Location of phosphorylated residues in human respiratory syncytial virus phosphoprotein

J. Navarro, 1 C. López-Otin 2 and N. Villanueva 1*

1 Servicio Virología, Centro Nacional de Microbiología, Virología e Inmunología Sanitarias, Instituto de Salud Carlos III, Majadahonda, 28220 Madrid and 2 Departamento de Biología Funcional, Área de Bioquímica y Biología Molecular, Facultad de Medicina, Universidad de Oviedo 33006, Oviedo, Spain

The phosphoprotein (P protein) from human respiratory syncytial virus Long strain, labelled in vivo with [32P]orthophosphate, was purified from virions or virus-infected human epithelial (Hep-2) cells. The main phosphorylated amino acid found was serine. The determination of the N-terminal sequence of unphosphorylated and phosphorylated fragments of P protein obtained after chemical or enzymic treatments suggested that some or all of the six serines present at positions 116, 117, 119, 143, 156 and 161 are the major phosphorylated residues, although a modification in serine residues at positions 86, 94 and 99 can not be ruled out.

Paramyxoviruses, such as respiratory syncytial (RS) virus, contain a single RNA molecule of negative polarity as the genome. By analogy with other family members, such as Sendai virus and Newcastle disease virus, it is thought that the genomic RNA is found tightly associated only with the nucleocapsid (NC) protein NP. This core structure is the template for RNA synthesis. The viral polymerase activity responsible for RNA synthesis seems to be a combination of the L and P proteins, which are associated with NCs less tightly than the NP protein (Kolakofsky & Roux, 1987). As in the case of rhabdoviruses (Emerson & Schubert, 1987), the L protein is considered to be the actual polymerase, whereas the precise role of the P protein in RNA synthesis remains undefined. The vesicular stomatitis virus (VSV) NS protein is considered to be the counterpart of the paramyxovirus P protein, and the degree of NS phosphorylation has been correlated, in vitro, with an increase in RNA polymerase activity (Kingsford & Emerson, 1980). However, the inhibition of kinase activity displayed by NCs did not affect its transcription activity (Massey et al., 1990). The inhibition of NS phosphorylation is correlated also, in vivo, with a decrease in viral VSV RNA synthesis (Müller-Decker et al., 1987).

In contrast with VSV, the lack of P protein phosphorylation in vivo does not affect NC function for RNA synthesis (Villanueva et al., 1991).

A step towards understanding how P protein phosphorylation is involved in virus development was to analyse and locate the phosphorylated residues. Here we describe their mapping within the P protein amino acid sequence previously deduced (López et al., 1988) for RS virus Long strain (Fig. 1). The deduced amino acid sequence of the P protein of the Long strain of RS virus (López et al., 1988) (Fig. 1) shows the existence of 23 serine, 18 threonine and four tyrosine residues which may be phosphorylated.

Fig. 1. Amino acid sequence of the P protein, deduced from the nucleotide sequence of the Long strain (López et al., 1988), is shown using the one letter code. The putative cleavage sites found upon digestion with chymotrypsin (** *), formic acid (****) and trypsin (********) are shown.
A subconfluent monolayer of Hep-2 cells was infected with RS virus Long strain at 1 p.f.u./cell. Twenty-four h post-infection (p.i.) the medium was removed and the cells were incubated with phosphate-deprived Dulbecco's MEM (DMEM) containing 0.1 volume of the corresponding complete medium, 2.5% dialysed foetal bovine serum and 100 μCi/ml H$_3$PO$_4$ (Amersham). After 48 h the cells were scraped off and the medium was clarified by low speed centrifugation. The supernatant was precipitated twice with polyethylene glycol 6000 (Merck) at a final concentration of 6%. The precipitated material was collected by centrifugation in a minifuge for 20 min at 4 °C and resuspended in TNE buffer (50 mM-Tris-HCl pH 7.5, 100 mM-NaCl and 10 mM-EDTA). This material was pelleted through a 33% sucrose cushion in TNE buffer at 240000 g for 2 h at 4 °C. The pellet was resuspended in the same buffer, and was considered to contain virions. The cell pellets, after two washes with phosphate-buffered saline, were resuspended in 10 mM-Tris-HCl pH 7.5, 0.15 M-NaCl, 5 mM-EDTA, 1% Triton X-100 and 1% sodium deoxycholate. The supernatants were collected after centrifugation in a minifuge for 15 min at 4 °C.

Labelled P protein was purified from cytoplasmic extracts or extracellular viral particles by immunoaffinity chromatography using the monoclonal antibody RS/IP (Garcia-Barreno et al., 1989). Protein samples purified by immunoaffinity chromatography were fractionated by 10% acrylamide gel electrophoresis (Studier, 1972), and the stained P protein band was cut off. Purified 32P-labelled P protein was hydrolysed with 6 M-HCl for 90 min at 110 °C and the phosphoamino acids were separated by one-dimensional electrophoresis on a paper layer using acetic acid, pyridine and water (50:5:45) at 900 V for 90 min. Phosphoserine, phosphothreonine and phosphotyrosine were included as markers and stained with ninhydrin. The position of 32P-labelled phosphoamino acids was determined by autoradiography (Fig. 2a), and shows that serine was the main phosphorylated residue, the modification observed for threonine and tyrosine being negligible. Additionally, chemical and enzymic treatments were carried out to cleave the P protein and locate the phosphoserine residues. The pieces of gel were either untreated or treated with 70% formic acid (Merck) or with 15 μg of trypsin (Sigma) or with 15 μg of chymotrypsin (Sigma)
for 18 h at 37 °C. The protein was eluted from gel pieces, vacuum-dried, resuspended in sample buffer and electrophoresed in a 20% acrylamide gel. The separated peptides were transferred to Immobilon paper by electrophoresis in the presence of 3 mM-Na₂CO₃-10 mM-NaHCO₃ (pH 9.9) in 20% methanol. The peptides were visualized by staining with amido black in 40% methanol, 10% acetic acid. The labelled peptides were localized by autoradiography. ³²P-labelled and, in some cases, unlabelled fragments were selected for sequencing. Identified peptides were sliced out, and the pieces of paper were mounted in the reaction chamber of a pulse liquid phase sequenator (Applied Biosystems) and subjected to five to seven cycles of automated Edman degradation (Hewick et al., 1981; Vanderkerckhove et al., 1985). The sequence determined for each band is shown in Table 1 by the single-letter code, as well as its apparent Mᵣ.

When the ³²P-labelled P protein was chemically digested by addition of formic acid, which preferentially cleaves at aspartate–proline bonds (Fontana & Gross, 1986) eight fragments, F1 to F8, were visualized by amido black staining (Table 1, Fig. 2b, lane 3). Some of them, such as F7, were found in very low amounts and probably were generated by non-specific acidic cleavage of peptide bonds. The N-terminal sequences of F2 and F6, the most strongly ³²P-labelled fragments, were the same, starting at residue 91 of the P protein. Fragments F1, F3 and F5 have blocked N termini residues, like that of the undigested P protein, suggesting that those fragments may contain the N-terminal end of the intact P protein. The N-terminal region of F4, F7 and F8 was also determined. The terminus of F4 corresponds to the glycine at position 165, that of F7 to asparagine at position 192 and that of fragment F8 corresponds to the proline residue present at position 34.

Since the P protein (241 residues, 33K) and its peptides, generated upon formic acid cleavage, show an anomalous migration on gel electrophoresis it is not possible to predict the location of the C terminus for these peptides. Nevertheless, after looking at the sequence of the P protein, indicated in Fig. 1, we found that the amount of methionine present in those peptides may give an approximation of their length. Hence, formic acid treatment was also done with [³⁵S]methionine-labelled P protein to determine the presence of methionine in the different fragments (Table 1). The digestion of ³²P-labelled P protein with trypsin (Fig. 2b, lanes 7 and 8) which cleaves peptide bonds at the C-terminal side of lysine and arginine residues (Wilkinson, 1986), generated only two major fragments, T1 and T2, both labelled, with apparent Mᵣ values of 15K and 14K. These fragments had the same N-terminal sequence, starting at glutamate residue number 104 of the P protein sequence. Chymotrypsin digestion of the ³²P-labelled P protein, which cleaves preferentially but not exclusively after aromatic residues (Wilkinson, 1986), resulted in different digestion patterns, depending on the incubation conditions. In one experiment two strongly labelled fragments, C1 and C2 of 20K and 15K, together with an unlabelled fragment, C3 of 9K, were observed (Fig. 2b, lanes 9 and 10). C3 started at position 29 of the P protein.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Fragment number</th>
<th>Apparent Mᵣ</th>
<th>Phosphate labelling*</th>
<th>N-terminal sequence</th>
<th>Starting residue (N terminus) of the peptide</th>
<th>[³⁵S]Methionine labelling*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formic acid</td>
<td>F1</td>
<td>27K</td>
<td>+ + +</td>
<td>Blocked</td>
<td>1</td>
<td>+ + +</td>
</tr>
<tr>
<td></td>
<td>F2</td>
<td>20K</td>
<td>+ + +</td>
<td>PIPS (PIPSD)</td>
<td>91</td>
<td>+ + +</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>17K</td>
<td>+</td>
<td>Blocked</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>15K</td>
<td>-</td>
<td>GIRDA</td>
<td>165</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>F5</td>
<td>14.5K</td>
<td>+</td>
<td>Blocked</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>F6</td>
<td>14K</td>
<td>+ + +</td>
<td>PIPS (PIPSD)</td>
<td>91</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>F7</td>
<td>10K</td>
<td>-</td>
<td>LIXAM</td>
<td>192</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>F8</td>
<td>3K</td>
<td>-</td>
<td>PKKKD</td>
<td>34</td>
<td>-</td>
</tr>
<tr>
<td>Trypsin</td>
<td>T1</td>
<td>15K</td>
<td>+</td>
<td>ETIETF</td>
<td>104</td>
<td>ND†</td>
</tr>
<tr>
<td>Chymotrypsin</td>
<td>C1</td>
<td>20K</td>
<td>+ + +</td>
<td>ND</td>
<td>104</td>
<td>ND ND§</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>14K</td>
<td>+ + +</td>
<td>ND</td>
<td>ND</td>
<td>ND ND§</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>9K</td>
<td>-</td>
<td>TSPKD</td>
<td>29</td>
<td>ND ND§</td>
</tr>
<tr>
<td></td>
<td>C4</td>
<td>12K</td>
<td>+ + +</td>
<td>KETIET</td>
<td>103</td>
<td>ND ND§</td>
</tr>
<tr>
<td></td>
<td>C5</td>
<td>5K</td>
<td>-</td>
<td>Blocked</td>
<td>1</td>
<td>ND ND§</td>
</tr>
<tr>
<td></td>
<td>C6</td>
<td>2.5K</td>
<td>-</td>
<td>ND</td>
<td>ND</td>
<td>ND ND§</td>
</tr>
</tbody>
</table>

* Symbols (+ + +, + +, +, −) indicate strength of labelling.
† ND, Not determined.
§ ND, Not determined.
In other experiments involving chymotrypsin digestion, a major labelled fragment, $C_4$ with an $M_r$ of 12K (Fig. 2b, lanes 11 and 12) was generated, the lysine at position 103 of the P protein being the N-terminal residue of this fragment. In this experiment two unlabelled fragments also appeared, $C_5$ and $C_6$, with apparent $M_r$ values of 5K and 2-5K. The N terminus of $C_5$ was blocked.

The specific activity of the whole P protein and its fragments was calculated from densitometric analyses of stained proteins on Immobilon paper and their corresponding autoradiographs. The specific activities of protein fragments were always higher than those of the whole P protein (data not shown). These results indicated that the labelled fragments contain most of the $^{32}P$ radioactivity present in the P protein. In addition fragments generated after different treatments were separated by 20% PAGE and the run was stopped before the dye front ran out. In these conditions we found that a peptide with 15 residues and with a calculated $M_r$ of 2K was included in the gel (see also the electrophoretic mobility of glucagon in Fig. 2). Thus, it seems that these analyses did not miss a high proportion of phosphorylated serine residues.

A summary of the results obtained is presented in Fig. 3. The main conclusion to be drawn from these data is that the major phosphorylation sites of the P protein are located, most probably, in a central region of the P protein, between residues 103 and 165. Thus, some or all serine residues at positions 116, 117, 119, 143, 156 and 161 are phosphorylated. Phosphorylation of those at positions 86, 94 and 99 seems to be unlikely but cannot be completely ruled out.

This conclusion is supported by the following results. (i) All strongly labelled fragments ($F_1$, $F_2$, $F_6$, $T_1$, $T_2$ and $C_4$) obtained after different treatments contained residues 103 to 165. All of these fragments had a higher specific activity than the whole P protein. (ii) All unlabelled fragments ($C_3$, $C_5$, $F_8$, $F_4$ and $F_7$) obtained after different treatments are located outside of residues 103 to 165. (iii) Unlabelled fragments $C_3$, $C_5$ and $F_8$, located at the N terminus of the phosphorylated central region, should contain serine residues at positions 23 and 30 (present in $C_5$), 39, 42, 45, 54, 58 and 60 (present in $F_8$ and $C_3$) and 86, 94 and 99 (probably present in $C_3$). Moreover, serine at position 86 should be contained in a formic acid digestion fragment with 26 residues, considering that no peptide bonds other than D–P or N–P were cleaved by formic acid treatment in this part of the P protein molecule. However, no such fragment was found. Serines at positions 94 and 99 should be included in a tryptic peptide with 12 residues and with a calculated $M_r$ of 1-5K that should be included in the gel shown in Fig. 2. These results suggest that those serines probably are not modified. Nevertheless these arguments do not absolute-

We are indebted to Dr I. Correas for her advice on previous experiments and Dr J. Avila for many useful discussions. We thank Dr R. Najera for his support, Dr B. Garcia-Barreno for providing the monoclonal antibody RS/1P, Dr E. Mendez and Dr J. A. Melero for constructive comments, Dr I. Outschoorn for English corrections, A.
del Pozo for the photographs, J. Colino and A. de la Fuente for their help in the typewriting, and E. Cubero, I. Garcia-Albert and R. Martinez for excellent technical assistance. J. Navarro is the recipient of a Fellowship from Comunidad Autónoma de Madrid.

References


(Received 20 June 1990; Accepted 13 February 1991)