

# Characterization of the interaction domains between the phosphoprotein and the nucleocapsid of human Metapneumovirus

## Running title: Metapneumovirus N-P interaction

Hortense Decool<sup>1</sup>, Benjamin Bardiaux<sup>2</sup>, Luis Checa Ruano<sup>2</sup>, Olivier Sperandio<sup>2,3</sup>, Jenna Fix<sup>1</sup>, Irina Gutsche<sup>4</sup>, Charles-Adrien Richard<sup>1</sup>, Monika Bajorek<sup>1</sup>, Jean-François Eléouët<sup>1#</sup>, Marie Galloux<sup>1#</sup>.

<sup>1</sup>Université Paris-Saclay, INRAE, UVSQ, VIM, 78350, Jouy-en-Josas, France.

<sup>2</sup> Institut Pasteur, Unité de Bioinformatique Structurale, CNRS UMR3528, 28 rue du Dr Roux, 75015, Paris, France.

<sup>3</sup> Inserm U973 MTi, 25 rue Hélène Brion 75013, Paris, France.

<sup>4</sup> University of Grenoble Alpes, CNRS, CEA, CNRS, IBS, Grenoble, France.

#Correspondance: [jean-francois.eleouet@inrae.fr](mailto:jean-francois.eleouet@inrae.fr); [marie.galloux@inrae.fr](mailto:marie.galloux@inrae.fr)

**Keywords :** HMPV, phosphoprotein, nucleoprotein, protein-protein interaction, structural modeling

20 **ABSTRACT**

21 Human metapneumovirus (HMPV) causes severe respiratory diseases in young children. The HMPV  
22 RNA genome is encapsidated by the viral nucleoprotein, forming an RNA-N complex ( $N^{Nuc}$ ), which  
23 serves as template for genome replication and mRNA transcription by the RNA-dependent RNA  
24 polymerase (RdRp). The RdRp is formed by the association of the large polymerase subunit (L), which  
25 has RNA polymerase, capping and methyltransferase activities, and the tetrameric phosphoprotein (P).  
26 P plays a central role in the RdRp complex by binding to  $N^{Nuc}$  and L, allowing the attachment of the L  
27 polymerase to the nucleocapsid template. During infection these proteins concentrate in cytoplasmic  
28 inclusion bodies (IBs) where viral RNA synthesis occurs. By analogy to the closely related pneumovirus  
29 respiratory syncytial virus (RSV), it is likely that the formation of IBs depends on the interaction between  
30 P and  $N^{Nuc}$ . However, the HMPV P- $N^{Nuc}$  interaction still remains to characterize. Here, we finely  
31 characterized the binding domains involved in HMPV P and  $N^{Nuc}$  interaction by studying binding between  
32 recombinant proteins, combined with the use of a functional assay of the polymerase complex activity  
33 and the study of the recruitment of these proteins to IBs by immunofluorescence. We show that the last  
34 6 C-terminal residues of HMPV P are necessary and sufficient for binding to  $N^{Nuc}$ , that P binds the N-  
35 terminal domain of N ( $N_{NTD}$ ), and identified conserved N residues critical for the interaction. Our results  
36 allowed to propose a structural model of the HMPV P- $N^{Nuc}$  interaction.

37

38 **IMPORTANCE**

39 Like respiratory syncytial virus (RSV), human metapneumovirus (HMPV) is a leading cause of severe  
40 respiratory infections in children but also affects human populations of all ages worldwide. Nowadays,  
41 no vaccine or efficient antiviral treatments are available for these two pneumoviruses. A better  
42 understanding of the molecular mechanisms involved in viral replication could help the design or  
43 discovery of specific antiviral compounds. In this work we have investigated the interaction between two  
44 major viral proteins involved in HMPV RNA synthesis, the N and P proteins. We finely characterized  
45 their domains of interaction, in particular a pocket on the surface of the N protein that could be used as  
46 a template for the design of compounds interfering with N-P complexes and viral replication.

47

48 **INTRODUCTION**

49 Pneumonia is the leading cause of death among children younger than 5 years worldwide, and severe  
50 pneumonia is more frequently caused by viruses than bacteria (1). After the closely related respiratory  
51 syncytial virus (RSV), human metapneumovirus (HMPV) is recognized to be one of the most important  
52 cause of viral bronchiolitis and pneumonia in young children, causing 7 to 19% of all cases of acute  
53 respiratory tract infections (1, 2). HMPV infects mainly newborn, children, elderly and  
54 immunocompromised individuals worldwide. This virus was first reported in 2001 from Dutch children  
55 with acute lower respiratory tract illness, and serological studies have revealed that virtually every child  
56 has been exposed to HMPV by the age of 5 years (3). The clinical features of HMPV infection display  
57 as mid-to-upper respiratory tract infection, and can be severe enough to cause life-threatening  
58 bronchiolitis and pneumonia. As yet, there is no effective treatment or licensed vaccine for HMPV.

59 Together with RSV, HMPV belongs to the *Pneumoviridae* family in the order *Mononegavirales* (4).  
60 HMPV is an enveloped virus that forms pleomorphic or filamentous virions. The virus genome is  
61 composed of a negative-sense single-stranded RNA of approximately 13.3 kb in size which encodes  
62 eight genes in the following order: 3'-N-P-M-F-M2-SH-G-L-5' (5, 6). The M2 gene of HMPV contains two  
63 overlapping open reading frames (ORFs), encoding for M2-1 and M2-2 proteins which precise functions  
64 during HMPV replication remain unclear.

65 The HMPV genome is encapsidated by multiple copies of the nucleoprotein (N) forming helical  
66 nucleocapsids (N<sup>Nuc</sup>). This N<sup>Nuc</sup> serves as template for genome replication and mRNA transcription by  
67 the viral polymerase complex formed by the large polymerase subunit (L) and its main cofactor the  
68 phosphoprotein (P). After virus binding to the cell surface and virus-cell membrane fusion, mediated by  
69 surface glycoproteins (F and G), nucleocapsids are released into the cytoplasm. Replication and  
70 transcription of the viral genome take place within viro-induced cytoplasmic inclusions named inclusion  
71 bodies (IBs) (7). These structures can be observed upon expression of P and N proteins alone (8), and  
72 it was recently shown for RSV that the interaction between P and N<sup>Nuc</sup> is critical for the formation of IBs  
73 (9).

74 Among the components of the polymerase complex, P plays a pivotal role as a cofactor of the L  
75 polymerase but also as a molecular hub between viral and cellular partners. HMPV P, 294 amino acid  
76 residues in length, forms homo-tetramers. The atomic structure of the coiled-coil oligomerization domain  
77 (residues 171-194) was resolved by crystallography (10). Small angle X-ray scattering (SAXS) studies

78 indicated that the flanking N-terminal (residues 1-170) and C-terminal (residues 195-294) regions  
79 (named PNT and PCT respectively) are intrinsically disordered, some of them, such as residues 195-  
80 237, having  $\alpha$ -helical propensity (10, 11). More recently, the structure of the LP complex was resolved  
81 by cryo-EM (12). It revealed a tentacular arrangement of P, with each of the four protomers adopting a  
82 distinct conformation, demonstrating a “folding-upon-partner-binding” mechanism. Depending on the  
83 protomer, the L-binding region involved regions 171-236, 172-217, 170-231 or 169-264. On the other  
84 hand, by binding to the nucleocapsid, P mediates the attachment of the L protein to the  $N^{Nuc}$  template  
85 for viral RNA synthesis.

86 By analogy to RSV (13), the C-terminal domain of HMPV P is also thought to be the  $N^{Nuc}$  binding  
87 domain, but this has not been shown yet. It is noteworthy that the C-terminal extremity of P supposed  
88 to bind to  $N^{Nuc}$  was not visible in the crystal structure of the LP complex, indicating that this region is  
89 disordered in the absence of  $N^{Nuc}$ . Furthermore, the encapsidation of neosynthesized genome or  
90 antigenome necessitates a pool of monomeric, RNA-free N, termed  $N^0$ , which is kept in an unassembled  
91 state through an interaction with P which plays the role of molecular chaperone, until delivery to the sites  
92 of viral RNA synthesis. The crystal structures of recombinant HMPV N protein (394 residues) expressed  
93 in *E. coli* and purified either as an RNA-free monomeric N in complex with the N-terminal residues of P,  
94 or as rings of 10 N protomers complexed to RNA were resolved (14). In both states, the structures show  
95 that N presents two globular domains ( $N_{NTD}$  and  $N_{CTD}$ ) separated by a flexible linker that forms the RNA  
96 groove, and N- and C-arms. In the oligomeric state, the N- and C-arms play a key role in the interaction  
97 between N protomers and oligomerization, the N-arm binding to the flank of the  $N_{i+1}$  protomer and the  
98 C-arm binding atop the  $N_{i+1}$  protomer ( $i$  corresponding to the middle subunit of three adjacent N  
99 protomers). In the  $N^0$  state, the important conformational changes consist in packing of the C-arm of N  
100 in the RNA groove, impairing RNA binding.

101 In this work, based on the structural data of HMPV P and N proteins and their strong homologies  
102 with RSV N and P proteins, we finely characterized the binding domains involved in HMPV P- $N^{Nuc}$   
103 interaction. By combining biochemical and functional cellular assays, coupled with a rational mutational  
104 approach, we identified residues of P and N critical for their interaction. Our data show that the last C-  
105 terminal residues of P bind to the  $N_{NTD}$ . These results allowed to establish a structural model of the  
106 interaction which could be used for the rational design of antivirals targeting the  $N^{Nuc}$ -P interaction of  
107 HMPV.

108

109 **RESULTS**

110 **The HMPV PCT domain allows to purify N-RNA rings expressed in bacteria**

111 By analogy to RSV, it is thought that the N<sup>Nuc</sup> binding domain of HMPV P is located at its C-terminus.  
112 For RSV, it was shown that co-expression of the C-terminal disordered region of P (PCT, residues 161-  
113 241) fused to GST together with N in bacteria allows the purification of complexes formed by GST-PCT  
114 and ring shape N-RNA oligomers (13, 15). Similarly, when expressed alone in *E. coli*, HMPV N also  
115 forms decameric rings containing RNA (14). In a first attempt to characterize the N<sup>Nuc</sup> binding site of  
116 HMPV P, we thus decided to co-express recombinant HMPV GST-PCT (residues 200-294) and N  
117 proteins (from HMPV CAN 97-83 strain) in *E. coli*. The purified complexes were analyzed by SDS-PAGE  
118 stained with Coomassie blue. As shown on [figure.1A](#), N was co-purified with GST-PCT to > 95%  
119 homogeneity. The PCT was then separated from GST by thrombin cleavage and the solubilized complex  
120 was analyzed by size exclusion chromatography, following optical density (OD) at 220, 260 and 280 nm  
121 ([Fig. 1A](#)). The elution profile showed a major peak (P1) with a OD<sub>260nm</sub>/OD<sub>280nm</sub> ratio > 1 and an apparent  
122 mass of ~ 500 kDa, in agreement with the expected size of N-RNA decamers ([Fig. 1B](#)). A second peak  
123 of lower intensity (P2), with an apparent mass of ~ 10 kDa was detected only at 220 nm. This peak  
124 should correspond to PCT which does not present aromatic residues and a higher mass weight than  
125 predicted due to the elongated shape of this fragment. The fractions of P1 peak were pooled; analysis  
126 by SDS-PAGE stained with Coomassie blue of the sample showed the presence of a unique band  
127 corresponding to N ([Fig. 1B](#)). The separation of PCT and N upon gel filtration reveals the relatively low  
128 affinity of monomeric PCT for N. We then further analyzed the sample collected from the P1 peak by  
129 combining dynamic light scattering (DLS) and electron microscopy (EM) approaches. The profile of DLS  
130 measurement confirmed the homogeneity of the sample, with a single peak at 18 nm ([Fig 1C](#)), in  
131 agreement with the size of N-rings observed by EM ([Fig 1D](#)).  
132 Altogether, these results show that the GST-PCT fusion protein is sufficient to interact with  
133 nucleocapsid-like N-RNA rings.

134

135 **The last 6 residues of HMPV P constitute the minimum domain for N<sup>Nuc</sup> binding**

136 We thus investigated the minimal domain of HMPV P involved in N<sup>Nuc</sup> binding. By analogy with RSV for  
137 which the 9 C-terminal residues of P are sufficient to interact with N (13), a series of GST fused peptides

138 of 9 to 1 amino acid long derived from the C-terminal of P was generated. These constructs were co-  
139 expressed with N in *E.coli*, and their capacity to purify N was analyzed by SDS-PAGE and Coomassie  
140 blue staining. As shown on [figure. 2A](#), the minimal sequence required for N purification corresponds to  
141 the 6 last C-terminal residues of P. In parallel, alanine scanning was performed with the GST-P[285-  
142 294] construct to characterize the residues of P involved in the interaction with N<sup>Nuc</sup>. Again, GST-P  
143 constructs and N were co-expressed in bacteria, followed by purification by Glutathione-Sepharose  
144 beads affinity, and co-purification of N was analyzed by SDS-PAGE and Coomassie blue staining. Only  
145 the four I289A, Y290A, L292A and M294A substitutions abrogated the interaction with N ([Fig. 2B](#)). In  
146 order to further investigate the potential role of residues Q291 and I293 of P in the interaction with N,  
147 the impact of the double mutations Q291A/I293A inserted into GST-PCT on N binding was also tested.  
148 These mutations did not affect N binding, confirming that these two residues are not directly involved in  
149 the interaction ([Fig. 2C](#)). These results confirm that the last residues of P are directly involved in the  
150 interaction with N. Of note, our results suggest that HMPV P-N interaction mainly involves hydrophobic  
151 interaction. As both hydrophobic and acidic residues of RSV P were previously shown to be critical for  
152 the interaction with N ([Fig. 2D](#)), these results suggest that the binding of HMPV P on N differs from RSV  
153 and involves specific interactions. We thus tested the capacity of each PCT fragments to pulldown N  
154 proteins. [Figure 2E](#) shows that HMPV PCT and RSV PCT did not allow to purify RSV N and HMPV N  
155 respectively. This last result confirms that HMPV and RSV N and P proteins cannot cross-react, and  
156 that pneumoviruses P-N interactions are specific to each virus.

157

### 158 **Search for the P binding site on the surface of HMPV N**

159 Although HMPV and RSV P-N<sup>Nuc</sup> interactions have their own specificity and cannot cross-react, the  
160 strong homologies between N proteins suggest that the PCT binding site on N protein could be partially  
161 conserved. For RSV, the P binding site on N<sup>Nuc</sup> is located at the surface of the N<sub>NTD</sub>. The residues of  
162 RSV N<sub>NTD</sub> involved in the interaction with P, which are partially conserved between HMPV and RSV  
163 ([Fig. 3A](#)), were shown to form a well-defined pocket constituted of hydrophobic and positively charged  
164 residues (16, 17). Corresponding residues of HMPV N<sub>NTD</sub> also form a pocket at the surface of the domain  
165 ([Fig. 3B](#)). In a first attempt to characterize the HMPV P binding domain on N, we thus tested the capacity  
166 of P to interact with the N<sub>NTD</sub>. The N<sub>NTD</sub> was expressed in *E. coli* and purified using a 6xHis tag, showing  
167 that this domain is soluble when expressed alone ([Fig. 3C](#)). In parallel, the GST-P or GST-PΔM294

168 (deletion of the last C-terminal residue of P) proteins were expressed in bacteria either alone and purified  
169 using the GST tag, or co-expressed with N<sub>NTD</sub> followed by purification by 6xHis tag. Analysis of the  
170 samples by SDS-PAGE showed that N<sub>NTD</sub> allowed to co-purify the wild type P, and that deletion of the  
171 last residue of P was sufficient to impair the interaction (Fig. 3C). We then substituted by alanine the  
172 residues of the HMPV N<sub>NTD</sub> identified by analogy with RSV, i.e. residues L46, L47, E50, Y53, D128,  
173 R132, M135, R151, P152, and S153 (Fig. 3B). Again, the capacity of GST-P to co-purify N mutant  
174 proteins was tested. Whereas mutants L46A, L47A, and P152 could still interact with P, only a weak  
175 band of N was detected for mutants E50A, M135A, R151A and S153A, and mutations Y53A, D128A,  
176 R132A abrogated the interaction (Fig. 3D). Of note, it was previously shown that similar punctual  
177 mutations did not impact RSV N folding and solubility, suggesting that these mutations on HMPV N only  
178 affect the capacity of P to interact with N. These results thus reveal a critical role of hydrophobic  
179 residues, but also of negatively and positively charged residues of N<sub>NTD</sub> in P binding. Overall, these data  
180 confirm that the P binding site on N presents strong homologies between HMPV and RSV.

181

## 182 **Validation of P-N<sup>Nuc</sup> binding domains in eukaryotic cells**

183 As P-N<sup>Nuc</sup> interaction is required for viral polymerase activity, we then studied the impact of selected  
184 mutations of P (M294A) or N (R132A and R151A) on HMPV replication/transcription using a bicistronic  
185 subgenomic minigenome, pGaussia/Firefly. This construct contains the Gaussia and Firefly luciferase  
186 genes under the control of gene start and gene stop sequences, as well as Leader and Trailer  
187 sequences of HMPV genome (Fig. S1). Briefly, BSRT7 cells were transfected with plasmids coding for  
188 N (N<sub>WT</sub>, N<sub>R132A</sub> or N<sub>R151A</sub>), P (P<sub>WT</sub> or P<sub>M294A</sub>), L, and M2-1 proteins, the plasmid pGaussia/Firefly  
189 minigenome, and pSV-β-Gal (used to normalize transfection efficiencies). In this system, expression of  
190 luciferases depends on the HMPV polymerase complex activity which can thus be quantified by  
191 luminescence measurement. As shown on figure 4A, mutations M294A of P or R132A of N abrogated  
192 HMPV polymerase activity and R151A induced a strong decrease on its activity, with only 30% of activity  
193 compared to the control condition, although the levels of expression of these proteins were similar to  
194 those of wild type proteins (Fig. 4B). These results correlate with those of pulldown assays, the mutant  
195 N<sub>R151A</sub> that still presented a weak interaction with P displaying a residual polymerase activity.  
196 In HMPV-infected cells, P-N<sup>Nuc</sup> interaction induces the formation of inclusion bodies (8). For RSV, we  
197 have shown previously that IBs correspond to viral factories where viral RNA synthesis occurs (18). We

198 then assessed the impact of these mutations on the capacity of N and P to form pseudo-IBs when co-  
199 expressed in cells. To facilitate N detection, N-GFP fusion protein was generated to allow its detection  
200 by epifluorescence. Based on previous studies on RSV P-N<sup>Nuc</sup> interaction, the GFP was fused to the C-  
201 terminus of N and this construct was co-expressed with wild-type N at a ratio of 1:1. Cells were thus co-  
202 transfected with plasmids encoding N-GFP, N (both WT or mutants R132A or R151A), and P (WT or  
203 mutant M294A). The P protein was detected by immunolabelling using rabbit anti-P antibody, and nuclei  
204 were stained with Hoechst 33342. When expressed together in BSRT7 cells, WT N-GFP/N and P were  
205 found to co-localize into cytoplasmic inclusions similar to IBs observed upon HMPV infection (7, 8) (Fig.  
206 4C). On the contrary, when expressing P<sub>M294A</sub> in the presence of N<sub>WT</sub> or P<sub>WT</sub> in the presence of N<sub>R132A</sub>,  
207 P presented a diffuse cytoplasmic distribution and N protein was mostly detected in cytoplasmic  
208 aggregates (Fig. 4C). These observations suggest that in the absence of N-P interaction, overexpressed  
209 N presents a strong tendency to aggregate. Finally, co-expression of the mutant N<sub>R151A</sub> with P still  
210 allowed to observe IBs, although less numerous and smaller than those observed in the presence of  
211 WT proteins. This last result correlates with previous results showing that mutation R151A of N  
212 drastically reduced but did not abrogate N-P interaction and polymerase activity. Altogether, these  
213 results confirm that the residues of P and N previously identified as critical for N<sup>Nuc</sup>-P interaction *in vitro*  
214 are also critical for the formation of IBs and polymerase activity in cells.

215

## 216 **Molecular modelling of PCT-N<sub>NTD</sub> interaction**

217 To gain insight into the possible binding mode of P with N<sub>NTD</sub>, we built models of the last 6 residues of  
218 P (peptide I289-YQLI-M294 named P6) bound to N, by analogy with the PCT-N<sub>NTD</sub> complex of RSV. For  
219 RSV, only the last 2 residues of P (D240-F241) were resolved in the PCT-N<sub>NTD</sub> X-ray structures (17),  
220 and are bound to a hydrophobic pocket of N similar to the one we characterized at the surface of HMPV  
221 N. Using comparative modeling and refinement with HADDOCK (19), we obtained three clusters of  
222 HMPV P6/N<sub>NTD</sub> complex structures representing possible binding modes of P6 (Table S1). The first two  
223 clusters in terms of HADDOCK-score were the most populated and displayed similar Root-mean-square  
224 deviation (RMSD) among their members (considering only the last three residues of P, i.e. residues  
225 L292-I293-M294, called P3). In those two clusters, the C-terminal carboxyl group of M294 is involved in  
226 an ionic interaction with the guanidinium of either R131 or R152 side-chains of N. Additionally, only L292  
227 and M294 of P3 consistently display hydrophobic contacts with N. This result is consistent with our

228 mutation data showing that I293 of P is not required for binding to N, contrary to L292 and M294. The  
229 10 best-scoring structures from cluster 1 and 2 are shown in [Fig. S2](#), superimposed with the RSV  
230 P2/N<sub>NTD</sub> structure (residues D240-F241 of P). Owing to the best HADDOCK-score and non-bonded  
231 energies of cluster 2, we considered this cluster as most representative for the potential binding mode  
232 of HMPV P3 to N<sub>NTD</sub>. Interestingly, more than half of the structures from this cluster display a hydrogen-  
233 bond between the backbone oxygen of I293 of P3 with the side-chain of R151 of N, while the side-chain  
234 of R132 of N forms a salt-bridge with the C-terminal carboxyl group of M294 of P. Thus, the interactions  
235 determined by molecular simulation correlate with the experimental data showing the importance of  
236 R132 and R151 of N and of M294 of P for PCT-N interaction.

237 The interactability propensity of each modeled N<sub>NTD</sub> conformation was further profiled with our in-house  
238 tool InDeep in order to select a subset of N<sub>NTD</sub> models being more likely to interact with the P3 peptide.  
239 On this subset of N<sub>NTD</sub> conformations, the hydrophobic channel of InDeep was also used to locate  
240 regions where hydrophobic moieties of P3 are expected to interact at the N<sub>NTD</sub> surface. Three  
241 hydrophobic patches are detected on the N<sub>NTD</sub> surface, that present a good match with hydrophobic  
242 moieties of P3 structures belonging to cluster 2 of HADDOCK: i) one occupied by L292 of P3; ii) another  
243 occupied by M294 of P3; and iii) the last one located in an unoccupied region of the N pocket ([Fig.5A](#)).  
244 Based on experimental and modelling information, the structure shown in [Figure 5](#) represents a  
245 consistent model of P3/N<sub>NTD</sub> complex. The N<sub>NTD</sub> conformation of this structure is predicted to be  
246 favorable to protein-protein interaction by InDeep, and the P3 conformation, belonging to cluster 2 of  
247 HADDOCK, binds to N<sub>NTD</sub> with a combined effect of electrostatic/polar interactions and hydrophobic  
248 contacts in agreement with the mutational experiments. In addition, a procedure of virtual alanine  
249 scanning using the same InDeep approach confirmed some of the mutagenesis results, such as the  
250 importance of R132, M135, and R151 of N for the interaction with P. It also suggests that residue L139  
251 of N may be key for the N<sub>NTD</sub> patch for partners' binding.

252 These analyses reveal that despite the different physico-chemical nature of residues involved in the  
253 interaction, PCT binding at the N surface is structurally similar between HMPV and RSV.

254

## 255 **DISCUSSION**

256 Beyond its interest in better understanding the molecular mechanisms of virus replication, the study of  
257 protein-protein interactions required for Pneumoviruses' polymerase complex activity also presents

258 interest for the development of specific antiviral compounds. The activity of this complex depends on  
259 multiple highly specific protein-protein interactions that have no cellular counterparts. Among those,  
260 N<sup>Nuc</sup>-P interaction which is critical for viral polymerase activity constitutes a singular potential target for  
261 antivirals. We have previously described RSV N<sup>Nuc</sup>-P interaction at the molecular level, showing that the  
262 C-terminus of P binds to a well-defined pocket at the surface of N (9, 13, 16). We also made the proof  
263 of concept that small chemical compounds can target this pocket of N and inhibit viral replication in cells  
264 (17). More recently, Sá et al. showed that flavonoids derived molecules could also specifically bind to N  
265 pocket and block its interaction with P (20). Given the strong homologies between RSV and HMPV N  
266 and P proteins, it could be expected that a similar strategy could be used to block HMPV replication. It  
267 is noteworthy that ideally, the discovery of compounds that could inhibit replication of these two closely  
268 related viruses at the same time would be much more advantageous.

269 In the present study, we investigated the domains involved in HMPV P-N<sup>Nuc</sup> interaction using pulldown  
270 assays between recombinant proteins (either truncated or mutated proteins) expressed in *E. coli*. As for  
271 RSV, the HMPV N protein was purified as nanorings formed by N and bacterial RNA (Fig 1). Our results  
272 revealed that mechanisms of binding between P and N<sup>Nuc</sup> are similar between RSV and HMPV, the C-  
273 terminus of P interacting with a pocket located on the surface of the N N-terminal globular domain N<sub>NTD</sub>.  
274 However, it appeared that P and N amino acid residues involved in these interactions have different  
275 contributions/roles between RSV and HMPV, and that the P-N<sup>Nuc</sup> interaction is specific of each virus.  
276 More specifically, we showed that the minimal domain of HMPV P required for N binding is constituted  
277 by the last 6 C-terminal residues of P which are mostly hydrophobic residues (IYQLIM), and determined  
278 that only Q291 and I293 are not critical for the interaction (Fig. 2). This P fragment is thus shorter than  
279 for RSV for which N binding minimal domain involves the 9 C-terminal residues of P and is composed  
280 of both hydrophobic and acidic residues (13, 16). We also identified 7 residues of N involved in P binding,  
281 *i.e.* residues E50, Y53, D128, R132, M135A, R151A and S153A, which form a well-defined pocket at  
282 the N<sub>NTD</sub> surface (Fig 3). Of note, the pivotal role of residues R132 and R151 of N for polymerase activity  
283 and IBs formation sustained *in vitro* binding assays (Fig. 4). Of most interest, residues R132 and R150  
284 of RSV N were also shown to play a key role in P-N<sup>Nuc</sup> interaction, the aliphatic part of the R132 side  
285 chain being involved in the stacking of the aromatic ring of the residue F241 of P. In order to better  
286 understand the binding mode of HMPV P to N, molecular modelling was thus performed and the best  
287 structural models support that R132 and R151 of N and of M294 of P are critical for PCT-N interaction

288 (Fig. 5). Nevertheless, molecular modelling was restricted to the binding of the last 3 C-terminal residues  
289 of P, and did not allow to predict the interaction of all the 6 critical residues of P on N. As also observed  
290 in the crystal structures of RSV PCT/N<sub>NTD</sub> complexes (17), the last C-terminal residue of P seems to  
291 drive the binding to a well-defined N<sub>NTD</sub> pocket and this primary binding could allow transient contacts  
292 with upstream P residues outside of the N pocket.

293 In conclusion, our data confirm the strong structural homologies between HMPV and RSV P-N  
294 complexes but also highlight some singularities. These data suggest that the HMPV P binding pocket  
295 at the N<sub>NTD</sub> surface could represent a new target for the rational design of antivirals.

296

297

## 298 MATERIALS AND METHODS

### 299 *Plasmid constructions*

300 The N and P sequences used for bacterial expression were derived from HMPV CAN97-83 strain. The  
301 sequence of N or N<sub>NTD</sub> (residues 30-255 of N) were cloned into pET-28a(+) vector (Novagen) at  
302 Ncol/Xhol or BamHI/Xhol restriction sites respectively, to allow bacterial expression of full-length N or  
303 N<sub>NTD</sub> domain with a C-terminal or N-terminal 6xHis tag. The sequence of P or PCT (residues 200-294  
304 of P), were cloned into pGEX-4T3 vector, at BamHI and SmaI sites to express in bacteria GST-P and  
305 GST-PCT fusion proteins. For the construction GST-P[285-294], complementary antiparallel  
306 oligonucleotides were hybridized and inserted at the BamHI/SmaI sites in the pGEX-4T-3 vector.

307 A HMPV minigenome plasmid containing Gaussia/Firefly luciferases was designed and synthesized by  
308 Genscript, cloned in pUC57 vector, and containing the trailer, leader, gene start (GS) and gene end  
309 (GE) sequences derived from HMPV CAN97-83 strain (see Fig S1). The first ORF of this  
310 pGaussia/Firefly minigenome codes for the Firefly luciferase and the second one codes for the Gaussia  
311 luciferase. The plasmids pP, pN, pL, and pM2-1 corresponding to the sequences of NL99-1 HMPV strain  
312 cloned into pCite vector. For expression of N-GFP fusion protein, the mGFP gene was amplified by PCR  
313 and cloned in frame at the 3' end of N at the EcoRI site of pN, using In-fusion HD cloning kit (Takara  
314 Bio). Point mutations were introduced in the P and N sequences by site-directed mutagenesis using the  
315 Q5 site-directed mutagenesis Kit (New England Biolabs). Sequence analysis was carried out to check  
316 the integrity of all constructs. All the oligonucleotides sequences are available on request.

317

318 **Antibodies**

319 Antisera used in this study included polyclonal rabbit antisera raised against recombinant HMPV N and  
320 P expressed in bacteria. A mouse monoclonal anti- $\beta$ -tubulin (Sigma) and secondary antibodies  
321 directed against mouse and rabbit Ig G coupled to HRP (P.A.R.I.S) were also used for immunoblotting.  
322 Secondary antibody directed against rabbit Ig G coupled to Alexafluor-594 (Invitrogen) was used for  
323 immunofluorescence experiments.

324

325 **Cell culture and transfections**

326 BHK-21 cells (clone BSRT7/5) constitutively expressing the T7 RNA polymerase (21) were grown in  
327 Dulbecco Modified Essential Medium (Lonza) supplemented with 10% fetal calf serum (FCS), 2 mM  
328 glutamine, and antibiotics. The cells were grown at 37°C in 5% CO<sub>2</sub>, and transfected using  
329 Lipofectamine 2000 (Invitrogen) as described by the manufacturer.

330

331 **Minigenome replication assay**

332 BSRT7/5 cells at 90% confluence in 48-well dishes were transfected using Lipofectamine 2000  
333 (Invitrogen) with a plasmid mixture containing 0.125  $\mu$ g of pGaussia/Firefly minigenome, 0.125  $\mu$ g of  
334 pN, 0.125  $\mu$ g of pP (WT and mutants), 0.06  $\mu$ g of pL, 0.03  $\mu$ g of pM2-1, as well as 0.03  $\mu$ g of pSV- $\beta$ -Gal  
335 (Promega) to normalize transfection efficiencies. Cells were harvested at 24 h post-transfection and  
336 lysed in Firefly lysis buffer (30 mM Tris [pH 7.9], 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol [DTT], 1% [vol/vol]  
337 Triton X-100, and 15% [vol/vol] glycerol). The Firefly luciferase activity was determined for each cell  
338 lysate with an Infinite 200 Pro (Tecan, Männedorf, Switzerland) and normalized based on  $\beta$ -  
339 galactosidase ( $\beta$ -Gal) expression. Transfections were done in triplicate, and each independent  
340 transfection experiment was performed three times. For proteins expression analysis, cells were lysed  
341 in Laemmli buffer and analyzed by Western blotting (WB) using anti-N, anti-P, and anti-tubulin antibodies  
342 according to standard protocols.

343

344 **Fluorescence microscopy**

345 Immunofluorescence microscopy was performed with cells grown on coverslips and previously  
346 transfected with pN-GFP, pN (both WT or mutants R132A and R151A, at a ratio 1:1), and pP (WT or  
347 mutant M294A). At 24 h post transfection, cells were fixed with 4% paraformaldehyde (PFA) for 25 min,

348 made permeable, and blocked for 30 min with PBS containing 0.1% Triton X-100 and 3% bovine serum  
349 albumin (BSA). Cells were then successively incubated for 1 h at room temperature with primary and  
350 secondary antibody mixtures diluted in PBS containing 0.3% BSA. For nucleus labeling, cells were  
351 exposed to Hoechst 33342 stain (Invitrogen) during incubation with secondary antibodies. Coverslips  
352 were mounted with ProLong Gold antifade reagent (Invitrogen) and observed with an inverted  
353 fluorescence microscope (Zeiss Axiovision). Images were processed with ZEN software (Zeiss) and  
354 ImageJ software.

355

356 ***Expression and purification of recombinant proteins***

357 *E. coli* BL21 bacteria (DE3) (Novagen) transformed with pGEX-P plasmids (WT or mutant) or pET-N<sub>NTD</sub>  
358 were grown at 37°C for 8 hours in 100 ml of Luria Bertani (LB) medium containing 100µg/ml ampicillin  
359 or 50µg/ml kanamycin, respectively. Bacteria transformed with pET-N-derived plasmids together with  
360 pGEX-P derived plasmids were grown in LB medium containing ampicillin (100 µg/ml) and kanamycin  
361 (50 µg/ml). The same volume of LB was then added and protein expression was induced by adding  
362 80µg/ml isopropyl-β-D-thio-galactoside (IPTG) to the medium. The bacteria were incubated for 15 hours  
363 at 28°C and then harvested by centrifugation. For purification using the GST-tag, bacteria were re-  
364 suspended in lysis buffer (50 mM Tris-HCl pH 7.8, 60 mM NaCl, 1 mM EDTA, 2 mM DTT, 0.2% Triton  
365 X-100, 1 mg/ml lysozyme) supplemented with complete protease inhibitor cocktail (Roche), incubated  
366 for 1 hour on ice, sonicated, and centrifuged at 4°C for 30 min at 10,000g. Glutathione-Sepharose 4B  
367 beads (GE Healthcare) were added to clarified supernatants and incubated at 4°C for 3 hours. Beads  
368 were then washed two times in lysis buffer and three times in PBS 1X, then stored at 4°C in an equal  
369 volume of PBS. To isolate GST-free P fragments or N rings, beads containing bound proteins were  
370 incubated with thrombin (Novagen) overnight at 20°C. Purified recombinant N proteins were loaded onto  
371 a Superdex 200 16/30 column (GE Healthcare) and eluted in 20 mM Tris/HCl pH 8.5, 150 mM NaCl.  
372 For purification of N<sub>NTD</sub>-6xHis fusion protein purification, bacterial pellets were re-suspended in lysis  
373 buffer (20 mM Tris-HCl pH8, 500 mM NaCl, 0.1% TritonX-100, 10 mM imidazole, 1 mg/ml lysozyme)  
374 supplemented with complete protease inhibitor cocktail (Roche). After sonication and centrifugation,  
375 lysates were incubated 30 min with chelating Sepharose Fast Flow beads charged with Ni<sup>2+</sup> (GE  
376 Healthcare). Beads were then successively washed in the washing buffer (20 mM Tris-HCl, pH 8, 500  
377 mM NaCl) containing increasing concentration of imidazole (25, 50, and 100 mM), and proteins were

378 eluted in the same buffer with 500 mM imidazole. For SDS-PAGE analysis, samples were prepared in  
379 Laemmli buffer, denatured 5 min at 95 °C, separated on 10% polyacrylamide gel and detected by  
380 Coomassie brilliant blue.

381

382 ***Dynamic light scattering (DLS)***

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

388

389 ***Negative stain electron microscopy observations of recombinant nucleoproteins***

390 Three microliters of sample were applied to the clean side of carbon on a carbon–mica interface and  
391 stained with 2% sodium silicotungstate. Micrographs were recorded on a FEI Tecnai T12 microscope  
392 operated at 120 kV with a Gatan Orius 1000 camera. Images were recorded at a nominal magnification  
393 of 23 000  $\times$  resulting in a pixel size of 2.8 Å.

394

395 ***Molecular modelling***

396 In the X-ray structures of HMPV N<sup>Nuc</sup> (PDB 5fvc) or N<sup>0</sup>-P (PDB 5fvd), the N<sub>NTD</sub> residues L110 to M113,  
397 corresponding to a long beta-hairpin close to the PCT binding site, are not modeled. We used Modeller  
398 (22) to first construct a complete model HMPV N<sub>NTD</sub> (residues 32 to 251) using as templates the X-ray  
399 structures of HMPV N<sup>Nuc</sup> (residues 32 to 109 and 114 to 251) and RSV N<sub>NTD</sub> (PDB 4uc6) for the missing  
400 beta-hairpin (residues 110 to 113). Next, using Modeller again, we generated 100 models of HMPV N<sub>NTD</sub>  
401 complexed with the last six residues of P (peptide I289-YQLI-M294 called P6) using the RSV N<sub>NTD</sub>  
402 structure in complex with the last two residues of RSV P (PDB 4uc9), *i.e.* D240-F241 (called P2), as  
403 template for HMPV P6 by aligning the last 2 residues of HMPV-P and RSV-P. To compensate for  
404 possible alternate conformations of exposed residues in apo and holo-form of N<sub>NTD</sub>, no template was  
405 used for R132 and R151 side-chains. Next, P6-N<sub>NTD</sub> models where M294 side-chain of P is pointing  
406 towards the N<sub>NTD</sub> hydrophobic pocket (by analogy with the binding of F241 of P on N RSV) were  
407 selected, to obtain an ensemble 41 P6-N<sub>NTD</sub> models. Finally, the models were submitted to a water

408 refinement procedure using the refinement interface of the HADDOCK2.2 server (23). Each input P6-  
409 N<sub>NTD</sub> complex model was refined 20 times to obtain 820 refined models. The best 400 models with the  
410 lowest HADDOCK-score were clustered based on similarity of the L292-I293-M294 residues of P, called  
411 P3 (I289-Y290-Q291 were not considered for clustering since they appear very floppy in the refined  
412 models and had no structural templates during homology modeling). Models were first superposed on  
413 their N<sub>NTD</sub> backbone atoms and the RMSD was computed on the backbone atoms of L292-I293-M294  
414 of P. Clustering was performed using the HADDOCK tool *cluster\_struct* (24) with a 2 Å cut-off and a  
415 minimal cluster size of 10.

416 In addition, the 820 P6-N<sub>NTD</sub> conformations generated by HADDOCK were profiled using InDeep  
417 software in order to select the most suitable N<sub>NTD</sub> conformations to bind P3. InDeep is a deep-learning  
418 model based on a FCN (Fully Convolutional Network) trained from known structures of protein  
419 complexes capable of predicting protein-protein interaction interfaces. Each N<sub>NTD</sub> pocket conformation  
420 is placed within a 3D grid composed of 1 Å<sup>3</sup> voxels. The model computes a probability value for each  
421 voxel representing its interaction propensity. The predictions were performed on the putative P6 binding  
422 site on N<sub>NTD</sub> surface. An “interactability-score” of the pocket is computed by taking the mean of the 150  
423 best voxels probabilities. This value represents the binding propensity of a given N<sub>NTD</sub> pocket  
424 conformation. Then, on the 200-top conformations with the highest “interactability-score” the  
425 hydrophobic channel of InDeep was applied in order to locate the most probable hydrophobic patches  
426 on the N<sub>NTD</sub> pocket where P3 hydrophobic side chains could interact. In addition, once a set most  
427 promising conformations was identified, a virtual alanine scanning was performed using the same tool  
428 but by mutating one by one the residues surrounding the patch of interaction of P3. The results are  
429 provided as a set of correlations, one for each residue, highlighting the impact of the simulated mutation  
430 on the interactability of the N<sub>NTD</sub>.

431

432

433

#### 434 **ACKNOWLEDGMENTS**

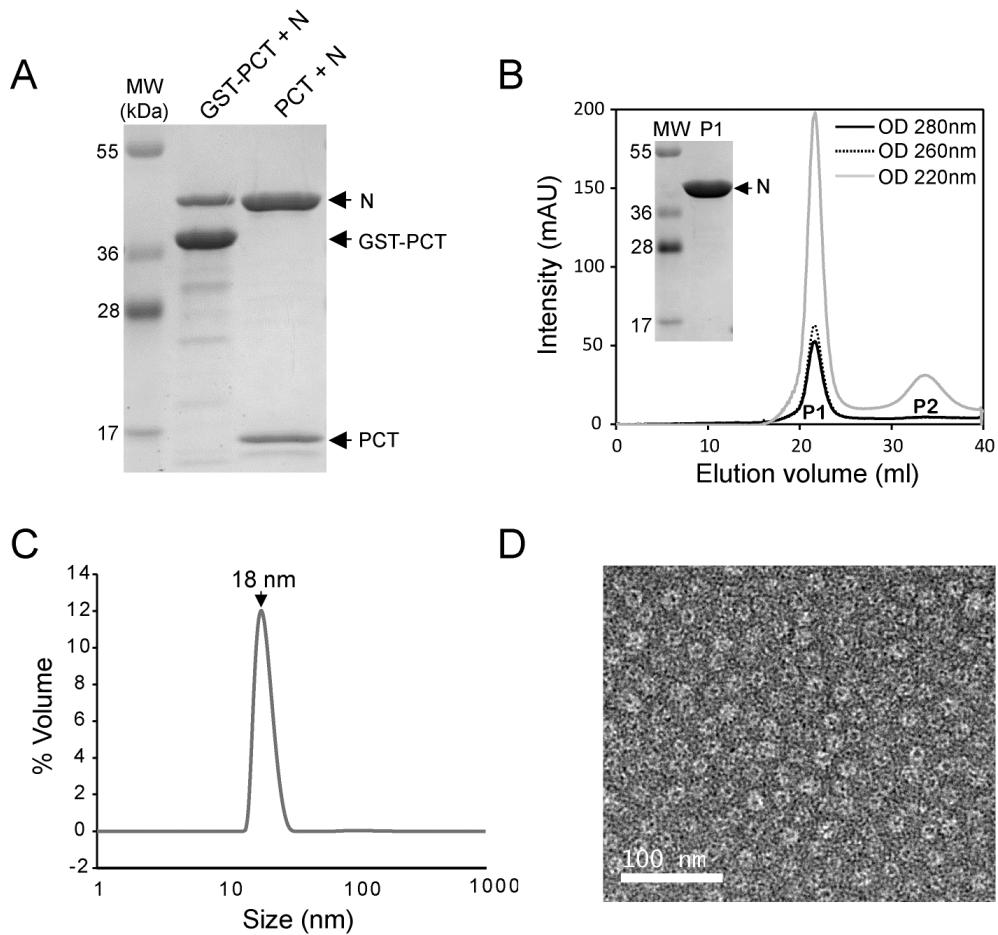
435 We thank the immunologist’s team of Seppic laboratory (Maisons-Alfort, France) and the animal facilities  
436 of Anses, Maisons-Alfort for rabbit immunization and collect of anti-N and anti-P antisera, P. Collins

437 (NIH, Boston) and S. Biacchesi (INRAE, Jouy-en-Josas) for the sequences of HMPV CAN97-83 N and  
438 P genes, B. van den Hoogen (Erasmus MC, Rotterdam) for the plasmids of HMPV NL99-1 strain.  
439 This work was carried out with the financial support of the French Agence Nationale de la Recherche,  
440 specific program ANR Decrisp n° ANR-19-CE11-0017.

441  
442 We declare that we have no conflicts of interest with the contents of this article.  
443 H.D., M.G. and J.-F.E. designed experiments. H.D., M.B. performed molecular, and cellular assays.  
444 B.B., L.C.R. and O.S. performed molecular simulations. I.G performed electron microscopy  
445 observations. C.-A.R. performed gel filtration and *in vitro* study. M.G. wrote the paper with contributions  
446 from all authors, and J.-F.E. and M.G. edited the manuscript. All authors commented on the manuscript.

447  
448  
449

450 **Figure legends**

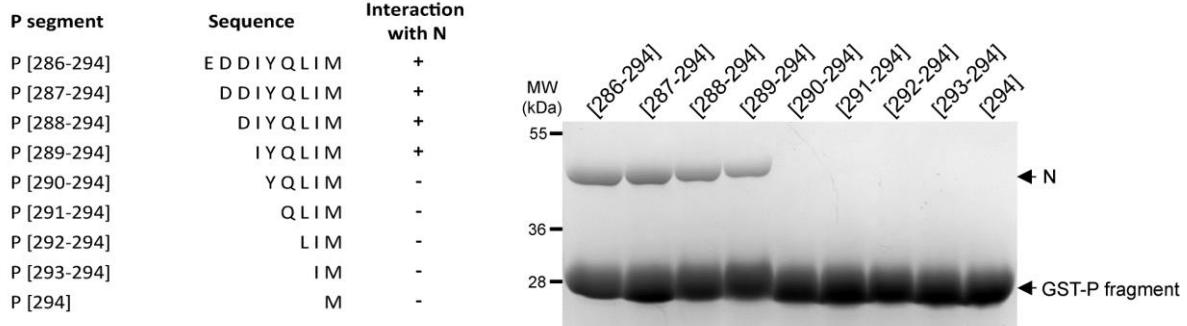


451

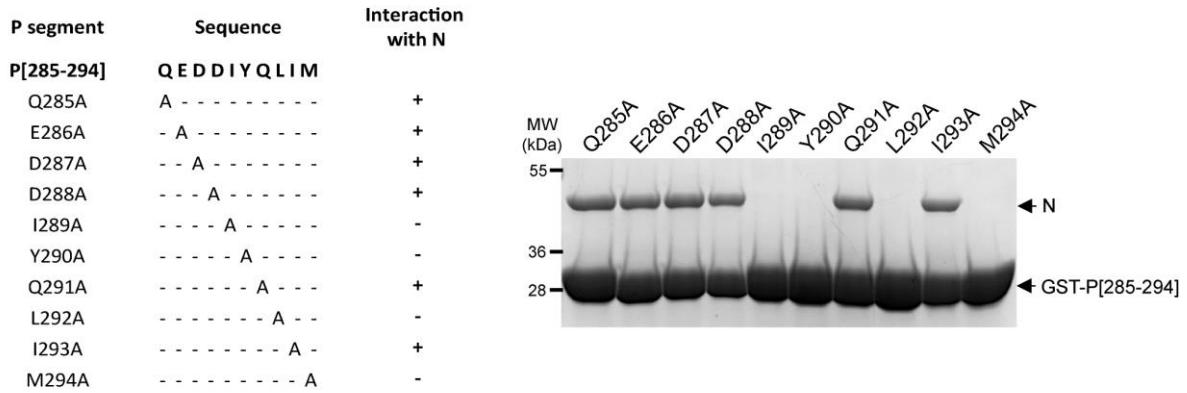
452 **Figure 1:** Purification of HMPV N-RNA rings using GST-PCT. A/ GST-PCT and N proteins were co-  
453 expressed in *E. coli*, followed by purification using the GST tag. The product of purification was analyzed  
454 by SDS-PAGE and Coomassie blue staining (left lane). The sample was then incubated in the presence  
455 of thrombin in order to cleave GST (remaining on beads) and isolate PCT-N in the supernatant (right  
456 lane), B/ Gel filtration profile of purified PCT-N complex. The curves corresponding to OD spectra at 220  
457 nm, 260 nm, and 280 nm are presented. P1 and P2 indicate the two peaks detected. The fractions  
458 corresponding to P1 were pooled and the sample was analyzed by SDS-PAGE colored with Coomassie  
459 blue, C/ Dynamic light scattering (DLS) analysis of the purified N protein, showing a homogenous peak  
460 at 18 nm, corresponding to N oligomers. D/ Image of purified HMPV N-RNA rings as observed by  
461 negative stain electron microscopy. Bar, 100 nm.

462

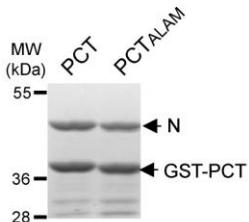
A



B



C



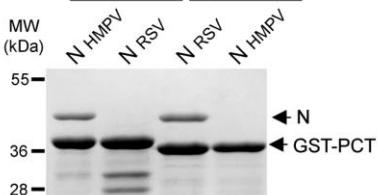
463

D

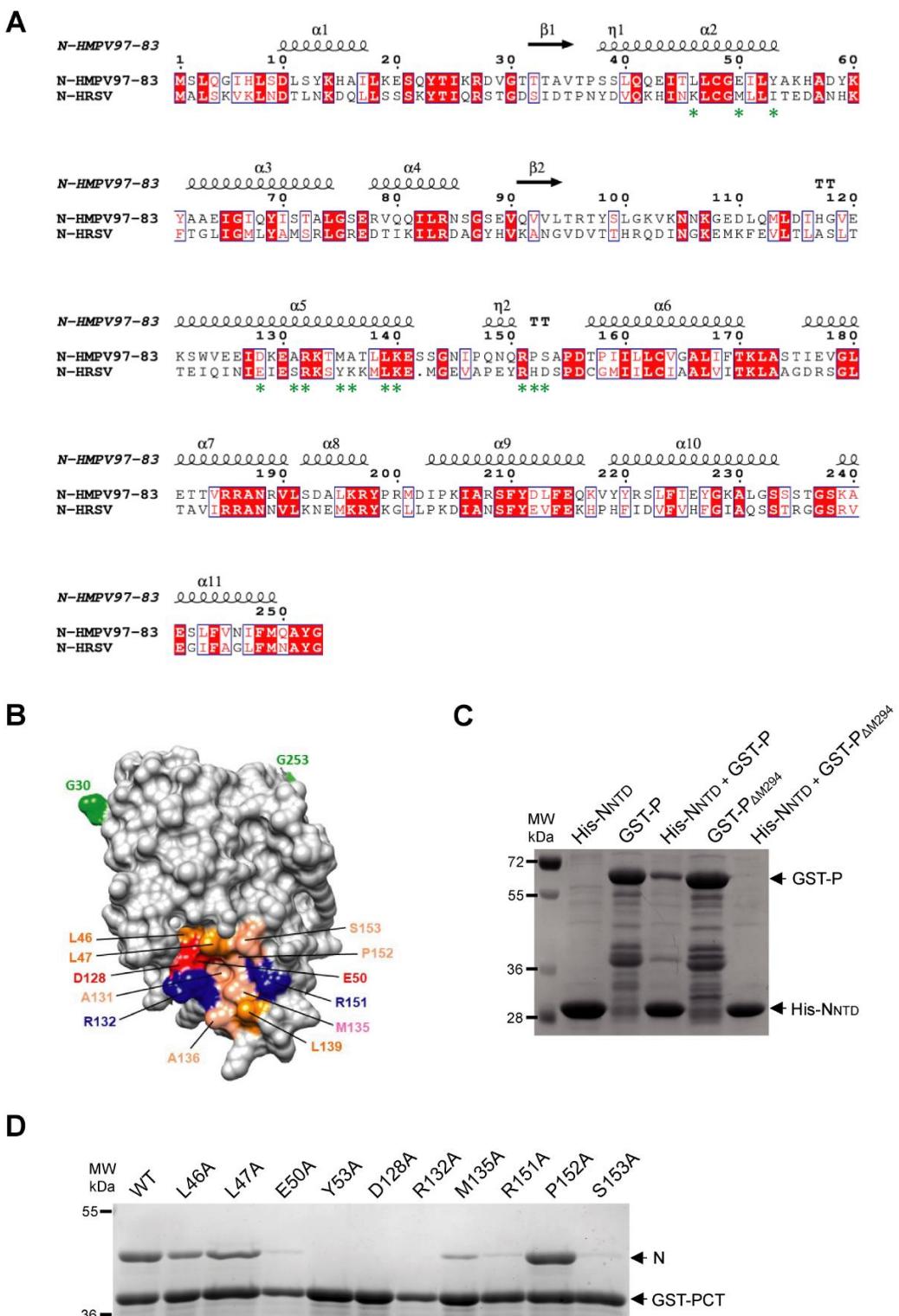
**HMPV**      **EDDIYQLIM**  
**RSV**      **DNDLISLEDF**

E

**HMPV**      **GST-PCT**      **RSV**      **GST-PCT**



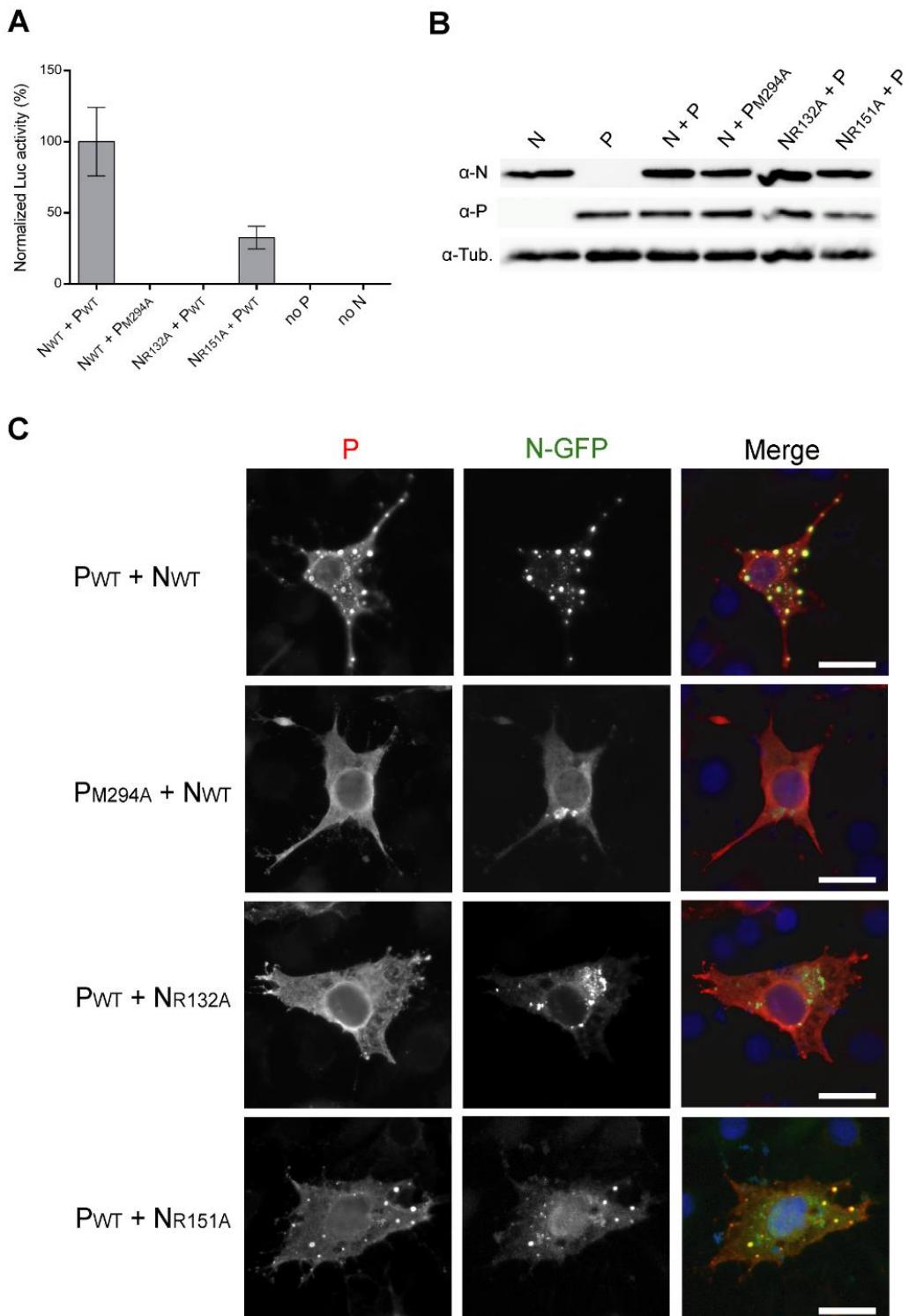
464 **Figure 2:** Identification of the residues of P involved in N binding. A/ GST-P fragments and B/ GST-  
465 P[285-294] (WT or mutants) (sequences indicated on the left) were co-expressed with N in bacteria,  
466 followed by purification using the GST tag. The products of purification were analyzed by SDS-PAGE  
467 and Coomassie blue staining (right). C/ GST-PCT or GST-PCT<sub>ALAM</sub> corresponding to the double  
468 substitution of residues Q291 and I293 residues of P by alanine were co-expressed with N. The products  
469 of purification by GST were analyzed by SDS-PAGE and Coomassie blue staining. D/ Sequence  
470 alignment of the 9 last C-terminal residues of HMPV and RSV P proteins. The residues critical for the  
471 interaction with the N protein are indicated in bold. E/ HMPV and RSV GST-PCT (residues P(200-294)  
472 for HMPV and P(161-241) for RSV) constructs were co-expressed with either HMPV or RSV N proteins.  
473 Copurification of N proteins by GST-PCT fragment was analyzed by SDS-PAGE and Coomassie blue  
474 staining.  
475



476 **Figure 3: Search for the PCT binding site on HMPV N.** A/ Amino acid sequence alignments between  
 477 the N-terminal domains of N (residues 1-252/253) proteins of HMPV (strain CAN 97-83) and human  
 478 RSV (RSV, strain Long VR-26). Invariant residues are highlighted in white font on a red background.  
 479 The secondary structure elements observed in the crystal structure of HMPV N protein (11) are indicated  
 480 above the sequence. Green asterisks below the sequence indicate the residues constituting the P  
 481 binding pocket of RSV N. Uniprot accession codes: NCAP\_HRSVA (Human RSV); A1DZS3\_9MONO  
 482 (Human MPV). Sequences were aligned with Clustal W and treated with ESPript 3. B/ Surface  
 483

484 representation of HMPV N<sub>NTD</sub> (from Gly30 to Gly 253, indicated in green) showing the potential P binding  
485 pocket, with acidic amino acids colored in red, basic residues in blue, and hydrophobic residues in  
486 orange, C/ GST-P or GSP-P<sub>ΔM294</sub> and N<sub>NTD</sub>-6xHis proteins were co-expressed in *E. coli*, followed by  
487 purification using the 6xHis tag. GST-P or GSP-P<sub>ΔM294</sub> alone were purified using the GST-tag. The  
488 products of purification were analyzed by SDS-PAGE and Coomassie blue staining, D/ GST-PCT and  
489 mutant N proteins were co-expressed in *E. coli*, followed by purification using the GST tag. The product  
490 of purification was analyzed by SDS-PAGE and Coomassie blue staining.  
491

492

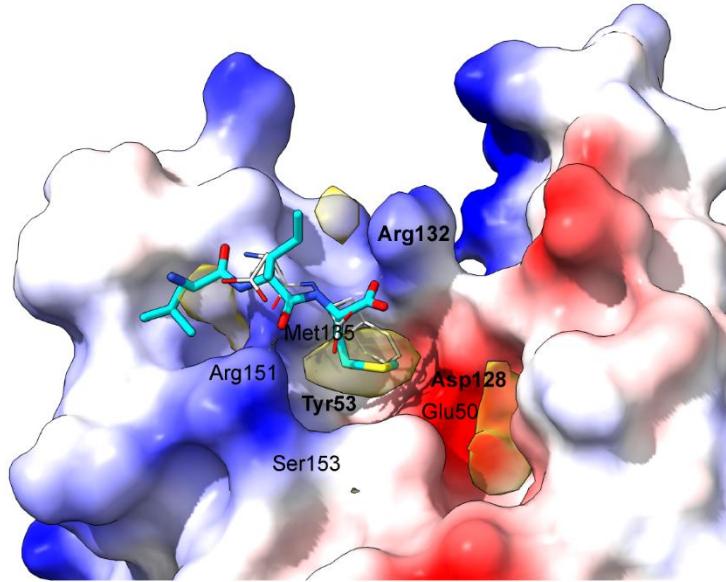


493

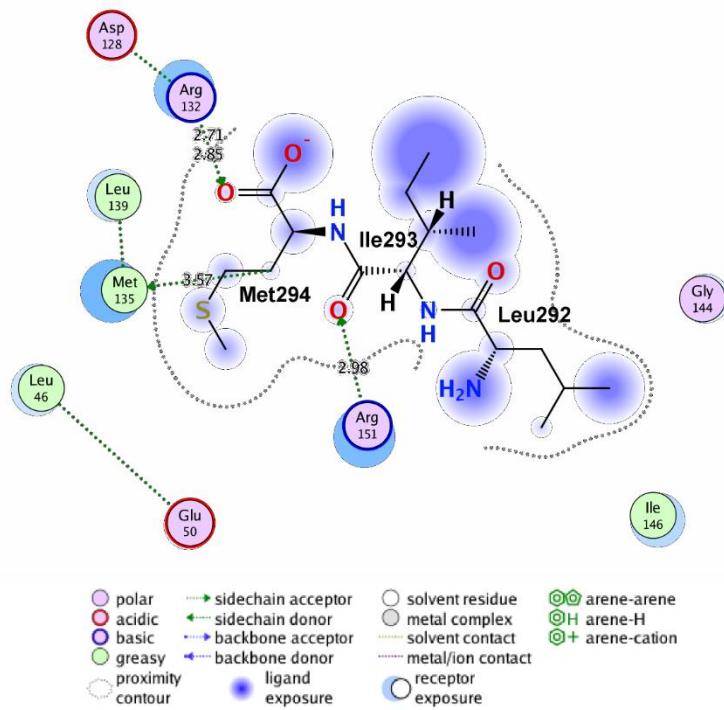
494 **Figure 4:** The residues involved in P-N<sup>Nuc</sup> interaction are critical for the polymerase activity and the  
495 formation of inclusion bodies. A/ HMPV polymerase activity in the presence of P and N mutants.  
496 BSRT7/5 cells were transfected with plasmids encoding the P (WT or M294A mutant), N (WT, or R132A  
497 and R151A mutants), L, and M2-1 proteins, the pGaussia/Firefly minigenome, together with pCMV-βGal  
498 for transfection standardization. Viral RNA synthesis was quantified by measuring the Firefly luciferase  
499 activity after cell lysis 24 h after transfection. Each luciferase minigenome activity value was normalized  
500 based on β-galactosidase expression and is the average of three independent experiments performed

501 in triplicate. Error bars represent standard deviations calculated based on three independent  
502 experiments made in triplicate. B/ Western blot showing the expression of N protein variants in BSRT7/5  
503 cells. C/ Impact of P and N mutations on the formation of HMPV cytoplasmic IBs. N-GFP and P (or  
504 mutants) proteins were coexpressed in BSRT7/5 cells; cells were then fixed 24 h post-transfection,  
505 labeled with anti-P antibody and the distribution of viral proteins was observed by fluorescence  
506 microscopy. Nuclei were stained with Hoechst 33342. Scale bars, 10  $\mu$ m.  
507

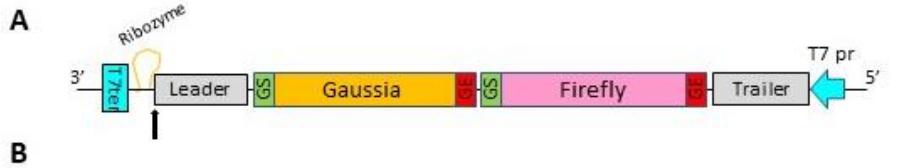
A



B



508  
509  
510 **Figure 5:** Model of HMPV P3/N<sub>NTD</sub> complex from HADDOCK refinement. A/ The N<sub>NTD</sub> protein is colored  
511 by electrostatic surface and the P3 peptide (residues L292-I293-M294 of P) is shown as sticks. The  
512 residues of N<sub>NTD</sub> critical for the interaction with P are indicated, with residues for which mutation to  
513 alanine abrogate P binding are labeled in bold. Yellow Surface shows InDeep hydrophobic channel  
514 prediction where hydrophobic contacts are expected to occur at the N<sub>NTD</sub> surface. B/ Interaction diagram  
515 between P3 peptide (Leu292-Ile293-Met294) and N<sub>NTD</sub>. Legend indicated below the diagram.  
516  
517



518

519 **Supplementary Figure 1:** Minigenome construct. A/ Schematic representation of the elements  
520 constituting the minigenome sequence. The sequences of the leader and trailer (grey), gene start (GS,  
521 green) and gene end (GE, red) are derived from HMPV strain CAN98-87. The full sequence is framed  
522 by T7 promoter (T7pr, blue arrow) and terminator (T7ter, blue) sequences. The black arrow indicates  
523 the cleavage site induced by the presence of the Ribozyme (yellow). B/ Nucleotide sequence of the  
524 HMPV minigenome. The elements are highlighted according to the color code used in A/. Sequences  
525 of Firefly and Gaussia luciferase are indicated by uppercase letters. The enzymatic restriction sites are  
526 indicated in bold.

527

528

529

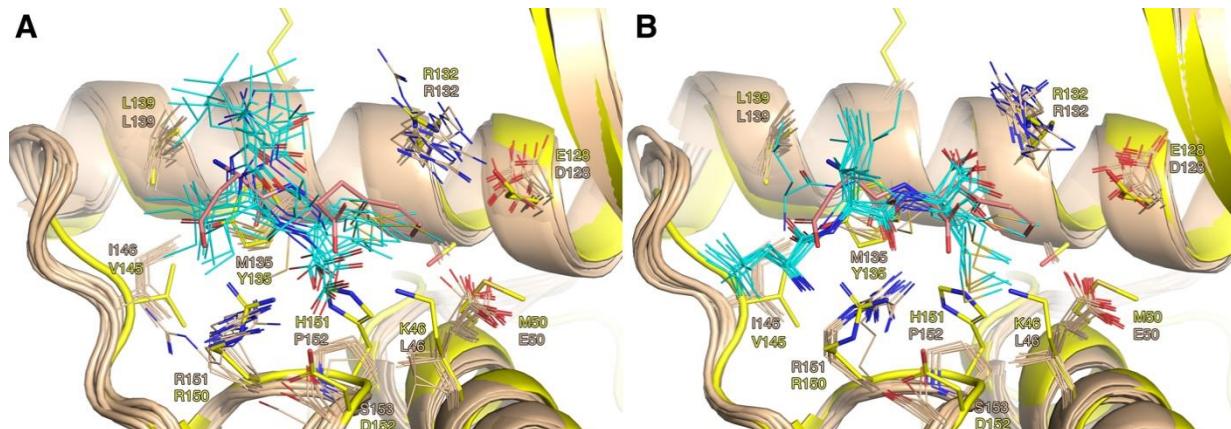
530 **Supplementary Table 1.** Cluster statistics of HADDOCK refined models obtained for HMPV P6/N<sub>NTD</sub>  
531 complex

	Cluster 1	Cluster 2	Cluster 3
HADDOCK score (all)	-47.3 +/- 4.3	-52.5 +/- 7.9	-45.5 +/- 3.4
HADDOCK score (top10 <sup>a</sup> )	-57.8 +/- 2.8	-69.3 +/- 4.0	-49.9 +/- 2.8
Cluster size	198	134	36
Cluster RMSD (Å)	0.95 +/- 0.28	0.99 +/- 0.29	1.12 +/- 0.56
RMSD from the overall lowest-energy structure Å	1.69 +/- 0.18	0.99 +/- 0.29	1.47 +/- 0.24
Van der Waals energy	-15.8 +/- 3.2	-19.3 +/- 3.4	-15.8 +/- 3.0
Electrostatic energy	-83.5 +/- 26.1	-87.1 +/- 20.0	-77.1 +/- 12.3
Buried Surface Area (Å <sup>2</sup> )	620.4 +/- 101.1	673.7 +/- 102.8	606.1 +/- 104.4
<i>Ionic/polar interactions frequencies</i>			
(N)R132--(P)M294 salt-bridge	2.0 %	76.9 %	2.8 %
(N)R151--(P)M294 salt-bridge	75.3 %	21.6 %	75.0 %
(N)R132-N <sub>η</sub> 1/2--(P)I293-O H-bond	0 %	-	0 %
(N)R151-N <sub>η</sub> 1/2--(P)I293-O H-bond	-	57.5 %	-
# Hydrophobic inter-molecular contacts <sup>b</sup>			
(P)L292	1.0 +/- 1.2	1.0 +/- 1.4	0.0 +/- 0.2
(P)I293	0.3 +/- 0.8	0.3 +/- 0.6	0.4 +/- 0.8
(P)M294	3.6 +/- 2.3	4.2 +/- 2.2	3.3 +/- 1.7

<sup>a</sup> 10 lowest-energy structure of the cluster ; <sup>b</sup> average number of contacts between Carbon or Sulfur atoms within 3.9 Å

532  
533

534



535  
536

537 **Supplementary Figure 2.** 10 best-scoring structures of clusters 1 (A) and 2 (B) from HADDOCK  
538 refinement of HMPV P3 (residues L292-I293-M294 of P)/N<sub>NTD</sub> complex superimposed with RSV  
539 P2/N<sub>NTD</sub>. N<sub>NTD</sub> is colored in beige (HMPV) or yellow (RSV), and HMPV P3 peptide in cyan (HMPV) or  
540 RSV P2 peptide (corresponding to residues D240-F241 of P) in pink. Side-chains of residues in contact  
541 with the P peptide in HADDOCK models are also shown, along with their RSV equivalent.

542

543

544

545

546

547

548

549 **References**

550

551 1. Pneumonia Etiology Research for Child Health Study G. 2019. Causes of severe  
552 pneumonia requiring hospital admission in children without HIV infection from Africa  
553 and Asia: the PERCH multi-country case-control study. *Lancet* 394:757-779.

554 2. Schildgen V, van den Hoogen B, Fouchier R, Tripp RA, Alvarez R, Manoha C, Williams J,  
555 Schildgen O. 2011. Human Metapneumovirus: lessons learned over the first decade.  
556 *Clin Microbiol Rev* 24:734-54.

557 3. van den Hoogen BG, de Jong JC, Groen J, Kuiken T, de Groot R, Fouchier RA, Osterhaus  
558 AD. 2001. A newly discovered human pneumovirus isolated from young children with  
559 respiratory tract disease. *Nat Med* 7:719-24.

560 4. Afonso CL, Amarasinghe GK, Banyai K, Bao Y, Basler CF, Bavari S, Bejerman N, Blasdell  
561 KR, Briand FX, Briese T, Bukreyev A, Calisher CH, Chandran K, Cheng J, Clawson AN,  
562 Collins PL, Dietzgen RG, Dolnik O, Domier LL, Durrwald R, Dye JM, Easton AJ, Ebihara  
563 H, Farkas SL, Freitas-Astua J, Formenty P, Fouchier RA, Fu Y, Ghedin E, Goodin MM,  
564 Hewson R, Horie M, Hyndman TH, Jiang D, Kitajima EW, Kobinger GP, Kondo H, Kurath  
565 G, Lamb RA, Lenardon S, Leroy EM, Li CX, Lin XD, Liu L, Longdon B, Marton S, Maisner  
566 A, Muhlberger E, Netesov SV, Nowotny N, et al. 2016. Taxonomy of the order  
567 Mononegavirales: update 2016. *Arch Virol* 161:2351-60.

568 5. van den Hoogen BG, Bestebroer TM, Osterhaus AD, Fouchier RA. 2002. Analysis of the  
569 genomic sequence of a human metapneumovirus. *Virology* 295:119-32.

570 6. Herfst S, de Graaf M, Schickli JH, Tang RS, Kaur J, Yang CF, Spaete RR, Haller AA, van  
571 den Hoogen BG, Osterhaus AD, Fouchier RA. 2004. Recovery of human  
572 metapneumovirus genetic lineages a and B from cloned cDNA. *J Virol* 78:8264-70.

573 7. Cifuentes-Munoz N, Brantlie J, Slaughter KB, Dutch RE. 2017. Human  
574 Metapneumovirus Induces Formation of Inclusion Bodies for Efficient Genome  
575 Replication and Transcription. *J Virol* 91:e01282-17.

576 8. Derdowski A, Peters TR, Glover N, Qian R, Utley TJ, Burnett A, Williams JV, Spearman  
577 P, Crowe JE. 2008. Human metapneumovirus nucleoprotein and phosphoprotein  
578 interact and provide the minimal requirements for inclusion body formation. *J Gen  
579 Virol* 89:2698-2708.

580 9. Galloux M, Risso-Ballester J, Richard CA, Fix J, Rameix-Welti MA, Eleouet JF. 2020.  
581 Minimal Elements Required for the Formation of Respiratory Syncytial Virus  
582 Cytoplasmic Inclusion Bodies In Vivo and In Vitro. *mBio* 11:e01202-20.

583 10. Leyrat C, Renner M, Harlos K, Grimes JM. 2013. Solution and crystallographic  
584 structures of the central region of the phosphoprotein from human metapneumovirus.  
585 *PLoS One* 8:e80371.

586 11. Renner M, Paesen GC, Grison CM, Granier S, Grimes JM, Leyrat C. 2017. Structural  
587 dissection of human metapneumovirus phosphoprotein using small angle x-ray  
588 scattering. *Sci Rep* 7:14865.

589 12. Pan J, Qian X, Lattmann S, El Sahili A, Yeo TH, Jia H, Cressey T, Ludeke B, Noton S,  
590 Kalocsay M, Fearn R, Lescar J. 2020. Structure of the human metapneumovirus  
591 polymerase phosphoprotein complex. *Nature* 577:275-279.

592 13. Tran TL, Castagne N, Bhella D, Varela PF, Bernard J, Chilmonczyk S, Berkenkamp S,  
593 Benhamo V, Grznarova K, Grosclaude J, Nespolous C, Rey FA, Eleouet JF. 2007. The  
594 nine C-terminal amino acids of the respiratory syncytial virus protein P are necessary

595 and sufficient for binding to ribonucleoprotein complexes in which six ribonucleotides  
596 are contacted per N protein protomer. *The Journal of general virology* 88:196-206.

597 14. Renner M, Bertinelli M, Leyrat C, Paesen GC, Saraiva de Oliveira LF, Huiskonen JT,  
598 Grimes JM. 2016. Nucleocapsid assembly in pneumoviruses is regulated by  
599 conformational switching of the N protein. *eLife* 5:e12627.

600 15. Tawar RG, Duquerroy S, Vonrhein C, Varela PF, Damier-Piolle L, Castagne N, MacLellan  
601 K, Bedouelle H, Bricogne G, Bhella D, Eleouet JF, Rey FA. 2009. Crystal structure of a  
602 nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science*  
603 326:1279-83.

604 16. Galloux M, Tarus B, Blazevic I, Fix J, Duquerroy S, Eleouet JF. 2012. Characterization of  
605 a viral phosphoprotein binding site on the surface of the respiratory syncytial  
606 nucleoprotein. *Journal of virology* 86:8375-87.

607 17. Ouizougun-Oubari M, Pereira N, Tarus B, Galloux M, Lassoued S, Fix J, Tortorici MA,  
608 Hoos S, Baron B, England P, Desmaele D, Couvreur P, Bontems F, Rey FA, Eleouet JF,  
609 Sizun C, Slama-Schwok A, Duquerroy S. 2015. A Druggable Pocket at the  
610 Nucleocapsid/Phosphoprotein Interaction Site of Human Respiratory Syncytial Virus.  
611 *Journal of virology* 89:11129-43.

612 18. Rincheval V, Lelek M, Gault E, Bouillier C, Sitterlin D, Blouquit-Laye S, Galloux M,  
613 Zimmer C, Eleouet JF, Rameix-Welti MA. 2017. Functional organization of cytoplasmic  
614 inclusion bodies in cells infected by respiratory syncytial virus. *Nature Communications*  
615 8:563.

616 19. Dominguez C, Boelens R, Bonvin AM. 2003. HADDOCK: a protein-protein docking  
617 approach based on biochemical or biophysical information. *J Am Chem Soc* 125:1731-  
618 7.

619 20. Sa JM, Piloto JV, Cilli EM, Tasic L, Fossey MA, Almeida FCL, Souza FP, Caruso IP. 2020.  
620 Hesperetin targets the hydrophobic pocket of the nucleoprotein/phosphoprotein  
621 binding site of human respiratory syncytial virus. *J Biomol Struct Dyn*  
622 doi:10.1080/07391102.2020.1835717:1-13.

623 21. Buchholz UJ, Finke S, Conzelmann KK. 1999. Generation of bovine respiratory syncytial  
624 virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture,  
625 and the human RSV leader region acts as a functional BRSV genome promoter. *Journal of virology*  
626 73:251-9.

627 22. Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial  
628 restraints. *J Mol Biol* 234:779-815.

629 23. van Zundert GCP, Rodrigues J, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond  
630 ASJ, van Dijk M, de Vries SJ, Bonvin A. 2016. The HADDOCK2.2 Web Server: User-  
631 Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* 428:720-725.

632 24. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. 1999. Peptide  
633 Folding: When Simulation Meets Experiment. *Angewandte Chemie International Edition*  
634 38: 236-240.