Characterization of the Interaction Domains between the Phosphoprotein and the Nucleoprotein of Human Metapneumovirus

Hortense Decool, Benjamin Bardiaux, Luis Checa Ruano, Olivier Sperandio, Jenna Fix, Irina Gutsche, Charles-Adrien Richard, Monika Bajorek, Jean-François Eléouët, Marie Galloux

Characterization of the interaction domains between the phosphoprotein and the nucleoprotein of human Metapneumovirus

Running title: Metapneumovirus N-P interaction

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

[1]Université Paris-Saclay, INRAE, UVSQ, VIM, 78350, Jouy-en-Josas, France.
[2] Institut Pasteur, Université de Paris, CNRS UMR3528, Unité de Bioinformatique Structurale, 28 rue du Dr Roux, 75015, Paris, France.
[3] Inserm U973 MTi, 25 rue Hélène Brion 75013, Paris, France.
[4] University of Grenoble Alpes, CNRS, CEA, CNRS, IBS, Grenoble, France.

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

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

ABSTRACT

Human metapneumovirus (HMPV) causes severe respiratory diseases in young children. The HMPV RNA genome is encapsidated by the viral nucleoprotein (N), forming an RNA-N complex (N^{Nuc}), which serves as template for genome replication and mRNA transcription by the RNA-dependent RNA polymerase (RdRp). The RdRp is formed by the association of the large polymerase subunit (L), which has RNA polymerase, capping and methyltransferase activities, and the tetrameric phosphoprotein (P). P plays a central role in the RdRp complex by binding to N^{Nuc} and L, allowing the attachment of the L polymerase to the N^{Nuc} template. During infection these proteins concentrate in cytoplasmic inclusion bodies (IBs) where viral RNA synthesis occurs. By analogy to the closely related pneumovirus respiratory syncytial virus (RSV), it is likely that the formation of IBs depends on the interaction between HMPV P and N^{Nuc}, which has not been demonstrated yet. Here, we finely characterized the binding P- N^{Nuc} interaction domains by using recombinant proteins, combined with a functional assay for the polymerase complex activity, and the study of the recruitment of these proteins to IBs by immunofluorescence. We show that the last 6 C-terminal residues of HMPV P are necessary and sufficient for binding to N^{Nuc}, that P binds to the N-terminal domain of N (N_{NTD}), and identified conserved N residues critical for the interaction. Our results allowed to propose a structural model for the HMPV P-N^{Nuc} interaction.

IMPORTANCE

Human metapneumovirus (HMPV) is a leading cause of severe respiratory infections in children but also affects human populations of all ages worldwide. Nowadays, no vaccine or efficient antiviral treatments are available for this pneumovirus. A better understanding of the molecular mechanisms involved in viral replication could help the design or discovery of specific antiviral compounds. In this work we have investigated the interaction between two major viral proteins involved in HMPV RNA synthesis, the N and P proteins. We finely characterized their domains of interaction, and identified a pocket on the surface of the N protein, a potential target of choice for the design of compounds interfering with N-P complexes and inhibiting viral replication.

INTRODUCTION

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

Together with respiratory syncytial virus (RSV), HMPV belongs to the *Pneumoviridae* family in the order *Mononegavirales* (4). HMPV is an enveloped virus that forms pleomorphic or filamentous virions. The virus genome is composed of a negative-sense single-stranded RNA of approximately 13.3 kb in size, which encodes 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 overlapping open reading frames (ORFs), encoding for M2-1 and M2-2 proteins which precise functions during HMPV viral cycle remain unclear. The HMPV genome is encapsidated by multiple copies of the nucleoprotein (N) forming helical nucleocapsids (N^{Nuc}). This N^{Nuc} serves as template for genome replication and mRNA transcription by the viral polymerase complex formed by the large polymerase subunit (L) and its main cofactor the phosphoprotein (P). After virus binding to the cell surface and virus-cell membrane fusion mediated by the fusion protein F, nucleocapsids are released into the cytoplasm. Replication and transcription of the viral genome take place within viro-induced cytoplasmic inclusions named inclusion bodies (IBs) (7). These structures can be observed upon expression of P and N proteins alone (8), and it was recently shown for RSV that the interaction between P and N^{Nuc} is critical for the formation of IBs (9).

Among the components of the polymerase complex, P plays a pivotal role as a cofactor of the L polymerase, but also as a molecular hub between viral partners. HMPV P, 294 amino acid residues in length, forms homo-tetramers. The atomic structure of the coiled-coil oligomerization domain (residues 171-194) was resolved by crystallography (10). Small angle X-ray scattering (SAXS) studies indicated that the flanking N-terminal (residues 1-170) and C-terminal (residues 195-294) regions (named PNT

and PCT respectively) are intrinsically disordered, some of them, such as residues 195-237, having α-helical propensity (10, 11). More recently, the structure of the LP complex was resolved by cryo-EM (12). It revealed a tentacular arrangement of P, with each of the four protomers adopting a distinct conformation, demonstrating a "folding-upon-partner-binding" mechanism. Depending on the protomer, the L-binding region involved regions encompassing residues 171-236, 172-217, 170-231 or 169-264. On the other hand, by binding to N, P mediates the attachment of the L protein to the N^{Nuc} template for viral RNA synthesis.

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

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

RESULTS

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

HMPV and RSV P proteins present a conserved structural organization, with a central tetramerization domain flanked by N- and C-terminal highly disordered regions (Fig. 1A). Co-expression of N with the C-terminal residues 161-241 of RSV P fused to GST in bacteria allowed the purification of ring shape N-RNA oligomers, showing that the C-terminal disordered region of RSV P (PCT) is involved in N^{Nuc} binding (13, 15). Similarly, when expressed alone in *E. coli*, HMPV N also forms decameric rings containing RNA (14). In a first attempt to characterize the N binding site of HMPV P, we thus co-expressed recombinant HMPV GST-PCT (residues 200-294) and N proteins (from CAN 97-83 strain) in *E. coli*. The purified complexes were analyzed by SDS-PAGE stained with Coomassie blue. As shown on figure.1B, N was co-purified with GST-PCT to > 95% homogeneity. The PCT was then separated from GST by thrombin cleavage (Fig. 1B) and the solubilized complex was analyzed by size exclusion chromatography, following optical density (OD) at 220, 260 and 280 nm. The elution profile showed a major peak (P1) with a OD_{260nm}/OD_{280nm} ratio > 1 and an apparent mass of ~ 500 kDa, in agreement with the expected size of N-RNA decamers (Fig. 1C). A second peak of lower intensity (P2), with an apparent mass of ~ 10 kDa was detected only at 220 nm. This peak should correspond to PCT, which does not present aromatic residues and a higher apparent molecular weight than predicted, likely due to the elongated shape of this fragment. The fractions of P1 peak were pooled, and sample analysis by SDS-PAGE stained with Coomassie blue showed the presence of a unique band corresponding to N (Fig. 1C). The separation of PCT and N upon gel filtration reveals the relatively low affinity of monomeric PCT for N. We then further analyzed the sample collected from the P1 peak by combining dynamic light scattering (DLS) and electron microscopy (EM) approaches. The DLS measurement profile confirmed the homogeneity of the sample, with a single peak at 18 nm (Fig 1D), while the observation of purified N by EM confirmed the presence N-rings (Fig 1E). These N-RNA rings are thus similar to those purified by Renner *et al*., who determined the 3D structure of N-RNA rings with a diameter close to 16 nm (14).

Altogether, these results show that the HMPV GST-PCT fusion protein is sufficient to interact with nucleocapsid-like N-RNA rings.

The last 6 residues of HMPV P constitute the minimum domain for N^{Nuc} binding

We then investigated the minimal domain of HMPV P involved in N-RNA rings binding. The 9 C-terminal residues of RSV P were previously shown to be sufficient to interact with N (13). Sequence alignment of the C-terminal extremity of pneumoviruses P proteins reveals the presence of two distinct motifs, separated by a highly acidic stretch in the case of Metapneumovirus P proteins (Fig. 2A). Of particular interest, the last 9 C-terminal residues of P proteins form a partially conserved motif. To investigate the role of these C-terminal residues of HMPV P in N binding, a series of GST fused peptides of 9 to 1 amino acid long derived from the C-terminal of P was generated. These constructs were co-expressed with N in *E.coli*, and their capacity to purify N was analyzed by SDS-PAGE and Coomassie blue staining. As shown on figure. 2B, the minimal sequence required for N purification corresponds to the 6 last C-terminal residues of P. In parallel, alanine scanning was performed with the GST-P[285-294] construct to characterize the residues of P involved in the interaction with N-RNA rings. Again, GST-P constructs and N were co-expressed in bacteria, followed by purification by Glutathione-Sepharose beads affinity, and co-purification of N was analyzed by SDS-PAGE and Coomassie blue staining. Only the four P substitutions I289A, Y290A, L292A and M294A abrogated the interaction with N (Fig. 2C). In order to further investigate the potential role of P residues Q291 and I293 in the interaction with N, a double substitution Q291A/I293A was inserted in GST-PCT. These two mutations did not affect N binding, confirming that these residues may not be directly involved in the interaction (Fig. 2D). These results confirm that the last residues of P are directly involved in the interaction with N, and suggest that HMPV P-N interaction mainly involves hydrophobic interactions. Of particular interest, the last 7 C-terminal residues of the P proteins of HMPV strains NL 99/1 (B1 subgroup), CAN 98-75 (B2 subgroup) and CAN 97-83 (A2 subgroup) are conserved, thus confirming their critical role. As both hydrophobic and acidic residues of RSV P were previously shown to be critical for the interaction with N (Fig. 2E), these results suggest that the binding of HMPV P on N differs from RSV and involves specific interactions. We thus tested the capacity of each PCT fragments to pulldown heterologous N proteins. Figure 2F shows that HMPV and RSV GST-PCT constructs did not allow to purify heterologous N. This last result confirms that HMPV and RSV N and P proteins cannot cross-interact, and that pneumoviruses P-N interactions are specific to each genus.

Search for the P binding site on the surface of HMPV N

Although HMPV and RSV P-N interactions have their own specificity and cannot cross-react, the strong homologies between N proteins suggest that the PCT binding site on N protein could be partially conserved. For RSV, the P binding site on N^{Nuc} is located at the surface of the helical nucleocapsid, on the N_{NTD} (16). The residues of RSV N_{NTD} involved in the interaction with P, which are partially conserved between HMPV and RSV (Fig. 3A), were shown to form a well-defined pocket constituted of hydrophobic and positively charged residues (16, 17). Corresponding residues of HMPV N_{NTD} also form a pocket exposed at the surface of N-RNA rings (Fig. 3B). In a first attempt to characterize the HMPV P binding domain on N, we thus tested the capacity of P to interact with the N_{NTD}. The N_{NTD} was expressed in *E. coli* and purified using a 6xHis tag, showing that this domain is soluble when expressed alone (Fig. 4A). In parallel, the GST-P or GST-$P_{\Delta M294}$ (deletion of the last C-terminal residue of P) proteins were expressed in bacteria, either alone and purified using the GST tag, or co-expressed with N_{NTD} followed by purification by 6xHis tag. Analysis of the samples by SDS-PAGE showed that N_{NTD} allowed to co-purify the wild type P, and that deletion of the last residue of P was sufficient to impair the interaction (Fig. 4A). Of note, the presence of numerous bands was observed in the samples containing GST-P alone or incubated in the presence of N_{NTD}. These bands migrated with a lower apparent mass weight compared to GST-P protein, suggesting a high propensity of recombinant GST-P to be cleaved and degraded. The HMPV N_{NTD} residues identified by analogy with RSV, i.e. residues L46, L47, E50, D128, R132, M135, R151, P152, and S153 (Fig. 4B), were then substituted by alanine. We first assessed the impact of these mutations on the expression and solubility of N proteins fused to a C-terminal poly-His tag expressed in *E. coli*. Analysis of purified N proteins by SDS -PAGE showed that none of these mutations induced a defect of production or solubility of N (Fig. 4C). The capacity of GST-P to co-purify mutated N proteins was then tested. Whereas substitutions L46A, L47A, and P152A did not disrupt N-P interaction P, E50A, D128A, R132A, M135A, R151A and S153A substitutions strongly affected or abrogated N-P interaction (Fig. 4D, upper panel). Of note, expression of N mutants in bacteria lysates was assessed by western blot and confirmed that all the N were expressed (Fig 4D, lower panel). These results thus revealed a critical role of hydrophobic residues, but also of negatively and positively charged residues of N_{NTD} in

P binding. Overall, these data confirm that the P binding site on N presents strong homologies between HMPV and RSV.

Validation of P-N^{Nuc} binding domains in eukaryotic cells

As P-N^{Nuc} interaction is required for viral polymerase activity, we then studied the impact of mutations on P and N on HMPV replication/transcription using a bicistronic minigenome, pGaussia/Firefly. This construct contains the Gaussia and Firefly luciferase genes under the control of HMPV gene start and gene stop sequences, as well as Leader and Trailer sequences of HMPV genome (Fig. S1). Briefly, BSRT7 cells were transfected with plasmids coding for N (WT or mutants), P (WT or mutants), L, and M2-1 proteins, the plasmid pGaussia/Firefly minigenome, and pSV-β-Gal (used to normalize transfection efficiencies). In this system, the expression of luciferase depends on the HMPV polymerase complex activity which can thus be quantified by luminescence measurement. As shown on figure 5A, mutations I289A, Y290A, L292A and M294A of P abrogated the polymerase activity, and only a residual activity was detected for the mutant Q291A. On the contrary, the P mutation I293A induced only a slight decrease of the polymerase activity. On the other hand, nearly all the mutations of on N induced a decrease of ≥ 50% of the polymerase activity compared to wild type N, with only mutations D128A and R132A totally impairing the activity (Fig 5A, right panel). For all these P and N mutants, the levels of protein expression were similar to those of wild type proteins as assessed by western blot (Fig.5B).

We then assessed if the impact of N and P mutations on polymerase activity could correlate with a defect in IBs morphogenesis. Knowing that co-expression of N and P is sufficient to induce the formation of pseudo-IBs (8), we used an approach based on fluorescent proteins expression. First, in order to investigate the impact of P mutations, cells were co-transfected with plasmids encoding P (WT or mutants) together with plasmids encoding N and GFP-N (at a ratio 2:1). A similar strategy used for RSV showed that, although GFP-N fusion protein had a dominant negative effect on RSV RNA synthesis, co-expression of GFP-N with wild type N partially restored polymerase activity (18). In our HMPV minigenome assay, co-expression of N with GFP-N at a ratio 2:1 allowed to rescue almost 70% of the polymerase activity (Fig. 5C). Cells transfected with P and N: GFP-N were fixed 24 h post-transfection and labelled with anti-P antibody. As shown on figure 6, although P labelling revealed a diffuse cytoplasmic localization of the protein, wild type P was found to co-localize with GFP-N into

cytoplasmic inclusions similar to IBs observed upon HMPV infection (7, 8). No impact on IBs morphology was detected upon alanine substitution of P residue I293. On the contrary, Q291A substitution seemed to induce a defect in IBs formation compared to wild type condition, as smaller IBs were observed, and a total loss of IBs formation was observed when mutating residues I289, Y290, L292 and M294 of P (Fig. 6B and Table 1).

In parallel, the impact of N mutations on IBs morphogenesis was investigated by co-transfecting N (WT or mutants) with P carrying BFP inserted between residues S61 and S62 (NL99-1 strain), in a region poorly conserved among HMPV strains and predicted as naturally intrinsically disordered. This P-BFP construct was functional since 65% of the HMPV polymerase activity was recovered when tested in our minigenome assay (Fig. 5D). Transfected cells were fixed before immunolabelling with anti-N antibody. Pseudo IBs were detected upon co-expression of P-BFP and WT N, but also upon alanine substitution of most of the residues (i.e. L46, L47, E50, M135, P252, and S153) (Fig. 7 and Table 1). As previously observed for the mutant Q291A of P, the mutation R151A of N seemed to induce a defect of IBs formation compared to the wild type condition, and no IBs were observed when expressing N mutants D128A and R132A (Fig. 7 and Table 1). Altogether, these results reveal a good correlation between the polymerase functioning and the capacity of N and P to interact and form IBs. Besides, these data confirm that the residues of P and N previously identified as critical for N^{Nuc}-P interaction *in vitro* are also critical for polymerase activity and the formation of IBs and in cells.

Molecular modeling of PCT-N_{NTD} interaction

To gain insight into the possible binding mode of P with N_{NTD}, we built models of the last 6 residues of P (peptide I289-YQLI-M294 named P6) bound to N, by analogy with the PCT-N_{NTD} complex of RSV. For RSV, only the last 2 residues of P (D240-F241) were resolved in the PCT-N_{NTD} X-ray structures (16), and are bound to a hydrophobic pocket of N similar to the one we characterized at the surface of HMPV N. Using comparative modeling and refinement with HADDOCK (19), we obtained three clusters of HMPV P6/N_{NTD} complex structures representing possible binding modes of P6 (Table S1). The first two clusters in terms of HADDOCK-score were the most populated and displayed similar Root-mean-square deviation (RMSD) among their members (considering only the last three residues of P, i.e. residues L292-I293-M294, called P3). In those two clusters, the C-terminal carboxyl group of M294 is involved in an ionic interaction with the guanidinium of either R131 or R152 side-chains of N.

Additionally, only L292 and M294 of P3 consistently display hydrophobic contacts with N. This result is consistent with our mutation data showing that I293 of P is not required for binding to N, contrary to L292 and M294. The 10 best-scoring structures from cluster 1 and 2 are shown in figure 9, superimposed with the RSV P2/N_{NTD} structure (residues D240-F241 of P). Owing to the best HADDOCK-score and non-bonded energies of cluster 2, we considered this cluster as most representative for the potential binding mode of HMPV P3 to N_{NTD}. Interestingly, more than half of the structures from this cluster display a hydrogen-bond between the backbone oxygen of I293 of P3 with the side-chain of R151 of N, while the side-chain of R132 of N forms a salt-bridge with the C-terminal carboxyl group of M294 of P. Thus, the interactions determined by molecular simulation correlate with the experimental data showing the importance of R132 and R151 of N and of M294 of P for PCT-N interaction.

The interactability propensity of each modeled N_{NTD} conformation was further profiled with our in-house tool InDeep (20) in order to select a subset of N_{NTD} models being more likely to interact with the P3 peptide. On this subset of N_{NTD} conformations, the hydrophobic channel of InDeep was also used to locate regions where hydrophobic moieties of P3 are expected to interact at the N_{NTD} surface. Three hydrophobic patches are detected on the N_{NTD} surface, that present a good match with hydrophobic moieties of P3 structures belonging to cluster 2 of HADDOCK: i) one occupied by L292 of P3; ii) another occupied by M294 of P3; and iii) the last one located in an unoccupied region of the N pocket (Fig.8A). Based on experimental and modeling information, the structure shown in figure 8 represents a consistent model of P3/N_{NTD} complex. The N_{NTD} conformation of this structure is predicted to be favorable to protein-protein interaction by InDeep, and the P3 conformation, belonging to cluster 2 of HADDOCK, binds to N_{NTD} with a combined effect of electrostatic/polar interactions and hydrophobic contacts in agreement with the mutational experiments. In addition, a procedure of virtual alanine scanning using the same InDeep approach confirmed some of the mutagenesis results, such as the importance of R132, M135, and R151 of N for the interaction with P. It also suggests that residue L139 of N may be key for the N_{NTD} patch for partners' binding.

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

DISCUSSION

Beyond its interest in better understanding the molecular mechanisms of virus replication, the study of protein-protein interactions required for Pneumoviruses' polymerase complex activity also presents interest for the development of specific antiviral compounds. The activity of this complex depends on multiple highly specific protein-protein interactions that have no cellular counterparts. Among those, N^{Nuc}-P interaction which is critical for viral polymerase functioning constitutes a singular potential target for antivirals. We have previously described RSV N^{Nuc}-P interaction at the molecular level, showing that the C-terminus of P binds to a well-defined pocket at the surface of N (9, 13, 17). We also made the proof of concept that small chemical compounds can target this pocket of N and inhibit viral replication in cells (16). More recently, Sá et al. showed that flavonoids derived molecules could also specifically bind to N pocket and block its interaction with P (21). Given the strong homologies between RSV and HMPV N and P proteins, it could be expected that a similar strategy could be used to block HMPV replication. It is noteworthy that ideally, the discovery of compounds that could inhibit replication of these two closely related viruses at the same time would be much more advantageous.

In the present study, we investigated the domains involved in HMPV P-N^{Nuc} interaction using pulldown assays between recombinant proteins (either truncated or mutated proteins) expressed in *E. coli*. As for RSV, the HMPV N protein was purified as nanorings formed by N and bacterial RNA, which can be considered as N^{Nuc}-like complexes (Fig 1). Our results revealed that the mechanisms of binding between P and N-RNA rings are similar between RSV and HMPV, the C-terminus of P interacting with a pocket located on the surface of the N N-terminal globular domain N_{NTD}. However, it appeared that P and N amino acid residues involved in these interactions have different contributions/roles between RSV and HMPV, and that the P-N^{Nuc} interaction is specific of each virus. Our results correlate with these data showing that pneumoviruses' polymerase complexes activity necessitates genus-specific combination of L, P, N, and M2-1 (22, 23). More specifically, we showed that the minimal domain of HMPV P required for N binding is constituted by the last 6 C-terminal residues of P which are mostly hydrophobic residues (IYQLIM), and determined that only Q291 and I293 are not critical for the interaction *in vitro* (Fig. 2). Analysis of the impact of mutations on polymerase activity and IBs' formation confirmed these results, but also revealed that residue Q291 of P is critical for the polymerase activity. This P fragment is thus shorter than for RSV for which N binding minimal domain involves the 9 C-terminal residues of P and is composed of both hydrophobic and acidic residues (13, 17). Interestingly, sequence analysis of the C-terminal residues of Orthopneumoviruses and

Metapneumoviruses P proteins highlights the conservation of residues critical for N-P interaction (Sup 10A), confirming a strong conservation of the interaction mode for each virus. We also identified 6 residues of N involved in P binding, *i.e*. residues E50, D128, R132, M135A, R151A and S153A, which form a well-defined pocket at the N_{NTD} surface (Fig 4). Using combined approaches, we revealed a correlation between the impact of the mutations on N-P interaction, and the level of polymerase activity (Table 1). However, only mutations of residues D128 and R132 abrogated IBs' formation and polymerase functioning. Although residues of N proteins involved in P binding are poorly conserved between RSV and HMPV, sequence analysis of N proteins of Orthopneumoviruses and Metapneumoviruses shows that 3 essential residues for N-P interaction present the same physico-chemical properties (132R/R, 135Y/M and 151R/R, Fig 10B). Of most interest, residues R132 and R150 of RSV N were also shown to play a key role in P-N^{Nuc} interaction, the aliphatic part of the R132 side chain being involved in the stacking of the aromatic ring of the residue F241 of P. In order to better understand the binding mode of HMPV P to N, molecular modeling was performed and the best structural models support that R132 and R151 of N and of M294 of P are critical for PCT-N interaction (Fig. 7). Nevertheless, molecular modeling was mainly targeted towards the binding of the last 3 C-terminal residues of P, and did not allow to reliably predict the interaction of all the 6 critical residues of P on N. As also observed in the crystal structures of RSV PCT/N_{NTD} complexes (16), the last C-terminal residue of P seems to drive the binding to a well-defined N_{NTD} pocket and this primary binding could allow transient contacts with upstream P residues outside of the N pocket. Remarkably, the acidic stretch in RSV PCT upstream of the main N binding domain was shown to contribute also to N binding (13, 24). Compared to RSV P, HMPV PCT is characterized by the presence of an 18 residues long sequence upstream the last 10 C-terminal residues, mainly composed of acidic residues (Fig. 2A), which could, as for RSV, contribute to HMPV P-N interaction. This sequence, located between L and N binding sites of P could also be involved in the modulation of the interaction of P with L and N.

In conclusion, our data confirm the strong structural homologies between HMPV and RSV P-N complexes but also highlight some singularities. Altogether, our study suggests that the HMPV P binding pocket at the N_{NTD} surface could represent a new target for the rational design of antivirals.

MATERIALS AND METHODS

Plasmid constructions

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

A HMPV minigenome plasmid containing Gaussia/Firefly luciferases was designed and synthesized by Genscript, cloned in pUC57 vector, and containing the trailer, leader, gene start (GS) and gene end (GE) sequences derived from HMPV CAN97-83 strain (see Fig S1). The first ORF of this pGaussia/Firefly minigenome codes for the Firefly luciferase and the second one codes for the Gaussia luciferase. The plasmids pP, pN, pL, and pM2-1 corresponding to the sequences of NL99-1 HMPV strain cloned into pCite vector. For expression of GFP-N fusion protein, the mGFP gene was amplified by PCR and cloned in frame at the 3' end of N at the EcoRI site of pN, using In-fusion HD cloning kit (Takara Bio). For expression of P-BFP fusion protein, a BamHI cleavage site was inserted between position S61 and S62 of P by site-directed mutagenesis. The BFP gene was then amplified by PCR and cloned at the BamHI site of modified pP, using In-fusion HD cloning kit (Takara Bio). Point mutations were introduced in the P and N sequences by site-directed mutagenesis using the Q5 site-directed mutagenesis Kit (New England Biolabs). Sequence analysis was carried out to check the integrity of all constructs. All the oligonucleotides sequences are available on request.

Antibodies

Antisera used in this study included polyclonal rabbit antisera raised against recombinant HMPV N and P expressed in bacteria. A mouse monoclonal anti-β-tubulin (Sigma) and secondary antibodies directed against mouse and rabbit Ig G coupled to HRP (P.A.R.I.S) were also used for immunoblotting. Secondary antibodies directed against rabbit Ig G coupled to Alexafluor-594 or to Alexafluor-488 (Invitrogen) were used for immunofluorescence experiments.

Cell culture and transfections

BHK-21 cells (clone BSRT7/5) constitutively expressing the T7 RNA polymerase (25) were grown in Dulbeco Modified Essential Medium (Lonza) supplemented with 10% fetal calf serum (FCS), 2 mM glutamine, and antibiotics. The cells were grown at 37˚C in 5% CO2, and transfected using Lipofectamine 2000 (Invitrogen) as described by the manufacturer.

Minigenome replication assay

BSRT7/5 cells at 90% confluence in 48-well dishes were transfected using Lipofectamine 2000 (Invitrogen) with a plasmid mixture containing 0.125 µg of pGaussia/Firefly minigenome, 0.125 µg of pN, 0.125 µg of pP (WT and mutants), 0.06 µg of pL, 0.03 µg of pM2-1, as well as 0.03 µg of pSV-β-Gal (Promega) to normalize transfection efficiencies. Cells were harvested at 24 h post-transfection and lysed in Firefly lysis buffer (30 mM Tris [pH 7.9], 10 mM MgCl2, 1 mM dithiothreitol (DTT), 1% [vol/vol] Triton X-100, and 15% [vol/vol] glycerol). The Firefly luciferase activity was determined for each cell lysate with an Infinite 200 Pro (Tecan, Männedorf, Switzerland) and normalized based on β-galactosidase (β-Gal) expression. Transfections were done in triplicate, and each independent transfection experiment was performed three times. The significance of the differences between wild type condition and mutants was determined using Student's t-test (GraphPad Prism software). Significance is represented: ***p < 0.001; ****p < 0.001; ns, not significant. For proteins expression analysis, cells were lysed in Laemmli buffer and analyzed by Western blotting (WB) using anti-N, anti-P, and anti-tubulin antibodies according to standard protocols.

Fluorescence microscopy

Immunofluorescence microscopy was performed with cells grown on coverslips and co-transfected with either pP (WT or mutant) and pN and pGFP-N (ratio pN:pGFP-N of 2:1), or pP-BFP and pN (WT or mutants). At 24 h post transfection, cells were fixed with 4% paraformaldehyde (PFA) for 25 min, made permeable, and blocked for 30 min with PBS containing 0.1% Triton X-100 and 3% bovine serum albumin (BSA). Cells were then successively incubated for 1 h at room temperature with primary (rabbit anti-N or rabbit anti-P) and secondary antibodies mixtures diluted in PBS containing 0.3% BSA. For nucleus labeling, cells were exposed to Hoechst 33342 stain (Invitrogen) during incubation with secondary antibodies. Coverslips were mounted with ProLong Gold antifade reagent (Invitrogen) and observed with a Nikon TE200 microscope equipped with a CoolSNAP ES2

(Photometrics, Tucson, USA) camera. Images were processed with ZEN software (Zeiss) and ImageJ software.

Expression and purification of recombinant proteins

E. coli BL21 bacteria (DE3) (Novagen) transformed with pGEX-P plasmids (WT or mutant) or pET-N_{NTD} were grown at 37°C for 8 hours in 100 ml of Luria Bertani (LB) medium containing 100µg/ml ampicillin or 50µg/ml kanamycin, respectively. Bacteria transformed with pET-N-derived plasmids together with pGEX-P derived plasmids were grown in LB medium containing ampicillin (100 µg/ml) and kanamycin (50 µg/ml). The same volume of LB was then added and protein expression was induced by adding 80µg/ml isopropyl-ß-D-thio-galactoside (IPTG) to the medium. The bacteria were incubated for 15 hours at 28°C and then harvested by centrifugation. For purification using the GST-tag, bacteria were re-suspended in lysis buffer (50 mM Tris-HCl pH 7.8, 60 mM NaCl, 1 mM EDTA, 2 mM DTT, 0.2% Triton X-100, 1 mg/ml lysozyme) supplemented with complete protease inhibitor cocktail (Roche), incubated for 1 hour on ice, sonicated, and centrifuged at 4°C for 30 min at 10,000g. Glutathione-Sepharose 4B beads (GE Healthcare) were added to clarified supernatants and incubated at 4°C for 3 hours. Beads were then washed two times in lysis buffer and three times in PBS 1X, then stored at 4°C in an equal volume of PBS. To isolate GST-free P fragments or N rings, beads containing bound proteins were incubated with thrombin (Novagen) overnight at 20°C. Purified recombinant N proteins were loaded onto a Superdex 200 16/30 column (GE Healthcare) and eluted in 20 mM Tris/HCl pH 8.5, 150 mM NaCl.

For purification of 6xHis-N and N_{NTD}-6xHis fusion proteins purification, bacterial pellets were re-suspended in lysis buffer (20 mM Tris-HCl pH8, 500 mM NaCl, 0.1% TritonX-100, 10 mM imidazole, 1 mg/ml lysozyme) supplemented with complete protease inhibitor cocktail (Roche). After sonication and centrifugation, lysates were incubated 30 min with chelating Sepharose Fast Flow beads charged with Ni^{2+} (GE Healthcare). Beads were then successively washed in the washing buffer (20 mM Tris-HCl, pH 8, 500 mM NaCl) containing increasing concentration of imidazole (25, 50, and 100 mM). For SDS-PAGE analysis of proteins on beads, samples were prepared in Laemmli buffer, denatured 5 min at 95 °C, separated on 10% plolyacrylamide gel and detected by Coomassie brilliant blue.

Dynamic light scattering (DLS)

Size measurement of purified N oligomers 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).

Negative stain electron microscopy observations of recombinant nucleoproteins

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 FEI 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.

Molecular modeling

In the X-ray structures of HMPV N-RNA rings (PDB 5fvc) or N^0-P (PDB 5fvd), the N_{NTD} residues L110 to M113, corresponding to a long beta-hairpin close to the PCT binding site, are not modeled. We used Modeller (26) to first construct a complete model HMPV N_{NTD} (residues 32 to 251) using as templates the X-ray structures of HMPV N-RNA rings (residues 32 to 109 and 114 to 251, PDB) and RSV N_{NTD} (PDB 4uc6) for the missing beta-hairpin (residues 110 to 113). Next, using Modeller again, we generated 100 models of HMPV N_{NTD} complexed with the last six residues of P (peptide I289-YQLI-M294 called P6) using the RSV N_{NTD} structure in complex with the last two residues of RSV P (PDB 4uc9), *i.e.* D240-F241 (called P2), as template for HMPV P6 by aligning the last 2 residues of HMPV-P and RSV-P. To compensate for possible alternate conformations of exposed residues in apo and holo-form of N_{NTD}, no template was used for R132 and R151 side-chains. Next, P6-N_{NTD} models where M294 side-chain of P is pointing towards the N_{NTD} hydrophobic pocket (by analogy with the binding of F241 of P on N RSV) were selected, to obtain an ensemble 41 P6-N_{NTD} models. Finally, the models were submitted to a water refinement procedure using the refinement interface of the HADDOCK2.2 server (27). Each input P6-N_{NTD} complex model was refined 20 times to obtain 820 refined models. The best 400 models with the lowest HADDOCK-score were clustered based on similarity of the L292-I293-M294 residues of P, called P3 (I289-Y290-Q291 were not considered for clustering since they appear very floppy in the refined models and had no structural templates during

homology modeling). Models were first superposed on their N_{NTD} backbone atoms and the RMSD was computed on the backbone atoms of L292-I293-M294 of P. Clustering was performed using the HADDOCK tool *cluster_struc* (28) with a 2Å cut-off and a minimal cluster size of 10.

In addition, the 820 P6-N_{NTD} conformations generated by HADDOCK were profiled using InDeep software (20) in order to select the most suitable N_{NTD} conformations to bind P3. InDeep is a deep-learning model based on a FCN (Fully Convolutional Network) trained from known structures of protein complexes capable of predicting protein-protein interaction interfaces (https://ippidb.pasteur.fr/targetcentric/) (20). Each N_{NTD} pocket conformation is placed within a 3D grid composed of 1 $Å^3$ voxels. The model computes a probability value for each voxel representing its interaction propensity. The predictions were performed on the putative P6 binding site on N_{NTD} surface. An "interactability-score" of the pocket is computed by taking the mean of the 150 best voxels probabilities. This value represents the binding propensity of a given N_{NTD} pocket conformation. Then, on the 200-top conformations with the highest "interactability-score" the hydrophobic channel of InDeep was applied in order to locate the most probable hydrophobic patches on the N_{NTD} pocket where P3 hydrophobic side chains could interact. In addition, once a set most promising conformations was identified, a virtual alanine scanning was performed using the same tool but by mutating one by one the residues surrounding the patch of interaction of P3. The results are provided as a set of correlations, one for each residue, highlighting the impact of the simulated mutation on the interactability of the N_{NTD}.

ACKNOWLEDGMENTS

We thank the immunologist's team of Seppic laboratory (Maisons-Alfort, France) and the animal facilities of Anses, Maisons-Alfort for rabbit immunization and collect of anti-N and anti-P antisera, P. Collins (NIH, Boston) and S. Biacchesi (INRAE, Jouy-en-Josas) for the sequences of HMPV CAN97-83 N and P genes, B. van den Hoogen (Erasmus MC, Rotterdam) for the plasmids of HMPV NL99-1 strain.

This work was carried out with the financial support of the French Agence Nationale de la Recherche, specific program ANR DecRisP n° ANR-19-CE11-0017.

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

H.D., M.G. and J.-F.E. designed experiments. H.D., M.B. performed molecular, and cellular assays. B.B., L.C.R. and O.S. performed molecular simulations. I.G performed electron microscopy observations. C.-A.R. performed gel filtration and *in vitro* study. 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. 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.
2. Schildgen V, van den Hoogen B, Fouchier R, Tripp RA, Alvarez R, Manoha C, Williams J, Schildgen O. 2011. Human Metapneumovirus: lessons learned over the first decade. Clin Microbiol Rev 24:734-54.
3. van den Hoogen BG, de Jong JC, Groen J, Kuiken T, de Groot R, Fouchier RA, Osterhaus AD. 2001. A newly discovered human pneumovirus isolated from young children with respiratory tract disease. Nat Med 7:719-24.
4. Afonso CL, Amarasinghe GK, Banyai K, Bao Y, Basler CF, Bavari S, Bejerman N, Blasdell 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.
5. van den Hoogen BG, Bestebroer TM, Osterhaus AD, Fouchier RA. 2002. Analysis of the genomic sequence of a human metapneumovirus. Virology 295:119-32.
6. Herfst S, de Graaf M, Schickli JH, Tang RS, Kaur J, Yang CF, Spaete RR, Haller AA, van den Hoogen BG, Osterhaus AD, Fouchier RA. 2004. Recovery of human metapneumovirus genetic lineages a and B from cloned cDNA. J Virol 78:8264-70.
7. Cifuentes-Munoz N, Branttie J, Slaughter KB, Dutch RE. 2017. Human Metapneumovirus Induces Formation of Inclusion Bodies for Efficient Genome Replication and Transcription. J Virol 91:e01282-17
8. Derdowski A, Peters TR, Glover N, Qian R, Utley TJ, Burnett A, Williams JV, Spearman P, Crowe JE. 2008. Human metapneumovirus nucleoprotein and phosphoprotein interact and provide the minimal requirements for inclusion body formation. J Gen Virol 89:2698-2708.
9. 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:e01202-20..
10. Leyrat C, Renner M, Harlos K, Grimes JM. 2013. Solution and crystallographic structures of the central region of the phosphoprotein from human metapneumovirus. PLoS One 8:e80371.
11. Renner M, Paesen GC, Grison CM, Granier S, Grimes JM, Leyrat C. 2017. Structural dissection of human metapneumovirus phosphoprotein using small angle x-ray scattering. Sci Rep 7:14865.

12. Pan J, Qian X, Lattmann S, El Sahili A, Yeo TH, Jia H, Cressey T, Ludeke B, Noton S, Kalocsay M, Fearns R, Lescar J. 2020. Structure of the human metapneumovirus polymerase phosphoprotein complex. Nature 577:275-279.
13. Tran TL, Castagne N, Bhella D, Varela PF, Bernard J, Chilmonczyk 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.
14. 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
15. 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.
16. Ouizougun-Oubari M, Pereira N, Tarus B, Galloux M, Lassoued S, Fix J, Tortorici MA, Hoos S, Baron B, England P, Desmaele D, Couvreur P, Bontems F, Rey FA, Eleouet JF, Sizun C, Slama-Schwok A, Duquerroy S. 2015. A Druggable Pocket at the Nucleocapsid/Phosphoprotein Interaction Site of Human Respiratory Syncytial Virus. J Virol 89:11129-43.
17. 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.
18. 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. Nature Communications 8:563.
19. Dominguez C, Boelens R, Bonvin AM. 2003. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc 125:1731-7.
20. Vincent Mallet, Luis Checa Ruano, Alexandra Moine Franel, Michael Nilges, Karen Druart, Guillaume Bouvier, Olivier Sperandio. 2021. InDeep : 3D fully convolutional neural networks to assist in silico drug design on protein-protein interactions. bioRxiv doi: https://doi.org/10.1101/2021.07.28.453974
21. Sa JM, Piloto JV, Cilli EM, Tasic L, Fossey MA, Almeida FCL, Souza FP, Caruso IP. 2020. Hesperetin targets the hydrophobic pocket of the nucleoprotein/phosphoprotein binding site of human respiratory syncytial virus. J Biomol Struct Dyn 20:1-13.
22. de Graaf M, Herfst S, Schrauwen EJA, Choi Y, van den Hoogen BG, Osterhaus A, Fouchier RAM. 2008. Specificity and functional interaction of the polymerase complex proteins of human and avian metapneumoviruses. J Gen Virol 89:975-983.
23. Marriott AC, Smith JM, Easton AJ. 2001. Fidelity of leader and trailer sequence usage by the respiratory syncytial virus and avian pneumovirus replication complexes. J Virol 75:6265-72.
24. Shapiro AB, Gao N, O'Connell N, Hu J, Thresher J, Gu RF, Overman R, Hardern IM, Sproat GG. 2014. Quantitative investigation of the affinity of human respiratory syncytial virus phosphoprotein C-terminus binding to nucleocapsid protein. Virol J 11:191.
25. Buchholz UJ, Finke S, Conzelmann KK. 1999. Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader region acts as a functional BRSV genome promoter. J Virol 73:251-9.
26. Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779-815.
27. van Zundert GCP, Rodrigues J, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bonvin A. 2016. The HADDOCK2.2 Web

Server: User-Friendly Integrative Modeling of Biomolecular Complexes. J Mol Biol 428:720-725.
28. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. 1999. Peptide Folding: When Simulation Meets Experiment. Angewandte Chemie International Edition 38: 236-240.
29. Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. Genome Res 14:1188-90).

609 **Figure legends**

610

611 **Figure 1:** Purification of HMPV N-RNA rings using GST-PCT. A/ Schematic representation of HMPV
612 and RSV P proteins. The oligomerization domain is represented by a rectangle and disordered N- and
613 C-terminal domains (PNT and PCT respectively) by lines. PCT is indicated. Residues flanking the
614 structural domains are indicated by numbers. B/ GST-PCT and N proteins were co-expressed in *E.*
615 *coli*, followed by purification using the GST tag. The product of purification was analyzed by SDS-
616 PAGE and Coomassie blue staining (left lane). The sample was then incubated in the presence of
617 thrombin in order to cleave GST (remaining on beads) and isolate PCT-N in the supernatant (right
618 lane), C/ Gel filtration profile of purified PCT-N complex. The curves corresponding to OD spectra at
619 220 nm, 260 nm, and 280 nm are presented. P1 and P2 indicate the two peaks detected. The
620 fractions corresponding to P1 were pooled and the sample was analyzed by SDS-PAGE colored with
621 Coomassie blue, D/ Dynamic light scattering (DLS) analysis of the purified N protein, showing a
622 homogenous peak at 18 nm, corresponding to N oligomers. E/ Image of purified HMPV N-RNA rings
623 as observed by negative stain electron microscopy. Bar, 100 nm.

624

625 **Figure 2:** Identification of the residues of P involved in N binding. A/ Sequence alignment of
626 *Pneumoviridae* P C-terminus (PCT), from the Orthopneumovirus and Metapneumovirus genus: hRSV-
627 A/-B (human RSV strains A and B, Uniport entriesPHOSP_HRSVA and PHOSP_HRSVB,
628 respectively), oRSV (ovine RSV, Uniprot entry PHOSP_ORSVW), bRSV (bovine RSV, Uniprot entry
629 PHOSP_BRSVA), mPV (murine pneumonia virus, Uniprot entry PHOSP_MPV15), aMPV-A/-C (avian
630 metapneumovirus type A and C, Uniprot entries Q65032 and PHOSP_AMPV1, respectively) and
631 hMPV (human metapneumovirus strain CAN97-83, Uniprot entry PHOSP_HMPVC). Sequence
632 numbering and secondary structures are indicated on top for hRSV and at the bottom for hMPV. B/
633 GST-P fragments and C/ GST-P[285-294] (WT or mutants) (sequences indicated on the left) were co-
634 expressed with N in bacteria, followed by purification using the GST tag. The products of purification
635 were analyzed by SDS-PAGE and Coomassie blue staining (right). D/ GST-PCT or GST-PCT$_{ALAM}$
636 corresponding to the double substitution of residues Q291 and I293 residues of P by alanine were co-
637 expressed with N. The products of purification by GST were analyzed by SDS-PAGE and Coomassie
638 blue staining. E/ Sequence alignment of the 9 last C-terminal residues of HMPV and RSV P proteins.

639 The residues critical for the interaction with the N protein are indicated in red bold. F/ HMPV and RSV
640 GST-PCT (residues P[200-294] for HMPV and P[161-241] for RSV) constructs were co-expressed
641 with either HMPV or RSV N proteins. Copurification of N proteins by GST-PCT fragment was analyzed
642 by SDS-PAGE and Coomassie blue staining
643
644 **Figure 3:** Search for the PCT binding site on HMPV N. A/ Amino acid sequence alignments between
645 the N-terminal domains of N (residues 1-252/253) proteins of HMPV (strain CAN 97-83) and human
646 RSV (HRSV, strain Long VR-26). Invariant residues are highlighted in white font on a red background.
647 The secondary structure elements observed in the crystal structure of HMPV N protein (11) are
648 indicated above the sequence. Green asterisks below the sequence indicate the residues constituting
649 the P binding pocket of RSV N. Uniprot accession codes: NCAP_HRSVA (Human RSV);
650 A1DZS3_9MONO (Human MPV). Sequences were aligned with Clustal W and treated with ESPript 3.
651 B/ Side view (left) and bottom view (right) of the HMPV N-RNA rings (PDB accession number 5fvc)
652 showing one N monomer coloured according to domains: N-arm in blue, C-arm in green, and N- and
653 C-terminal domains (N_{NTD} and N_{CTD}, respectively) are in yellow and red. The residues identified by
654 analogy with RSV as involved in P binding on N_{NTD} are indicated by black spheres.
655
656 **Figure 4:** Identification of the residues of N_{NTD} involved in P binding. A/ GST-P or GST-$P_{\Delta M294}$ and
657 N_{NTD}-6xHis proteins were co-expressed in *E. coli*, followed by purification using the 6xHis tag. GST-P
658 or GST-$P_{\Delta M294}$ alone were purified using the GST-tag. The products of purification were analyzed by
659 SDS-PAGE and Coomassie blue staining, B/ Surface representation of HMPV N_{NTD} (from Gly30 to Gly
660 253, indicated in green) showing the potential P binding pocket, with acidic amino acids colored in red,
661 basic residues in blue, and hydrophobic residues in orange, C/ The impact of N mutations on the
662 expression and solubility of the protein was assessed by expressing 6xHis-N protein (WT or mutants)
663 in *E. coli*, followed by purification and SDS-PAGE analysis. D/ GST-PCT and mutant N proteins were
664 co-expressed in *E. coli*, followed by purification using the GST tag. The products of purification were
665 analyzed by SDS-PAGE and Coomassie blue staining (upper panel), and the expression of N proteins
666 in bacteria was validated by western blot (lower panel).
667
668 **Figure 5:** The residues involved in P-N-RNA rings interaction are critical for the polymerase activity. A/
669 HMPV polymerase activity in the presence of P and N mutants. BSRT7/5 cells were transfected with

plasmids encoding the P (WT or mutants), N (WT, or mutants), L, and M2-1 proteins, the pGaussia/Firefly minigenome, together with pCMV-βGal for transfection standardization. Viral RNA synthesis was quantified by measuring the Gaussia luciferase activity after cell lysis 24 h after transfection. Each luciferase minigenome activity value was normalized based on β-galactosidase expression and is the average of three independent experiments performed in triplicate. Error bars represent standard deviations calculated based on two independent experiments made in quadruplicate. B/ Western blot showing the expression of P and N proteins (WT and mutants) in BSRT7/5 cells. C and D/ HMPV polymerase activity in the presence of pN:pGFP-N at a ratio 2:1, or P-BFP , respectively. Error bars represent standard deviations calculated based on two independent experiments made in triplicate. The statistical significance of differences was calculated using Student's *t* test. ***p < 0.001; ****p < 0.001; ns, not significant.

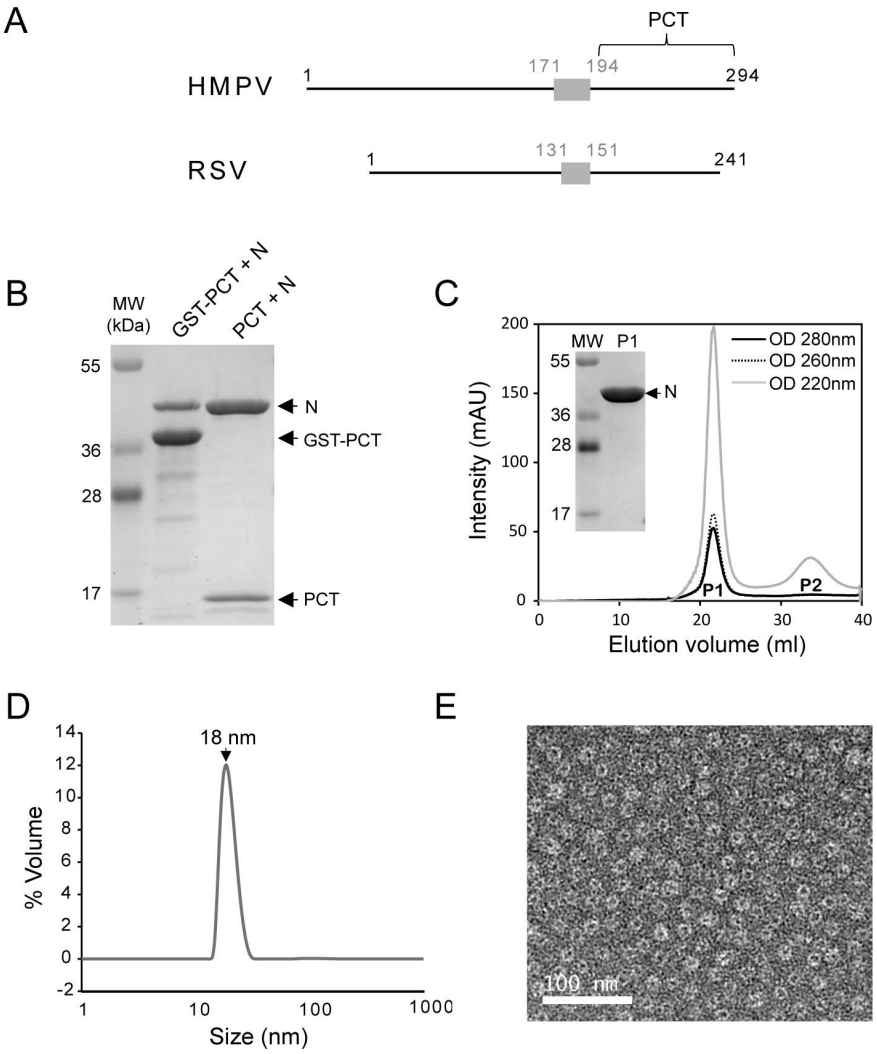
Figure 6: Impact of P mutations on the formation of HMPV cytoplasmic IBs. BSRT7/5 cells transfected with plasmids encoding for P (WT or mutants) and N + GFP-N (ratio 2:1) proteins were fixed 24 h post-transfection, labeled with anti-P antibody, and the distribution of viral proteins was observed by fluorescence microscopy. Nuclei were stained with Hoechst 33342. Scale bars, 10 µm.

Figure 7: Impact of N mutations on the formation of HMPV cytoplasmic IBs. BSRT7/5 cells transfected with plasmids encoding for P-BFP and N (WT or mutants) proteins were fixed 24 h post-transfection, labeled with anti-N antibody, and the distribution of viral proteins was observed by fluorescence microscopy. Scale bars, 10 µm.

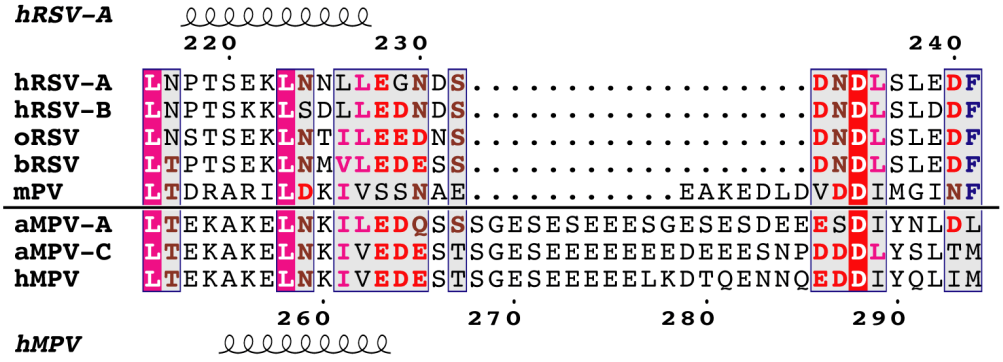
Figure 8: Model of HMPV P3/N_{NTD} complex from HADDOCK refinement. A/ The N_{NTD} protein is colored by electrostatic surface and the P3 peptide (residues L292-I293-M294 of P) is shown as sticks. The residues of N_{NTD} critical for the interaction with P are indicated, with residues for which mutation to alanine abrogate P binding are labeled in bold. Yellow Surface shows InDeep hydrophobic channel prediction where hydrophobic contacts are expected to occur at the N_{NTD} surface. B/ Interaction diagram between P3 peptide (Leu292-Ile293-Met294) and N_{NTD}. Legend indicated below the diagram.

Figure 9. 10 best-scoring structures of clusters 1 (A) and 2 (B) from HADDOCK refinement of HMPV P3 (residues L292-I293-M294 of P)/N_{NTD} complex superimposed with RSV P2/N_{NTD}. N_{NTD} is colored in beige (HMPV) or yellow (RSV), and HMPV P3 peptide in cyan (HMPV) or RSV P2 peptide (corresponding to residues D240-F241 of P) in pink. Side-chains of residues in contact with the P peptide in HADDOCK models are also shown, along with their RSV equivalent.

Figure 10. Conservation of pneumoviruses N and P residues. A/ Sequence logo of the last 9 C-terminal residues of P of Metapneumovirus (left) and Orthopneumovirus (right) sequences. P sequences were collected from the Interpro entry IPR003487 (464 unique sequences) and logo were created from non-redundant peptides sequences of the last 9 amino acids. B/ Sequence logo of the residues forming the P-binding site on N of the Metapneumovirus (left) and Orthopneumovirus (right) genus. N sequences were collected from the Interpro entry IPR004930 (359 unique sequences) and aligned. Classification as Ortho- and Metapneumovirus was based on Uniprot taxonomy (taxon ids 1868215 for OrthoPV and 162387 for MetaPV). Residue numbers are shown at the bottom, and grey markers indicate residues identified as essential for P-N interaction. Logo were generated with WebLogo (29).

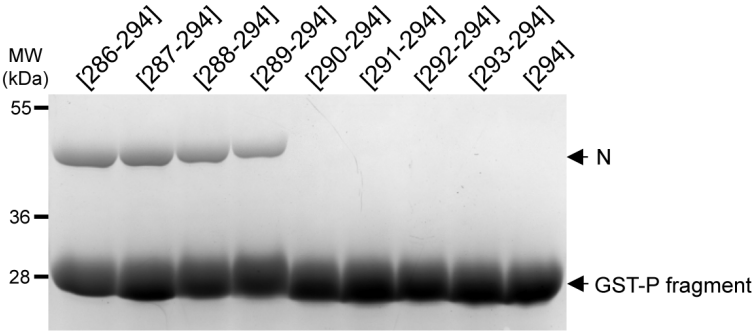


A



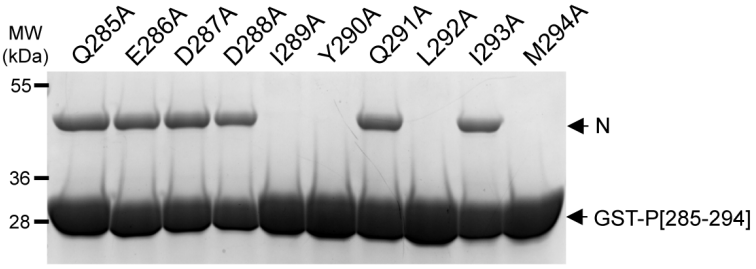
B

P segment	Sequence	Interaction with N
P [286-294]	E D D I Y Q L I M	+
P [287-294]	D D I Y Q L I M	+
P [288-294]	D I Y Q L I M	+
P [289-294]	I Y Q L I M	+
P [290-294]	Y Q L I M	-
P [291-294]	Q L I M	-
P [292-294]	L I M	-
P [293-294]	I M	-
P [294]	M	-

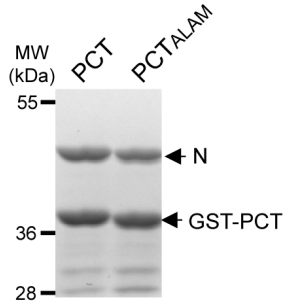


C

P segment	Sequence	Interaction with N
P[285-294]	**Q E D D I Y Q L I M**	
Q285A	A - - - - - - - - -	+
E286A	- A - - - - - - - -	+
D287A	- - A - - - - - - -	+
D288A	- - - A - - - - - -	+
I289A	- - - - A - - - - -	-
Y290A	- - - - - A - - - -	-
Q291A	- - - - - - A - - -	+
L292A	- - - - - - - A - -	-
I293A	- - - - - - - - A -	+
M294A	- - - - - - - - - A	-

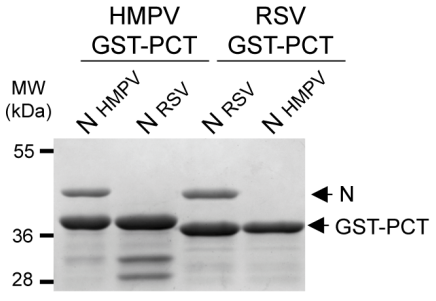


D

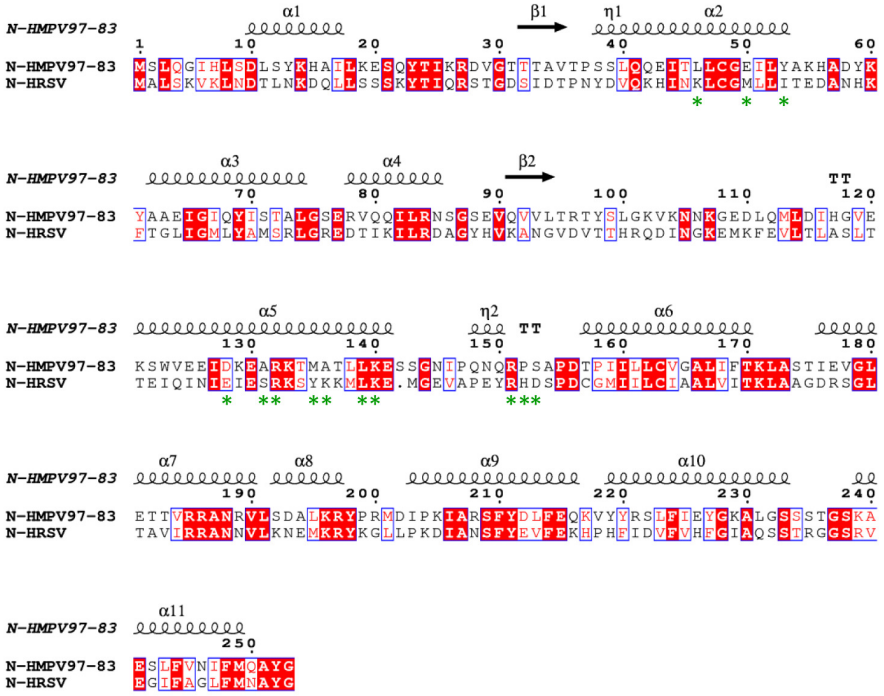


E

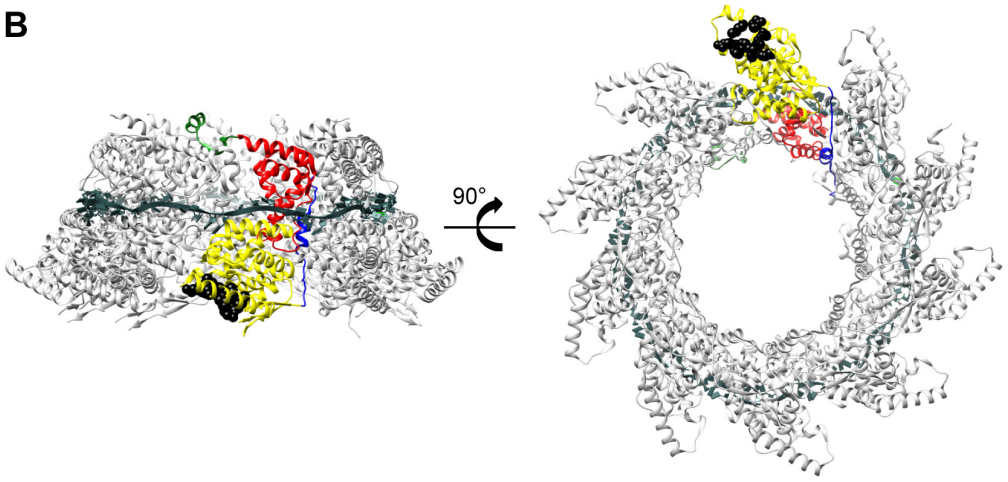
F



A



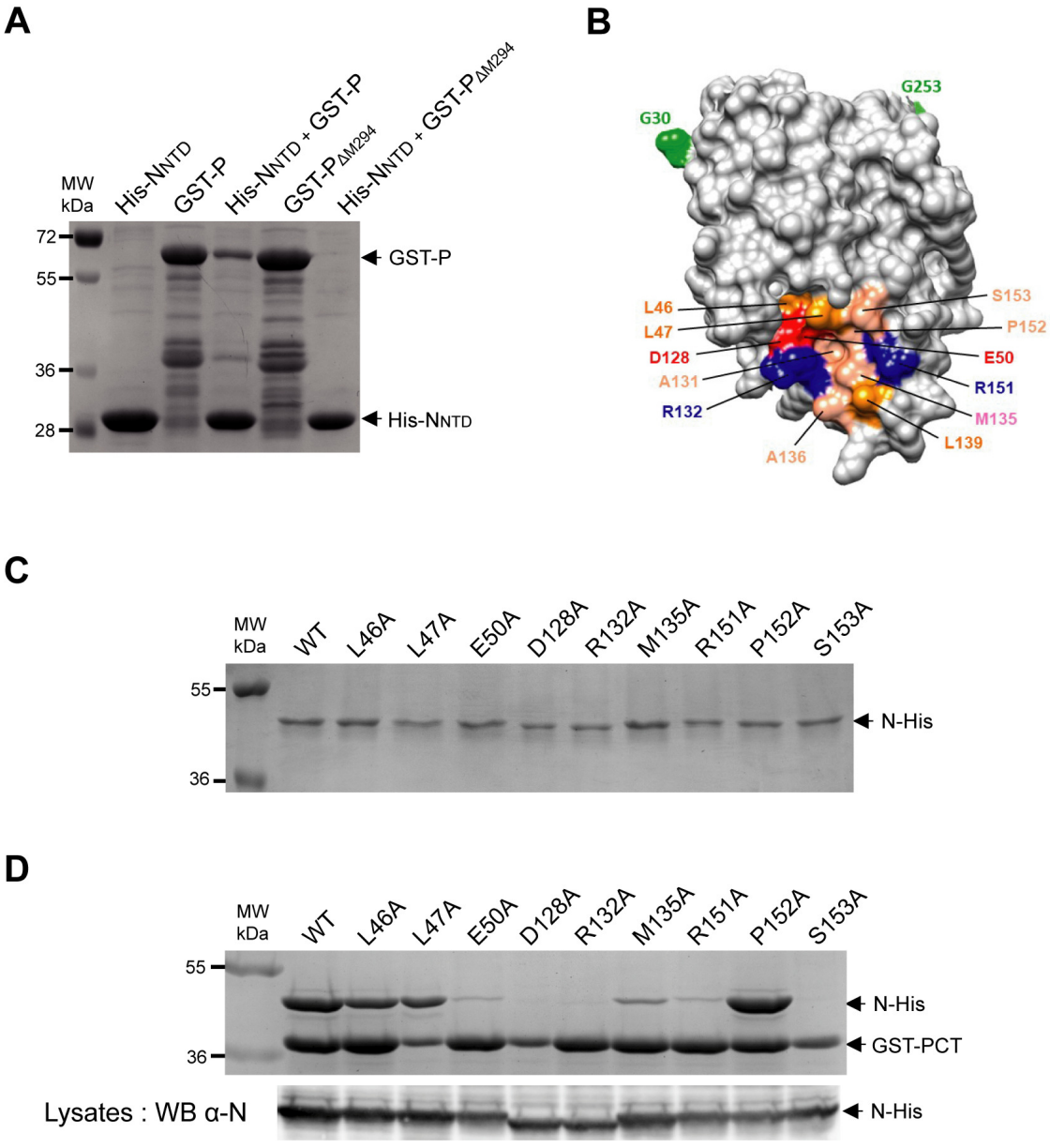
B

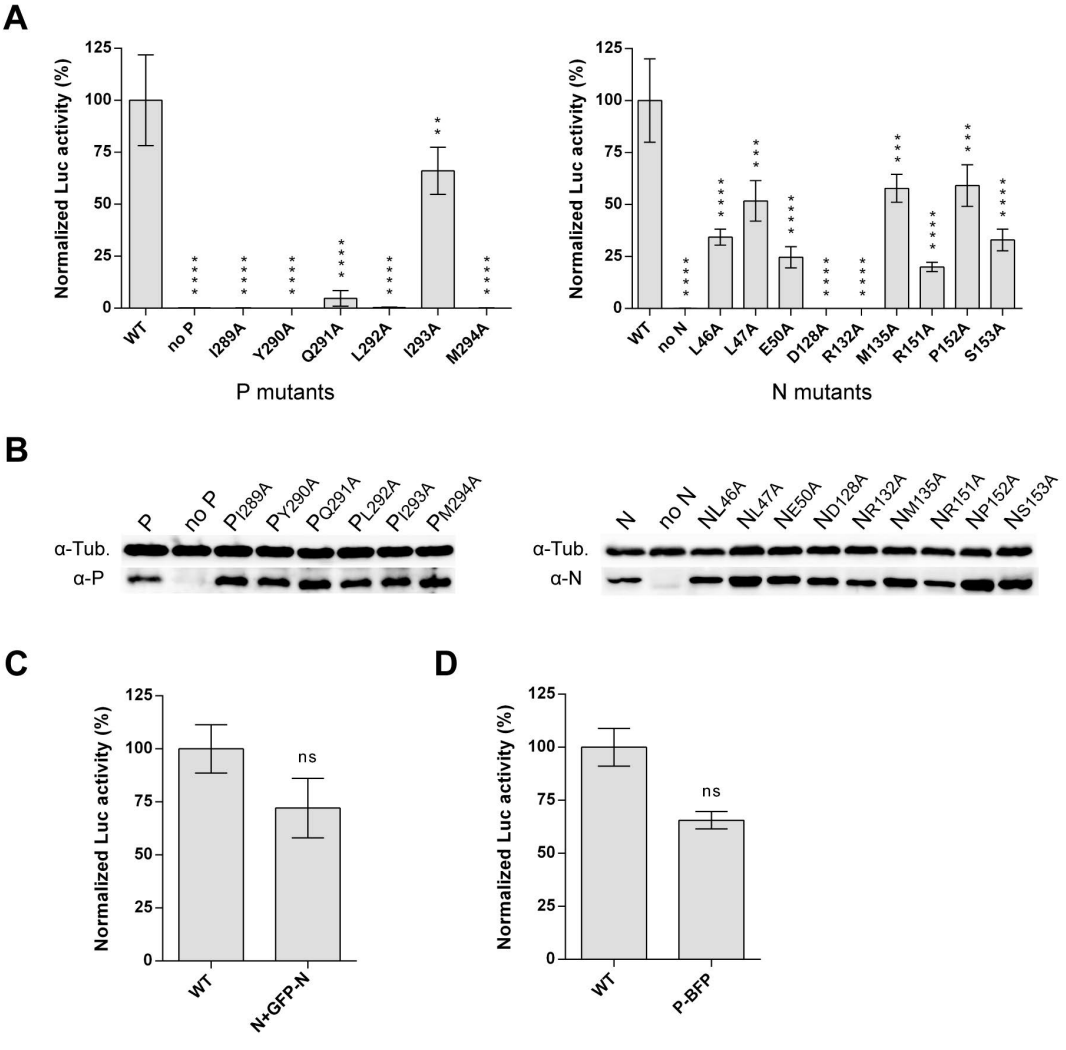


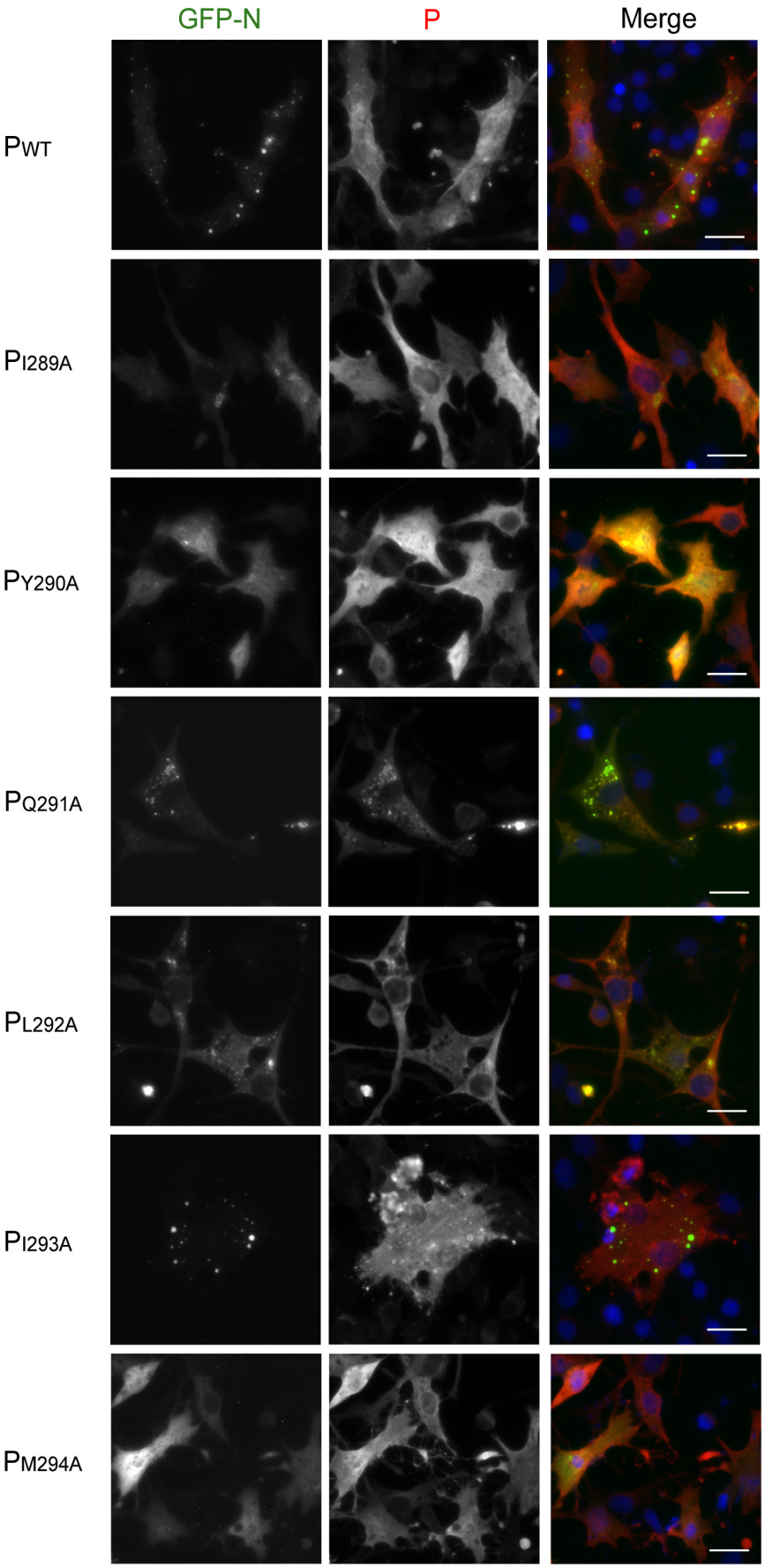
Journal of Virology

JVI

Journal of Virology







Journal of Virology

JVI

Journal of Virology

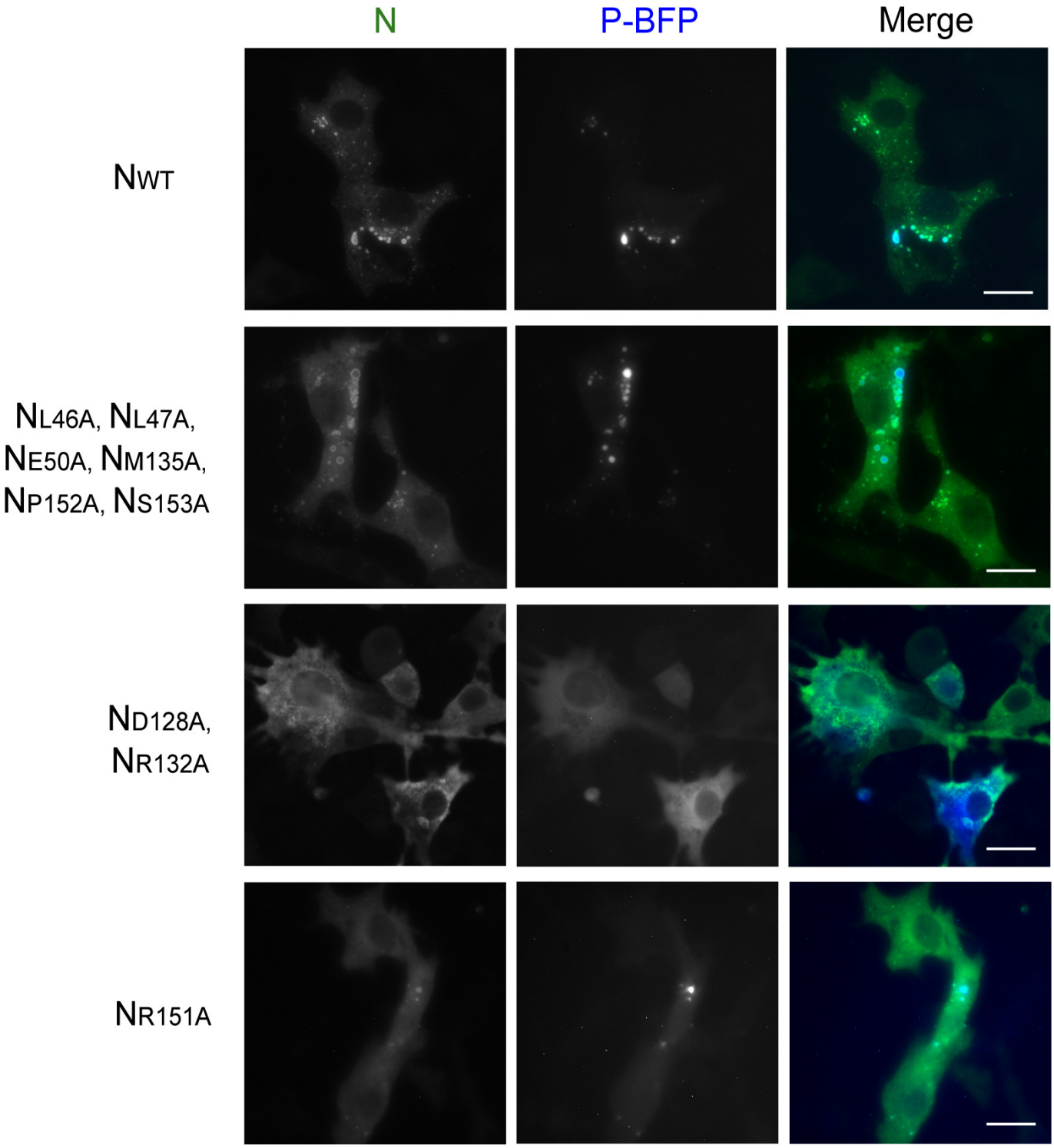
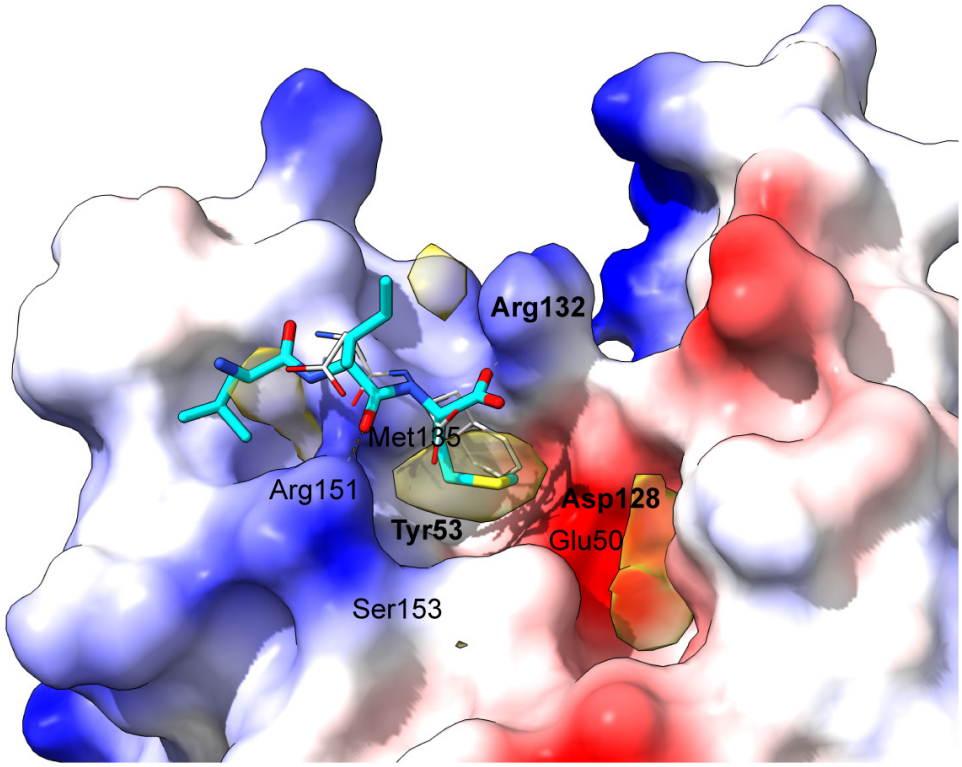


Table 1. Characterization of P and N mutants

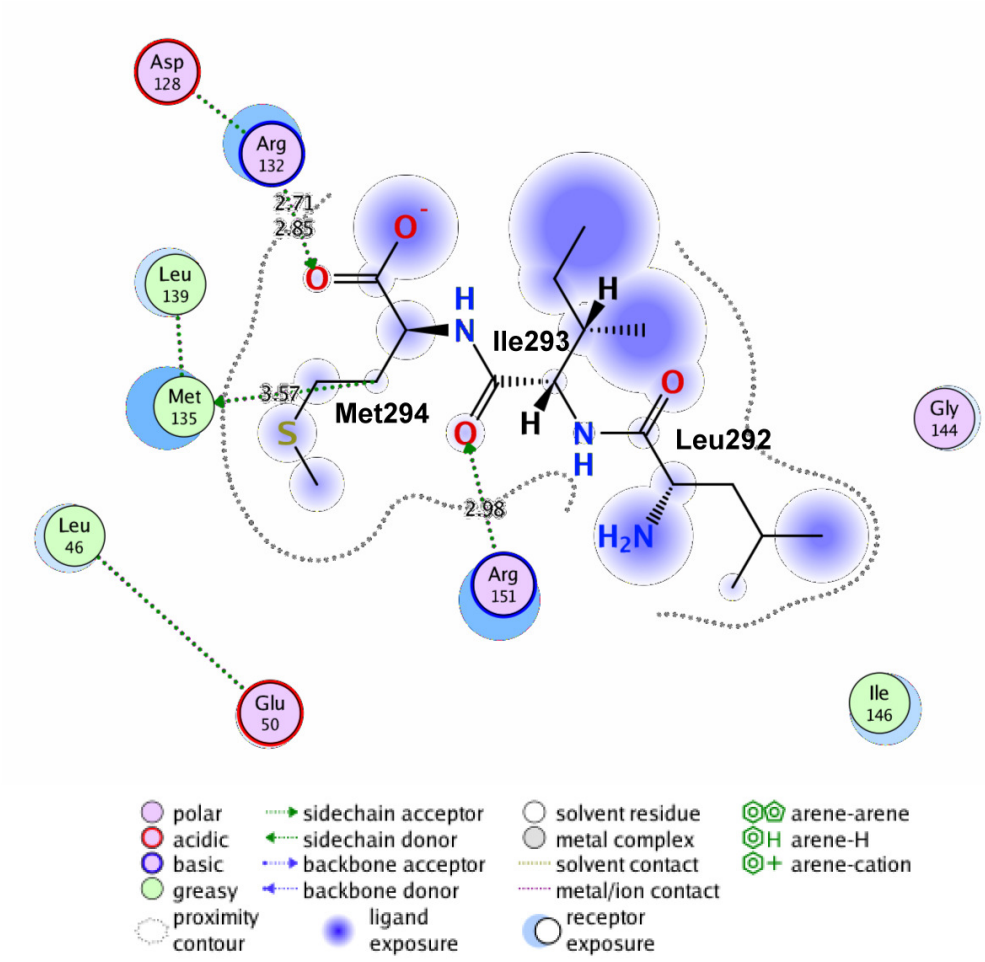
	Mutation	Pulldown of N by GST-PCT	Polymerase activity *	Observation of IBs
P mutants	I289A	-	-	-
	Y290A	-	-	-
	Q291A	+	5%	+/-
	L292A	-	-	-
	I293A	+	65%	+
	M294A	-	-	-
N mutants	L46A	+	35%	+
	L47A	+	50%	+
	E50A	+/-	25%	+
	D128	-	-	-
	R132A	-	-	-
	M135A	+/-	60%	+
	R151A	+/-	20%	+/-
	P152A	+	60%	+
	S153A	-	30%	+

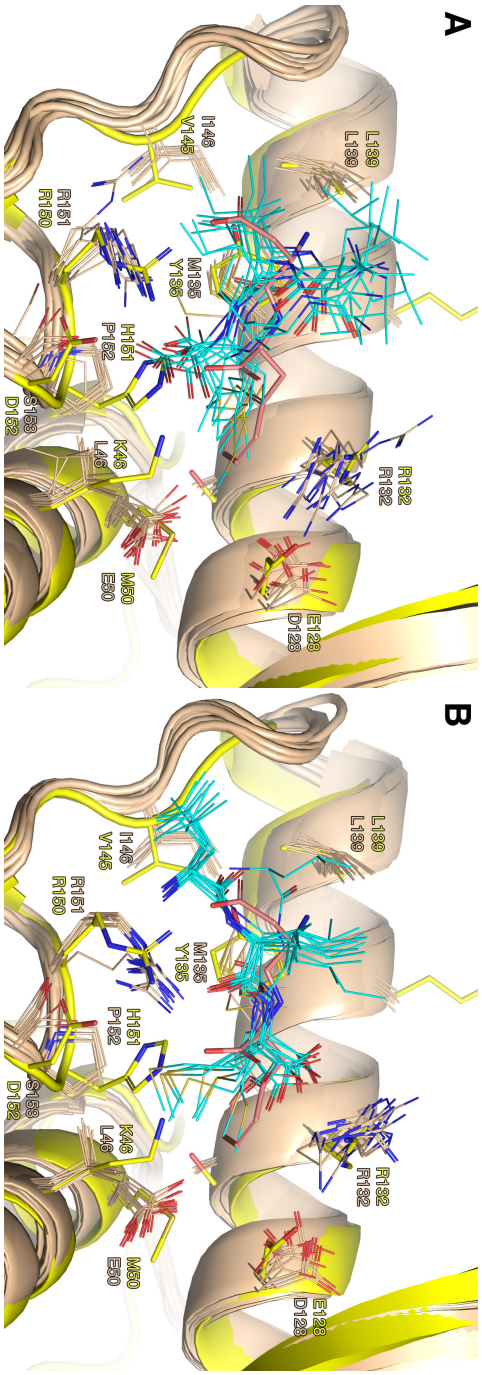
* Values correspond to rounded-up averages ; -, no polymerase activity.

A

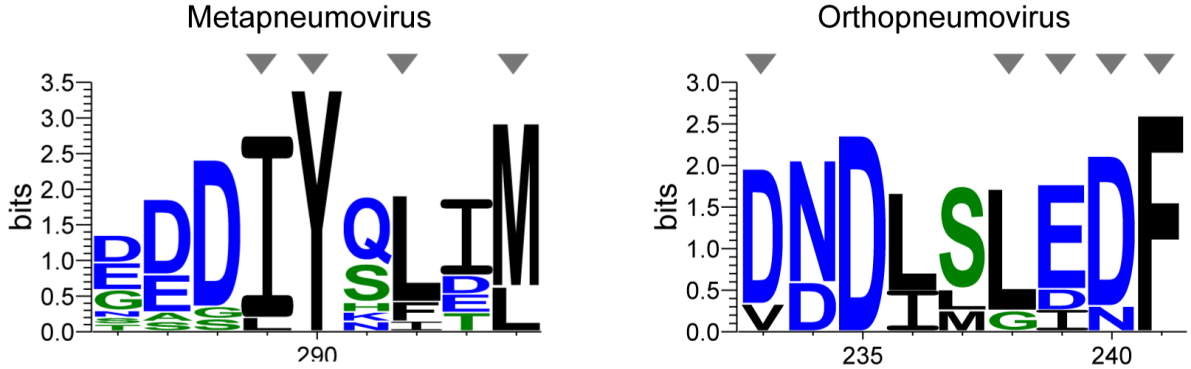


B



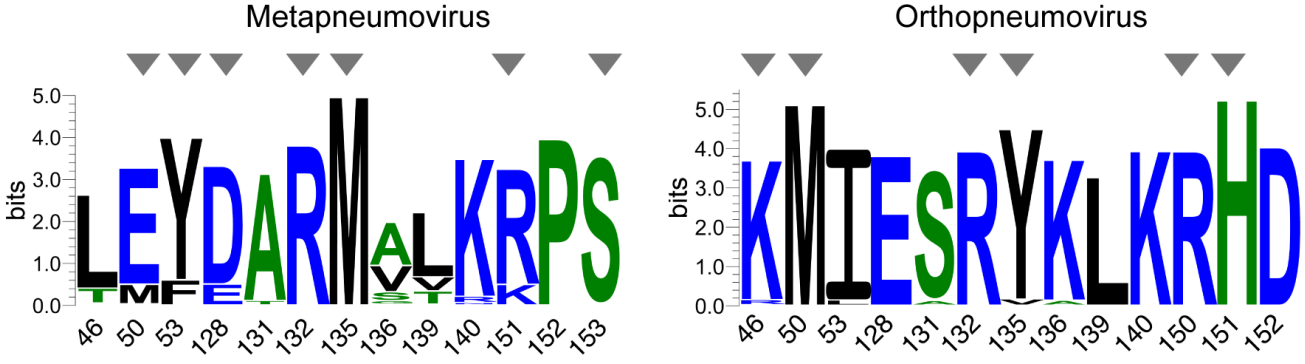


A



P C-terminal residues

B



N residues