



**HAL**  
open science

## **In vitro study of RNA encapsidation by the nucleoprotein of human Respiratory Syncytial Virus**

Lorène Gonnin, Charles-Adrien Richard, Irina Gutsche, Didier Chevret, Joris Troussier, Jean- Jacques Vasseur, Françoise Debart, Jean-François Eléouët, Marie Galloux

► **To cite this version:**

Lorène Gonnin, Charles-Adrien Richard, Irina Gutsche, Didier Chevret, Joris Troussier, et al.. In vitro study of RNA encapsidation by the nucleoprotein of human Respiratory Syncytial Virus. 2022. hal-03797501v1

**HAL Id: hal-03797501**

**<https://hal.science/hal-03797501v1>**

Preprint submitted on 20 Jul 2022 (v1), last revised 4 Oct 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# In vitro study of RNA encapsidation by the nucleoprotein of human Respiratory Syncytial Virus

**Running title: RNA encapsidation by the RSV nucleoprotein**

Lorène Gonnin<sup>1</sup>, Charles-Adrien Richard<sup>1</sup>, Irina Gutsche<sup>2</sup>, Didier Chevret<sup>1</sup>, Joris Troussier<sup>3</sup>, Jean-Jacques Vasseur<sup>3</sup>, Françoise Debart<sup>3</sup>, Jean-François Eléouët<sup>1#</sup>, Marie Galloux<sup>1#</sup>.

<sup>1</sup>Université Paris-Saclay, INRAE, UVSQ, VIM, Jouy-en-Josas, France.

<sup>2</sup>University of Grenoble Alpes, CEA, CNRS, IBS, Grenoble, France.

<sup>3</sup>IBMM, Université de Montpellier, ENSCM, CNRS, UMR 5247, Montpellier, France.

#Correspondance: jean-francois.eleouet@inrae.fr; marie.galloux@inrae.fr

**Keywords** : RSV, nucleoprotein, encapsidation, modified RNAs, liquid-liquid phase, viral factories

## ABSTRACT

Respiratory syncytial virus (RSV) has a negative-sense single-stranded RNA genome constitutively encapsidated by the viral nucleoprotein N, forming a helical nucleocapsid which is the template for viral transcription and replication by the viral polymerase L. Recruitment of L onto the nucleocapsid depends on the viral phosphoprotein P, which is an essential L cofactor. A prerequisite for genome and antigenome encapsidation is the presence of the monomeric, RNA-free, neosynthesised N protein, named N<sup>0</sup>. Stabilisation of N<sup>0</sup> depends on the binding of the N-terminal residues of P to its surface, that prevents N oligomerisation. However, the mechanism involved in the transition from N<sup>0</sup>-P to nucleocapsid assembly, and thus in the specificity of viral genome encapsidation, is still unknown. Furthermore, although the interaction between P and N complexed to RNA has been shown to be responsible for the morphogenesis of viral factories, where viral transcription and replication occur, the specific role of N oligomerisation and RNA in this process has not been elucidated. In the present study, using a chimeric protein between N and the first 40 N-terminal residues of P, we succeeded in purifying a recombinant N<sup>0</sup>-like protein competent for RNA encapsidation *in vitro*. Our results showed the importance of RNA length for stable encapsidation and revealed differences in encapsidation depending on the nature of the 5' end, without any specificity for RNA sequence. Finally, we showed that RNA encapsidation is crucial for the *in vitro* reconstitution of pseudo-viral factories.

## INTRODUCTION

The *Mononegavirales* order includes many human pathogenic viruses such as those responsible for rabies, measles, mumps, haemorrhagic fevers due to Ebola or Nipah viruses, and respiratory diseases induced by human respiratory syncytial virus (HRSV), metapneumovirus (HMPV), or parainfluenza viruses (PIV) (1). All these viruses have a negative-sense single-stranded RNA genome that is constitutively encapsidated by the nucleoprotein N, forming the nucleocapsid (NC). This NC serves as a template for the viral polymerase L responsible for both viral transcription and replication (2, 3). For the vast majority of these viruses, transcription and replication are cytoplasmic, occurring in viral factories which are viro-induced organelles called inclusion bodies (IBs) (4-8).

Among *Mononegavirales*, HRSV is the prototype of the *Pneumoviridae* family and the *Orthopneumovirus* genus (9, 10). HRSV is the most common cause of acute lower respiratory infections in young children worldwide (11-13). It is also recognised as a major cause of severe respiratory

infections in immunocompromised and elderly people (14-16). Currently, no vaccine is available, and the only specific treatment against HRSV is prophylactic, consisting of injection of humanised monoclonal antibodies directed against the fusion protein F (Palivizumab, Sinagis) (17). However, the efficacy of this treatment is controversial and its high cost limit its prescription to at-risk children. The HRSV genome is approximately 15.2 kb long, and contains 10 genes encoding 11 proteins (2). Replication and transcription rely on four of these proteins: the RNA-dependent RNA polymerase L which exhibits all the enzymatic activities required for viral replication and transcription, its cofactor the phosphoprotein P responsible for recruitment of L on the NC, the nucleoprotein N involved in genome and antigenome encapsidation, and the transcription factor M2-1 that has been described as an “antiterminating” factor during transcription and interacts with P and viral mRNAs.

During transcription, the polymerase L recognises the *leader* sequence at the 3' end of the genome, and the polymerase is guided by the *gene start* (*gs*) and *gene end* (*ge*) sequences flanking each gene in the viral genome. The neosynthesised viral mRNAs are capped and methylated at the 5' end and polyadenylated at the 3' end by the polymerase. It is noteworthy that the synthesis of all viral mRNAs depends on the presence of the M2-1 protein, which is required for L processivity along the genome (18, 19). The mechanism of action of M2-1 is still poorly understood. However, it has recently been shown that RSV mRNAs concentrate with M2-1 in specific sub compartments of IBs called IBAGs (for IBs associated granules) (8), suggesting that the M2-1-mRNA interaction may promote the release of mRNAs from the polymerase L.

On the contrary, replication of the viral genome only depends on L, P, and the NC. In that case, the polymerase recognises the *leader* sequence at the 3' end of the genome, the RNA synthesis starts at the first nucleotide and the polymerase proceeds through the *ge* and *gs* sequences to the end of the genome (or antigenome) (20, 21). Genomes and antigenomes, which have a triphosphate 5' end (22, 23), are directly encapsidated by the N protein during their synthesis, forming helical NC (24). The structure of N expressed in *E.coli* and purified as N-RNA rings composed of 10 protomers revealed that each N protomer interacts with 7 nucleotides (25). The RNA binds within a groove formed by the interface between the two globular domains of N ( $N_{\text{NTD}}$  and  $N_{\text{CTD}}$ , respectively) that are separated by a flexible hinge region. The N protein also possesses two N- and C-terminal extensions (N- and C-arms) that are involved in N oligomerisation: the N-arm of the  $N_i$  protomer binds to the  $N_{i-1}$  protomer, whereas the C-arm of  $N_i$  binds to the top of the  $N_{\text{CTD}}$  of the  $N_{i+1}$  protomer (24, 25). RNA encapsidation therefore

depends on both direct interaction with RNA and the ability of N to oligomerise, two mechanisms that are closely coupled. When expressed alone, the *Mononegavirales* N proteins all show a strong tendency to encapsidate cellular RNAs. This implies that viral replication depends on the ability to maintain a pool of RNA-free monomeric N, termed N<sup>0</sup>, available for the specific encapsidation of genomic and antigenomic RNAs. It is now well established that all these viruses share a common mechanism, with the neosynthesised N being maintained in the N<sup>0</sup> form by P which acts as a molecular chaperone (26-34). Specifically for HRSV, and by analogy to the N<sup>0</sup>-P complex of the HMPV whose crystal structure has been solved (29), the binding of the 28 N-terminal residues of P to the N<sub>CTD</sub> surface overlaps with the binding sites of both the N- and C-arms of the N<sub>i+1</sub> and N<sub>i-1</sub> protomers in the oligomeric form, preventing self-oligomerisation of N (28, 35). Notably, N<sup>0</sup> is characterised by a rotation of N<sub>NTD</sub> relative to N<sub>CTD</sub> compared to the oligomeric form and a stacking of the C-terminal arm of N into the positively charged RNA groove, that blocks RNA binding (28). However, the mechanism involved in the switch from N<sup>0</sup> to N-RNA, which regulates specific encapsidation of genomes and antigenomes, remains poorly characterised.

In addition to their critical role in viral polymerase function, N and P have been shown to be the scaffold proteins responsible for IBs' morphogenesis (36, 37). It has recently been shown that IBs are liquid organelles formed by liquid-liquid phase separation induced by N-P interaction (37, 38). Importantly, the P protein (241 residues long), which forms tetramers through its central oligomerisation domain (residues 131-151) flanked by highly disordered N and C-terminal domains (P<sub>NTD</sub> and P<sub>CTD</sub> respectively) that are involved in multiple protein-protein interactions, plays a central role in regulating polymerase activity as well as in IBs' formation. The P<sub>NTD</sub> interacts with N<sup>0</sup> (35), M2-1 (39), the cellular phosphatase PP1 (39) and the viral matrix protein M responsible for virion assembly (40), whereas P<sub>CTD</sub> is involved in the interaction with L (41, 42) and NC (43, 44), and is critical for IBs' formation (37). Both the length of the P<sub>CTD</sub> and its interaction with N complexed to RNA are required for efficient IBs' morphogenesis. However, the importance of N oligomerisation and of RNAs in the morphogenesis of IBs remains to be characterised.

In this study we designed a fusion protein between the full-length N and the peptide derived from the first 40 N-terminal residues of P (P40) responsible for chaperone activity on N<sup>0</sup>. This strategy allowed the purification of a recombinant monomeric and RNA-free N-P40 chimeric protein. The resulting N-P40 was competent for *in vitro* RNAs encapsidation, the length of RNAs being critical for stable

oligomerisation, and invariably leading to the formation of N-RNA rings under the conditions tested. Furthermore, using this N-P40 protein, we revealed that RNA encapsidation is critical for the morphogenesis of pseudo-IBs *in vitro*.

## RESULTS

### Production of a N<sup>0</sup>-like recombinant protein

Recently, a strategy based on co-expression of P<sub>NTD</sub> with His-tagged N in *E. coli* allowed the purification of a N<sup>0</sup>-P<sub>NTD</sub> complex competent for encapsidation (45). In the present study, we used a strategy based on the expression of a fusion protein between His-tagged full-length N and the 40 N-terminal residues of P (P40) in *E. coli* (Fig. 1A). This approach was previously used to determine the crystal structure of the HMPV N<sup>0</sup>-P complex (29). After purification of this chimeric protein in the presence of 1 M NaCl, the sample was analysed by size exclusion chromatography, following the optical densities (OD) at 260 and 280 nm, and compared to the profile of purified wild-type N (N<sub>wt</sub>) forming N-RNA rings (44). The elution profile of N-P40 showed two peaks. The major P1 peak with OD<sub>260nm</sub>/OD<sub>280nm</sub> ratio > 1 and apparent mass of ~ 500 kDa (estimated from the Superdex 200 calibration profile), of slightly smaller size than N-RNA rings, that could correspond to N-RNA oligomers or aggregates. A minor P2 peak presented OD<sub>260nm</sub>/OD<sub>280nm</sub> ratio < 1 and apparent mass of ~ 50 kDa, as expected for monomeric RNA-free N-P40 protein (Fig. 1B). The fractions of the P2 peak were pooled, and sample analysis by SDS-PAGE stained with Coomassie blue showed the presence of a single band with a molecular weight close to 50 kDa, consistent with the expected mass of the N-P40 protein (Fig. 1B, inset). The sample was then dialysed to 300 mM NaCl, and concentrated up to 1.5 mg/ml. Dynamic light scattering (DLS) analysis of the sample revealed the presence of a major protein peak near 6.7 nm in diameter that should correspond to monomers, and a minor peak of oligomers near 17 nm in diameter (Fig 1C). These data revealed that in our conditions N-P40 tended to aggregate or oligomerise upon concentration, and led us to work with a protein concentration of 1 mg/ml. We finally analysed the secondary and tertiary structures of N-P40 by circular dichroism (CD), using as previously N<sub>wt</sub> purified as N-RNA rings as a control. The far-UV spectra of both N-P40 and N<sub>wt</sub> showed peaks at 208 and 222 nm, typical of secondary structures with predominant  $\alpha$ -helical content (Fig 1D). The peak intensity shift observed for N-P40 revealed a lower

helical content compared to N<sub>wt</sub>. This observation correlates with the insertion of linkers in the N-P40 constructs, which are expected to be unfolded. The near-UV CD allows the detection of signals from aromatic residues engaged in a rigid chiral environment, consistent with the presence of a tertiary structure. The detection of a broad signal around 280 nm for the N-P40 protein can be attributed to the signal from Tyr and Trp residues, and validates that the protein has a tertiary structure (Fig 1E). In comparison, the spectrum obtained for N<sub>wt</sub> revealed an intense positive signal between 260 and 280 nm, which can be attributed to the presence of RNA masking the signals from the aromatic residues of N.

Overall, these results show that N-P40 can be purified as a monomeric and RNA-free protein and that the fusion of P40 does not have a major impact on N folding. This latter observation correlates with previous results showing that binding of P40 to N does not induce major conformational changes in N (28).

### **N-P40 is competent for encapsidation of RNAs**

We then assessed the capacity of the N-P40 protein to encapsidate RNA *in vitro*. As each N<sub>wt</sub> protomer has been shown to interact with 7 nucleotides within purified N-rings formed of 10 or 11 protomers (25), we first tested the ability of N-P40 to encapsidate a 70-nucleotide long (70-mer) RNA or DNA of similar sequence. The RNA sequence that we used corresponded to the trailer sequence of the RSV genome (Table 1). Encapsidation of oligonucleotides was analysed using a band shift assay on native acrylamide and agarose gels, using N<sub>wt</sub> as a control. As shown in Figure 2A, the N<sub>wt</sub> complexed to RNA, and thus negatively charged, clearly migrated within the gel, while only slight migration was observed for purified N-P40. A clear band shift was observed for N-P40 incubated in the presence of RNA, whereas the protein seemed to poorly interact with DNA. In parallel, the samples were observed by negative stain electron microscopy (ns-EM). As expected, images of N-P40 alone showed only very small species (Fig. 2C). While N-P40 incubated with 70-mer RNA formed ring-shaped N-RNA assemblies, only aggregates were observed for the sample of N-P40 incubated in the presence of DNA (Fig.2C). Using the same approach, we confirmed that N-P40 was also capable of encapsidating shorter 14-mer RNAs and forming N-RNA rings (Fig 2B and C). Furthermore, we observed that N-P40 similarly encapsidated 14-mer RNAs of different sequences (Fig 2B and Table 1), suggesting the lack of sequence specificity for encapsidation.

Our results show that, although fused to N, the P40 peptide is not sufficient to prevent RNA encapsidation *in vitro*. Furthermore, while N-P40 specifically encapsidated RNA over DNA, no specificity in RNA sequence was observed under the conditions tested. It is noteworthy that, as previously observed by Gao *et al.* (45), *in vitro* encapsidation did not result in the formation of helical nucleocapsids but only in N-RNA rings.

### Investigation of the minimal RNA length required for encapsidation

Using the sequence derived from the HRSV trailer sequence, we further studied the minimal RNA length required for encapsidation, using 5- to 11-mer oligonucleotides (Table 1). As shown in Figure 3A, no band shift of N-P40 was detected on native acrylamide and agarose gels in the presence of 5- and 6-mer RNAs. Similar N-P40 band shifts were observed in the presence of 10- and 11-mer oligonucleotides, compared with N-P40 incubated with 14-mer RNAs. Finally, intermediate band shift profiles were observed for N-P40 incubated in the presence of 7-, 8-, and 9-mer RNAs. Yet, according to ns-EM observations, incubation of N-P40 with 7-mer and 11-mer RNAs both led to the formation of N-RNA rings, similar to the ones observed with the 14-mer and 70-mer RNAs (Fig 3B). To better understand these results, we next investigated the stability of the N-P40-RNA complexes by incubating the samples overnight in the presence of RNase A followed by dialysis. Migration analysis of the samples on a native agarose gel showed no band shift for N-P40 incubated with 6- to 10-mer RNAs, suggesting RNA digestion, whereas a band shift was still observed for N-P40 in the presence of 11-mer and 14-mer RNAs (Fig 3C). Further analysis of the samples was performed by measuring the  $OD_{260}/OD_{280}$  ratio, and by investigating structural changes of N-P40 by DLS and near-UV CD. Before RNase A treatment, all samples displayed an  $OD_{260}/OD_{280}$  ratio close to 1.2, similar to the value measured for  $N_{wt}$  (Table 2). After RNase A treatment,  $OD_{260}/OD_{280}$  ratios were  $< 1$  for N-P40 samples incubated with 6- to 10-mer RNAs (Table 2), validating RNA digestion upon RNase treatment. Of note, the  $OD_{260}/OD_{280}$  ratio of these samples remained higher than that obtained for purified N-P40. For N-P40 incubated in the presence of 11- and 14-mer RNAs, the  $OD_{260}/OD_{280}$  ratios were close to 1. These data suggested that although the RNAs were partially digested by RNase A, N-P40 was still associated with RNAs, preventing complete RNA degradation. In parallel, determination of the hydrodynamic diameter ( $D_h$ ) of the samples by DLS revealed that N-P40 incubated in the presence of 6- to 10-mer RNAs displayed a  $D_h$  between



6.7 and 9 nm after RNase treatment, close to the Dh obtained under the same conditions for N-P40 alone, whereas those of N-P40 incubated with 11- and 14-mer RNAs were 14.2 and 14.4 nm, respectively, similar to the Dh of N<sub>wt</sub> (Table 2). Furthermore, the near-UV CD spectra showed that only the N-P40 sample incubated with 11-mer RNA exhibited a positive peak around 260 nm, whereas the spectra with 7- to 10-mer RNAs were similar to purified N-P40 alone or displayed only a weak signal around 260 nm (Fig. 3D). These data correlate with OD and Dh measurements, suggesting that RNase treatment induced degradation of RNAs shorter than 11-nucleotides long, and N rings disassembly. Altogether, these results reveal that stable RNAs encapsidation by N-P40 depends on RNA length, with the minimal length to stabilise N-P40-RNA complexes being 11-mer RNAs.

### Impact of 5' end modification of RNAs on encapsidation

One of the differences between RNAs lies in post-transcriptional modifications, notably of the 5' end. Genomic and antigenomic RNAs have been described as having a 5' triphosphate (PPP) end (22, 23), whereas viral mRNAs are capped by L during their synthesis (3). Therefore, we investigated whether post-transcriptional modifications of the 5' of RNAs could impact the ability of N to encapsidate RNAs. As previously, N-P40 was incubated in the presence of synthetic 7- to 11-mer RNAs with either a 5'-PPP or a N7 and 2'-O methylated cap structure (<sup>m</sup>GpppG<sub>m</sub>) (Table 1), followed by native agarose gel electrophoresis analysis. Our data showed that the presence of PPP or <sup>m</sup>GpppG<sub>m</sub> at the 5' end inhibited the ability of N to encapsidate 7- and 8-mer RNAs, compared to 5' OH RNAs (Fig. 3A and 4A). However, N-P40 was still able to encapsidate 9-, 10- and 11-mer RNAs, and no difference was observed between 5' end PPP or capped RNAs. Interestingly, independently of the length and the capping of the RNAs, N-P40-RNA rings could always be detected by ns-EM (Fig 4B), suggesting that this method of observation may stabilize ring-shaped assemblies and highlights the importance of using multiple complementary techniques for the analysis of RNA encapsidation.

As before, we further investigated the impact of RNAs 5' end modification on the stability of N-P40-RNA complexes upon RNase A treatment. Surprisingly, although band shifts were still observed (Fig. 4C), N-P40 samples incubated with 5' PPP and 5' <sup>m</sup>GpppG<sub>m</sub> RNAs displayed OD<sub>260</sub>/OD<sub>280</sub> ratios of 0.9 and 1, respectively, and DLS analysis revealed the presence of different populations (not shown). Furthermore, while the near-UV CD spectrum of N-P40 incubated with 5' <sup>m</sup>GpppG<sub>m</sub> RNAs was similar to the spectrum

of N-P40 complexed with 5'OH RNA, the spectrum of N-P40 incubated with 5' PPP RNAs clearly differed from those of N-P40 alone or incubated with 5' OH RNAs (Fig. 4D).

These results revealed that modification of the 5' end of RNAs only partially impairs encapsidation but is not a critical factor for the specificity of encapsidation, although 5' end modifications delay RNAs encapsidation maybe by creating steric hindrance.

### **Role of RNA encapsidation on pseudo-IBs morphogenesis *in vitro***

We have previously shown that *in vitro* co-incubation of purified recombinant fluorescent N-RNA rings (mCherry-N<sub>wt</sub>) with P can lead to the formation of liquid droplets that exhibit properties similar to cellular IBs formed during infection. We also identified the key properties of P required for pseudo-IBs morphogenesis, i.e. oligomerisation, P<sub>CTD</sub> length and flexibility, and interaction of the C-terminus of P with N (37). However, this study did not decipher the role of N oligomerisation and of the presence of RNAs in this process. Furthermore, although our data suggested that the N/P ratio may be important for optimal IBs formation, these results remained preliminary, lacking a systematic study of the impact of N/P ratio and concentration on the morphogenesis of IBs. As phase separation is known to be facilitated by increased protein concentration (46-50), we here decided to limit the protein concentrations to avoid spontaneous pseudo-IBs formation. Using a fixed minimum concentration of 3.5  $\mu$ M for mCherry-N<sub>wt</sub> or P, the addition of the second partner at a 1:1 ratio allowed us to observe the formation of small droplets of 1-5  $\mu$ m in diameter (Fig. 5). A fixed low P concentration limited the size and number of droplets that remained similar whatever the concentration of mCherry-N (Fig.5A). On the other hand, in a fixed low N concentration condition, the number and size of droplets increased proportionally to the addition of P until reaching a plateau for a P/N ratio  $\geq$  4 (Fig.5B). These results correlate with our previous observations (37), and revealed that P concentration is a limiting factor for IBs formation.

To study the impact of N-P40 oligomerisation in the presence of RNAs on pseudo-IBs morphogenesis, we then investigated whether pseudo-IBs could form when N-P40 was incubated with a 70-mer RNA in the presence of P-BFP. Since no droplets were observed with a 3  $\mu$ M concentration of N-P40 in the presence of an excess of P-BFP (15  $\mu$ M) and 70-mer RNA at 50  $\mu$ M, we then screened for the optimal conditions allowing pseudo-IBs formation, with a minimum concentration of proteins. Pseudo-IBs of 10-15  $\mu$ m in diameter were observed in the presence of 3.5  $\mu$ M of P-BFP with a 4-fold excess of N-P40

(13.5  $\mu\text{M}$ ) and 70-mer RNAs (50 $\mu\text{M}$ ). Of note, pseudo-IBs of only 2-3  $\mu\text{m}$  in diameter were observed when co-incubating P-BFP with N<sub>wt</sub> (RNA-N rings) in the same conditions (Fig. 6A). It is noteworthy that no droplets were detected when incubating P-BFP in the presence of either RNAs or N-P40 alone (Fig. 6A). We then assessed the impact of RNA length on pseudo-IBs morphogenesis by incubating N-P40 with RNAs of various lengths before the addition of P-BFP, under the same conditions. As shown in figure 6B, the addition of 14-mer RNA resulted in the formation of droplets similar to those observed in the presence of 70-mer RNA. More interestingly, droplet size decreased when N-P40 was incubated with 11-mer and 8-mer RNAs, and only aggregates were observed in the presence of 6-mer RNA. Furthermore, whereas RNase A treatment of N-P40 incubated with 11-mer RNA still allowed to observe droplets, only aggregates were detected for the sample with 8-mer RNA (Fig. 6C). This result correlates with data showing the instability of N oligomerisation in the presence of short RNAs. Finally, pseudo-IBs similar to those obtained in the presence of 11-mer RNA with 5' OH were observed in the presence of 11-mer RNAs with 5' PPP and <sup>m</sup>GpppG<sub>m</sub> (Fig. 6D).

Altogether, our data revealed that co-incubation of P with monomeric N does not allow pseudo-IBs morphogenesis *in vitro*, and that the addition of short RNAs is not sufficient to induce phase separation. Overall, these results showed that N oligomerization in the presence of RNA is critical for the morphogenesis of pseudo-IBs.

## DISCUSSION

Like for the majority of viruses belonging to the *Mononegavirales* order, HRSV replication and transcription take place within cytoplasmic viral factories called IBs that concentrate the viral and cellular proteins required for these activities, as well as neosynthesised antigenomic and genomic RNAs and viral mRNAs. After initial transcription of the viral genome, the encapsidation of antigenomic and genomic RNAs depends on the presence of neosynthesised N<sup>0</sup> protein. However, given the strong tendency of N to oligomerise on RNAs, some particular mechanisms might regulate the specificity of viral RNA encapsidation, and thus the transition from N<sup>0</sup> to N-RNA. Here, using a chimeric protein composed of full-length N and P40, a peptide corresponding to the 40 N-terminal residues of P, we showed that this construct is competent for RNA encapsidation *in vitro*, but without sequence specificity. This first observation showed that the presence of P40 is not sufficient to prevent RNA encapsidation

by N, suggesting the existence of another factor that could regulate the transition from monomeric N<sup>0</sup> to N-RNA assembly in infected cells, most likely by enhancing the stability of N<sup>0</sup>. In agreement with a recent publication by Gao *et al.* (45), our results also confirmed that RNA as short as 7-mer can be encapsidated, and that N oligomerisation on RNAs always leads to the formation of N-RNA rings *in vitro*. We showed that the minimum length for stable encapsidation was 11 nucleotides, revealing that RNA length is critical for the stability of N-RNA oligomers. However, we cannot exclude that the presence of P40 fused to the C-terminus of N may partially alter N oligomerisation. In contrast to measles virus for which helical NC can form *in vitro* using recombinant N protein and 6-mer RNAs (4), our data suggest that additional factors such as viral or cellular partners or post-translational modifications may be required for proper HRSV N oligomerisation along RNA to obtain helical NCs.

Currently, the specificity of encapsidation of viral genomic and antigenomic RNAs by N is unknown. The difference between the 5' ends of mRNAs (capped) and genomic and antigenomic RNAs (PPP) could explain this mechanism. However, we found that the presence of either PPP or <sup>m</sup>GpppG<sub>m</sub> at the 5' end instead of OH only partially impedes the ability of N to encapsidate RNAs, most probably by creating steric hindrance. In light of the recent discovery of subcompartments called IBAGs within IBs, that specifically concentrate M2-1 and viral mRNAs (8), we hypothesise that the exclusion of viral mRNAs complexed to M2-1 from the polymerase complex could contribute, at least in part, to the segregation of mRNAs and thereby specificity of genome encapsidation by N. Our study highlighted the absence of encapsidation of short 5' PPP RNAs. This observation suggests that the 5' PPP end of genomes and antigenomes is not encapsidated and thus should be exposed and detected by sensors of the cellular innate immune system such as RIG-I. Recently, cryo-electron microscopy studies of Newcastle disease virus, Sendaï, Nipah, and Cetacean morbillivirus have revealed heterogeneity in NC assembly, with the presence of double-headed helical structures (51-54). Such an assembly could hide the 5' end of antigenomes and genomes, thus preventing 5' PPP detection by cellular sensors. However, further investigation will be needed to decipher the accessibility and protection of the 5' end of viral genomes. Finally, we investigated the importance of P, N oligomerisation, and RNAs in the morphogenesis of pseudo-IBs *in vitro*. The tetrameric RSV P protein has intrinsically disordered regions involved in multiple transient interactions known to be critical for phase separation, and is a major scaffolding component of IBs (37). When P was mixed with a monomeric N<sup>0</sup>-like protein (N-P40) and RNAs too short to be encapsidated, no phase separation was observed, indicating that the presence of RNA-N oligomers is

critical for pseudo-IBs formation. Interestingly, the length of encapsidated RNAs seems to modulate the size of pseudo-IBs. These observations correlate with the fact that RNAs have been shown to play a major role in liquid-liquid phase separation mechanisms (48) (47). Further characterisation of IBs morphogenesis in the presence of RSV helical NC should be of great interest, as the size and high flexibility of these structures could modulate IBs morphogenesis. Of note, a recent cryo-electron tomography study of RSV infected cells has revealed the presence of N rings in filamentous RSV particles (55), showing for the first time the heterogeneity of N assembly in the cell. The role of these structures during viral cycle i.e. IBs morphogenesis, control of the innate immune response to infection, and virion assembly remains to be determined.

## **MATERIALS AND METHODS**

**Plasmid constructs.** The gene sequence of the nucleoprotein N was amplified without stop codon by PCR using Pfu DNA polymerase (Stratagene, Les Ulis, France) and cloned in the pET28a+ vector, at BamHI and XhoI sites. The oligonucleotides used were: 5'-GCCGCCGATCCATGGCTCTTAGAAAG TCAAGTTG-3' (BamHI) and 5'-GAGGAGCTCGAGAAG CTCTACATCATTATCTTTTGG-3' (XhoI). The sequence coding TEV cleavage site was then introduced at the BamHI site after annealing of the oligonucleotides 5'-GATCCGAGAACCTATATTTCCAGG-3' and 5'-GATCCCTGGAAATATAGGTTCT CG-3'. Finally, the sequence coding for the 40 N-terminal residues of P was PCR amplified using the oligonucleotides 5'-GAGGAGCTCGAGGGTAGCGGTAGCGGTAGCGGTAGCATGGAAAAGTTTGC TCCTGAATTCC-3' (containing the sequence coding for 4 Gly/Ser, underlined) and 5'-GAG GAGCTCGAGTTAGATACTATCTTTTTTCTTCCCATC-3' (with a stop codon at the 3' end, underlined), and inserted at the XhoI site. The viral sequences derived from the human RSV strain ATCC VR-26 (GenBank accession number: AY911262.1). The final plasmid allows the expression of the fusion protein 6xHis-N-4xgs-P40 containing a TEV cleavage site upstream of the N sequence. For expression and purification of recombinant mCherry-N, N, and P-BFP proteins, the previously described pET-mCherry-N, pET-N and pET-P-BFP plasmids were used (35, 37).

**Expression and purification of the recombinant proteins.** The *Escherichia coli* BL21 (DE3) bacteria strain (Novagen, Madison, WI) were transformed by the pET-N-P40 plasmid. Cultures were grown at 37 °C in 2xYT medium containing 100 µg/ml kanamycin. After 8 h, an equal volume of 2xYT medium containing 100 µg/ml kanamycin was added to the cultures, and protein expression was induced by the addition of 80 µg/ml isopropyl β-d-1-thiogalactopyranoside (IPTG) overnight at 28 °C. Bacteria were then harvested by centrifugation. Pellets were resuspended in lysis buffer (20 mM Tris-HCl, pH 8, 500 mM NaCl, 0.1% Triton X-100, 10 mM imidazole, 4 mM benzamidine, 1 mg/ml lysozyme and complete protease inhibitor cocktail (Roche, Mannheim, Germany)). After incubation on ice for 30 min, the lysates were sonicated, benzonase (Novagen) (final concentration 5 U/mL) and RNase A (final concentration of 1 U/ml)) were added to the lysate, followed by incubation for 30 min at room temperature, finally NaCl was added (up to a concentration of 1 M). After centrifugation at 10,000 *g* for 30 min at 4 °C, the lysates were incubated with chelating Sepharose Fast Flow beads (GE Healthcare) charged with Ni<sup>2+</sup>, for 1 h at 4 °C. Beads were washed with washing buffers (20 mM Tris-HCl, pH 8, 1 M NaCl) with increasing imidazole concentrations (10, 50 and 100 mM), before elution of the protein using 600 mM imidazole. Purified proteins were then loaded on a Hi-Load 16/600 Superdex 200 column (GE Healthcare) and eluted in 20 mM Tris-HCl, pH 8.5, 1 M NaCl. Then, the purified protein was dialysed against 10 mM Tris-HCl, pH 8.5, with decreasing NaCl concentrations (500 and 300 mM) buffers, for 4 h at each step, at 4 °C. Finally, the purified protein was concentrated using a centrifugal concentrator with a MWCO of 10 kDa (vivaspin turbo 4, Sartorius). Wild type mCherry-N, N, and P-BFP proteins were expressed and purified as previously described (37).

**Synthesis of RNA substrates.** RNA sequences were chemically synthesized on a solid support using an ABI 394 automated synthesizer with 2'-O-pivaloyloxymethyl 3'-O-phosphoramidite ribonucleosides and with 2'-O-methyl 3'-O-phosphoramidite adenosine to obtain <sup>7m</sup>GpppA<sub>m</sub>-RNA (ChemGenes Corp., USA) (56). After RNA assembly, depending on the desired 5'-end of the RNA, three different processes were applied. For 5'-OH RNA sequences, the solid support was directly subjected to a basic treatment (28% aqueous ammonia solution at r.t for 3 h) to recover the crude RNA material. In the other two cases, the 5'-hydroxyl group was phosphorylated and the resulting *H*-phosphonate derivative was oxidized and activated to a phosphoroimidazolidate derivative to react with pyrophosphate or GDP, yielding pppRNA (57) or GpppRNA (58), respectively. After deprotection and release from the solid support by the same

ammonia treatment as for 5'-OH RNA, all RNA substrates were purified by IEX-HPLC (> 95% pure) and their identity was confirmed by MALDI-TOF spectrometry. *N7*-methylation of the cap structure (Gppp) was performed using human (guanine-*N7*)-MTase to obtain <sup>7m</sup>Gppp-RNA (58).

**RNA-N-P40 complexes formation.** RNA oligonucleotides (15 μM) and purified N-P40 protein (10 μM) in 20 mM Tris-HCl, pH 8.5, 300 mM NaCl buffer were co-incubated for 30 min at room temperature. The presence of RNA was assessed by measuring the OD<sub>260nm</sub>/OD<sub>280nm</sub> absorption ratio. For RNase treatment, samples were incubated overnight at 4°C in the presence of RNase A (PureLink™, Invitrogen), then dialysed against 10 mM NaP, 300 mM NaF, pH 8.5, for 3 h at 4 °C. After dialysis, the OD<sub>260nm</sub>/OD<sub>280nm</sub> absorption ratio was measured.

**Band shift on native acrylamide and agarose gels.** 50% sucrose loading buffer was added to the samples before loading either on native 4% acrylamide SDS-PAGE gel or on native 1% agarose gel. For acrylamide native gel analysis, migration was performed in 0.2 x TBE (pH 8.0 for 2 h at 200 V at 4 °C), and gels were stained with Coomassie blue. For agarose native gel analysis, migration was performed in 1X Tris-Glycine buffer during 1 h 30 at 80 V before staining with amido black 10B.

**Dynamic light scattering (DLS).** Size measurement of purified N-P40 alone or incubated with RNAs was performed at 20 °C using a helium-neon laser wavelength of 633 nm and detection angle of 173° with a Zetasizer Nano (Malvern). Ten measurements were made, with an acquisition time of 10 s for each measurement. Hydrodynamic diameters ( $D_H$ ) were calculated using the DLS software provided by the instrument manufacturer. The results were presented as size distribution (nm).

**Circular dichroism spectroscopy.** CD experiments were performed on a J-810 spectropolarimeter (Jasco, Tokyo, Japan) in a thermostated cell holder at 20 °C. Spectra of N-P40 and RNA-N-P40 complexes were measured after dialysis against 10 mM NaP, 300 mM NaF, pH 8.5 buffer, and at concentrations of 10 μM. Far-UV spectra (190 – 260 nm) were recorded in a 0.5 mm path-length quartz cell using a bandwidth of 2 nm and an integration time of 1 s. Near-UV spectra (260 – 320 nm) were recorded in a 10 mm path-length quartz cell using a bandwidth of 2 nm and an integration time of 1 s. Each spectrum was the average of 3 scans, with a scan rate of 100 nm/min and 50 nm/min for far-UV and near-UV spectra, respectively. Correction by subtracting the signal from the buffer was made and

the spectra were smoothed with the fast Fourier transform (FFT) filter (Jasco Software, Tokyo, Japan), data were treated as previously described (35).

**Negative stain electron microscopy observations of N-P40-RNA complexes.** Three microliters of sample were applied to the clean side of carbon on a carbon–mica interface and stained with 2% sodium silicotungstate. Micrographs were recorded on a Thermofisher Scientific Tecnai T12 microscope operated at 120 kV with a Gatan Orius 1000 camera. Images were recorded at a nominal magnification of 23 000 × resulting in a pixel size of 2.8 Å.

**In vitro assay of pseudo-IBs formation.** As previously described (37), mCherry-N and P proteins, or P-BFP and wild type N or N-P40 recombinant proteins alone or in the presence of RNAs in 20 mM Tris/HCl pH 8.5, 150 mM NaCl buffer were co-incubated at different P/N molecular ratio on glass slides, and the molecular-crowding agent Ficoll was added on the droplets of solution. Then coverslips were laid on the droplets. For RNase treatment, N-P40 recombinant protein was pre-incubated in presence of 8mer and 11-mer RNAs, then incubated half an hour at room temperature in the presence of RNase A (PureLink™, Invitrogen). Pseudo-IBs were then observed with a Nikon TE200 inverted microscope equipped with a Photometrics CoolSNAP ES2 camera. Images were processed using MetaVue software (Molecular Devices) and ImageJ software.

## ACKNOWLEDGMENTS

This work was carried out with the financial support of the French Agence Nationale de la Recherche, specific programs ANR DecRisP n° ANR-19-CE11-0017 and ANR RSVFact n° ANR-21-CE15-0030-02. Protein purification benefited from the purchase of a gel filtration system funded by the Région Ile de France (DIM- OneHealth 2018). For electron microscopy observations, this work used the EM platform of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). The electron microscope facility is supported by the Auvergne-Rhône-Alpes Region, the Fondation Recherche Médicale (FRM), the fonds FEDER, and the GIS-Infrastructures en Biologie Santé et Agronomie (IBISA).



We declare that we have no conflicts of interest with the contents of this article.

L.G., M.G. and J.-F.E. designed experiments. L.G. and D.C. performed molecular and protein expression. L.G. performed protein purification, *in vitro* encapsidation assay and biophysical characterisation. C.-A.R. performed gel filtration and *in vitro* phase separation assay. J. T., J.-J. V. and F. D. synthesised RNAs. I.G performed electron microscopy observations. M.G. wrote the paper with contributions from all authors, and J.-F.E. and M.G. edited the manuscript. All authors commented on the manuscript.

## REFERENCES

1. Rima B, Collins P, Easton A, Fouchier R, Kurath G, Lamb RA, Lee B, Maisner A, Rota P, Wang LF. (2018). Problems of classification in the family Paramyxoviridae. *Arch. Virol.* **163**, 1395-1404.
2. Collins PL, Crowe JE. (2007). Respiratory Syncytial Virus and Metapneumovirus. In *Fields Virology*, pp 1601-1646 Edited by D M Knipe & P M Howley Philadelphia: Lippincott Williams & Wilkins Fifth edition:1601-46.
3. Fearn R, Plemper RK. (2017). Polymerases of paramyxoviruses and pneumoviruses. *Virus Res.* **234**, 87-102.
4. Guseva S, Milles S, Jensen MR, Salvi N, Kleman JP, Maurin D, Ruigrok RWH, Blackledge M. (2020). Measles virus nucleo- and phosphoproteins form liquid-like phase-separated compartments that promote nucleocapsid assembly. *Sci. Adv.* **6**:eaaz7095.
5. Heinrich BS, Maliga Z, Stein DA, Hyman AA, Whelan SPJ. (2018). Phase Transitions Drive the Formation of Vesicular Stomatitis Virus Replication Compartments. *mBio* **9**(5):e02290-17.
6. Nikolic J, Le Bars R, Lama Z, Scrima N, Lagaudriere-Gesbert C, Gaudin Y, Blondel D. 2017. Negri bodies are viral factories with properties of liquid organelles. *Nat. Commun.* **8**:58.
7. Zhou Y, Su JM, Samuel CE, Ma D. (2019). Measles Virus Forms Inclusion Bodies with Properties of Liquid Organelles. *J. Virol.* **93**(21):e00948-19.
8. Rincheval V, Lelek M, Gault E, Bouillier C, Sitterlin D, Blouquit-Laye S, Galloux M, Zimmer C, Eleouet JF, Rameix-Welti MA. (2017). Functional organization of cytoplasmic inclusion bodies in cells infected by respiratory syncytial virus. *Nat. Commun.* **8**:563.
9. Afonso CL, Amarasinghe GK, Banyai K, Bao Y, Basler CF, Bavari S, Bejerman N, Blasdel KR, Briand FX, Briese T, Bukreyev A, Calisher CH, Chandran K, Cheng J, Clawson AN, Collins PL, Dietzgen RG, Dolnik O, Domier LL, Durrwald R, Dye JM, Easton AJ, Ebihara H, Farkas SL, Freitas-Astua J, Formenty P, Fouchier RA, Fu Y, Ghedin E, Goodin MM, Hewson R, Horie M, Hyndman TH, Jiang D, Kitajima EW, Kobinger GP, Kondo H, Kurath G, Lamb RA, Lenardon S, Leroy EM, Li CX, Lin XD, Liu L, Longdon B, Marton S, Maisner A, Muhlberger E, Netesov SV, Nowotny N, et al. (2016). Taxonomy of the order Mononegavirales: update 2016. *Arch. Virol.* **161**, 2351-60.
10. Rima B, Collins P, Easton A, Fouchier R, Kurath G, Lamb RA, Lee B, Maisner A, Rota P, Wang L, Ictv Report C. (2017). ICTV Virus Taxonomy Profile: Pneumoviridae. *J. Gen. Virol.* **98**, 2912-2913.
11. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, Polack FP, Balsells E, Acacio S, Aguayo C, Alassani I, Ali A, Antonio M, Awasthi S, Awori JO, Azziz-Baumgartner E, Baggett HC, Baillie VL, Balmaseda A, Barahona A, Basnet S, Bassat Q, Basualdo W, Bigogo G, Bont L, Breiman RF, Brooks WA, Broor S, Bruce N, Bruden D, Buchy P, Campbell S, Carosone-Link P, Chadha M, Chipeta J, Chou M, Clara W, Cohen C, de Cuellar E, Dang DA, Dash-Yandag B,

- Deloria-Knoll M, Dherani M, Eap T, Ebruke BE, Echavarria M, de Freitas Lazaro Emediato CC, Fasce RA, Feikin DR, Feng L, et al. (2017). Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *Lancet* **390**, 946-958.
12. Nair H, Nokes DJ, Gessner BD, Dherani M, Madhi SA, Singleton RJ, O'Brien KL, Roca A, Wright PF, Bruce N, Chandran A, Theodoratou E, Sutanto A, Sedyaningsih ER, Ngama M, Munywoki PK, Kartasmita C, Simoes EA, Rudan I, Weber MW, Campbell H. (2010). Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet* **375**, 1545-55.
  13. Pneumonia Etiology Research for Child Health Study G. (2019). Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. *Lancet* **394**, 757-779.
  14. Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. (2005). Respiratory syncytial virus infection in elderly and high-risk adults. *N. Engl. J. Med.* **352**, 1749-59.
  15. Asner S, Stephens D, Pedulla P, Richardson SE, Robinson J, Allen U. (2013). Risk factors and outcomes for respiratory syncytial virus-related infections in immunocompromised children. *Pediatr. Infect. Dis. J.* **32**, 1073-6.
  16. Shah JN, Chemaly RF. (2011). Management of RSV infections in adult recipients of hematopoietic stem cell transplantation. *Blood* **117**, 2755-63.
  17. Collins PL, Melero JA. (2011). Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years. *Virus Res.* **162**, 80-99.
  18. Collins PL, Hill MG, Camargo E, Grosfeld H, Chanock RM, Murphy BR. (1995). Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5' proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11563-7.
  19. Fearn R, Collins PL. (1999). Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J. Virol.* **73**, 5852-64.
  20. Noton SL, Fearn R. (2015). Initiation and regulation of paramyxovirus transcription and replication. *Virology* 479-480:545-54.
  21. Noton SL, Tremaglio CZ, Fearn R. (2019). Killing two birds with one stone: How the respiratory syncytial virus polymerase initiates transcription and replication. *PLoS Pathog.* 15:e1007548.
  22. Hefti E, Bishop DH. (1975). The 5' nucleotide sequence of vesicular stomatitis viral RNA. *J. Virol.* **15**, 90-6.
  23. Hefti E, Bishop DH. (1975). The 5' sequence of VSV viral RNA and its in vitro transcription product RNA. *Biochem. Biophys. Res. Commun.* **66**, 785-92.
  24. Bakker SE, Duquerroy S, Galloux M, Loney C, Conner E, Eleouet JF, Rey FA, Bhella D. (2013). The respiratory syncytial virus nucleoprotein-RNA complex forms a left-handed helical nucleocapsid. *J. Gen. Virol.* **94**, 1734-1738.
  25. Tawar RG, Duquerroy S, Vonrhein C, Varela PF, Damier-Piolle L, Castagne N, MacLellan K, Bedouelle H, Bricogne G, Bhella D, Eleouet JF, Rey FA. (2009). Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science* **326**, 1279-83.
  26. Leyrat C, Yabukarski F, Tarbouriech N, Ribeiro EA, Jr., Jensen MR, Blackledge M, Ruigrok RW, Jamin M. (2011). Structure of the vesicular stomatitis virus N(0)-P complex. *PLoS Patho.* 7:e1002248.
  27. Yabukarski F, Lawrence P, Tarbouriech N, Bourhis JM, Delaforge E, Jensen MR, Ruigrok RW, Blackledge M, Volchkov V, Jamin M. (2014). Structure of Nipah virus unassembled nucleoprotein in complex with its viral chaperone. *Nature Struct. Mol. Biol.* **21**, 754-9.
  28. Esneau C, Raynal B, Roblin P, Brule S, Richard CA, Fix J, Eleouet JF, Galloux M. (2019). Biochemical characterization of the respiratory syncytial virus N(0)-P complex in solution. *J. Biol. Chem.* **294**, 3647-3660.

29. Renner M, Bertinelli M, Leyrat C, Paesen GC, Saraiva de Oliveira LF, Huiskonen JT, Grimes JM. (2016). Nucleocapsid assembly in pneumoviruses is regulated by conformational switching of the N protein. *eLife* **5**:e12627.
30. Aggarwal M, Leser GP, Kors CA, Lamb RA. (2018). Structure of the Paramyxovirus Parainfluenza Virus 5 Nucleoprotein in Complex with an Amino-Terminal Peptide of the Phosphoprotein. *J. Virol.* **92**(5):e01304-17.
31. Zhu T, Song H, Peng R, Shi Y, Qi J, Gao GF. (2017). Crystal Structure of the Marburg Virus Nucleoprotein Core Domain Chaperoned by a VP35 Peptide Reveals a Conserved Drug Target for Filovirus. *J. Virol.* **91**(18):e00996-17.
32. Guryanov SG, Liljeroos L, Kasaragod P, Kajander T, Butcher SJ. (2015). Crystal Structure of the Measles Virus Nucleoprotein Core in Complex with an N-Terminal Region of Phosphoprotein. *J. Virol.* **90**, 2849-57.
33. Kirchdoerfer RN, Abelson DM, Li S, Wood MR, Saphire EO. (2015). Assembly of the Ebola Virus Nucleoprotein from a Chaperoned VP35 Complex. *Cell reports* **12**, 140-9.
34. Dong X, Wang X, Xie M, Wu W, Chen Z. (2022). Structural Basis of Human Parainfluenza Virus 3 Unassembled Nucleoprotein in Complex with Its Viral Chaperone. *J. Virol.* **96**:e0164821.
35. Galloux M, Gabiane G, Sourimant J, Richard CA, England P, Moudjou M, Aumont-Nicaise M, Fix J, Rameix-Welti MA, Eleouet JF. (2015). Identification and Characterization of the Binding Site of the Respiratory Syncytial Virus Phosphoprotein to RNA-Free Nucleoprotein. *J. Virol.* **89**, 3484-96.
36. Garcia J, Garcia-Barreno B, Vivo A, Melero JA. (1993). Cytoplasmic inclusions of respiratory syncytial virus-infected cells: formation of inclusion bodies in transfected cells that coexpress the nucleoprotein, the phosphoprotein, and the 22K protein. *Virology* **195**, 243-7.
37. Galloux M, Risso-Ballester J, Richard CA, Fix J, Rameix-Welti MA, Eleouet JF. (2020). Minimal Elements Required for the Formation of Respiratory Syncytial Virus Cytoplasmic Inclusion Bodies In Vivo and In Vitro. *mBio* **11**, (5):e01202-20.
38. Risso-Ballester J, Galloux M, Cao J, Le Goffic R, Hontonnou F, Jobart-Malfait A, Desquesnes A, Sake SM, Haid S, Du M, Zhang X, Zhang H, Wang Z, Rincheval V, Zhang Y, Pietschmann T, Eleouet JF, Rameix-Welti MA, Altmeyer R. (2021). A condensate-hardening drug blocks RSV replication in vivo. *Nature* **595**, 596-599.
39. Richard CA, Rincheval V, Lassoued S, Fix J, Cardone C, Esneau C, Nekhai S, Galloux M, Rameix-Welti MA, Sizun C, Eleouet JF. (2018). RSV hijacks cellular protein phosphatase 1 to regulate M2-1 phosphorylation and viral transcription. *PLoS Pathog.* **14**:e1006920.
40. Bajorek M, Galloux M, Richard CA, Szekely O, Rosenzweig R, Sizun C, Eleouet JF. (2021). Tetramerization of Phosphoprotein is Essential for Respiratory Syncytial Virus Budding while its N Terminal Region Mediates Direct Interactions with the Matrix Protein. *J. Virol.* **11**(8):1225.
41. Sourimant J, Rameix-Welti MA, Gaillard AL, Chevret D, Galloux M, Gault E, Eleouet JF. (2015). Fine mapping and characterization of the L-polymerase-binding domain of the respiratory syncytial virus phosphoprotein. *J. Virol.* **89**, 4421-33.
42. Gilman MSA, Liu C, Fung A, Behera I, Jordan P, Rigaux P, Ysebaert N, Tcherniuk S, Sourimant J, Eleouet JF, Sutto-Ortiz P, Decroly E, Roymans D, Jin Z, McLellan JS. (2019). Structure of the Respiratory Syncytial Virus Polymerase Complex. *Cell* **179**, 193-204 e14.
43. Tran TL, Castagne N, Bhella D, Varela PF, Bernard J, Chilmoneczyk S, Berkenkamp S, Benhamo V, Grznarova K, Grosclaude J, Nespoulos C, Rey FA, Eleouet JF. (2007). The nine C-terminal amino acids of the respiratory syncytial virus protein P are necessary and sufficient for binding to ribonucleoprotein complexes in which six ribonucleotides are contacted per N protein protomer. *J. Gen. Virol.* **88**, 196-206.
44. Galloux M, Tarus B, Blazevic I, Fix J, Duquerroy S, Eleouet JF. (2012). Characterization of a viral phosphoprotein binding site on the surface of the respiratory syncytial nucleoprotein. *J. Virol.* **86**, 8375-87.
45. Gao Y, Cao D, Ahn HM, Swain A, Hill S, Ogilvie C, Kurien M, Rahmatullah T, Liang B. (2020). In vitro trackable assembly of RNA-specific nucleocapsids of the respiratory syncytial virus. *J. Biol.*

- Chem. **295**, 883-895.
46. Banani SF, Lee HO, Hyman AA, Rosen MK. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell. Biol.* **18**, 285-298.
  47. Alberti S, Gladfelter A, Mittag T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419-434.
  48. Alberti S, Saha S, Woodruff JB, Franzmann TM, Wang J, Hyman AA. (2018). A User's Guide for Phase Separation Assays with Purified Proteins. *J. Mol. Biol.* **430**, 4806-4820.
  49. Shimobayashi SF, Ronceray P, Sanders DW, Haataja MP, Brangwynne CP. (2021). Nucleation landscape of biomolecular condensates. *Nature* **599**, 503-506.
  50. Brocca S, Grandori R, Longhi S, Uversky V. (2020). Liquid-Liquid Phase Separation by Intrinsically Disordered Protein Regions of Viruses: Roles in Viral Life Cycle and Control of Virus-Host Interactions. *Int. J. Mol. Sci.* **21**(23):9045.
  51. Song X, Shan H, Zhu Y, Hu S, Xue L, Chen Y, Ding W, Niu T, Gu J, Ouyang S, Shen QT, Liu ZJ. (2019). Self-capping of nucleoprotein filaments protects the Newcastle disease virus genome. *Elife* **8**:e45057.
  52. Zhang N, Shan H, Liu M, Li T, Luo R, Yang L, Qi L, Chu X, Su X, Wang R, Liu Y, Sun W, Shen QT. (2021). Structure and assembly of double-headed Sendai virus nucleocapsids. *Commun. Biol.* **4**:494.
  53. Ker DS, Jenkins HT, Greive SJ, Antson AA. (2021). CryoEM structure of the Nipah virus nucleocapsid assembly. *PLoS Pathog.* **17**:e1009740.
  54. Zinzula L, Beck F, Klumpe S, Bohn S, Pfeifer G, Bollschweiler D, Nagy I, Plietzko JM, Baumeister W. 2021. Cryo-EM structure of the cetacean morbillivirus nucleoprotein-RNA complex. *J Struct Biol* **213**:107750.
  55. Conley MJ, Short JM, Burns AM, Streetley J, Hutchings J, Bakker SE, Power BJ, Jaffery H, Haney J, Zanetti G, Murcia PR, Stewart M, Fearn R, Vijayakrishnan S, Bhella D. (2022). Helical ordering of envelope-associated proteins and glycoproteins in respiratory syncytial virus. *EMBO J.* **41**:e109728.
  56. Parey N, Baraguey C, Vasseur JJ, Debart F. (2006). First evaluation of acyloxymethyl or acylthiomethyl groups as biolabile 2'-O-protections of RNA. *Org. Lett.* **8**, 3869-72.
  57. Zlatev I, Lavergne T, Debart F, Vasseur JJ, Manoharan M, Morvan F. (2010). Efficient solid-phase chemical synthesis of 5'-triphosphates of DNA, RNA, and their analogues. *Org. Lett.* **12**, 2190-3.
  58. Thillier Y, Decroly E, Morvan F, Canard B, Vasseur JJ, Debart F. (2012). Synthesis of 5' cap-0 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human (guanine-N(7))-methyl transferase. *RNA* **18**, 856-68.

Table 1: RNAs sequences

RNAs length	5' modification	RNA sequence *
70-mer #		5'ACGCGAAAAAAUGCGUACAACAAACUUGCAUAAACCAA AAAAUGGGGCAAUAAGAAUUUGAUAAAGUAC3'
14-mer (a)		5'-ACGCGAAAAAUGC-3'
14-mer (b)		5'-GGGCGAAAAAUGC-3'
14-mer (c)		5'-GUGCGAAAAAUGC-3'
11-mer		5'-ACGCGAAAAA-3'
10-mer	OH	5'-ACGCGAAAAA-3'
9-mer		5'-ACGCGAAAA-3'
8-mer		5'ACGCGAAA-3'
7-mer		5'-ACGCGAA-3'
6-mer		5'-ACGCGA-3'
5-mer		5'-ACGCG-3'
11-mer		5'-ACGCGAAAAA-3'
10-mer		5'ACGCGAAAAA-3'
9-mer	PPP	5'ACGCGAAAA-3'
8-mer		5'ACGCGAAA-3'
7-mer		5'-ACGCGAA-3'
11-mer		5'-ACGCGAAAAA-3'
10-mer		5'ACGCGAAAAA-3'
9-mer	Cap (7m+2'O-Gppp)	5'ACGCGAAAA-3'
8-mer		5'ACGCGAAA-3'
7-mer		5'-ACGCGAA-3'

\* All the sequences correspond to HRSV leader sequence except those of the 14-mer b and c indicated in italics.

# Corresponding 70-mer DNA was used in the study.

Table 2: Characterisation of 5' OH RNA-N-P40 complexes after RNase A treatment

Proteins and RNAs	OD <sub>260nm</sub> /OD <sub>280nm</sub>	Dh (nm) *
Nwt #	1.24	15.75
N-P40	0.67	7.9
N-P40 + 6-mer RNA	0.79	8.9
N-P40 + 7-mer RNA	0.9	8.7
N-P40 + 8-mer RNA	0.88	6.7
N-P40 + 9-mer RNA	0.87	7.6
N-P40 + 10-mer RNA	0.94	9
N-P40 + 11-mer RNA	1.03	14.2
N-P40 + 14-mer RNA	1	14.4

\* Hydrodynamic diameter determined by DLS. Data representative of 2 independent experiments. # Purified as N-RNA rings.

## Figure legends

**Figure 1:** Purification and characterisation of recombinant N-P40 fusion protein. A/ Amino acid sequence of the N-P40 fusion protein. The N-terminal 6xHis tag is in green, the N sequence is in blue and the sequence of P40 is in red. Amino acids in black correspond to additional residues and linkers. B/ Gel filtration profile of purified N-P40 (red) and N<sub>wt</sub> (black). The curves corresponding to OD spectra at 260 nm and 280 nm are presented as dash and solid lines respectively. P1 and P2 indicate the two peaks detected for the gel filtration profile of N-P40. The fractions corresponding to P2 were pooled and the sample was analysed by SDS-PAGE coloured with Coomassie blue. The asterisk indicates the peak corresponding to the C-terminal fragment of P used for N<sub>wt</sub> purification (44) C/ Dynamic light scattering

(DLS) analysis of the purified N-P40 protein, showing a homogenous peak close to 6.7 nm, corresponding to N-P40 monomer and a minor peak of oligomers. D and E/ Far-UV (D) and near-UV CD (E) spectra of N<sub>wt</sub> corresponding to N-RNA rings in black, and N-P40 in grey.

**Figure 2:** The N-P40 fusion protein is competent for RNA encapsidation. A/ Analysis of N-P40 migration alone or incubated in the presence of 70-mers RNA or DNA (see 70-mer sequence in Table 1) by native polyacrylamide (left) or agarose (right) gel electrophoresis. Polyacrylamide and agarose gels were stained with Coomassie blue or amido black respectively, as indicated. N<sub>wt</sub> corresponds to purified recombinant RNA-N rings, used as control. B/ Analysis of 14-mer RNAs (see sequences Table 1) encapsidation by N-P40 on native acrylamide (left) or agarose (right) gels. C/ Images of N-P40 alone or incubated in the presence on 70-mer RNA, 70-mer DNA, or 14-mer RNA, as observed by ns-EM. Scale bar, 50 nm.

**Figure 3:** Formation of stable RNA-N complexes by N-P40 depends on RNA length. A/ Analysis of N-P40 migration alone or incubated in the presence of 5- to 11-mer RNAs, compared to 14-mer RNAs, by native polyacrylamide (left) or agarose (right) gel electrophoresis. Polyacrylamide and agarose gels were stained with Coomassie blue or amido black respectively, as indicated. B/ Images of N-P40-RNA complexes formed upon incubation of N-P40 in the presence of 7- and 11-mer RNAs with 5' OH, as observed by ns-EM. Scale bar, 50 nm. C/ Analysis of N-P40 migration alone or incubated in the presence of 5- to 11-mer RNAs, after treatment with RNase A, by native agarose gel electrophoresis. Gel was stained with amido black. D/ Near-UV CD spectra of N-P40 alone or incubated in the presence of 7- to 11-mer RNAs, after treatment with RNase A.

**Figure 4:** Impact of RNA 5' end modifications on the ability of N-P40 to encapsidate RNAs. A/ Analysis of N-P40 migration alone or incubated in the presence of 7- to 11-mer RNAs with 5' PPP (left), or 5' <sup>m</sup>GpppG<sub>m</sub> (right) by native agarose gel electrophoresis. Gel was stained with amido black. B/ Images of N-P40-RNA complexes formed upon incubation of N-P40 in the presence of 7- and 11-mer RNAs with 5' PPP or 5' <sup>m</sup>GpppG<sub>m</sub>, as observed by ns-EM. Scale bar, 50 nm. C/ Comparison of the migration of N-P40 alone or incubated in the presence of 11-mer RNAs with 5' OH, PPP or <sup>m</sup>GpppG<sub>m</sub> after RNase A treatment by native agarose gel electrophoresis, gel stained with amido black. D/ Near-UV CD spectra

of N-P40 alone or incubated in the presence of 11-mer RNAs with 5' OH, PPP or  $m^7GpppG_m$  after RNase A treatment.

**Figure 5:** Impact of  $N_{wt}$  and P concentration and ratio on pseudo-IBs' formation. Recombinant mCherry- $N_{wt}$  (corresponding to mCherry-N-RNA rings) and P proteins were co-incubated in the presence of 15% Ficoll and the formation of droplets was observed by fluorescence microscopy. A/ P concentration was fixed to 3.5  $\mu$ M and mCherry- $N_{wt}$  was added up to ratio mCherry- $N_{wt}$  / P of 5, B/ mCherry- $N_{wt}$  concentration was fixed to 3.5  $\mu$ M and P concentration increased up to ratio P / mCherry- $N_{wt}$  of 5. Scale bar, 10  $\mu$ m.

**Figure 6:** N oligomerization is needed for the formation of pseudo-IBs. A/ Recombinant P-BFP (3.5  $\mu$ M) was incubated in the presence of either N-P40 preincubated with 70-mer RNA (13.5 and 50  $\mu$ M respectively),  $N_{wt}$  (13.5  $\mu$ M), 70-mer RNA (50  $\mu$ M), or N-P40 (13.5  $\mu$ M). B/ Recombinant P-BFP (3.5  $\mu$ M) was incubated in the presence N-P40 (13.5  $\mu$ M) pre-incubated with 14-mer, 11-mer (5' end OH), 8-mer, or 6-mer RNAs (50  $\mu$ M). C/ Recombinant P-BFP (3.5  $\mu$ M) was incubated in the presence N-P40 (13.5  $\mu$ M) pre-incubated with 8-mer and 11-mer RNAs (50  $\mu$ M) followed by treatment with RNase A, and observed by fluorescence microscopy. D/ Recombinant P-BFP (3.5  $\mu$ M) was incubated in the presence N-P40 (13.5  $\mu$ M) pre-incubated with 5' PPP or  $m^7GpppG_m$  11-mer RNAs (50  $\mu$ M). For each condition, the formation of droplets was analysed by fluorescence microscopy. Scale bar, 10  $\mu$ m.