Biochemical characterization of the respiratory syncytial virus P–P and P–N protein complexes and localization of the P protein oligomerization domain

Nathalie Castagné,1 Alexandra Barbier,1 Julie Bernard,1 Human Rezaei,1 Jean-Claude Huet,2 Céline Henry,2 Bruno Da Costa1 and Jean-François Éléouët1

Correspondence
Jean-François Éléouët
eleouet@jouy.inra.fr
Unité de Virologie et Immunologie Moléculaires1 and Unité de Biochimie et Structure des Protéines2, INRA, 78350 Jouy-en-Josas, France

Received 20 November 2003
Accepted 22 January 2004

The RNA-dependent RNA polymerase complex of respiratory syncytial virus (RSV) is composed of the large polymerase (L), the phosphoprotein (P), the nucleocapsid protein (N) and the co-factors M2-1 and M2-2. The P protein plays a central role within the replicase–transcriptase machinery, forming homo-oligomers and complexes with N and L. In order to study P–P and N–P complexes, and the role of P phosphorylation in these interactions, the human RSV P and N proteins were expressed in E. coli as His-tagged or GST-fusion proteins. The non-phosphorylated status of recombinant P protein was established by mass spectrometry. GST-P and GST-N fusion proteins were able to interact with RSV proteins extracted from infected cells in a GST pull-down assay. When co-expressed in bacteria, GST-P and His-P were co-purified by glutathione-Sepharose affinity, showing that the RSV P protein can form oligomers within bacteria. This result was confirmed by chemical cross-linking experiments and gel filtration studies. The P oligomerization domain was investigated by a GST pull-down assay using a series of P deletion constructs. This domain was mapped to a small region situated in the central part of P (aa 120–150), which localized in a computer-predicted coiled-coil domain.

When co-expressed in bacteria, RSV N and P proteins formed a soluble complex that prevented non-specific binding of N to bacterial RNA. Therefore, RSV P protein phosphorylation is not required for the formation of P–P and N–P complexes, and P controls the RNA binding activity of N.

INTRODUCTION

Respiratory syncytial virus (RSV) belongs to the genus Pneumovirus of the family Paramyxoviridae (Collins et al., 2001). It is an enveloped virus and its genome is composed of a single-stranded non-segmented negative-sense RNA, approximately 15 kb in length, which is encapsidated by the nucleocapsid (N) protein. The resulting RNase A-resistant helical ribonucleoprotein complex is the template for transcription and replication of the genome and interacts with the polymerase complex. This RNA-dependent RNA polymerase complex is composed of the L (large polymerase) protein, the P (phospho) protein, the RNA-binding, transcription processivity factor M2-1 (Collins et al., 2001) and M2-2, which works as a replication/transcription regulator (Bermingham & Collins, 1999). Direct interaction between P and M2-1 was recently demonstrated (Mason et al., 2003). N, P and L are the minimal components for viral RNA replication (Grosfeld et al., 1995; Yu et al., 1995).

The mechanisms of specific recognition and encapsidation of the viral genomic RNA by the nucleocapsid protein are not well understood. The encapsidation signals are contained in the leader (Le) and Trailer (Tr) regions, which are very similar and situated at the 3′ ends of the genomic and antigenomic RNAs, respectively (Collins et al., 2001). When expressed alone in insect cells, the RSV N proteins form nucleocapsid-like structures, due to non-specific binding to cellular RNA (Meric et al., 1994; Bhella et al., 2002). Likewise, when expressed alone in E. coli, the N protein binds RNA in a non-specific manner (Murphy et al., 2003). Thus, the nucleocapsid proteins of these viruses have intrinsic RNA-binding properties that might be controlled or modulated by other factors. For the paramyxoviruses Sendai virus and measles virus, it has been proposed that prior to nucleocapsid assembly, N associates with the P protein, acting as a chaperone for N and preventing non-specific binding to cellular RNA (Curran et al., 1995;
Sphnner et al., 1997). Co-expression of the rabies virus N and P proteins in insect cells allowed the purification of a N–P soluble complex that does not contain RNA (Mavrikas et al., 2003). Thus, the N protein might be targeted specifically to the Le and Tr nascent chains as an N–P complex.

The RSV P protein can interact with L (Khattar et al., 2001a), N (Garcia et al., 1993; Samal et al., 1993; Garcia-Barreno et al., 1994; Mazumder et al., 1994; Villanueva et al., 1994; Sanchez-Seco et al., 1995; Dupuy et al., 1999). The paramyxovirus P protein is thought to exert at least two different functions: (i) to position the L molecule (Navarro et al., 1991), M2-1 (Mason et al., 2003) and with itself, forming homo-tetramers (Asenjo & Villanueva, 2000). It is a 241 aa protein that is phosphorylated essentially by cellular casein kinase II on several serine residues located in the middle (positions 116, 117 and 119) and in the C terminus (positions 253 and 257) of the molecule (Navarro et al., 1991; Mazumder et al., 1994; Villanueva et al., 1994; Sanchez-Seco et al., 1995; Dupuy et al., 1999). The paramyxovirus P protein is phosphorylated and transcriptional activity (Barik et al., 1995; Schumacher et al., 1996; Asenjo & Villanueva, 2000; Villanueva et al., 2000). However, structural modifications of the P protein, rather than loss of phosphorylation of these residues, could explain the apparent requirement for phosphorylation (Lu et al., 2002). It is not known whether association of RSV P with N or with L depends on its oligomeric state and phosphorylation.

In the present work, we have expressed the RSV N and P proteins in E. coli to study P–P and P–N complexes. The ability of unphosphorylated P to form oligomers and to associate with N was investigated, and the P oligomerization domain was mapped to a short amino acid stretch in the central part of P.

### METHODS

Plasmid constructions. Random-primed cDNA synthesis was done using Superscript II (Gibco-BRL) and 1 µg of total cytoplasmic RNA isolated from Hep-2 cells infected with the long strain of RSV (Villanueva et al., 2000). The cDNAs encoding P and N proteins were amplified by PCR using high fidelity PfuTurbo Polymerase (5 U; Stratagene), 100 ng of each primer listed in Table 1, and then cloned in pGEX-4T3 expression vector (Pharmacia). PCR-amplified DNA was digested by BamHI and inserted at the BamHI/Smal sites into pGEX-4T3 vector (Pharmacia). The N and P genes were subsequently subcloned into the pET-28b(+) vector (Novagen) at the BamHI/XhoI sites.

### Table 1. Oligonucleotides for PCR and site-directed mutagenesis

The introduced restriction enzyme cleavage sites are underlined. RSV-specific sequences are in bold. Introduced stop codons are in italic.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5′→3′</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P5′Bam+</td>
<td>GAGGGATCCATCATGGAAGGAAATTTTGCTCCTG</td>
<td>5′ primer for P</td>
</tr>
<tr>
<td>P′-3′</td>
<td>CTGTGGTTGTTGTTGTTGAAGGCGA</td>
<td>3′ primer for P, PA50N and PA161N</td>
</tr>
<tr>
<td>N5′Bam+</td>
<td>GAGGGATCCATGCTCTTTAACAAAGTGCAAGTTG</td>
<td>5′ primer for N</td>
</tr>
<tr>
<td>N3′</td>
<td>TTAACCTCAAGCCCTACATTTATCTTGG</td>
<td>3′ primer for N</td>
</tr>
<tr>
<td>P50Bam+</td>
<td>GAGGGATCCACAAAAGGACCGCTTATACATCG</td>
<td>5′ primer for PA50N</td>
</tr>
<tr>
<td>P90Bam+</td>
<td>GAGGGATCCACACCGCTATACAGGTGATAT</td>
<td>5′ primer for P90–170</td>
</tr>
<tr>
<td>P100Bam+</td>
<td>GAGGGATCCACTATACAGGTGATAT</td>
<td>5′ primer for P100–166</td>
</tr>
<tr>
<td>P120Bam+</td>
<td>GAGGGATCCATGCTATACAGGTGATAT</td>
<td>5′ primer for P120–150</td>
</tr>
<tr>
<td>P161Bam+</td>
<td>GAGGGATCCATGCTAGGCTAGGTTGATAT</td>
<td>5′ primer for PA161N</td>
</tr>
<tr>
<td>P150−</td>
<td>GTGAAGCATCTCTAGTATTC</td>
<td>3′ primer for P120–150</td>
</tr>
<tr>
<td>P166−</td>
<td>CTATACCTACATCCTGATG</td>
<td>3′ primer for P100–166</td>
</tr>
<tr>
<td>P170−</td>
<td>CCATGGGATCTCTTATACAT</td>
<td>3′ primer for P90–170</td>
</tr>
<tr>
<td>PA230C+</td>
<td>CCGTTGGAGGAAGGAAATTTGAAGGTGACATG</td>
<td>+ sense primer for PA230C</td>
</tr>
<tr>
<td>PA230C−</td>
<td>GATCATTGCACTCTATACAGGTGACATG</td>
<td>antisense primer for PA230C</td>
</tr>
<tr>
<td>PA182C+</td>
<td>TGAATGAAAGAAATCTGCAAAGGATCTTTATAGG</td>
<td>+ sense primer for PA182C</td>
</tr>
<tr>
<td>PA182C−</td>
<td>GATCATTGCACTCTATACAGGTGACATG</td>
<td>antisense primer for PA182C</td>
</tr>
<tr>
<td>PA161C+</td>
<td>GCAAGTGGCAGGAGACTACAGGTGATGATG</td>
<td>+ sense primer for PA161C</td>
</tr>
<tr>
<td>PA161C−</td>
<td>CTATACCTACATCCTGATG</td>
<td>antisense primer for PA161C</td>
</tr>
<tr>
<td>PA120C+</td>
<td>GAATCTAGCTATTCTATATAGGTTGATAG</td>
<td>+ sense primer for PA120C</td>
</tr>
<tr>
<td>PA120C−</td>
<td>CTGATCATTTATCTTATTAGAAGATAGCTAG</td>
<td>antisense primer for PA120C</td>
</tr>
</tbody>
</table>
The C-terminal deletion mutants of the P protein were generated by introduction of stop codons using the *Pfu* DNA polymerase with the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer’s instructions, with the primers listed in Table 1. N-terminal deletion mutants and internal P domains were obtained by PCR using *Pfu* DNA polymerase (Stratagene) and primers listed in Table 1, and then subcloned into pGEX-4T3 vector at the BamHII/Smal sites. The integrity of all constructs was assessed by DNA sequencing.

**Expression and purification of GST-fusion proteins.** *E. coli* BL21(DE3) cells transformed with pGEX plasmids were grown at 37°C for 8 h in 125 ml of Luria–Bertani (LB) medium containing 100 µg ampicillin ml⁻¹. The same volume of LB was then added and protein expression was induced by adding 80 µg IPTG ml⁻¹ to the medium. Bacteria were grown at 25°C and harvested by centrifugation 15 h after induction. Bacterial pellets were resuspended in 10 ml of lysis buffer (50 mM Tris/HCl pH 7-8, 60 mM NaCl, 1 mM EDTA, 2 mM DTT, 0-2 % Triton X-100, 10 mM MgSO₄, 1 mM CaCl₂) supplemented with complete protease inhibitor cocktail (Roche) and incubated for 1 h on ice. RQ1 RNase-free DNase (Roche) was added to the lysates (final concentration 1 U ml⁻¹) and incubated at 25°C for 30 min. Lysates were spun at 10 000 g for 30 min at 4°C. Glutathione–Sepharose 4B beads (Pharmacia) were added to the clarified supernatants (100 µl of beads for 250 ml of lysis buffer) and washed three times with lysis buffer and then stored at 4°C.

Cultures of recombinant *E. coli* BL21 harbouring the two plasmids pGEX and pET were grown in 125 ml of LB containing ampicillin (100 µg ml⁻¹) and kanamycin (50 µg ml⁻¹), and harvested as described above.

To separate RSV recombinant proteins from GST, beads were incubated with biotinylated thrombin as described by the manufacturer (Novagen). Biotinylated thrombin was removed by the Thrombin Cleavage Capture kit as described by the manufacturer (Novagen).

**In vitro translation of N and P proteins.** [³⁵S]Methionine-labelled N and P proteins were translated in vitro using the TNT kit (Promega). Reactions were carried out using 1 µg of pET-N or pET-P plasmid in a 50 µl transcription/translation reaction containing 2 µl of translation grade [³⁵S]methionine (1000 µCi ml⁻¹ (37 MBq ml⁻¹); ICN).

**GST pull-down assays.** Mock- or RSV-infected Hep-2 cells [24 h post infection (p.i.)] grown in 75 cm² dishes were incubated with 50 µl of Promix [1000 µCi ml⁻¹ (37 MBq ml⁻¹); ICN] in 5 ml of culture medium. Cells were incubated overnight at 37°C, washed twice in 1× PBS and lysed for 1 h at 4°C in 4 ml of lysis buffer containing 10 mM Tris/HCl pH 7-2, 350 mM NaCl, 0-5 mM EGTA, 0.5 mM EDTA, 1 % Triton X-100, 1 mM DTT, 20 % glycerol, supplemented with complete protease inhibitor cocktail (Roche). Lysates were clarified by centrifugation at 10 000 g for 15 min at 4°C. Ten microlitres of the prepared glutathione–Sepharose 4B beads bound to GST-fusion proteins were mixed with cell lysates (250 µl) and rotated end-over-end for 4 h at 4°C. For assays employing *in vitro* translated proteins (IVT), beads were incubated with 5 µl of ³⁵S-labelled IVT reactions in 50 µl of bacteria lysis buffer described above. The beads were collected by centrifugation for 5 min at 500 g at 4°C and washed three times with lysis buffer. Pull-down eluates were heat denatured in Laemmli buffer and run on 12 % SDS-polyacrylamide gels, fixed and stained in solution containing 20 % ethanol, 10 % acetic acid and 0.25 % Coomassie brilliant blue R250, washed in fixing solution (20 % ethanol and 10 % acetic acid), dried and exposed by autoradiography.

**Generation of antisera.** Polyclonal antisera were prepared by immunizing rabbits three times at 2 week intervals using purified GST-fusion proteins (100 µg) for each immunization. The first and second immunizations were administered subcutaneously in 1 ml Freund’s complete and Freund’s incomplete adjuvant (Difco), respectively. The third immunization was done intramuscularly in Freund’s incomplete adjuvant. Animals were bled 10 days after the third immunization.

**Mass spectrometry analysis.** A 10 µl volume of recombinant protein was used (concentration 1-5 mg ml⁻¹). The sample was desalted on ZipTip C4 (Millipore) and eluted with 1 µl of 70 % (v/v) acetonitrile/0-3 % (v/v) trifluoroacetic acid. The sample was directly spotted onto the MALDI plate and dried at room temperature: 0-5 µl of 2,5-dihydroxybenzoic acid (DHB; 10 mg ml⁻¹) in water was added. Mass spectra were acquired on a Voyager-DE-STR time-of-flight mass spectrometer (Applied Biosystems) equipped with a nitrogen laser emitting at λ = 337 nm (Laser Science). The accelerating voltage used was 25 kV. All spectra were recorded in the positive reflector mode with a delayed extraction of 1700 ns and a 94 % grid voltage. The spectra were calibrated using an external calibration: cytochrome c, [M+H]+ = 12362.0 Da; trypsin [M+H]+ = 23464.5 Da.

**Amino acid sequencing.** Electrophoretic samples were transferred onto PVDF membrane by passive absorption. The bands of interest were excised, dried in a Speed-vac for 30 min, and the gel pieces were rehydrated in the initial volume of excised band (2 % SDS, 0-2 M Tris/HCl pH 8-5) for 1 h. After rehydration, 5 vols of water were added and then a piece of pretwet (methanol) 4 × 4 mm PVDF membrane (Problott Applied Biosystems) was added to the solution. At the end of this transfer time (2 days at room temperature), the membrane was washed five times with 1 ml 10 % methanol with vortexing. N-terminal Edman sequencing was performed on an Applied Biosystems Procise 494HT with reagents and methods recommended by the manufacturer.

**Cross-linking assays.** GST-fusion proteins bound to glutathione–Sepharose beads were extensively washed with 1× PBS and cleaved with thrombin as described by the manufacturer (Novagen). The resulting proteins were incubated for 1 h at room temperature with increasing concentrations of freshly diluted glutaraldehyde in 1× PBS. Reactions were stopped by the addition of lysine at 200 mM for 30 min at room temperature.

**RESULTS**

**Expression and purification of GST-N and GST-P fusion proteins.** The cDNAs encoding the RSV N and P proteins were amplified by RT-PCR and cloned into pGEX-4T3 *E. coli* expression vector. Expression of the resulting GST-fusion recombinant proteins was induced with IPTG, and bacterial lysates were analysed by SDS-PAGE. Overproduced polypeptides were observed with apparent molecular masses of 67 (GST-N) and 60 kDa (GST-P) (Fig. 1A). GST, GST-P and GST-N proteins were affinity purified on glutathione–Sepharose 4B beads. Each protein was purified to > 95 % homogeneity as estimated by Coomassie brilliant blue staining (Fig. 1B). The recombinant GST-P and GST-N fusion proteins were used to immunize rabbits to obtain anti-P and anti-N polyclonal antibodies.

**GST-P and GST-N purified recombinant proteins can interact with the RSV polymerase complex**

A GST pull-down assay utilizing GST-P and GST-N fusion proteins, or GST as a negative control, was performed to test
whether GST-N and GST-P produced in E. coli were able to interact with the RSV polymerase complex. Glutathione-Sepharose-bound GST, GST-N and GST-P proteins were incubated overnight at 4°C with 35S-labelled cellular extracts from RSV- or mock-infected Hep-2 cells. After extensive washing, SDS-PAGE analysis of the resulting complexes revealed that GST-N and GST-P, but not GST, were able to pull down RSV-specific proteins (Fig. 2A). Four radioactive pulled down polypeptides, indicated by numbers 1 to 4 on Fig. 2(A), had apparent molecular masses of >94, 40, 35 and 25 kDa, corresponding to those expected for L, N, P and M2-1 RSV proteins, respectively. Very few polypeptides were pulled down by GST-P or GST-N from uninfected Hep-2 cells, compared with RSV-infected cells, indicating that these interactions were virus-specific. To confirm that GST-N directly interacted with RSV P proteins, and that GST-P could interact with N and P, a pull-down assay was done by using 35S-labelled in vitro-translated N and P proteins (Fig. 2B). 35S-Labelled P protein was pulled down by both GST-P (lane 2) and GST-N (lane 3), and 35S-labelled N protein was pulled-down by GST-P (lane 5). The phosphoprotein of the infectious haematopoietic necrosis virus (IHNV), a fish rhabdovirus, was used as a control for N–P and P–P interaction specificity. The in vitro translated, 35S-labelled IHNV P protein (kindly provided by Michel Brémont, INRA Jouy-en-Josas), was not pulled down by either GST-N or GST-P (lanes 6, 7). These results showed that the GST-N and GST-P proteins purified from bacteria were able to interact with their viral counterparts extracted from infected cells.

**RSV P* recombinant protein is not phosphorylated**

In order to analyse the phosphorylation status of N and P expressed in E. coli; RSV proteins were separated from GST by thrombin cleavage. The resulting recombinant proteins were further designated P* and N*. Beads containing bound GST and insoluble products were collected together by centrifugation for 15 min at 10,000 g. The efficiency of thrombin cleavage and solubility of P* and N* proteins were analysed by SDS-PAGE. The P protein was efficiently released after thrombin cleavage as a soluble protein migrating with an apparent mass of 35 kDa, as expected (Fig. 3A). In contrast, cleavage of GST-N was poorly efficient (5–10 %), and cleavage products were not soluble (Fig. 3B). These results indicated that the thrombin cleavage site was probably hidden by N in the GST-N fusion protein, and that N* might form non-soluble aggregates, possibly in association with bacterial RNA.

---

**Fig. 1.** Expression and purification of GST, GST-N and GST-P proteins in E. coli. Proteins were denatured in Laemmli buffer, resolved by SDS-PAGE on 12% polyacrylamide gels and detected by staining with Coomassie brilliant blue. (A) Total bacterial lysates before (-) and after (+) induction of protein expression with IPTG, 100 μl bacterial culture equivalent per lane; the GST-N protein is indicated by a symbol (●). (B) Purification of GST-fusion proteins with glutathione–Sepharose beads. Two microlitres of beads/wash buffer [1/1 (v/v)] were loaded per lane. m, Protein molecular size standards.

---

**Fig. 2.** GST-P and GST-N bind HRSV polymerase complex proteins. Pull-down assays were conducted with GST, GST-P or GST-N (2 μg per lane) immobilized on glutathione–Sepharose 4B beads. (A) Pull-down with 35S-labelled Hep-2 cells lysates before (-) or 24 h after (+) infection by RSV. The pulled-down radioactive proteins are indicated by arrows. (B) Pull-down with 35S-labelled in vitro-translated RSV P (lanes 1–3) and N (lanes 4 and 5) proteins, or IHNV P protein (lanes 6–8). Lane 8, 1 μl of IVT IHNV P protein product was loaded. The glutathione–Sepharose-adsorbed fusion proteins used for pull-down assays are indicated at the top. The radio-labelled, in vitro-translated proteins used for pull-down assays are indicated at the bottom. Pull-down eluates were separated by 12% SDS-PAGE and exposed to film for autoradiography. Protein molecular mass standards are indicated.
In order to determine whether the recombinant P* protein presents any post-translational modifications, its molecular mass was measured by MALDI-TOF (Fig. 3C). After cleavage by thrombin, the recombinant P* has three additional amino acid residues (GSI) at its N terminus, and a predicted mass of \([M+H]^+=27406.2\) Da. The mass of P* measured by MALDI-TOF was \([M+H]^+=27407.6\) Da. As the molecular mass of a phosphate group is 80 Da, and since there was no significant difference between the predicted and the measured masses, these results showed that the RSV P protein is neither phosphorylated nor post-translationally modified in E. coli.

**Unphosphorylated RSV P* recombinant protein forms oligomers**

To determine whether the RSV P protein expressed in E. coli was able to form oligomers, we co-expressed GST-P and His-P in bacteria by using a double antibiotics selection protocol. The ampicillin + kanamycin-resistant colonies were amplified in LB containing both antibiotics, and expression of proteins was induced with IPTG. After purification of GST-P with glutathione–Sepharose beads, the purified proteins were analysed by SDS-PAGE. As shown in Fig. 4(A), two polypeptides were co-purified by glutathione–Sepharose affinity, with apparent masses of \(\pm 50\) and \(\pm 40\) kDa, as expected for GST-P and His-P, respectively. The identity of GST-P and His-P was determined by mass spectrometry and N-terminal sequencing. This experiment showed that the unphosphorylated form of RSV P protein formed stable oligomers in bacteria.

We also investigated the oligomeric status of P*. First, the homogeneity of the P* protein preparation was analysed by size-exclusion chromatography. The purified P* protein was loaded on a Zorbax GF-450 column at 1 mg ml\(^{-1}\). The expected masses for P trimers and tetramers are 71 and 108 kDa, respectively. The protein was eluted as a single major peak with an apparent mass between aldolase (158 000) and albumin (67 000), and appeared highly homogeneous (Fig. 4B). This result is consistent with P* being mostly present as a trimer or a tetramer in the preparation. In a second step, the oligomeric status of P* was investigated by chemical cross-linking. The purified P* protein was incubated with increasing amounts of glutaraldehyde at final concentrations of 0·005 to 0·25 % for 1 h at room temperature. A control reaction for glutaraldehyde cross-linking was performed with GST. The cross-linked products were analysed by SDS-PAGE and stained with Coomassie blue (Fig. 4C). P* protein was cross-linked into slower-migrating forms (indicated by II, III and IV in Fig. 4C). As expected, GST was not cross-linked by glutaraldehyde. Based on their apparent molecular mass, species II (70 kDa) and III (140 kDa) had mobilities consistent with those expected for P dimers and tetramers, respectively. The larger multimer (IV) had a molecular mass slightly above 250 kDa and might correspond to a higher oligomeric form. Thus, the recombinant P* protein was able to assemble efficiently into oligomers.

To determine whether some other oligomeric forms of P*, such as trimers, were present in the samples but in lower amounts, the P* protein was cross-linked with glutaraldehyde at a final concentration of 0·025 % and analysed by Western blot using an anti-P rabbit polyclonal serum, as this technique is more sensitive than Coomassie blue staining. The strength of P–P interactions was estimated by incubating the P* protein with increasing concentrations from 0 to 1 M of NaCl. As shown in Fig. 4(D), only two major complexes migrating with apparent masses of 70 and 140 kDa were again detected. These results suggest that the P complex might be constituted of two dimers that assemble into tetramers. Formation or stability of P dimers and tetramers were not inhibited by high concentrations of NaCl (1 M).
Prediction of the coiled-coil oligomerization domain of P

RSV P protein sequence was submitted to a coiled-coil prediction program that determined two potential regions situated approximately between aa 120–150 and 175–215, respectively (Fig. 5A). Generally, coiled-coils are alpha-helices which have hydrophobic interfaces and a hydrophilic exterior. Every first and fourth amino acid of a 7 residue repeat in the helix is hydrophobic. As shown in Fig. 5(B), the first predicted domain (aa 129–152) presents a cyclic (3 or 4 aa) hydrophobic residues repetition. This was not found in the second computer-predicted region (175–215). Interestingly, sequence alignment of phosphoproteins of several pneumoviruses revealed that the putative first coiled-coil domain is part of a highly conserved domain (Fig. 5C). This conserved domain is situated in the central part of P proteins and seems to be limited in C terminus by a conserved GP (positions 158–159) that is incompatible with an alpha-helix structure. The other parts of the molecule show low sequence similarities between these pneumoviruses.

Mapping of the P oligomerization domain

In order to experimentally characterize the P oligomerization domain, a series of N-terminal, C-terminal and internal deletions mutants were made throughout the 241 aa of P protein fused to GST in the pGEX-4T3 vector (Fig. 6A). GST-P deletion mutants were co-expressed in E. coli with the His-P protein. As replication rates of pET and pGEX plasmids are different, this experiment should not be regarded as a quantitative one. After protein purification by glutathione–Sepharose beads and analysed by SDS-PAGE stained with Coomassie brilliant blue (Fig. 4A), the absorbance signal of the eluate from the Agilent ZORBAX Bio Series GF-450 column loaded with P*. Sample (100 µl) was run in PBS. The following size markers (Amersham Pharmacia Biotech) were used to calibrate the column (in parentheses, apparent Mr): 1, thyroglobulin (669 000); 2, catalase (232 000); 3, aldolase (158 000); 4, bovin serum albumin (67 000); 5, chymotrypsinogen A (25 000); 6, ribonuclease A (13 700). (C) Coomassie blue-stained PAGE of cross-linking of P* protein with increasing concentrations of glutaraldehyde (left) compared with GST (right). (D) Western blot analysis of P* cross-linked with 0.25% glutaraldehyde (noted Glut.) in the presence of increasing concentrations of NaCl.
Fig. 5. Computer predictions on the HRSV Long strain P protein sequence. (A) Probability of coiled-coil formation in RSV P protein, scored with the algorithm of Lupas et al. (see http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_lupas.html). The program was run three times with windows of 14, 21 and 28. (B) Helical net diagram of the putative coiled-coil domain of P. The aligned hydrophobic residues are indicated in bold letters. (C) Alignment of P protein of HRSV, human metapneumovirus (HMPV), avian pneumovirus (APV) and pneumonia virus of mice (PVM) with the CLUSTALW (1.74) multiple sequence alignment program. Stars indicate amino acids identities, points similarities. The first coiled-coil putative domain shown in (B) is underlined. Positions of amino acid residues surrounding this region are numbered for RSV P protein.
PA120C (Fig. 6), indicating that the C terminus of the P protein is not necessary for oligomerization. Truncation of 50 aa at the N-terminal end of P (mutant PA50N) still allowed co-purification of His-P, when PA161N lost this capacity, suggesting that the oligomerization domain might be located between aa 50–160. To localize precisely the P–P binding domain, a series of constructs with both N- and C-terminal deletions were tested. The P90–170, P100–166 and P120–150 mutants retained P-binding capacity. Thus, terminal deletions were tested. The P90–170, P100–166 and P120–150 mutants retained P-binding capacity. These data demonstrated that the RSV P internal domain 120–150 completely abolished the GST-P/His-P interaction into bacteria, whereas a domain restricted to the amino acid stretch 120–150 retained the binding activity. These data demonstrated that the RSV P internal domain 120–150 that was sufficient for P oligomerization.

### RSV N and P recombinant proteins form a soluble complex in E. coli

In order to determine whether RSV N and P proteins could interact in bacteria and form a soluble N–P complex in the absence of phosphorylation, the two proteins were co-expressed in E. coli in two different combinations, GST-N plus His-P and GST-P plus His-N. In both cases, purification of GST-fusion proteins by glutathione–Sepharose 4B affinity resulted in the co-purification of two polypeptides (Fig. 7A). The apparent molecular masses of the purified polypeptides corresponded to those expected for GST-N (69 kDa) and His-P (38 kDa) (lane 1) and GST-P (61 kDa) and His-N (45 kDa) (lane 2). Therefore, co-expression of RSV N and P proteins allowed the formation of N–P complex(es) that co-purified. The solubility of N–P complexes was analysed by centrifugation of proteins after thrombin cleavage that might eliminate both GST and His-tag. Pellets and supernatants were analysed by SDS-PAGE and stained with Coomassie blue. As shown in Fig. 7B, thrombin cleavage of GST-P+His-N resulted in the release of a soluble N–P complex designated N0–P that was present in the supernatant. The two proteins present in the soluble fraction and stained with Coomassie blue migrated with apparent masses of 35 and 43 kDa, as expected for P and N, respectively. However, GST-N+His-P was not correctly cleaved by thrombin. Very low amounts of proteins were present in the soluble fraction and proteins were shorter than expected, indicating that non-specific cleavage occurred. The use of lower concentrations of thrombin did not improve the efficiency of the cleavage. Thus, the GST-P+His-N combination was used for further studies.

The composition and stoichiometry of the N0–P complex were then investigated by chemical cross-linking followed by Western blot analysis. Proteins were incubated with increasing amounts of glutaraldehyde at final concentrations of 0 to 10% for 1 h at room temperature. Western blotting was performed using anti-N and anti-P specific rabbit anti-sera. As shown in Fig. 7C, N and P were detected by the corresponding anti-sera in a specific manner. Only one band, corresponding to a complex with an apparent mass of about 75 kDa, was detected by both antibodies. It may correspond to a N0–P complex with a 1:1 stoichiometry for which the expected mass is 43 + 35 = 78 kDa in SDS-PAGE. A slower migrating band (70 kDa) was also revealed by anti-P serum only, which might correspond to P dimers. For higher glutaraldehyde concentration (0–1%), P aggregation was observed (XP on Fig. 7C). These results confirmed that purified proteins from bacteria expressing GST-P+His-N were heterogeneous and composed of at least two complexes that were identified as an N–P complex and P dimers.

### Bacterial RNA is co-purified with GST-N, but not with GST-N+His-P or GST-P+His-N complexes

When the GST-N fusion protein was cut by thrombin or eluted from glutathione–Sepharose by 50 mM glutathione, the protein was essentially present in the pellet after centrifugation for 1 min at 10 000 g, indicating that it was not soluble, possibly because of the association of N with bacterial RNA. In order to determine whether RNA was...
present in the preparations, 10 μg of GST-N, GST-N + His-P or GST-P + His-N purified protein complexes adsorbed on glutathione–Sepharose beads was heat denatured for 5 min at 100 °C, and run on a 2 % agarose gel stained with ethidium bromide. As shown in Fig. 7(D), presence of RNA was revealed for the GST-N protein, whereas only trace amounts of RNA were visible for GST-N + His-P and GST-P + His-N. The presence of bacterial RNA was never detected for GST or GST-P by this method (data not shown). These results showed that unspecific binding of N0 to bacterial RNA is prevented by the presence of P within the same bacteria, allowing soluble N0–P complex formation.

**DISCUSSION**

In this report, we have expressed the RSV P and N proteins in *E. coli*, in order to study (i) the role of protein phosphorylation in P–P and P–N interactions, (ii) to map protein interaction domains, and (iii) to isolate a soluble N0–P complex for further N0–P–RNA interaction and structural studies. For that purpose, we used an approach based on the screening of positive molecular interactions.

We investigated the role of phosphorylation in P oligomerization, firstly without modifying the amino acid sequence of the protein. In infected cells, about 80 % of P phosphorylation involves Ser-232, and the remainder involves Ser-116, -117, -119 and -237 (Barik et al., 1995; Sanchez-Seco et al., 1995). Mazumder et al. (1994) reported that P purified from bacteria was not phosphorylated and chromatographed with an apparent size of 120–150 kDa, suggesting that the P protein behaved as a trimer or a tetramer. Asenjo & Villanueva (2000) reported that recombinant RSV P protein purified from *E. coli* was unable to oligomerize. However, mutation of the five phosphorylated serines on RSV P protein did not abolish oligomer formation in Hep-2 cells. Thus, it was postulated that a transitory phosphorylation of P protein could occur and play a role in oligomerization. In this report, we clearly show that when expressed in bacteria the RSV P recombinant protein is able to oligomerize. Mass spectrometry analysis confirmed that recombinant P protein purified from *E. coli* was not phosphorylated. Gel filtration studies suggested that the purified P protein was homogeneous, with an apparent mass compatible with P being a tetramer. Chemical cross-linking experiments and Western blotting analysis of the cross-linked products confirmed the presence
of tetramers. At lower glutaraldehyde concentrations, complexes corresponding to P dimers were also found. Because trimers were not detected, the presence of P dimers might reflect a structural suborganization of the tetramers. As high molecular mass complexes were not observed by gel filtration, these complexes might only be formed during glutaraldehyde cross-linking. These results demonstrate that P phosphorylation is not required for P oligomerization. Furthermore, the P domain comprising aa 120–150, which does not contain phosphorylated serine, was still able to oligomerize. Thus, the role of RSV P protein phosphorylation might reside in other function(s). In vitro reconstituted transcription experiments suggested that P phosphorylation is involved in transcriptional activity of RSV P protein (Mazumder & Barik, 1994; Barik et al., 1995). Other data, obtained by mutagenesis of serines, suggested that P phosphorylation could play a role in P–L stabilization during transcription elongation activity of the viral polymerase (Dupuy et al., 1999). However, in vitro phosphorylation of P protein did not result in transcriptionally active P protein (Dupuy et al., 1999). Recently, Lu et al. (2002) suggested, by using reverse genetics, that P phosphorylation could be involved in virus budding rather than in virus replication. Finally, it was shown that P–M2-1 interaction is independent of P phosphorylation (Mason et al., 2003). All the data, together with our results, indicate that P phosphorylation is not needed for P–P, P–N and P–M2-1 interactions, and that RSV replication occurs without P phosphorylation. Even though P–L interactions have not been extensively investigated, we suggest that P phosphorylation could be required in other aspects of the virus-cycle. It could be involved in the formation of ‘frozen’ polymerase complex before the release of virions from infected cells, or necessary for interaction of the P protein with the Matrix (M) protein, to initiate the formation of new RSV particles.

The use of positive interaction screenings allowed us to map the P oligomerization domain to aa 120–150. Comparison between several members of the Paramyxoviridae family showed that the P protein is not well conserved at the amino acid level (not shown). However, sequence alignment restricted to four different members of the Pneumovirus genus revealed that this domain is situated in, and represents the major part of, a well conserved region in the middle of the molecule, the other parts showing low or no sequence similarities. Amino acid sequence comparison with the paramyxovirus Sendai virus phosphoprotein for which the atomic structure has been determined as being a tetramer (Tarbouriech et al., 2000), did not reveal any similarities (data not shown). However, these two domains could have a comparable secondary coiled-coil structure.

The role of P phosphorylation for N–P interaction has been investigated by using reverse genetics (Lu et al., 2002). N–P interaction was reduced to 40% when Ser-116, -117, -119 (central region), and -232 and -237 (C-terminal region) were mutated, suggesting that P phosphorylation could be important for N–P interactions. Our results showed that N0–P complex formation was very efficient in bacteria in the absence of phosphorylation. Thus, the reduced N–P interaction observed with recombinant viruses could be due to structural changes caused by serine mutations rather than lack of P phosphorylation.

The nucleocapsid protein of RSV associates with RNA with no sequence specificity when expressed alone, forming nucleocapsid-like structures (Bhella et al., 2002). When GST–N from bacteria was purified the fusion protein was poorly soluble after thrombin cleavage or elution by glutathione, and contained high amounts of RNA. Purified GST–N–RNA complexes were not displaced by addition of soluble P* protein (not shown). In order to maintain N in a soluble form (designated N0), we co-expressed N and P proteins in bacteria, and an N0–P complex was purified. This complex was soluble and did not contain high amounts of bacterial RNA. The same results were obtained recently with rabies virus N and P proteins co-expressed in insect cells, using recombinant baculoviruses (Mavrakis et al., 2003). Our experiments confirm that non-specific binding of RSV N protein to RNA is prevented by P, and that binding of N to RNA is strong and probably irreversible. The specificity of RNA-binding of the N0–P complex to viral RNA sequences will be investigated in future experiments using this purified N0–P complex.

ACKNOWLEDGEMENTS

We thank Mabel Berois for providing the Hep-2 cell line and the HRSV Long strain, Lauriane Renault for excellent technical help, Bernard Delmas, Isabelle Nobiron (INRA, Jouy-en-Josas) and Felix Rey (CNRS, Gif-sur-Yvette) for useful discussions, Michel Brémond (INRA, Jouy-en-Josas) for advice and providing the IHNV P protein encoding plasmid.

REFERENCES


Grosfeld, H., Hill, M. G. & Collins, P. L. (1995). RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. *J Virol* 69, 5677–5686.


