Identification and Characterization of the Binding Site of the Respiratory Syncytial Virus Phosphoprotein to RNA-Free Nucleoprotein

Marie Galloux, Gaëlle Gabiane, Julien Sourimant, Charles-Adrien Richard, Patrick England, Mohammed Moudjou, Magali Aumont-Nicaise, Jenna Fix, Marie-Anne Rameix-Welti, Jean-François Eléouët

Unité de Virologie et Immunologie Moléculaires (UR892), INRA, Jouy-en-Josas, France; EA3647—EPIM, UFR des Sciences de la Santé Simone Veil—UVSQ, Montigny-Le-Bretonneux, France; Institut Pasteur, Plate-forme de Biophysique des Macromolécules et de leurs Interactions, Paris, France; CNRS-UMR3528, Paris, France; IBBMC, Université Paris-XI, Orsay, France; AF-HP, Laboratoire de Microbiologie, Hôpital Ambroise Paré, Boulogne-Billancourt, France

ABSTRACT

The RNA genome of respiratory syncytial virus (RSV) is constitutively encapsidated by the viral nucleoprotein N, thus forming a helical nucleocapsid. Polymerization of N along the genomic and antigenomic RNAs is concomitant to replication and requires the preservation of an unassembled monomeric nucleoprotein pool. To this end, and by analogy with Paramyxoviridae and Rhabdoviridae, it is expected that the viral phosphoprotein P acts as a chaperone protein, forming a soluble complex with the RNA-free form of N (N0-P complex). Here, we have engineered a mutant form of N that is monomeric, is unable to bind RNA, and could thus mimic the N0 monomer. We used this N mutant, designated Nmono, as a substitute for N0 in order to characterize the P regions involved in the N0-P complex formation. Using a series of P fragments, we determined by glutathione S-transferase (GST) pulldown assays that the N and C termini of P are able to interact with Nmono. We analyzed the functional role of amino-terminal residues of P by site-directed mutagenesis, using an RSV polymerase activity assay based on a human RSV minireplicon, and found that several residues were critical for viral RNA synthesis. Using GST pulldown and surface plasmon resonance assays, we showed that these critical residues are involved in the interaction between P[1-40] peptide and Nmono in vitro. Finally, we showed that overexpression of the peptide P[1-29] can inhibit the polymerase activity in the context of the RSV minireplicon, thus demonstrating that targeting the N0-P interaction could constitute a potential antiviral strategy.

IMPORTANCE

Respiratory syncytial virus (RSV) is the leading cause of lower respiratory tract illness in infants. Since no vaccine or efficient antiviral treatment is available against RSV, it is essential to better understand how the viral machinery functions in order to develop new antiviral strategies. RSV phosphoprotein P, the main RNA polymerase cofactor, is believed to function as a chaperon protein, maintaining N as a nonassembled, RNA-free protein (N0) competent for RNA encapsidation. In this paper, we provide the first evidence, to our knowledge, that the N terminus of P contains a domain that binds specifically to this RNA-free form of N. We further show that overexpression of a small peptide spanning this region of P can inhibit viral RNA synthesis. These findings extend our understanding of the function of RSV RNA polymerase and point to a new target for the development of drugs against this virus.
nucleoprotein during replication to encapsidate neosynthesized single-stranded genomic RNA (–RNA) and antigenic RNA (+RNA). This suggests that N has to be prevented from binding to cellular RNAs but also from self-oligomerizing. To this avail, and by analogy with other paramyxoviruses and rhadoviruses, the RSV P protein is thought to play the role of a chaperone protein to maintain N in an unassembled and RNA-free form named N0, the capsid protein protomer (5). It is also expected that N0-P interaction involves domain(s) of P different from those involved in the NC-P interaction and that a specific N0 binding domain exists in the N-terminal region of RSV P. Recently, bioinformatics analysis of Mononegavirales P proteins revealed the presence of a motif within the first 40 amino acids (aa) of RSV P that could be involved in the specific interaction with the N protein (6).

For RSV, an N0-P complex had not yet been isolated and characterized. The RSV P protein is composed of 241 amino acids and constitutes the shortest P protein among paramyxoviruses. Human RSV P (strain Long) is phosphorylated mainly at S232 (7, 8), whereas the close homolog bovine RSV P ([1-29] inhibits viral RNA synthesis. The RSV P protein is composed of 241 amino acids and constitutes the shortest P protein among paramyxoviruses. Hu- man RSV P (strain Long) is phosphorylated mainly at S232 (7, 8), and other minor phosphorylation sites have been identified (S30, S39, S45, T46, S54, T108, S116, S117, S119, S237) (9, 10). The precise role of phosphorylation for P activity remains unclear, since (i) unphosphorylated P is competent for oligomerization and binding to N (11, 12) and (ii) substitution of all the phosphorylated residues has little effect on viral transcription and replication (13, 14). However, P phosphorylation could play a role in NC encapsidation in viral particles (10, 13). Although there is no sequence similarity between the P proteins of Pneumovirinae and other related families and subfamilies in the Mononegavirales order, all share similar organization, structure, and functions. RSV P forms homotetramers of elongated shape and has a modular organization with a protease-resistant central domain containing a predicted coiled coil (aa 120 to 160) involved in oligomerization, flanked by two intrinsically disordered regions (aa 1 to 120 and 161 to 241) (11, 15–17). The C-terminal region of P (aa 231 to 241) is involved in the interaction with the NC (12, 18). Some experiments performed with the close homolog bovine RSV (BRSV) P protein have suggested that the aa 161 to 180 region could also participate in N-P interactions (19) and that the N terminus of P (residues 1 to 40) interacts with N (20).

Here, our aim was to determine and characterize the region(s) of P capable of binding to N0 and involved in the formation of a N0-P complex. We used the crystallographic structure of the decameric N-RNA rings (21) to rationally design N mutations to disrupt the interaction with RNA and thus destabilize the oligomeric organization of N-RNA rings. This work resulted in the purification of a recombinant double mutant NK170A/R185A (N0) protein that (i) does not interact with nucleic acids, (ii) is monomeric, and (iii) still interacts with the P protein. We showed that the N-terminal residues (aa 1 to 29) of P are sufficient to bind to the monomeric N, and we identified residues of P critical for this interaction. Our results strongly suggest that the P binding domain on N0 is a molecular recognition element (MoRE). Furthermore, using an RSV minireplicon, we found that overexpression of the peptide P[1-29] inhibits viral RNA synthesis.

**MATERIALS AND METHODS**

**Plasmid constructions.** All the viral sequences were derived from the human RSV strain Long, ATCC VR-26 (Genbank accession number AY911262.1). The full-length P gene, the sequences of P with N-terminal deletions, or internal domains of P were PCR amplified using Pfu DNA polymerase (Stratagene, Les Ulis, France) (primer sequences available on request) and cloned into pGEX-4T-3 at BamHI-XhoI sites to engineer the pGEX-P plasmids. The C-terminal deletion mutants of P were obtained by introducing stop codons at the appropriate site in the coding sequence of pGEX-P, and point mutations K170A and R185A were introduced into pET-N-His by site-directed mutagenesis, using the Quikchange site-directed mutagenesis kit (Stratagene). pET-N and pET-NK170A/R185A were used to produce N-derived proteins with a C-terminal poly-His tag. Sequence analysis was carried out to check the integrity of all the constructions.

Plasmids for eukaryotic expression of the HRSV proteins N, P, M2-1, and L, designated pN, pP, pM2-1, and pL, have been described previously (12, 22). The pM/Luc subgenomic replicon, which contains the firefly luciferase (Luc) gene under the control of the M/Sh gene start sequence was derived from the pM/Sh subgenomic replicon (23) and has been described previously (24). The replication-defective minigenome PmΔ/Luc was constructed by deleting 24 nucleotides from the trailer region of the plasmid pM/Luc (25, 26). Point mutations were introduced in pP by site-directed mutagenesis as described above. The plasmid pP[1-29] was generated by replacing residue 30 of P by a stop codon. The sequence coding for mCherry was fused in frame upstream of P[1-29] in order to generate the plasmid pmCherry-P[1-29].

**Antibodies.** The following primary antibodies were used for immunoblotting: a mouse monoclonal anti-N protein (Serotec, Oxford, United Kingdom), rabbit anti-P and anti-N antisera, previously described (11), and a mouse monoclonal anti-β-tubulin (Sigma). Secondary antibodies directed against mouse and rabbit IgG coupled to horseradish peroxidase (HRP; P.A.R.I.S., Compiègne, France) were used for immunoblotting experiments.

**Expression and purification of recombinant proteins.** Escherichia coli BL21(DE3) bacteria (Novagen, Madison, WI) transformed with pGEX-P fragment plasmids were grown at 37°C for 8 h in 100 ml of Luria-Bertani (LB) medium containing 100 μg/ml ampicillin. Bacteria transformed with pET-N-derived plasmids alone or together with pGEX-P-derived plasmids were grown in LB medium containing kanamycin (50 μg/ml) or ampicillin and kanamycin, respectively. The same volume of LB was then added, and protein expression was induced by adding 80 μg/ml isopropyl-β-D-thiogalactoside (IPTG) to the medium. The bacteria were incubated for 15 h at 28°C and then harvested by centrifugation. For GST fusion protein purification, bacterial pellets were resuspended in lysis buffer (50 mM Tris-HCl [pH 7.8], 60 mM NaCl, 1 mM EDTA, 2 mM dithiothreitol [DTT], 0.2% Triton X-100, 1 mg/ml lysozyme) supplemented with complete protease inhibitor cocktail (Roche, Mannheim, Germany), incubated for 1 h on ice, sonicated, and centrifuged at 4°C for 30 min at 10,000 × g. Glutathione-Sepharose 4B beads (GE Healthcare, Uppsala, Sweden) were added to clarified supernatants and incubated at 4°C for 15 h. Beads were then washed twice in lysis buffer and three times in 1× phosphate-buffered saline (PBS) and stored at 4°C in an equal volume of PBS. To isolate GST-free P[1-40]–N170A/R185A complex, beads containing bound complex were incubated with biotinylated thrombin (Novagen) for 16 h at 20°C. Thrombin was then removed using the cleavage capture kit according to the manufacturer’s instructions (Novagen). For purification of 6×His fusion proteins, bacterial pellets were resuspended in lysis buffer (20 mM Tris-HCl [pH 8], 500 mM NaCl, 0.1% Triton X-100, 10 mM imidazole, 1 mg/ml lysozyme) supplemented with complete protease inhibitor cocktail (Roche). After sonication and centrifugation, lysates were incubated for 1 h with chelating Sepharose Fast Flow beads charged with Ni2+ (GE Healthcare). Finally, beads were successively washed in the washing buffer (20 mM Tris-HCl, pH 8, 500 mM NaCl) containing increasing concentrations of imidazole (25, 50, and 100 mM), and proteins were eluted in the same buffer with 500 mM imidazole. Purified recombinant N proteins were loaded onto a Sephacryl S-200 HR 16/30 column (GE Healthcare) and eluted in 20 mM Tris-HCl (pH 8.5)–150 mM NaCl. The presence of RNA was determined by measuring the optical density at 260 nm.
200 nm (OD260)/OD290 nm absorption ratio, which is 2 for pure RNA and 0.57 for nucleic acid-free proteins.

Cross-linking analysis. Purified recombinant N proteins were dialyzed against 20 mM HEPES (pH 8.5)–150 mM NaCl and centrifuged for 10 min at 100,000 X g. Samples containing 5 μg of protein were incubated with increasing amounts of ethylene glycol disuccinate (EGS) for 30 min at 20°C. Reactions were stopped by the addition of glycine to a final concentration of 50 mM. The cross-linked products were analyzed by SDS-PAGE and silver staining.

Pulldown assays. GST or GST-P fragment fusion proteins fixed on beads were incubated in the presence of 30 μg of purified recombinant N proteins, in a final volume of 200 μl in 20 mM Tris (pH 8.5)–50 mM NaCl. After overnight incubation under agitation at 4°C, the beads were extensively washed with 20 mM Tris (pH 8.5)–50 mM NaCl, boiled in 30 μl Laemmli buffer, and analyzed by SDS-PAGE and Coomassie blue staining.

Cell culture and transfections. BHK-21 cells (clone BSRT7/5) constitutively expressing the T7 RNA polymerase (27) were grown in Dulbecco modified essential medium (Lonza, Cologne, Germany) supplemented with 10% fetal calf serum (FCS), 2 mM glutamine, and antibiotics. Cells were transfected using Lipofectamine 2000 (Invitrogen, Cergy-Pontoise, France) as described by the manufacturer.

Minigenome replication assay (minireplicon). Cells at 90% confluence in 24-well dishes were transfected with Lipofectamine 2000 (Invitrogen) with a plasmid mixture containing 0.25 μg of pMI/Luc, 0.25 μg of pN, 0.25 μg of pF, 0.125 μg of pL, and 62.5 ng of pRSV-β-Gal (where β-Gal is β-galactosidase; Promega) to normalize transfection efficiencies. Transfections were done in duplicate, and each independent transfection was performed three times. Cells were harvested 24 h posttransfection and then lysed in luciferase lysis buffer (30 mM Tris [pH 7.9], 10 mM MgCl2, 1 mM DTT, 1% Triton X-100, and 15% glycerol). Luciferase activities were determined for each cell lysate with an Anthos Lucy 3 luminometer (Bio Advance, Bussy Saint Martin, France) and normalized based on β-Gal expression.

Isothermal titration calorimetry assays. Thermodynamic analysis of the P-N (wild type and N70A/R185A) interactions was performed with a MicroCal ITC200 microcalorimeter (Microcal, Northampton, MA) at 20°C. For the interaction between wild-type N and P, samples were dialyzed against 20 mM Tris-HCl (pH 8.5)–150 mM NaCl. The P concentration in the microcalorimeter cell (200 μl) was 60 μM, and a total of 20 injections of 2 μl of N solution (concentration, 300 μM) were carried out at 180-s intervals, with stirring at 1,000 rpm. For the interaction between Nwt and NK170A/R185A, for which the observed OD260/OD280 was 0.6 (compared to 1.44 for N-RNA rings), confirming the absence of RNA. Purified recombinant wild-type N and N70A/R185A were then loaded onto an S200 column. The Nwt was found to elute from the column in single peak with an apparent molecular mass of ~500 kDa, as expected for N-RNA rings (Fig. 1B). The N70A/R185A eluted mainly in one peak with an apparent molecular mass of ~50 kDa, as expected for the monomeric form of N (44 kDa). The gel filtration profile of N70A/R185A also revealed the presence of minor peaks corresponding to higher molecular masses likely due to partial aggregation (Fig. 1B). To confirm the monomeric state of purified N70A/R185A recombinant protein, we performed cross-linking assays. Purified N-RNA rings and N70A/R185A proteins were cross-linked by increasing concentrations of EGS. As revealed in Fig. 1C, both forms of N were purified to homogeneity (>95%) and migrated in SDS-polyacrylamide gel with an expected apparent molecular mass of 44 kDa. For N-RNA rings, bands migrating with apparent masses of ~80 kDa, ~120 kDa, and higher molecular mass (>300 kDa) appeared upon increasing EGS concentration. These masses are compatible with N dimers, trimers, or higher order oligomers.
ers, and bigger oligomers, respectively. Under the same conditions, no cross-linked products were observed for the \(N^{K170A/R185A}\) protein, and we thus concluded that this mutant was present as a monomer. Finally, the folding of purified \(N^{K170A/R185A}\) recombinant protein was analyzed by circular dichroism (CD). As shown in Fig. 1D, far-UV CD spectra of N-RNA rings and \(N^{K170A/R185A}\) were similar and presented one positive peak at 190 nm and two negative peaks at 208 and 222 nm, typical of secondary structures with a mainly \(\alpha\)-helical content. These results thus show that no major secondary structure change occurs in the mutated \(N^{K170A/R185A}\) variant with respect to the N-RNA state.

Altogether, our data show that the \(N^{K170A/R185A}\) recombinant protein (i) does not interact with nucleic acids, (ii) is monomeric, and (iii) presents a secondary structure similar to that of the wild-type N. We thus used this mutant, which was designated \(N^{mono}\), as an N\(^0\) substitute for further experiments.

**Stoichiometry and thermodynamics of the P-N interactions followed by ITC.** We investigated the interaction between P and N-RNA rings or \(N^{mono}\), using isothermal titration calorimetry (ITC). This approach can give access to the stoichiometry of complexes, their equilibrium constants \((K_D)\), and the variation of enthalpy \((\Delta H)\) and entropy \((\Delta S)\) associated with complex formation (29). For the couple N-RNA/P, the P protein was loaded into the calorimeter sample cell and titrated with N-RNA rings. Due to the propensity of \(N^{mono}\) to aggregate at high concentration, the corresponding titration for the \(N^{mono}/P\) couple was performed by loading \(N^{mono}\) into the calorimeter sample cell and P in the titration syringe. Binding isotherm titrations of both couples are exothermic, as shown by the negative heat flow patterns (Fig. 2, upper panels). Integration of the heat flow revealed that N-RNA and P form a 0.5:1 stoichiometric complex with a \(K_D\) of \(2.5 \times 10^5 \text{M}^{-1}\), whereas P binds to \(N^{mono}\) with a stoichiometry of 1:1 and a \(K_D\) of \(0.8 \times 10^5 \text{M}^{-1}\) (Fig. 2). Taking into account the tetrameric state of P, the stoichiometries determined for both P-N couples suggest that one tetramer of P could interact with two N protomers in the context of N-RNA, whereas the N\(^0\)-P complex would consist of one P tetramer bound to four N\(^0\). These data also revealed that the N-RNA/P interaction is driven by both enthalpic and entropic components \((\Delta H = -3.25 \text{ kcal/mol}, \text{ and } -T \Delta S = -4.25 \text{ kcal/mol})\). As for the \(N^{mono}\)-P interaction, it is driven by a strong favorable enthalpy \((\Delta H = -14.43 \text{ kcal/mol})\), partially compensated by an unfavorable entropic con-
distribution ($-\Delta S = 6.47 \text{ kcal/mol}$). The last observation suggests that the Nmono–P interaction could involve some degree of structural reorganization.

**N- and C-terminal extremities of P interact with Nmono in solution.** Next, to identify the domain(s) of P capable of binding to Nmono, we performed GST pulldown assays of purified Nmono protein using serial deletions of P fused with GST. Based on previous mapping of the oligomerization domain of P between residues $\sim 120$ and $\sim 160$ (11, 16, 17, 30), different fragments of P corresponding to the N- and C-terminal domains of P including or not the oligomerization domain were fused to GST (Fig. 3A). GST-P fusion proteins bound to glutathione-Sepharose beads were then incubated with Nmono purified separately (Fig. 3B, left panel). Analysis of the interactions between GST-P proteins and Nmono revealed the presence of two binding sites on P (Fig. 3B, right panel). The first, the C-terminal domain of P alone or associated with the oligomerization domain (GST-P[127-241] and GST-P[161-241] constructs) interacted with Nmono, and deletion of the last 12 C-terminal residues of P (GST-P[127-229]) was sufficient to abrogate this interaction. Second, Nmono was also capable of interacting with GST-P[1-163], including the N-terminal domain of P and the oligomerization domain. Of note, the interaction between Nmono and this construct could not be detected at first, because GST-P[1-126] and N have a similar apparent molecular weight when analyzed by SDS-PAGE. However, this interaction was validated after separating P[1-126] and the GST tag by thrombin cleavage (Fig. 3C). Finally, no N-P interaction was observed with GST-P[60-126], suggesting that the second Nmono binding domain is located upstream of residue 60 of P (Fig. 3B). To confirm these results, and since the last 10 C-terminal amino acids of P (aa 231 to 241) were previously shown to be implicated in the interaction with N-RNA (12, 21), we compared GST pulldown of either N-RNA rings or Nmono with GST-P[1-230] and GST-P[231-241]. As shown in Fig. 3D, GST-P[231-241] interacted with both N-RNA and Nmono, confirming that the last 10 C-terminal residues of P are sufficient to interact with both forms of N. On the other hand, GST-P[1-230] interacted only with Nmono.

Altogether, these results suggest the presence of two N binding sites on P, the first one located in the C-terminal domain of P (PCTD) and capable of binding to both N-RNA and Nmono and the second one located in the N-terminal domain of P (PNTD) and specific for Nmono.

**Determination of the minimal N-terminal domain of P required for interaction with Nmono.** We then focused on the characterization of the minimal N-terminal domain of P required for Nmono binding. To this end, Nmono was coexpressed in bacteria with N-terminal fragments of P (PNTD) fused with GST, and the ability of Nmono to copurify with GST–PNTD proteins was evaluated by SDS-PAGE and Coomassie blue staining (Fig. 3E). Our data showed that Nmono copurified with GST-P[1-29] and GST-P[1-40] but not with GST-P[1-10], GST-P[1-20], GST-P[10-30], or GST-P[10-40]. These results revealed that an Nmono-specific binding site is located between amino acids 1 and 29 of P. It is noteworthy that residues 1 to 10, if not sufficient to interact with Nmono, are required for the interaction. Finally, we observed that Nmono also copurified with GST-P[231-241] used as a control but
less efficiently than with GST-P[1-29] and GST-P[1-40], in a reproducible way.

The last observation led us to determine whether the N- and C-terminal domains of P can bind concomitantly to Nmono. We performed GST pulldown assays with GST-P[161-241] using either separately purified N-RNA rings, Nmono, or the P[1-40]/Nmono complex. Analysis of the interactions by SDS-PAGE revealed that GST-P[161-241] can interact with both N-RNA rings and Nmono, although the last interaction seemed weak (Fig. 3F). On the contrary, no interaction was observed between GST-P[1-40]/Nmono complex and the preformed P[1-40]/Nmono complex. These data revealed that binding of P[1-40] to the monomeric N inhibited the binding of the C terminus of P. However, the N[31-252] domain, which was previously shown to contain the binding site for PCTD (18), did not copurify with GST-P[1-40] (not shown). It is thus unlikely that PNTD and PCTD bind to the same pocket on N.

FIG 3 Deletion mapping of Nmono binding domain(s) on P. (A) Schematic illustration of the wild-type and the truncated RSV P proteins used in this study. The oligomerization domain of P is represented as a gray box, and numbers indicate amino acid positions. Deletion mutants of P harboring a GST tag at the N terminus were purified on glutathione-Sepharose beads and incubated in the presence of Nmono. For each deletion mutant, the ability to interact with Nmono is summarized on the right. (B, C, D) GST-P-derived proteins were analyzed by SDS-PAGE before or after incubation with Nmono (B, C) or N-RNA rings (D), as indicated. (C) For GST-P[1-126], the presence of N0 after pulldown was verified after thrombin cleavage of GST tag (P, purification products; S and B, supernatant and beads after thrombin cleavage, respectively) (E, F) Identification of the minimal domain of P involved in the interaction with Nmono. (E) Analysis of copurification products of Nmono with P[231-241] and N-terminal fragments of P (GST-PΔA). (F) Study of the fixation of both N and C termini of P on Nmono. Recombinant N-RNA rings, Nmono, or P[1-40]+Nmono complex were incubated with GST or GST-P[161-241] bound to beads, and interactions were analyzed by SDS-PAGE and Coomassie blue staining.
In conclusion, these results show that the first 29 N-terminal and the last 10 C-terminal residues of P can independently bind to the RNA-free N monomer. However, although two distinct P binding sites exist for Nmono in solution, the interaction with the PCTD appears to be weaker, and binding of the PNTD inhibits the interaction with the PCTD.

Effect of N-terminal mutations of P on RNA polymerase activity. Based on these results, we investigated the potential role of N-terminal residues (aa 1 to 29) of P on the polymerase complex activity by site-directed mutagenesis using an RSV plasmid-based minireplicon system (24). Preventing the formation of an N0-P complex competent for genomic or antigenomic RNA encapsidation would result in a decrease of N-RNA template formation and amplification and therefore also of mRNA transcription and expression of the luciferase (Luc) reporter. We replaced one by one residues E2 to F28 of P by alanines (except residues Ala5, Ala13, and Ala17), and we tested the functionality of these P mutants within the polymerase complex. As shown in Fig. 4A, among these 24 mutants, 11 displayed a reduction of Luc activity of about or less than 50% compared to the wild-type P protein. The E2A, F4A, F8A, G10A, and F20A substitutions had the strongest effect, with a reduction of 80 to 90% of polymerase activity. As assessed by Western blotting, all P mutants displaying a reduction of Luc activity were expressed in amounts similar to those of wild-type P protein in BSRT7/5 cells (Fig. 4B).

These data revealed that the 11 residues E2, F4, E7, F8, G10, F20, L21, E22, K25, and G26 are critical for polymerase activity, and could be directly involved in the interaction between P and the monomeric RNA-free N protein. These results also suggest the presence of two potential critical stretches of residues, one located between residues E2 and G10 and the second between residues F20 and G26.

Identification of N-terminal residues of P involved in the interaction with the Nmono. We therefore attempted to determine whether the drop in polymerase activity induced by substitutions in the E2-F28 region of P was correlated to a defect of interaction between P and monomeric N. Mutations of all the residues E2 to F28 were introduced into the plasmid expressing GST-P[1-40] to generate 24 mutant proteins. The Nmono mutant protein was co-expressed with the wild-type or mutant forms of GST-P[1-40] in E. coli and purified by using the GST tag. Purified complexes were analyzed by SDS-PAGE. Mutations F4A, F8A, G10A, F20A, L21A, and I24A totally or nearly abrogated the interaction of GST-P[1-40] with Nmono (Fig. 5). The substitutions E7A and K25A attenuated partially the interaction compared to the wild-type P[1-40]. The other mutations and more specifically E2A, E22A, and G26A, which were shown to induce a decrease of polymerase activity, did not significantly modify the interaction between P[1-40] and Nmono.

In order to confirm these results and to quantify the impact of
P mutations E2A-G10A and F20A-G26A on the interaction with Nmono, we then characterized the specific interaction between GST-P[1-40] (wild type or mutants) and Nmono by surface plasmon resonance (SPR). GST-P[1-40] proteins were captured on an anti-GST antibody surface, and serial dilutions of Nmono were injected. We first characterized the specific interaction between wild-type GST-P[1-40] and Nmono. The interaction was transient, with a very high dissociation rate and a $K_d$ of 4 to 5 μM (Fig. 6; see also Fig. 8). Similar affinities were obtained for mutants K3A, P6A, H9A, and S23A. Mutations F8A and F20A had the strongest effect on the interaction with Nmono, with calculated $K_d$ so of 1 μM and >100 μM, respectively. The affinity for Nmono was also affected by mutations of residues F4, E7, L21, and I24 (10 μM $< K_d$ $< 20$ μM) and to a lesser extent by mutations of residues G10, E22, and K25 (5 μM $< K_d$ $< 10$ μM). Finally, the affinity of GST-P[1-40] for Nmono was increased nearly 2-fold by mutations E2A and G26A, with $K_d$s of 1.9 μM and 2.6 μM, respectively.

Altogether, these results reveal that the 11 residues of P previously identified as critical for polymerase activity are directly involved in the interaction with the soluble monomeric RNA-free N protein.

**Overexpression of peptide P[1-29] inhibits viral RNA synthesis.** In our minireplicon system, replication and transcription of RNA by RSV RdRp rely on the correct encapsidation by N0 of both the minigenome and the antimimigenome and the formation of an RNA-N template. It is thus expected that inhibition of the formation of an N0-P complex, which is competent for RNA encapsidation, would result in a decrease of RNA synthesis as Luc expression. Likewise, it is expected that, by binding to N0, a short peptide corresponding to the N-terminal domain of P could compete with full-length P and inhibit the formation of an N0-P complex and so RNA replication and transcription. To evaluate whether P[1-29] could interfere with RSV RNA synthesis, a plasmid encoding P[1-29] was transfected in BSRT7/5 cells in the context of the minireplicon. RdRp activity assayed by measuring Luc activity was reduced in a dose-dependent manner by the P[1-}

---

**FIG 5** Identification of N-terminal residues of P involved in the interaction with Nmono. Copurification of Nmono with GST-P[1-40] mutants from E. coli. GST-P[1-40] constructs (wild type or mutants E2A to F28A) were coexpressed with Nmono protein in bacteria, and complexes were pulled down using the GST tag. Interaction of Nmono with P[1-40] mutants was analyzed by SDS-PAGE and Coomassie blue staining.

only the transcription activity. As previously observed (25, 26), replication. Hence, using this system, Luc activity should reflect trailer sequence, still efficient for transcription but defective for this complex.

Among Mononegavirales, the crystal structures of N0-P complexes have been resolved for one rhabdovirus (vesicular stomatitis virus [VSV]) and one paramyxovirus (Nipah virus), allowing to clarify the role of P as a chaperone protein (32, 33). For VSV, the complex was reconstituted in vitro after purification of a recombinant N protein deleted of its 21 N-terminal residues, which was monomeric although still capable of interacting with RNA (32). A second step consisted of displacing RNA by a peptide, P[1-60], corresponding to residues 1 to 60 of the P protein. The crystal structure of this complex revealed that the binding domain of P[1-60] on N partially overlaps the RNA binding site and that the N terminus of P blocks the backside groove of N, preventing N oligomerization. For Nipah virus, the structure of a complex between a monomeric N mutant with its N and C termini deleted and a peptide corresponding to the sequence of the 50 N-terminal residues of P was also obtained (33). In this complex, the N terminus of P was shown to interact only with the C-terminal domain of N. In this case, the binding of P should prevent N oligomerization and keep N0 in an open conformation ready to interact with RNA. Finally, although the N0-P complex of RSV has not yet been characterized, an RNA-free N protein has been isolated by deletion of the first 12 N-terminal residues (34).

In the present study, based on the crystallographic structure of RSV RNA-N rings (21) and on the hypothesis that preventing RNA binding would also prevent N oligomerization, we generated the RNA-free N mutant protein K170A/R185A. This recombinant N mutant protein was shown to be monomeric at concentrations of up to 40 μM (Fig. 1B and C). This result shows that disrupting N-RNA interactions can partially prevent RSV N self-oligomerization. This first observation supports the debated theory according to which N RNA binding and oligomerization are coupled (5, 35, 36). We have also shown that this RNA-free N protein presents a secondary structure similar to that of N assembled with RNA (Fig. 1D). Therefore, it is expected that the switch between N0 and N-RNA does not require major secondary structure changes. This

Strategy to obtain an RSV N0-like protein. By analogy with Paramyxovirinae and Rhabdoviridae, it is assumed that pneumovirus replication requires a constant supply of monomeric, unassembled N (N0) to encapsidate neosynthesized single-stranded genomic RNA (−RNA) and antigenic RNA (+RNA) (31). However, when expressed alone in mammalian cells, insect cells, or bacteria, nucleoproteins of Mononegavirales bind to cellular RNA and form only stable RNA-N complexes (for a review, see reference 5). The phosphoprotein P plays a critical role in maintaining N0 unassembled during infection. However, purification of N0-P complexes is tricky and the existence of this theoretical complex has never been shown for RSV or other pneumoviruses. Structural characterization of the RSV N0-P complex constitutes a major challenge. It could be useful to gain insight into the mechanisms of regulation that sustain the switch between N0 and N-RNA but also to develop a rational search for inhibitors targeting this complex.

29]-expressing plasmid (Fig. 7A). Similar results were obtained when using a plasmid encoding mCherry-P[1-29] (Fig. 7A), for which expression can be easily visualized by epifluorescence.

To test whether P[1-29] could interfere with RSV transcription, we used an altered minireplicon (pMΔ/luc) with a deleted trailer sequence, still efficient for transcription but defective for replication. Hence, using this system, Luc activity should reflect only the transcriptional activity. As previously observed (25, 26), Luc activity was reduced to less than one-half when using the pMΔ/luc vector compared to pM/luc (Fig. 7B). We then performed a minigenome replication-defective assay in the context of the plasmid expressing mCherry-P[1-29], using a pmCherry-P[1-29]/pP ratio of 1.5. Expression of the mCherry-P[1-29] peptide reduced the transcriptional activity to 50% (Fig. 7C). Thus, the short peptide P[1-29] is capable of inhibiting both RNA replication and transcription by RdRp in mammalian cells.

DISCUSSION

Strategy to obtain an RSV N0-like protein. By analogy with Paramyxovirinae and Rhabdoviridae, it is assumed that pneumovirus replication requires a constant supply of monomeric, unassembled N (N0) to encapsidate neosynthesized single-stranded genomic RNA (−RNA) and antigenic RNA (+RNA) (31). However, when expressed alone in mammalian cells, insect cells, or bacteria, nucleoproteins of Mononegavirales bind to cellular RNA and form only stable RNA-N complexes (for a review, see reference 5). The phosphoprotein P plays a critical role in maintaining N0 unassembled during infection. However, purification of N0-P complexes is tricky and the existence of this theoretical complex has never been shown for RSV or other pneumoviruses. Structural characterization of the RSV N0-P complex constitutes a major challenge. It could be useful to gain insight into the mechanisms of regulation that sustain the switch between N0 and N-RNA but also to develop a rational search for inhibitors targeting this complex.

Among Mononegavirales, the crystal structures of N0-P complexes have been resolved for one rhabdovirus (vesicular stomatitis virus [VSV]) and one paramyxovirus (Nipah virus), allowing to clarify the role of P as a chaperone protein (32, 33). For VSV, the complex was reconstituted in vitro after purification of a recombinant N protein deleted of its 21 N-terminal residues, which was monomeric although still capable of interacting with RNA (32). A second step consisted of displacing RNA by a peptide, P[1-60], corresponding to residues 1 to 60 of the P protein. The crystal structure of this complex revealed that the binding domain of P[1-60] on N partially overlaps the RNA binding site and that the N terminus of P blocks the backside groove of N, preventing N oligomerization. For Nipah virus, the structure of a complex between a monomeric N mutant with its N and C termini deleted and a peptide corresponding to the sequence of the 50 N-terminal residues of P was also obtained (33). In this complex, the N terminus of P was shown to interact only with the C-terminal domain of N. In this case, the binding of P should prevent N oligomerization and keep N0 in an open conformation ready to interact with RNA. Finally, although the N0-P complex of RSV has not yet been characterized, an RNA-free N protein has been isolated by deletion of the first 12 N-terminal residues (34).

In the present study, based on the crystallographic structure of RSV RNA-N rings (21) and on the hypothesis that preventing RNA binding would also prevent N oligomerization, we generated the RNA-free N mutant protein K170A/R185A. This recombinant N mutant protein was shown to be monomeric at concentrations of up to 40 μM (Fig. 1B and C). This result shows that disrupting N-RNA interactions can partially prevent RSV N self-oligomerization. This first observation supports the debated theory according to which N RNA binding and oligomerization are coupled (5, 35, 36). We have also shown that this RNA-free N protein presents a secondary structure similar to that of N assembled with RNA (Fig. 1D). Therefore, it is expected that the switch between N0 and N-RNA does not require major secondary structure changes. This
result is in agreement with those obtained for VSV and Nipah virus (32, 33), for which it was shown that compared to the oligomerized RNA-bound form of N, the N- and C-terminal globular domains of N are present similar structures. Finally, the RSV recombinant N170A/R185A mutant can still interact with the P protein. The last observation strongly suggests that residues K170 and R185, which are located in the RNA groove of N and are critical for RNA binding, are not critical for N-P interaction. Thus, a different set of residues should be involved in RNA binding and in P binding. Since bacterially expressed recombinant RSV P protein is not phosphorylated, these results also show that P phosphorylation is not required for N-P complex formation.

P-N stoichiometries. The interactions between P and N-RNA rings or Nmono protein were compared by isothermal titration calorimetry (Fig. 2). The calculated stoichiometry for both binding pairs suggests that one tetramer of P could interact with two N protomers in the context of NC, whereas it could interact with four N0. Recently, an analytical ultracentrifugation experiment and a small-angle X-ray scattering (SAXS) study of the entire VSV N0-P complex suggested that a dimer of P associates with only one N0 (36). Therefore, our results suggest that N0-P complexes of RSV and rhabdoviruses do not have the same stoichiometry. Isothermal titration calorimetry experiments also revealed that P presents similar affinities for both N-RNA and Nmono proteins. The calculated Kd for these two N-P interactions are in the micro-molar range. It is postulated that, as a general feature for Mononegavirales, RSV replication requires the switch from N0-P to N-RNA-P, P freeing N0 for RNA encapsidation just after its synthesis next to the replicase L-P. It is thus expected that such a mechanism requires low affinities between P and N. It is worth noting that, based on our results, we cannot exclude that a tetramer of P could concomitantly interact with N0 and NC.

An N-specific binding domain is located at the N terminus of P. In parallel, identification of the Nmono binding domain on P was investigated by either coexpressing or coexpressing GST-fused P fragments with N. Altogether, the data demonstrated that the RNA-free N monomer presents binding sites for both PCTD and PNTD and that the oligomerization domain of P is not required for these interactions (Fig. 3). Furthermore, the binding of PNTD to Nmono is specific to the N monomer, as no binding was observed with N-RNA rings. These results also revealed that the 10 C-terminal and the 29 N-terminal residues of P are sufficient to interact with Nmono. The interaction between Nmono and the C terminus of P correlates with recent data from Shapiro et al., who showed that the C terminus of P binds to a monomeric RNA-free N protein, depleted from the 12 first N-terminal residues (37). However, we have further shown that PNTD binds very poorly to Nmono and that the binding of the PNTD to Nmono impairs the binding of PCTD (Fig. 3F). The last observation correlates with ITC data, since the experimental titration curve of Nmono by P can be fitted well, assuming a single binding site (Fig. 2B). Two hypotheses could explain these observations. First, PCTD and PNTD could bind to the same site of N. However, the NTd truncated variant, which was previously shown to contain the PNTD binding site (18), did not copurify with PNTD (not shown). This result strongly suggests the existence of two distinct P binding sites on Nmono. This observation is supported by the fact that PNTD and PCTD do not present sequence homologies. It is thus unlikely that these domains interact with the same N binding site. However, we cannot exclude the possibility that the PNTD binding site partially overlaps the PCTD binding site on N. Furthermore, although the NNTD is a globular, alpha-helical domain that should not undergo major conformational changes between N0 and N-RNA forms, PNTD binding on Nmono could induce a local conformational change that renders the PCTD binding site inaccessible. In any case, our results concerning the RSV N0-PNTD interaction are in agreement with data obtained previously with BRSV showing a direct interaction between PNTD (residues 1 to 40) and N using the yeast two-hybrid system (20). Since P and N are highly conserved between BRSV and HRSV, it is likely that PNTD is the main N binding site for both viruses. Our results are also supported by data from the literature. Indeed, using sensitive sequence similarity search programs, Karlin and Belshaw (6) found that all Paramyxovirinae share a short conserved motif in the first 40 amino acids of P (aa 11 to 26), which was called the soyuz motif and should be involved in N0 binding. Conserved motifs were also found within Pneumovirinae (mir motif), Filoviridae (sputnik motif), and Rhabdoviridae. The presence of N0 binding domains within the N terminus of P was confirmed by experimental data for SeV (residues 33 to 41) (38), HPIV3 (first 40 N-terminal residues) (39), RABV (residues 4 to 40) (40), VSV (residues 11 to 30) (41), and Nipah virus (residues 1 to 35) (33).

The N-binding domain of P could fold upon interaction. Based on our results, we then focused on the characterization of the Nmono binding domain located at the N terminus of P and more specifically on the potential involvement of residues 1 to 29 of P in the Nmono-P interaction by Ala scanning (Fig. 8A). Since this interaction is required for maintaining N in an RNA-free form (N0) competent for RNA encapsidation and formation of an RNA-N template used for replication and transcription, we first used the functional HRSV minireplicon system to screen for residues critical for polymerase activity. This approach highlighted the critical role of 11 residues of P, located between residues E2-G10 and F20-G26 (Fig. 4A). None of these mutations fully abolished RNA synthesis. Pulldown and SPR experiments were then combined to determine whether the effect of mutations on viral RNA synthesis correlates with a defect in Nmono-P interaction. Using these approaches, we confirmed that the Ala substitution of residues E2 and G26 increased the affinity of Nmono for the N terminus of P (Fig. 8A). These data highlight a direct role of these 11 residues in the interaction with the putative N0. The localization of these critical residues suggests the presence of two main domains of interaction of P on N0, located between residues E2-G10 and F20-G26 and separated by a stretch of 9 residues. The structural study of P predicted the presence of an α-helix in the region D12-F20 (16), and as previously evoked, the N0 binding site was predicted to be located between residues 11 and 26 (6).

Taking into account these data and the periodicity (i, i + 3; i, i + 4) of the critical residues F20-G26 of P, we thus modeled a putative short α-helix between residues E11 and F28 (Fig. 8B). This representation shows that residues F20, L21, 124, and K25, for which mutation strongly affected the interaction with Nmono, should be located on the same side of the putative helix. It is worth noting that the unfavorable entropic contribution determined by ITC for the Nmono-P interaction also argues for the involvement of structural constraints during the establishment of the interaction, which could be explained by the folding of the N terminus of P. Such a mechanism of folding of intrinsically disordered regions.
upon interaction has already been described for other paramyxoviruses such as measles virus (reviewed in reference 42) and for the rhabdovirus VSV N0-P interaction. More precisely, the nuclear magnetic resonance (NMR) study of the N-terminal domain of VSV P protein revealed that it contains two transient helices (43), which were subsequently shown to constitute molecular recognition elements (MoRE) for the interaction with N0 (32). A MoRE is a small segment of a protein that is intrinsically disordered but folds upon binding to a molecular partner (44). Thus, the N-terminal domain of the RSV P protein, which is located in a predicted disordered region extending from residues 1 to 120, could contain such a MoRE.

![Sequence of the N terminus of P and summary of effects of point mutations on RdRp activity and P-N0 interactions.](image)

**FIG 8** Sequence of the N terminus of P and summary of effects of point mutations on RdRp activity and P-N0 interactions. (A) The sequence of the first 40 N-terminal residues of P is indicated at the top. Residues of the predicted mir motif are boxed. The location of the predicted α-helix is indicated by a blue rectangle below the sequence. The left-hand column indicates the point mutations. Right-hand columns summarize the impact of mutations on (i) the polymerase activity, (ii) *in vitro* copurification of GST-P[1-40] mutants with Nmono, and (iii) affinity of Nmono for GST-P[1-40] mutants; nd, not determined. Residues identified as critical for both polymerase and interaction with Nmono based on pulldown and SPR data are in red in the sequence. (B) Helical wheel representation (HeliQuest online program) of the putative α-helix located between residues 11 and 28 of P. Residues critical for N0-binding are indicated by a star, and the putative site of interaction with the Nmono is indicated by a black half-circle under the wheel. Positively charged residues are in blue, negatively charged residues in red, and hydrophobic residues in yellow.
The N0-binding site is a potential target for antiviral compounds. Because P plays a central role in the functioning of the RNA polymerase, this protein could be a target of choice for antiviral strategies aiming to destabilize the complex and inhibit viral replication. We have demonstrated that coexpression of the peptide P[1-29] encompassing the N0 binding site together with the minireplicon is sufficient to inhibit the replication activity of the polymerase complex, in a dose-dependent manner. This small peptide thus constitutes an interesting basis upon which to develop antiviral approaches. A proof of concept has already been developed for rabies virus (RABV) (45) and more recently for Nipah virus (33), for which peptides that mimic the N terminus of the P protein were shown to present an antiviral effect. In addition, a study describing the prevention of nasopulmonary infection by mucosal delivery of a peptide that interferes with the fusion step of RSV (46) demonstrates the power of similar antiviral strategies.

Conserved mechanisms among Mononegavirales for RNA encapsidation. Altogether, the data presented here demonstrate for the first time that RSV P can bind to an RNA-free, monomeric form of N, thus forming an N-P-like complex, and reinforce previous observations that the P proteins of Rhadoviridae, Paramyxovirinae, and Pneumovirinae share similar functions and modular organization, although they have no detectable sequence similarity. All the P proteins are oligomeric and contain intrinsically disordered regions upstream and downstream of the oligomerization domain. From a functional point of view, they interact with many partners, including the L polymerase, N-RNA, and N0, the main nucleocapsid (RNA-N) and N0 binding domains being located at the C and N termini of P, respectively. However, a study of the P-N interaction of MuV recently reopened the debate concerning the role of P as a chaperone (4). Using electron microscopy, the authors have indeed shown that incubation of PNTD with nucleocapsid (NC) induces uncoupling of MuV NC. They also showed that both PNTD and PCTD bind to the NC and that overexpression of PNTD in cells increases the polymerase activity. These two results are different from our data since RSV PNTD did not bind to N-RNA rings whereas PNTD and PCTD can both bind to Nmono and overexpression of PNTD led to a decrease of the RSV polymerase activity.

Finally there seems to be a fundamental difference between Paramyxovirinae and Rhadoviridae concerning the organization of the functional domains of the P protein. Indeed, the L binding site has also been localized in the N-terminal part of rhadovirus P protein (47, 48). For rabies virus, it was proposed that these two functional domains do not overlap but are juxtaposed, composed of residues 4 to 40 for N0 and 40 to 70 for L. In contrast, for Paramyxoviridae, the L main binding site has been localized at the C terminus of P (49, 50), but the N terminus of P could also bind to L (38). Although the L binding domain of RSV does not appear to be located in the N-terminal region of P but rather in either the oligomerization domain or the C-terminal region (19, 30), it would be worthwhile to revisit this issue, since it could be of importance to fully understand the mechanism of N0 release by P in the context of L and RNA synthesis.

ACKNOWLEDGMENTS

We thank John Barr and David Karlin for critical readings of the manuscript, Origène Nyangulé and Christina Sizun for helpful discussions, and Damien Vitour for providing polyclonal rabbit anti-N antibody. Molecular graphics images were produced using Pymol software.

This work was carried out with the financial support of the French Agence Nationale de la Recherche, specific program ANR Blanc 2011 “Bronchiolitase” (ANR-11-BSV8-0024).

REFERENCES


