Structure of the Vesicular Stomatitis Virus N⁰-P Complex

Cédric Leyrat¹, Filip Yabukarski¹,², Nicolas Tarbouriech¹, Euripedes A. Ribeiro Jr.¹,², Malene Ringkjøbing Jensen², Martin Blackledge², Rob W. H. Ruigrok¹, Marc Jamin¹*¹

UMI 3265 UJF-EMBL-CNRS Unit of Virus Host Cell Interactions, Grenoble, France, ²UMR 5075 CEA-CNRS-UJF, Institut de Biologie Structurale, Grenoble, France

Abstract

Replication of non-segmented negative-strand RNA viruses requires the continuous supply of the nucleoprotein (N) in the form of a complex with the phosphoprotein (P). Here, we present the structural characterization of a soluble, heterodimeric complex between a variant of vesicular stomatitis virus N lacking its 21 N-terminal residues (N₂¹) and a peptide of 60 amino acids (P₆₀) encompassing the molecular recognition element (MoRE) of P that binds RNA-free N (N⁰). The complex crystallized in a decameric circular form, which was solved at 3.0 Å resolution, reveals how the MoRE folds upon binding to N and competes with RNA binding and N polymerization. Small-angle X-ray scattering experiment and NMR spectroscopy on the soluble complex confirms the binding of the MoRE and indicates that its flanking regions remain flexible in the complex. The structure of this complex also suggests a mechanism for the initiation of viral RNA synthesis.

Introduction

Negative-sense RNA viruses include numerous major human pathogens such as influenza virus, rabies virus, measles virus and respiratory syncytial virus. The (−)RNA genome of these viruses is condensed by a viral nucleoprotein (N) into a helical nucleocapsid [¹], that associates with the polymerase complex and serves as the template for RNA replication and transcription [²]. Replication of the genome thus requires a continuous supply of N molecules to encapsidate both the (+)RNA intermediate copies and the newly synthesized (−)RNA genomes in single-stranded forms [³]. For non-segmented (−)RNA viruses of the Rhabdoviridae and Paramyxoviridae, N is assisted by the viral phosphoprotein (P); P binds to nascent RNA-free N, forming a N⁰-P complex (the superscript ⁰ denotes the absence of RNA) that prevents the polymerization of N and the non-specific encapsidation of host cell RNAs [⁴,⁵,⁶,⁷,⁸,⁹,¹⁰]. These processes are independent of each other [¹⁰] and, therefore, P has to fulfill two chaperone activities; blocking both RNA binding and self-assembly of N, P binds to nascent RNA-free N, forming a N⁰-P complex (the superscript ⁰ denotes the absence of RNA) that prevents the polymerization of N and the non-specific encapsidation of host cell RNAs [⁴,⁵,⁶,⁷,⁸,⁹,¹⁰]. These processes are independent of each other [¹⁰] and, therefore, P has to fulfill two chaperone activities; blocking both RNA binding and self-assembly of N, P

RNA complex is stabilized by multiple salt bridges between the sugar-phosphate backbone of the RNA and basic residues of N, by contacts between neighboring N molecules involving hydrophobic side-to-side interactions, mainly between adjacent N⁰-NTD, and by the exchange of N- and C-terminal sub-domains between adjacent N protomers (NNT-arm, aa 1–21 and NCT-loop, aa 340–375, respectively) [¹⁴,¹⁵]. Once formed, the N-RNA complex is stable and cannot be disassembled by full-length P [¹⁶]. However, on the basis of the N-RNA structure, we hypothesized that deletion of the NNT-arm may sufficiently destabilize the N-RNA complex so that P or a peptide fragment of P containing the MoRE that binds N⁰ could displace the RNA molecule.

In this study, we report the reconstitution of complexes between a recombinant N of VSV lacking the 21 N-terminal residues, N₂¹, and either full-length P dimer [¹¹] or a peptide encompassing the N⁰-binding MoRE of P [¹³], named here P₆₀, that comprises the first 60 amino acids of P, a two-amino acid linker and a C-terminal His₅-tag. The characterization by absorbance spectroscopy and size-exclusion chromatography (SEC) combined with static light scattering (MALLS) demonstrates that both N₂¹⁰-P₆₀ and N₂¹⁰-P₆₀ dimer complexes are free of RNA in solution, forming soluble heterodimers or heterotrimers, respectively. Therefore, P₆₀ fulfills both chaperone activities of full-length P. The heterodimer N₂¹⁰-P₆₀ crystallized, but under the crystallization conditions, it assembled into a circular decamer of heterodimers that comprises the first 60 amino acids of P, a two-amino acid linker and a C-terminal His₅-tag. The characterization by absorbance spectroscopy and size-exclusion chromatography (SEC) combined with static light scattering (MALLS) demonstrates that both N₂¹⁰-P₆₀ and N₂¹⁰-P₆₀ dimer complexes are free of RNA in solution, forming soluble heterodimers or heterotrimers, respectively. Therefore, P₆₀ fulfills both chaperone activities of full-length P. The heterodimer N₂¹⁰-P₆₀ crystallized, but under the crystallization conditions, it assembled into a circular decamer of heterodimers very similar to the previously crystallized decameric N-RNA ring. The crystal structure of the decameric form of the N₂¹⁰-P₆₀ complex reveals the molecular mechanisms by which the N²-binding MoRE of P attaches to N. NMR spectroscopy confirms that the MoRE of P binds to N in the N₂¹⁰-P₆₀ complex in solution as in the crystal structure and shows that the regions of
Author Summary

The negative sense RNA genome of the rabdoviruses is encapsidated by the nucleoprotein, and the replication of the genome requires a continuous supply of RNA-free, monomeric nucleoprotein (N0) to encapsidate the newly synthesized (+)RNA intermediate antigensomes and (−)RNA genomes. In this process, the viral phosphoprotein acts as a chaperone, forming a heterodimeric complex, named N0-P, which prevents nascent N molecules from self-assembling and from binding to cellular RNAs. We reconstructed the N0-P complex of the prototype rabdovirus, vesicular stomatitis virus, and characterized its structure by crystal X-ray diffraction and solution experiments. Our results show how the N-terminal region of the phosphoprotein folds upon binding to the RNA-free nucleoprotein and how it prevents the nonspecific encapsidation of host-cell RNA. This complex is soluble and heterodimeric, but by forcing it to polymerize into a crystal it associated into a circular decamer of heterodimers very similar to the previously crystallized decameric N-RNA ring. On the basis of our results, we propose a model that explains the role of the phosphoprotein in the encapsidation of newly synthesized RNA and in the initiation of RNA synthesis by the viral polymerase.

P flanking this MoRE remain flexible in the complex. Finally, these results suggest mechanisms for the encapsidation of newly synthesized RNA and for the initiation of RNA synthesis by the viral polymerase.

Results

Strategy for reconstituting the N211−P and N211−P60 complexes

Production of a mutant of N deleted of its 21 first N-terminal residues (N211) in Escherichia coli led to the formation of inclusion bodies and of poorly soluble complexes, which could not be purified. In order to improve the solubility of the N211 mutant, it was produced in E. coli in fusion with an N-terminal maltose binding protein (MBP) tag. The purified MBP-N211 formed soluble, oligomeric N-RNA complexes, which eluted next to the exclusion volume of a Superdex 200 column (Figures S1A and S1B in Text S1). The presence of RNA was demonstrated by the absorbance ratio at 280 nm and 260 nm of 1.05 (A280 nm/A260 nm) (Figure S1C in Text S1). The MBP-N211 monomer migrated as a single protein of about 100 kDa on a denaturing 4–20% gradient PAGE (Figure S1B in Text S1). Incubation of MBP-N211 at 20°C overnight in the presence of P60 resulted in the displacement of the bacterial RNA from N, the dissociation of the oligomeric N-RNA complexes and the formation of a new species that eluted at 14.1 mL (Figure S2A in Text S1). The analysis by SEC-MALLS indicated a weight-averaged molecular mass of 92±2 kDa in agreement with the calculated molecular mass of the MBP-N211−P60 complex (calculated mass: 88,326 Da (MBP-N211)+8,053 Da (P60) = 96,379 Da). The co-elution of MBP-N211 and P60 was confirmed by denaturing 4–20% gradient PAGE (Figure S2B in Text S1). The complex contained much less RNA as shown by the absorbance spectrum (A280 nm/A260 nm = 1.60) (Figure S2C in Text S1). After cleavage of the MBP tag with the TEV protease, the resulting N211−P60 complex was purified by Ni2+ chelate affinity chromatography followed by SEC. The complex of N211 with full-length P dimer [11] was then prepared by incubating the purified N211−P60 complex overnight with the intact P dimer.

Solution properties of the N211−P and N211−P60 complexes

The molecular mass of the N211−P dimer complex determined by SEC-MALLS was constant throughout the chromatographic peak indicating that the complex was monodisperse (Mw/Mn = 1.00±0.01), and the molecular mass of 104±4 kDa was consistent with that of a heterotrimer composed of one N211 and an intact P dimer (calculated mass: 45,377 Da (N211)+2×30,976 Da = 107,329 Da) (Figure 1) in accordance with the dimeric state of P in solution [11] and with a previous determination for the rabies virus N−P complex by native mass spectrometry [17]. The hydrodynamic radius (Rd) of 3.8±0.1 nm is about 1.5 fold larger than that for a globular particle of the same molecular mass (calculated Rd = 4.0 nm) reflecting the elongated shape of the complex and the existence of a long N-terminal disordered region (aa 1–106) [11,12].

The N211−P60 complex contained no RNA (Figure 2A), and its Rd of 3.2±0.1 nm and molecular mass of 53±3 kDa indicated a globular 1:1 complex (Figure 2B), which agrees with the fact that P60 does not contain the dimerization domain of P [12,10]. The radius of gyration (Rg) of the N211−P60 complex determined from SAXS data (2.7±0.1 nm) (Figure S3 and Table S1 in Text S1) was similar to that of a single N protomer extracted from the N-RNA complex (2.8 nm) [15], but the calculated curve of the extracted protein poorly fitted the experimental curve of N211−P60, probably because of the presence of P60 (Figures S4A and S4B in Text S1). Ab initio bead models reconstructed from SAXS data [19] (Figure 2C) could easily accommodate the structure of an isolated N protomer deleted of its NTD-arm, except for the NCTD-loop, which is likely to adopt a different conformation in solution (Figure 2D). Although the low resolution of the model precluded the precise localization of P60, the absence of an empty groove at the interface between NNTD and NCTD suggests that P60 could bind in this region. These results clearly show that P60 fulfills both chaperone functions of P in maintaining N0 soluble and free of RNA, and because the size and flexibility of full-length P render the N211−P complex unsuitable for X-ray crystallography and NMR studies, the N211−P60 complex was used for further structural characterization.

Crystal structure of a decameric N211−P60 complex

The N211−P60 complex crystallized at low pH as a decameric circular complex. The structure was solved at a resolution of 3.0 Å by molecular replacement using the structure of an N protomer derived from the N-RNA complex [15] (Table 1). The structure of N in the N211−P60 complex was almost identical to that of N in the N-RNA complex (rmsd = 0.96 Å) (Figure 3A and Figure S5 in Text S1) [15]. The complex contained no RNA, but instead, in each protomer, residues 6 to 35 of P60 were visible in a groove formed by residues of the hinge region of N (aa 200–300) at the junction between NNTD and NCTD (Figure 3B). Previous observations showed that the isolated N0-binding region of P transiently populates β-helical conformers in the region 2–12 and 25–31 [13], and that residues 11 to 30 of P are essential for forming the N0-P complex [7]. Upon binding to N211, the second fluctuating helix is stabilized and extends from amino acids 17 to 31 (Figure 3B). The theoretical SAXS curve calculated from the crystal structure of one protomer of the N211−P60 complex fits adequately the experimental curve of the soluble complex (Figure S4A and S4C in Text S1), and the structure is perfectly accommodated within the ab initio bead model (Figure S4D in Text S1) showing that the structure of the N211−P60 complex in solution is the same as that in the crystal.

The structure of the N211−P60 complex shows how the N0-binding MoRE of P prevents both the interaction with RNA and
the self-assembly of soluble RNA-free N. The binding site of P is different from that of RNA but both sites do overlap (Figures 4A, 4B and Figure S5 in Text S1), and the C-terminal turn of the α-helix of P60 (aa 27–31) together with the following residues (aa 32–35) block the RNA binding cavity (Figure 4A) and inhibit RNA binding. Concomitantly, the other extremity of the MoRE (aa 6–15) docks in a shallow groove on the backside of the N protomer which, in the multimeric N-RNA complex, is occupied by the NCTD arm of the adjacent N-protomer and whose bottom is made up of the NCTD-loop of the N\textsubscript{nt}-protomer (Figure 4C and 4D). With w.t. N, the N-terminal part of the MoRE of P will compete with the NNT-arm of a neighboring N molecule and therefore interferes with the polymerization of N in the absence of RNA.

The MoRE of VSV P (aa 6–35) binds to the central hinge region of N mainly through hydrophobic interactions. The amphipathic α-helix of P, together with residues 14 to 16, inserts into a hydrophobic groove of N. Tyr14 perfectly fits into a small hydrophobic pocket (Figure 5A). In addition, the complex is stabilized by three intermolecular salt bridges (Figure 5A). This region of N (aa 200–300), which also plays a central role in binding RNA is highly conserved among VSV serotypes as well as in rabies virus (RAV) [20,21] (Figure S6 in Text S1). The hinge region of VSV N exhibits 30% identity in amino acid sequence with that of RAV N, as compared with 21% and 13% for the N- and C-terminal lobes, respectively. Figure 5B shows that several hydrophobic residues lining the binding groove of P are conserved between VSV and RAV N, as well as Arg132, suggesting that a similar complex forms in RAV (Figure S6 in Text S1). In the N-RNA complex, the RNA molecule interacts with N through electrostatic interactions between phosphate groups of the RNA backbone and basic residues of the protein, while the bases of three nucleotides (nt. 5, 7 and 8) are docked onto an hydrophobic surface of the RNA binding groove [15], which is also part of the MoRE binding site (Figure 5B). However, none of the basic residues of VSV N directly contacting the RNA backbone in the N-RNA complex is involved in the interaction with the MoRE of P. A similar mode of RNA binding was observed in the RAV N-RNA complex [14], but only two arginines out of the six residues involved in direct interactions with phosphate groups in the VSV complex are conserved in the RAV complex [20].

The RNA binding groove of N is rich in basic residues forming a highly positive surface area (Figure 5C), while the backside of NCTD harbors a negative surface potential (Figure 5C). The MoRE of P (aa 6–35) has a bipolar distribution of charges with a positive pole at its N-terminus and a negative pole its C-terminus. In the N\textsubscript{A21}\textsuperscript{D}P\textsubscript{60} complex, the negative pole of P localizes in the RNA binding groove, while the positive pole docks on the backside of NCTD, modifying the distribution of electrostatic potentials on these surfaces of N and suggesting that electrostatics could play a role in orientating P before binding (Figures 5C and 5D). The crystallization is at pH 4.6 and it is likely that protonation of acidic groups reduces repulsion forces that keep the N\textsubscript{A21}\textsuperscript{D}P\textsubscript{60} complex in its isolated form at pH 7.0.

NMR spectroscopy

In the crystal structure of the N\textsubscript{A21}\textsuperscript{D}P\textsubscript{60} complex residues 6 to 13 and 32 to 35 of P\textsubscript{60} exhibit conformational heterogeneity (Figure S7 in Text S1), while residues 1 to 5 and 36 to 68 are not visible. To further characterize the conformational dynamics of these parts of P\textsubscript{60} in the soluble N\textsubscript{A21}\textsuperscript{D}P\textsubscript{60} complex, we used nuclear magnetic resonance (NMR) spectroscopy. Initially, spectra of \textsuperscript{15}N, \textsuperscript{13}C, \textsuperscript{2}H-labeled P\textsubscript{60} in complex with unlabeled N\textsubscript{A21} were recorded. In a complex of this size (53 kDa) NMR signals are significantly broadened, precluding their detection, but in the HSQC spectrum of the N\textsubscript{A21}\textsuperscript{D}P\textsubscript{60} complex, resonances corresponding to the last 28 amino acids of P\textsubscript{60} (aa 41–60+linker+His\textsubscript{6} tag) are clearly visible (Figure 6A), suggesting that this tail remains free and flexible in the complex. Comparison with the free peptide showed that most resonances superimpose. Small chemical shift differences were observed for residues Q\textsubscript{41} to G\textsubscript{44}, probably due to the proximity of the bulk complex, and for two aromatic residues (Y\textsuperscript{52}, F\textsuperscript{54}) suggesting weak interactions of these residues with N. The amide backbone \textsuperscript{15}N transverse relaxation rate constant (R\textsubscript{2}) is

Figure 1. Heterotrimeric VSV N\textsubscript{0}\textsuperscript{D}-P dimer complex in solution. The complex formed between N\textsubscript{0}\textsuperscript{D} and full-length P dimer in solution was analyzed by SEC-MALLS. The N\textsubscript{A21}\textsuperscript{D}-P complex elutes at 12.0 mL and the remaining N\textsubscript{A21}\textsuperscript{D}-P\textsubscript{60} complex elutes at 15.6 mL (line). The molecular mass of 104±4 kDa (crosses) indicates a 1:2 complex between one RNA-free N\textsubscript{A21} and two P molecules in accordance with the previous observation that P forms exclusively dimers in solution and with the N\textsuperscript{D}-P\textsubscript{2} complex determined for rabies virus.

doi:10.1371/journal.ppat.1002248.g001
sensitive to rapid fluctuations at the pico- to nanosecond time scale, as well as to chemical shift exchange on the micro- to millisecond time scale. Transverse relaxation of the visible resonances increased from the C-terminus to the region containing the bound helix, indicating a corresponding increase in rigidity of the backbone (Figure 6B). Very weak additional peaks up-field shifted in the amide proton dimension were detected in a 15N-1H TROSY spectrum further supporting folding of the helical element upon binding (data not shown).

In a second experiment, addition of sub-stoichiometric amounts of unlabeled N\textsubscript{D21} attached to MBP (to maintain solubility) to 15N-labeled P\textsubscript{60} resulted in an overall reduction in the intensity of the peaks in the HSQC spectrum, in proportion of the amount of added MBP-N\textsubscript{D21} indicating that a fraction of P\textsubscript{60} formed a complex with MBP-N\textsubscript{D21}. Spin relaxation measurements of equilibrium mixtures of free and bound peptide refine our understanding of the dynamics of the system. Systematically larger \(R_2\) values observed for residues 1 to 17 reveal additional contributions from chemical shift exchange (\(R_n\)) that are not present in the free form of the peptide and that increase significantly upon increasing the molar ratio of MBP-N\textsubscript{D21} to P\textsubscript{60} (Figure 5C). This indicates that the N-terminal part of P\textsubscript{60} experiences conformational exchange when in complex with N\textsubscript{D21} with an interconversion rate on the micro to millisecond timescale. In the crystal structure, residues 6 to 13 bind in place of the N\textsubscript{NT}-arm of the protomer N\textsubscript{i-1} and pack onto the N\textsubscript{CT}-loop of protomer N\textsubscript{i+1} (Figures 4C and 4D), but in the isolated form of the complex, the N\textsubscript{CT}-loop binding surface for these residues is missing. These results confirmed that in the solution, like in the crystal, residues 17 to 35 of P form a stable complex with N, while the flanking N- and C-terminal regions remain dynamic. The flanking C-terminal part (aa 40–60) behaves as a flexible tail and shows little evidence of interaction with N. The flanking N-terminal region (aa 1–16) interacts with N but undergoes conformational exchange. These flexible regions seem dispensable for the chaperone activities of P since a shorter peptide encompassing residues 7 to 40 of P is also capable of displacing bacterial RNA from N\textsubscript{D21} and of forming a soluble heterodimeric 1:1 complex with this protein (data not shown).
Discussion

In the current study, we have determined the structure of the N$_0$-binding MoRE of P bound to N and demonstrated that the regions of P flanking the MoRE conserve some flexibility in the N$_0$-P complex. Because the N$_0$-P complex is required for the replication of the virus [5,22], inhibition of its formation might represent an interesting target for blocking viral replication.  

The N$_{A21}^0$-P$_{60}$ complex as a model of the N$_0$-P complex

Our results demonstrate that the reconstituted N$_{A21}^0$-P$_{60}$ complex is a suitable model for the viral N$_0$-P complex in agreement with previous studies. Both the crystal structure and the NMR spectroscopy experiments show that the MoRE of P, which adopts a stable conformation upon binding to N, includes residues 6 to 35 and corresponds closely to the fragment that was previously identified as essential and sufficient for maintaining N$_0$ when expressed in a recombinant form [7].  

Previous studies revealed that VSV P is a dimeric and modular protein in which the N-terminal part (aa 1–106) is globally disordered [11,12,13,18,24,25]. The dimerization domain of P is localized in the central region of the protein (aa 107–177) [18] and is therefore not present in P$_{60}$, which is monomeric [13]. The stoichiometry of the N$_{A21}^0$-P$_{60}$ complex [1:1] shows that one MoRE of P is capable of binding one N molecule. The stoichiometry of the complex formed between N$_{A21}$ and full-length P [1:2] in the concentration range used here suggests that a single N$_{A21}$ is bound to P dimer. The remaining part of the P dimer is tethered to N$_{A21}$ through a flexible linker, in agreement with the large hydrodynamic radius measured for the N$_{A21}^0$-P dimer complex. In isolation, the N-terminal region of P contains two transient α helices (aa 2–12 and 25–30) [13]. In the crystal structure, the second helix is stabilized and extends from residue 17 to residue 35, whereas the first helix is not present. Residues 1 to 5 are not visible and residues 6 to 15 adopt different conformations in the different protomers of the circular complex. Because NMR spectroscopy reveals that this N-terminal region of P (aa 1–17) is in chemical shift exchange, it is possible that in solution it adopts different conformations bound in different orientations on the surface of N, but that only those docked into the backside groove of N allow the packing of the N$_{A21}^0$-P$_{60}$ complex into crystals and were thus selected during the crystallization process.  

The truncated form of N (N$_{A21}$) conserves the ability of self-association in the presence of RNA and, like w.t. N, forms oligomeric N-RNA complexes when expressed in bacteria. As assumed from the structure of the oligomeric N-RNA complex [15], the N$_{A21}$-arm stabilizes the multimeric N-RNA complexes by linking together adjacent N protomers. The N-RNA complex formed with w.t. N could not be dissociated by the addition of full-length P or of a fragment of P encompassing the N$_0$-binding region. However, the deletion of the N-terminal sub-domain destabilized the complex and allowed P$_{60}$ to displace the RNA molecule and disassemble the multimeric N-RNA complex. In a previous study, the co-expression of P with a similar variant of VSV N lacking the first 22 amino acids (N$_{A21}$) led to the production of complexes of different sizes containing N$_{A21}$ and P but not of N-RNA complexes, suggesting a role for the N-terminal region of N in the encapsidation of the RNA [10]. Assuming an equilibrium between the N$_0$-P complex and the multimeric N-RNA complex, with w.t. N, the stabilization brought by the N$_{A21}$-arm to the multimeric assembly would displace the equilibrium towards the formation of the N-RNA complex. In the absence of the N$_{A21}$-arm, the truncated N molecules assemble onto cellular RNAs as seen in our expression system, but in the presence of co-expressed P, like upon addition of P$_{60}$ to our purified MBP-N$_{A21}$-RNA complexes, the equilibrium is displaced towards the formation of the N$_0$-P complexes. The absence of N-RNA complex in cells co-expressing N$_{A21}$ and P may not result from a default of encapsidation but rather from the displacement of the equilibrium towards N$_0$-P.

Unexpectedly, the N$_{A21}^0$-P$_{60}$ complex failed to crystallize as a heterodimer but crystallized into circular decamers of heterodimers. With the exception of the missing N$_{A21}$-arm, the structure of N$_{A21}$ in the N$_{A21}^0$-P$_{60}$ complex is very similar to that of N in the decameric N-RNA complex, with less than 1 Å r.m.s.d. between the two structures. Different explanations why multimerization occurs under crystallization conditions can be proposed. Firstly, the N$_{A21}^0$-P$_{60}$ complex crystallized at pH 4.6 like VSV circular N-RNA complexes [15], while solution SEC-MALLS and SAXS experiments were performed at pH 7.5 and NMR experiments at pH 6.0. A modification of the electrostatic surface potential (Figures 5C and 5D) could affect the equilibrium between heterodimeric and multimeric N$_{A21}^0$-P$_{60}$ complex. No evidence of multimerization was, however, found in solution at pH 4.6 in the concentration range used for SEC-MALLS and SAXS experiments. Secondly, the ring-like structure of ten protomers appears as a favored organization of VSV N, likely reflecting on some geometrical and/or surface properties of the protein. VSV N forms ring-like structures mostly containing ten N subunits in the presence of non-specific RNA when expressed in a recombinant form.
The RNA can be removed from these ring-like structures without disrupting the multimeric assembly [27]. A single amino acid variant of VSV N that is no longer capable of binding RNA also crystallized into a decameric assembly of empty N molecules [10]. These circular N-RNA complexes are artifacts of the crystallization process because the actual nucleocapsid is very long and cannot form rings. However, in the virion, the nucleocapsid adopts a bullet-shaped structure composed of a trunk in which the nucleocapsid regularly spirals into superposed turns of 37.5 subunits of N and of a tip which is formed of seven turns containing varying numbers of subunits [1]. The upper turn of the bullet tip, which may represent the nucleation centre from which the particle assembles, resembles a decameric ring, suggesting that the assembly in ten members ring or spiral corresponds to an optimal side-by-side orientation between adjacent N subunits. The RNA-free N\(^{D_{21}}\)\(^{0}\)-P\(_{60}\) complex is capable of assembling into circular multimers, and it seems likely that an increase of the concentration of the N\(^{D_{21}}\)\(^{0}\)-P\(_{60}\) complex under the crystallization conditions together with a change in pH shift the equilibrium towards the multimers.

This raises the question of the effect of crystallization on the structure of the N\(^{D_{21}}\)\(^{0}\)-P\(_{60}\) complex. The SAXS curve calculated for N\(^{D_{21}}\)\(^{0}\)-P\(_{60}\) protomer extracted from the crystal structure perfectly reproduced the experimental curve of the soluble complex, while NMR spectroscopy clearly shows that the same segment of P (aa 17–35) is involved in a stable complex with N in solution and in the crystal, arguing that crystallization has no major effect on the structure of the more rigid part of the complex. In solution, the N-terminal part of the MoRE of P (aa 1–16) appears to be in conformational exchange and could thus exist in different conformers including those observed in the crystal in which residues 6 to 16 are docked onto the backside groove of N. Crystallization of the N\(^{D_{21}}\)\(^{0}\)-P\(_{60}\) complex may thus select the more compact conformers and therefore not reproduce the conformational diversity of this region that is found in solution. In addition, the high conservation rate of residues of N forming the binding surface for the MoRE of P, both within VSV serotypes and between the evolutionarily more distant VSV and RAV, supports the localization of the interface between the two proteins and hints at the formation of a similar complex in RAV.

**Mechanisms of chaperone activities**

The characterizations in solution indicate that the N\(^{D_{21}}\)\(^{0}\)-P\(_{60}\) complex is RNA-free heterodimer and that, therefore, P\(_{60}\) or a shorter fragment of P (aa 7–40) fulfill both chaperone activities of P. The crystal structure of the N\(^{D_{21}}\)\(^{0}\)-P\(_{60}\) complex clearly shows how the N\(^0\)-binding region of P inhibits RNA binding by filling the RNA-binding groove of N. In solution, this part of P also forms a stable complex with N as seen by NMR spectroscopy. The structure also suggests how P prevents the self-assembly of N in the absence of RNA. In the bound form observed in the crystal structure, the N-terminal extremity of the MoRE of P (aa 6–16) directly competes with the N\(_{NT}\) arm of a neighboring N molecule. Assuming this region of P fluctuates between bound and free forms in the soluble N\(^0\)-P complex, the free form may act as an entropic bristle, thereby also preventing the oligomerization of N. With full-length P dimer, the flexibility and the bulkiness of the remainder of the protein may also contribute to this effect by masking the binding interfaces for RNAs or other N molecules. In addition, the MoRE of P exhibits a bipolar distribution of charges, with a positive pole at its N-terminal extremity and a negative pole at its C-terminal extremity. Binding of the MoRE of P modifies the...
electrostatic surface potential of N, notably reducing the positive surface potential on one side of the molecule, and may thereby affect the side-by-side interaction with another N molecule.

From the results presented here, we propose a hypothesis for the encapsidation of a newly synthesized RNA molecule during viral genome replication. By forming a complex with P, a nascent N molecule is prevented from binding to host-cell RNA and is preserved in a soluble form. During RNA replication, N is transferred to a growing RNA molecule and P is released. Little is known about the mechanism of this reaction or about the role played by the polymerase complex in this process. Our results show that the NNT-arm stabilizes the multimeric N-RNA complex and therefore suggest that the multimeric N-RNA complex is more stable than the N0-P complex. The transfer of N from the N0-P complex to the growing N-RNA complex could simply be driven by a higher stability of the N-RNA complex. Upon transfer of N onto the RNA and release of P, the backside groove of N of the last added N molecule is liberated and becomes available for accepting the NNT-arm of the next incoming N molecule (Figure 7A). By blocking the backside groove of N, the N-terminal part of P ensures that N molecules do not assemble into empty N polymers but assemble only onto an RNA molecule. It is also noteworthy that, in VSV, a high affinity binding site for the L protein was localized in the second half of the N-terminal disordered region of P [7,28]. The dynamic nature of the N-terminal region of P and the proximity of the two binding sites may have significance for the mechanism of action of the transcription/replication machinery. The binding of N0 to P may prevent the simultaneous binding of L, or conversely, the simultaneous binding of N0 and L may modify the activity of the polymerase.

In addition to its role in RNA encapsidation, the binding of the N-terminal region of P to N may also provide a mechanism for the initiation of (+)RNA synthesis at the genome 3' end and of (−)RNA synthesis at the antigenome 3' end [29,30]. Encapsidated RNA genome and antigenome are completely covered with the nucleoprotein and are not accessible to the RNA polymerase. However, the first N molecule at the 3' extremity of nucleocapsids exposes its binding site for the N-terminal MoRE of P (Figure 7B). By binding to this surface, P may destabilize the N-RNA complex sufficiently to displace several nucleotides from the first N protomer and allow the polymerase access to the RNA.
Materials and Methods

Reconstitution of the N\textsubscript{A21}\textsuperscript{0}-P and N\textsubscript{A21}\textsuperscript{0}-P\textsubscript{60} complexes

The cDNAs encoding vesicular stomatitis virus nucleoprotein or a fragment of this protein deleted of the 21 N-terminal residues were amplified by PCR and introduced into the pET-M40 plasmid (EMBL) using Ncol and XhoI restriction sites. The resulting constructs code for chimeric proteins that comprise an N-terminal maltose binding protein tag (MBP) and a tobacco etch virus (TEV) cleavage site. The cDNA encoding the 60 first N-terminal amino acids of VSV P (P\textsubscript{60}) was amplified by PCR and cloned into the pET28a plasmid containing a C-terminal His\textsubscript{6}-tag.
and a linker of two amino-acids (EL) using NcoI and XhoI restriction sites. All constructions were checked by DNA sequencing.

The plasmids were transformed into *Escherichia coli* Rosetta (DE3) cells and the expression of the recombinant proteins was induced with 1 mM isopropyl-1-thio-β-d-galactopyranoside (IPTG) for 18 h at 16°C. Cells were harvested by centrifugation, suspended in buffer A (20 mM Tris–HCl, pH 7.5, 150 mM NaCl, 5 mM DTT) containing protease inhibitors (Complete EDTA-Free, Roche Diagnostics) and disrupted by sonication. The extract was centrifuged at 20,000 g during 30 min at 4°C and the supernatant was filtered (0.45 μm). The MBP-fusion proteins were purified by affinity chromatography on amylose resin (New England Biolabs) followed by size exclusion chromatography (SEC) on a Superdex S200 column (GE Healthcare) equilibrated in buffer A. P60 was purified by affinity chromatography on a Ni2+ resin column (Quiagen) followed by SEC on a Superdex S75 column (GE Healthcare) equilibrated in buffer A supplemented with 50 mM Ghu and 50 mM Arg. Samples for NMR spectroscopy were produced in M9 minimal medium containing MEM vitamins (Gibco). For producing 15N-labeled P60, the medium was supplemented with 1.0 g.L−1 of 15NH4Cl and 2.0 g.L−1 of unlabeled glucose, while for producing 15N, 13C, 2H-labeled P60 the minimal medium was prepared in D2O and supplemented with 1.0 g.L−1 of 15NH4Cl and 2.0 g.L−1 of 13C glucose.

The N60-P60 complex was prepared by incubating overnight at 4°C an excess of P60 with the MBP-N60-RNA complexes. The MBP-N60-P60 complex was purified by Ni2+ chelate affinity chromatography by using the His-tag present on P60 to remove the excess of free MBP-N60-RNA complex, followed by SEC on a Superdex S200 column equilibrated in buffer A and amylose affinity chromatography to eliminate unbound P60.

The MBP tag was removed by incubating the protein with the TEV protease overnight at 4°C. The N protein contains the additional N-terminal tripeptide GAM. The N60-P60 complex was then purified using a Ni2+ chelate affinity chromatography followed by SEC on a Superdex S200 column equilibrated in buffer A. This procedure yielded pure N60-P60 complex. The samples were checked by SDS-PAGE using denaturing 4–20% gradient PAGE (Biorad).

**Figure 6.** The N-terminal and C-terminal region of P60 flanking the MoRE exhibits conformational flexibility in the soluble complex. (A) Comparison of the 2D 1H-15N HSQC NMR spectra of free 15N, 13C, 2H-labelled P60 (blue) and in complex with N621 (red). Both spectra were recorded at 14.1 T and 25°C in 20 mM Tris-HCl, 150 mM NaCl, 50 mM Ghu, 50 mM Arg with 10% D2O adjusted to pH 6.0. The labels indicate the assignment of the resonances of the complex. (B) 15N R2 spin relaxation rates measured under the same conditions as (A) for the free 15N-labeled P60 (blue) and the N621-P60 complex (red). (C) 15N R2 spin relaxation rates measured at 14.1 T and 10°C of the free 15N-labeled P60 (blue), a mixture of 0.27 mM 15N-labeled P60 and 0.09 mM unlabeled MBP-N621 (green) and a mixture of 0.24 mM 15N-labeled P60 and 0.17 mM unlabeled MBP-N621 (red).
doi:10.1371/journal.ppat.1002248.g006
Size exclusion chromatography (SEC) combined with detection by multi-angle laser light scattering (MALLS) and refractometry: SEC-MALLS

SEC was performed with a Superdex S200 column (GE Healthcare) equilibrated in 20 mM Tris–HCl, pH 7.5, 150 mM NaCl. Separations were performed at 20°C with a flow rate of 0.5 mL min⁻¹. 50 μL of a protein solution at a concentration ranging from 2.7 to 8.0 mg mL⁻¹ were injected. On-line multi-angle laser light scattering (MALLS) detection was performed with a DAWN-EOS detector (Wyatt Technology Corp., Santa Barbara, CA) using a laser emitting at 690 nm. Protein concentration was measured on-line by refractive index measurements using a RI2000 detector (Schambeck SFD) and a refractive index increment dn/dc = 0.185 mL g⁻¹. Data were analyzed and weight-averaged molecular masses (Mw) were calculated using the software ASTRA V (Wyatt Technology Corp., Santa Barbara, CA) as described previously [11]. For size determination, the column was calibrated with proteins of known Stokes’ radius (Rs) [31].

Small angle X-ray scattering (SAXS) and ab initio modeling

SAXS data were collected at the European Synchrotron Radiation Facility (E.S.R.F., Grenoble, France) on beamline ID14-3. The sample-to-detector distance was 1 m and the wavelength of the X-rays was 0.931 Å. Samples were contained...
in a 1.9 mm wide quartz capillary. The exposition time was optimized for reducing radiation damage. Data acquisition was performed at 20°C. Data reduction was performed using the established procedure available at ID14-3 and buffer background runs were subtracted from sample runs.

The SAXS profile of the N\textsubscript{21}P\textsubscript{60} complex was recorded for scattering vectors, $q = \frac{4\pi \sin \theta}{\lambda}$, in the range 0.05 nm\textsuperscript{-1} < $q$ < 3.5 nm\textsuperscript{-1}. The profiles obtained at three different protein concentrations 2.7–8.0 mg.mL\textsuperscript{-1} had the same shape and were flat at low $q$ values indicating the absence of significant aggregation (Figure S4A). The radius of gyration and forward intensity at zero angle, I\textsubscript{0}, were in agreement with SEC-MALLS analysis, and the molecular mass of 65 kDa derived from the scattering curve of the 15N-P\textsubscript{60} complex. Each protomer includes residues 22 to 422 of N and residues 6 to 33 of P\textsubscript{60}. In some protomers, residues 34 and 35 could also be constructed.

### NMR spectroscopy

NMR experiments were performed on a Varian spectrometer operating at a 1H frequency of 600 MHz. All samples contained 20 mM Tris-HCl, 150 mM NaCl, 50 mM Glu, 50 mM Arg with 10% D\textsubscript{2}O adjusted to pH 6.0. The concentration of free 15N, 13C, 2H-labeled P\textsubscript{60} was 0.9 mM and the concentration of the N\textsubscript{21}P\textsubscript{60} complex in 0.1 mM sodium acetate buffer, pH 6.0. The concentration of free 15N, 13C, 2H-labeled P\textsubscript{60} was 0.28 mM. In the titration experiment, 15N-labeled P\textsubscript{60} was initially at 0.28 mM. MBP- N\textsubscript{21}P\textsubscript{60} was added at final concentrations of 0.09 mM or 0.17 mM. 15N-labeled P\textsubscript{60} complex was crystallized in space group P2\textsubscript{1}2\textsubscript{1}2. The crystallographic asymmetric unit contained five protomers of N\textsubscript{21}P\textsubscript{60} complex. Each protomer includes residues 22 to 422 of N and residues 6 to 33 of P\textsubscript{60}. The N\textsubscript{21}P\textsubscript{60} complex was crystallized in space group P2\textsubscript{1}2\textsubscript{1}2. The crystallographic asymmetric unit contained five protomers of N\textsubscript{21}P\textsubscript{60} complex. Each protomer includes residues 22 to 422 of N and residues 6 to 33 of P\textsubscript{60}. In some protomers, residues 34 and 35 could also be constructed.

### Supporting Information

**Text S1** This file contains seven additional figures named S1 to S7 and an additional table S1.

**Acknowledgments**

We thank Ivan Ivanov and Danielle Blondel for helpful discussions. We thank the Partnership for Structural Biology for the excellent structural biology environment. Crystallographic coordinates and structure factors are deposited in the Protein Data Bank with accession code 3PMK.

**Author Contributions**

Conceived and designed the experiments: CL, FY, MRJ MB RWHR MJ. Performed the experiments: CL, FY, NT MRJ MB. Analyzed the data: CL, FY, NT, MRJ MB RWHR MJ. Contributed reagents/materials/analysis tools: CL, FY, EAR MRJ MB. Wrote the paper: CL, MRJ MB RWHR MJ.
References

1. Ge P, Tiao J, Schein S, Green TJ, Luo M, et al. (2010) Cryo-EM model of the bullet-shaped vesicular stomatitis virus. Science 327: 689–693.
2. Arfein H, Davis NL, Wertz G, Schlueter M, Lazarzari RA (1985) Role of the nucleocapsid protein in regulating vesicular stomatitis virus RNA synthesis. Cell 41: 259–267.
3. Patton JT, Davis NL, Wertz GW (1984) N protein alone satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. J Virol 49: 309–309.
4. Peluso RW, Moyer SA (1988) Viral proteins required for the in vitro replication of vesicular stomatitis virus defective interfering particle genome RNA. Virology 162: 369–376.
5. Howard M, Wertz G (1989) Vesicular stomatitis virus RNA replication: a role for the NS protein. J Gen Virol 70(Pt 10): 2681–2694.
6. Masters PS, Banerjee AK (1986) Complex formation with vesicular stomatitis virus phosphoprotein NS prevents binding of nucleocapsid protein N to nonspecific RNA. J Virol 62: 2636–2638.
7. Chen M, Oginoh T, Banerjee AK (2007) Interaction of vesicular stomatitis virus P and N proteins: Identification of two overlapping domains at the N-terminus of P that are involved in N0-P complex formation and encapsidation of viral genome RNA. J Virol 81: 13478–13483.
8. Mavrikas M, Mehouas S, Real E, Iieni F, Blondel D, et al. (2006) Rabies virus chaperone: identification of the phosphoprotein peptide that keeps nucleoprotein soluble and free from non-specific RNA. Virology 349: 422–429.
9. Curran J, Marq JB, Kolokolov D (1995) An N-terminal domain of the Sendai paramyovirus P protein acts as a chaperone for the NP protein during the nascent chain assembly step of genome replication. J Virol 69: 849–855.
10. Zhang X, Green TJ, Tsao J, Qiu S, Luo M (2008) Role of intermolecular interactions of vesicular stomatitis virus nucleoprotein in RNA encapsidation. J Virol 82: 674–682.
11. Gérard FCA, Ribeiro E, Albertini A, Zaccar G, Ebel C, et al. (2007) Unphosphorylated Rhabdoviridae phosphoproteins form elongated dimers in solution. Biochemistry 46: 10329–10336.
12. Gérard FCA, Ribeiro EA, Leyrat C, Ivanov I, Blondel D, et al. (2009) Modular organization of rabies virus phosphoprotein. J Mol Biol 389: 979–996.
13. Leyrat C, Jensen MR, Ribeiro EA, Gérard F, Raugot R, et al. (2011) The N-binding region of the vesicular stomatitis virus phosphoprotein is globally disordered but contains transient α-helices. Prot Sci 20: 542–556.
14. Albertini AA, Wernmont AK, Muziol T, Raveli RB, Clapier CR, et al. (2006) Crystal structure of the rabies virus nucleoprotein-RNA complex. Science 313: 360–363.
15. Green TJ, Zhang X, Wertz GW, Luo M (2006) Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. Science 313: 357–360.
16. Ribeiro EA, Leyrat C, Gérard FC, Albertini AA, Fall C, et al. (2009) Binding of rabies virus polymerase cofactor to recombinant circular nucleoprotein-RNA complexes. J Mol Biol 394: 556–575.
17. Mavrikas M, Iieni F, Mazza C, Schoenin G, Ebel C, et al. (2003) Isolation and characterization of the rabies virus N-degP complex produced in insect cells. Virology 305: 406–414.
18. Ding H, Green TJ, Luo M, Cell 71: 4264–4271.
19. Castel G, Clotet M, Cai GJ, Pernaud C, Mehouas S, et al. (2009) Peptides that mimic the amino-terminal end of the rabies virus phosphoprotein have antiviral activity. J Virol 83: 10808–10820.
20. Ding H, Green TJ, Luo M (2004) Crystallization and preliminary X-ray analysis of a proteinase-Resistant domain within the phosphoprotein of vesicular stomatitis virus (Indiana). Acta Crystallogr D Biol Crystallogr 60: 2087–2090.
21. Ribeiro EA, Jr., Favier A, Gerard FC, Leyrat C, Brutsch B, et al. (2008) Solution structure of the C-terminal nucleoprotein-RNA binding domain of the vesicular stomatitis virus phosphoprotein. J Mol Biol 382: 525–538.
22. Green TJ, Macpherson S, Qiu S, Lebowitz J, Wertz GW, et al. (2000) Study of the assembly of vesicular stomatitis virus N protein: role of the P protein. J Virol 74: 9515–9524.
23. Green TJ, Rossow M, Tsao J, Kang J, Ge P, et al. (2010) Access of RNA encapsidated in the nucleocapsid of vesicular stomatitis virus. J Virol 85: 2714–2722.
24. Emerson SU, Schubert M (1987) Location of the binding domain for the RNA polymerase L and the ribonucleocapsid template within different halves of the NS phosphoprotein of vesicular stomatitis virus. Proc Natl Acad Sci U S A 84: 5653–5659.
25. Abraham G, Banerjee AK (1976) Sequential transcription of the genes of vesicular stomatitis virus. Proc Natl Acad Sci U S A 73: 1504–1508.
26. Emerson SU (1982) Reconstitution studies detect a single polymerase entry site on the vesicular stomatitis virus genome. Cell 31: 635–642.
27. Uversky VN (1995) Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. Biochemistry 32: 12928–12930.
28. Konarre PV, Volkov VV, Sololova A, M.H.J. K, Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J Appl Cryst 36: 1277–1282.
29. Semenyuk AV, Svergun D (1991) GNOM - a program package for small-angle scattering data processing. J Appl Cryst 24: 537–540.
30. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. J Appl Cryst 36: 860–864.
31. Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. J Appl Cryst 34: 33–41.
32. Powell HR (1999) The Rossmann Fourier autoindexing algorithm in MOSFLM. Acta Crystallogr D Biol Crystallogr 55: 1690–1695.
33. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. Acta Crystallogr D Biol Crystallogr 59: 1131–1137.
34. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. J Appl Cryst 40: 658–674.
35. Cowtan KD (2006) The Buccaneer software for automated model building. Acta Crystallogr D Biol Crystallogr 62: 1002–1011.
36. Emuley P, Lokhand B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66: 496–501.
37. Marshakov GN, Vazgin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240–253.
38. Lasowski RA, Macarthur MW, Moss DS, Thornton JM. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26: 283–291.
39. Farrow NA, Muhundiram R, Snger AU, Pascal SM, Kay CM, et al. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry 33: 5984–6003.