15, and 20 s, and blotted dry with #1 filter paper (Whatman). Images were collected on a Jeol 1400 equipped with a Gatan 4k × 4× CCD camera at 30,000× magnification (3.71 Å/pixel), a defocus range of −1 to −2 μm, and 20 $e^-/Å^2$ total dose using Leginon (62).

Cryo-EM samples were prepared using a C-Flat 4/1 400 mesh copper grid (EMS) with a thin layer of continuous carbon that were glow discharged in air for 10 s with 5 W of power. Purified L–P in Buffer A was diluted to 0.25 μM using Buffer A. For each sample, 3.5 μL was incubated with 0.05% glutaraldehyde for 5 min on ice in the dark. The sample was applied to a grid suspended in a Vitrobot operating at 4 °C with 100% humidity. After 90 s, the sample was blotted with 25 force for 4 s and immediately plunged into liquid ethane cooled to liquid nitrogen temperatures. Images were collected using semiautomated data collection in Leginon (62) on a JEOL 3200FS microscope at 200 kV equipped with a Gatan K2 direct detector and omega energy filter operating in superresolution mode at a magnification of 30,000× (0.597 Å per pixel), defocus range from −1.5 to −4.5 μm, and a dose of 2.02 e^-/A^2 per frame for 40 frames. Two datasets were collected on separately prepared samples with 1,077 and 2,607 micrographs, respectively.

Image Processing. For negative-stained samples, particles were picked using DogPicker, extracted, and two-dimensionally (2D) classified using iterative MSA/MRA topological alignment within the Appion data processing software (63–66). A particle stack of 96,043 particles with a box size of 80 × 80 pixels was subjected to iterative, multireference projection-matching 3D refinement using libraries from the EMAN2 software package, starting with a circular mask of 163 Å and increasing to 193 Å, 237 Å, and finally 282 Å (67).

Cryo-EM micrographs were binned by 2 and motion-corrected using Motioncor2 with dose-weighting (68). Particle coordinates were selected using gAutomatch (developed by Kai Zhang, Medical Research Council Laboratory of Molecular Biology, Cambridge, United Kingdom). gCTF v0.50 was used for per particle CTF estimation (69). Particles were extracted using Relion 1.4 and all subsequent processing was done in Relion 2 or Relion 3 (70–72). The initial set of 717,008 particles was extracted with a further binning by 2, resulting in a pixel size of 2.24 Å. We performed 142 rounds of 3D classification using the negative-stain volume as an initial reference, a circular radius of 300 Å, and a soft mask around the complex. Class 5 displayed signs of α-helices so particles corresponding to that class were unbinned by a factor of 2 (pixel size 1.12 Å) and subjected to automatic 3D refinement, which resulted in a reconstruction of 6.0 Å. Three-dimensional classification without alignment was performed to sort out compositional or conformational heterogeneity in the dataset for 100 iterations. Classes 1 and 4 both improved in resolution after classification and each class was then refined separately and postprocessed to 4.8 and 5.2 Å, respectively. Higher-order aberration and magnification anisotropy calculations were performed in Relion 3.1-beta and rerefinement and postprocessing of both classes resulted in final resolutions of 4.38 and 4.8 Å, respectively. All resolutions reported use the gold-standard Fourier shell correlation (FSC) at the 0.143 criterion (73).

Model Building. ***P-OD EM density.*** The crystal structure of P was rigid body fit into the density using the C-terminal half of the P-OD so as not to be influenced by the poor density at the N terminus using University of California, San Francisco (UCSF) Chimera (74).

P-XD. A homology model for the PIV5 P-XD was generated using the mumps P-XD crystal structure (26). Rigid body docking into the EM density was unambiguous.

L. A homology model of the PIV5 L protein was generated using the VSV L structure (17). This model was rigid body docked into the class 1 EM density using the RdRp and PRNTase domains only. The CD was rigid body docked into its position in the density requiring a small rotation away from the RdRp-PRNTase module. A homology model of the MTase-CTD module using the hMPV crystal structure was generated along with the homology model from VSV was rigid body docked into its novel position in the PIV5 density. The model of L was built manually in Coot (59) utilizing the homology model as a guide. Flexible loops and portions where the density was poor were left unmodeled if the C-α path was ambiguous. In many cases there are differences in secondary structure between PIV5 and VSV. In these cases, bulky side chains visible in the density and secondary structure prediction was utilized in order to determine the correct register of the polypeptide. The zinc binding sites identified in the VSV structure were conserved and displayed density, suggesting the presence of metal ions in our density. These sites were built according to the corresponding sites in VSV. Ramachandran outliers and other errors introduced in manual model building were fixed prior to multiple rounds of Phenix real-space refinement (75). The resulting complete model including P was refined using Phenix real-space refine.

For class004 (*SI Appendix*, Fig. S2*B*), the CD-MTD-CTD module was rigid body docked into the density and rerefined using Phenix real space refine (75). All visualization, figures, and movies were generated in UCSF Chimera X (76).

Data Availability. Coordinates for the model of PIV5 P solved by X-ray crystallography have been deposited in the Protein Data Bank (PDB) with ID code 6VAG. Electron density maps and coordinates for the models of PIV5 L–P solved by cryo-EM have been deposited in the Electron Microscopy Data

Bank (EMDB) with ID codes EMDB-21095 and EMDB-21096, and the Protein Data Bank with ID codes 6V85 and 6V86, for class001 and class004, respectively.

ACKNOWLEDGMENTS. We thank Dr. Jonathan Remis for assistance with microscope operation and data collection; Jason Pattie for computer support; Ryan Marcum and George Leser for comments on the manuscript; and the staff at the Structural Biology Facility of Northwestern University for technical support. This work was supported by a Cornew Innovation Award from the Chemistry of Life Processes Institute at Northwestern University (to Y.H.); a Catalyst Award by the Chicago Biomedical Consortium with support from the Searle Funds at The Chicago Community Trust (to Y.H.); Institutional Research Grant IRG-15-173-21 from the American Cancer Society (to Y.H.); and an H Foundation Core Facility Pilot Project Award (to Y.H.). Y.H. is supported by National Institute of General Medical Sciences Grant R01-GM135651, National Cancer Institute (NCI) Grant P01-CA092584, and a Pilot Project Award from NCI Grant U54-CA193419. R.A. is supported by the Molecular Biophysics Training Program from National Institute of General Medical Sciences/NIH (T32-GM008382). M.A. is an Associate, and R.A.L. is an Investigator of the Howard Hughes Medical Institute. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract DE-AC02-06CHI1357. Our work at Life Sciences Collaborative Access Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology TriCorridor program (Grant 085PI000817).

1. M. Alayyoubi, G. P. Leser, C. A. Kors, R. A. Lamb, Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E1792–E1799 (2015).
2. M. Aggarwal, G. P. Leser, C. A. Kors, R. A. Lamb, Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein in complex with an amino-terminal peptide of the phosphoprotein. *J. Virol.* **92**, 2242–2214 (2018).
3. B. D. Welch *et al*., Structure of the parainfluenza virus 5 (PIV5) hemagglutinin-neuraminidase (HN) ectodomain. *PLoS Pathog.* **9**, e1003534 (2013).
4. B. D. Welch *et al*., Structure of the cleavage-activated prefusion form of the parainfluenza virus 5 fusion protein. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16672–16677 (2012).
5. T. A. Poor *et al*., On the stability of parainfluenza virus 5 F proteins. *J. Virol.* **89**, 3438–3441 (2015).
6. R. A. Lamb, G. D. Parks, "Paramyxoviridae: The viruses and their replication" in *Fields Virology*, D. N. Knipe, B. N. Fields, P. M. Howley, Eds. (Lippincott, Williams, and Wilkins, Philadelphia, 2013), **vol. 1**, pp. 957–995.
7. S. U. Emerson, R. R. Wagner, L protein requirement for in vitro RNA synthesis by vesicular stomatitis virus. *J. Virol.* **12**, 1325–1335 (1973).
8. N. Hercyk, S. M. Horikami, S. A. Moyer, The vesicular stomatitis virus L protein possesses the mRNA methyltransferase activities. *Virology* **163**, 222–225 (1988).
9. D. M. Hunt, E. F. Smith, D. W. Buckley, Aberrant polyadenylation by a vesicular stomatitis virus mutant is due to an altered L protein. *J. Virol.* **52**, 515–521 (1984).
10. T. Ogino, M. Kobayashi, M. Iwama, K. Mizumoto, Sendai virus RNA-dependent RNA polymerase L protein catalyzes cap methylation of virus-specific mRNA. *J. Biol. Chem.* **280**, 4429–4435 (2005).
11. K. R. Qanungo, D. Shaji, M. Mathur, A. K. Banerjee, Two RNA polymerase complexes from vesicular stomatitis virus-infected cells that carry out transcription and replication of genome RNA. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 5952–5957 (2004).
12. S. Vidal, D. Kolakofsky, Modified model for the switch from Sendai virus transcription to replication. *J. Virol.* **63**, 1951–1958 (1989).
13. S. L. Noton, C. Z. Tremaglio, R. Fearns, Killing two birds with one stone: How the respiratory syncytial virus polymerase initiates transcription and replication. *PLoS Pathog.* **15**, e1007548 (2019).
14. S. L. Noton, R. Fearns, Initiation and regulation of paramyxovirus transcription and replication. *Virology* **479–480**, 545–554 (2015).
15. S. Plumet, W. P. Duprex, D. Gerlier, Dynamics of viral RNA synthesis during measles virus infection. *J. Virol.* **79**, 6900–6908 (2005).
16. M. A. Hoffman, A. K. Banerjee, Precise mapping of the replication and transcription promoters of human parainfluenza virus type 3. *Virology* **269**, 201–211 (2000).
17. B. Liang *et al*., Structure of the L protein of vesicular stomatitis virus from electron cryomicroscopy. *Cell* **162**, 314–327 (2015).
18. J. Pan *et al*., Structure of the human metapneumovirus polymerase phosphoprotein complex. *Nature* **577**, 275–279 (2020).
19. M. S. A. Gilman *et al*., Structure of the respiratory syncytial virus polymerase complex. *Cell* **179**, 193–204.e14 (2019).
20. M. S. Sidhu, J. P. Menonna, S. D. Cook, P. C. Dowling, S. A. Udem, Canine distemper virus L gene: Sequence and comparison with related viruses. *Virology* **193**, 50–65 (1993).
21. O. Poch, B. M. Blumberg, L. Bougueleret, N. Tordo, Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: Theoretical assignment of functional domains. *J. Gen. Virol.* **71**, 1153–1162 (1990).
22. R. Chandrika, S. M. Horikami, S. Smallwood, S. A. Moyer, Mutations in conserved domain I of the Sendai virus L polymerase protein uncouple transcription and replication. *Virology* **213**, 352–363 (1995).
23. G. C. Paesen *et al*., X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat. Commun.* **6**, 8749 (2015).
24. A. A. Rahmeh *et al*., Critical phosphoprotein elements that regulate polymerase architecture and function in vesicular stomatitis virus. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14628–14633 (2012).
25. B. Morin, P. J. Kranzusch, A. A. Rahmeh, S. P. J. Whelan, The polymerase of negative-stranded RNA viruses. *Curr. Opin. Virol.* **3**, 103–110 (2013).
26. R. L. Kingston, L. S. Gay, W. S. Baase, B. W. Matthews, Structure of the nucleocapsid-binding domain from the mumps virus polymerase; an example of protein folding induced by crystallization. *J. Mol. Biol.* **379**, 719–731 (2008).
27. R. L. Kingston, D. J. Hamel, L. S. Gay, F. W. Dahlquist, B. W. Matthews, Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8301–8306 (2004).
28. K. Johansson *et al*., Crystal structure of the measles virus phosphoprotein domain responsible for the induced folding of the C-terminal domain of the nucleoprotein. *J. Biol. Chem.* **278**, 44567–44573 (2003).
29. J. F. Bruhn *et al*., Crystal structure of the nipah virus phosphoprotein tetramerization domain. *J. Virol.* **88**, 758–762 (2014).
30. R. Cox *et al*., Structural and functional characterization of the mumps virus phosphoprotein. *J. Virol.* **87**, 7558–7568 (2013).
31. G. Communie *et al*., Structure of the tetramerization domain of measles virus phosphoprotein. *J. Virol.* **87**, 7166–7169 (2013).
32. H. Ding, T. J. Green, S. Lu, M. Luo, Crystal structure of the oligomerization domain of the phosphoprotein of vesicular stomatitis virus. *J. Virol.* **80**, 2808–2814 (2006).
33. N. Tarbouriech, J. Curran, R. W. Ruigrok, W. P. Burmeister, Tetrameric coiled coil domain of Sendai virus phosphoprotein. *Nat. Struct. Biol.* **7**, 777–781 (2000).
34. G. D. Parks, Mapping of a region of the paramyxovirus L protein required for the formation of a stable complex with the viral phosphoprotein P. *J. Virol.* **68**, 4862–4872 (1994).
35. B. Cevik *et al*., The phosphoprotein (P) and L binding sites reside in the N-terminus of the L subunit of the measles virus RNA polymerase. *Virology* **327**, 297–306 (2004).
36. S. M. Horikami, S. Smallwood, B. Bankamp, S. A. Moyer, An amino-proximal domain of the L protein binds to the P protein in the measles virus RNA polymerase complex. *Virology* **205**, 540–545 (1994).
37. D. E. Holmes, S. A. Moyer, The phosphoprotein (P) binding site resides in the N terminus of the L polymerase subunit of Sendai virus. *J. Virol.* **76**, 3078–3083 (2002).
38. A. G. Malur, S. K. Choudhary, B. P. De, A. K. Banerjee, Role of a highly conserved NH(2)-terminal domain of the human parainfluenza virus type 3 RNA polymerase. *J. Virol.* **76**, 8101–8109 (2002).
39. M. Chenik, M. Schnell, K. K. Conzelmann, D. Blondel, Mapping the interacting domains between the rabies virus polymerase and phosphoprotein. *J. Virol.* **72**, 1925–1930 (1998).
40. S. Smallwood, K. W. Ryan, S. A. Moyer, Deletion analysis defines a carboxyl-proximal region of Sendai virus P protein that binds to the polymerase L protein. *Virology* **202**, 154–163 (1994).
41. V. Du Pont, Y. Jiang, R. K. Plemper, Bipartite interface of the measles virus phosphoprotein X domain with the large polymerase protein regulates viral polymerase dynamics. *PLoS Pathog.* **15**, e1007995 (2019).
42. A. A. Rahmeh *et al*., Molecular architecture of the vesicular stomatitis virus RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20075–20080 (2010).
43. P. C. Jordan *et al*., Initiation, extension, and termination of RNA synthesis by a paramyxovirus polymerase. *PLoS Pathog.* **14**, e1006889 (2018).
44. P. Bi, B. Shu, P. Gong, Crystal structure of the coxsackievirus A16 RNA-dependent RNA polymerase elongation complex reveals novel features in motif A dynamics. *Virol. Sin.* **32**, 548–552 (2017).
45. M. Nishio *et al*., Human parainfluenza virus type 2 L protein regions required for interaction with other viral proteins and mRNA capping. *J. Virol.* **85**, 725–732 (2011).
46. C. Chu *et al*., Structure of the guanylyltransferase domain of human mRNA capping enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10104–10108 (2011).
47. J. B. Ruedas, J. Perrault, Putative domain-domain interactions in the vesicular stomatitis virus L polymerase protein appendage region. *J. Virol.* **88**, 14458–14466 (2014).
48. Y. Zhou *et al*., Structure and function of flavivirus NS5 methyltransferase. *J. Virol.* **81**, 3891–3903 (2007).
49. L. M. Bloyet *et al*., Regulation of measles virus gene expression by P protein coiled-coil properties. *Sci. Adv.* **5**, eaaw3702 (2019).
50. G. Tekes, A. A. Rahmeh, S. P. Whelan, A freeze frame view of vesicular stomatitis virus transcription defines a minimal length of RNA for 5′ processing. *PLoS Pathog.* **7**, e1002073 (2011).
51. D. Kolakofsky, P. Le Mercier, F. Iseni, D. Garcin, Viral DNA polymerase scanning and the gymnastics of Sendai virus RNA synthesis. *Virology* **318**, 463–473 (2004).
52. S. A. Krumm, M. Takeda, R. K. Plemper, The measles virus nucleocapsid protein tail domain is dispensable for viral polymerase recruitment and activity. *J. Biol. Chem.* **288**, 29943–29953 (2013).
53. P. R. Paul, D. Chattopadhyay, A. K. Banerjee, The functional domains of the phosphoprotein (NS) of vesicular stomatitis virus (Indiana serotype). *Virology* **166**, 350–357 (1988).
54. T. Ogino, T. J. Green, RNA synthesis and capping by non-segmented negative strand RNA viral polymerases: Lessons from a prototypic virus. *Front. Microbiol.* **10**, 1490 (2019).
55. J. Bibby, R. M. Keegan, O. Mayans, M. D. Winn, D. J. Rigden, AMPLE: A cluster-and-truncate approach to solve the crystal structures of small proteins using rapidly computed ab initio models. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 1622–1631 (2012).

56. R. M. Keegan *et al*., Exploring the speed and performance of molecular replacement with AMPLE using QUARK ab initio protein models. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 338–343 (2015).
57. A. J. McCoy *et al*., Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
58. P. V. Afonine *et al*., Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
59. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
60. V. B. Chen *et al*., MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
61. B. He, R. G. Paterson, C. D. Ward, R. A. Lamb, Recovery of infectious SV5 from cloned DNA and expression of a foreign gene. *Virology* **237**, 249–260 (1997).
62. C. Suloway *et al*., Automated molecular microscopy: The new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
63. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
64. N. R. Voss, C. K. Yoshioka, M. Radermacher, C. S. Potter, B. Carragher, DoG Picker and TiltPicker: Software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* **166**, 205–213 (2009).
65. G. C. Lander *et al*., Appion: An integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* **166**, 95–102 (2009).
66. T. Ogura, K. Iwasaki, C. Sato, Topology representing network enables highly accurate classification of protein images taken by cryo electron-microscope without masking. *J. Struct. Biol.* **143**, 185–200 (2003).
67. G. Tang *et al*., EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
68. S. Q. Zheng *et al*., MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
69. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
70. J. Zivanov *et al*., New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
71. D. Kimanius, B. O. Forsberg, S. H. Scheres, E. Lindahl, Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016).
72. S. H. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
73. R. Henderson *et al*., Outcome of the first electron microscopy validation task force meeting. *Structure* **20**, 205–214 (2012).
74. E. F. Pettersen *et al*., UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
75. P. V. Afonine *et al*., Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
76. T. D. Goddard *et al*., UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).