Phosphoprotein of Rinderpest Virus Forms a Tetramer Through Coiled Coil Region Important for Biological Function: A Structural Insight.

Abdur Rahaman¹, N. Srinivasan², N. Shamala³ and M. S. Shaila¹¶

¹Department of Microbiology and Cell Biology, ²Molecular Biophysics Unit and ³Department of Physics, Indian Institute of Science, Bangalore 560012, India.

¶Corresponding author

Address for correspondence:

Prof. M. S. Shaila
Department of Microbiology and Cell Biology
Indian Institute of Science, Bangalore 560012
Phone: 91-80-23600139/22932702
Fax: 91-80-23602697
Email: shaila@mcbl.iisc.ernet.in

Running title: Tetramerization of RPV P is important for its function
SUMMARY

Phosphoprotein (P) of negative sense RNA viruses functions as a transcription transactivator of the viral polymerase (L). We report here the characterization of oligomeric P protein of rinderpest virus (RPV) and provide a structural basis for its multimerization. By size exclusion chromatography and dynamic light scattering analyses we show that bacterially expressed P protein exists as an oligomer, thus excluding the role of phosphorylation in P protein oligomerization. Gel filtration analyses of various parts of P protein, also expressed in E. coli, revealed that the predicted coiled coil region in the C-terminus domain is responsible for P protein oligomerization. Dynamic light scattering analysis confirmed the oligomeric nature of the coiled coil region of P. Chemical crosslinking analysis suggested that the C terminal coiled coil region exists as a tetramer. The tetramer is formed by coiled coil interaction as shown by circular dichroism spectral analysis. Based on sequence homology, we propose a three dimensional structure of the multimerization domain of RPV P using the crystal structure for multimerization domain of sendai virus P as a template. Four-stranded coiled coil structure of the model is stabilized by a series of interactions predominantly between short non-polar side chains emerging from different strands. In an in vivo replication/transcription
system using a synthetic minigenome of RPV, we show that multimerization is essential for P protein function(s) and the multimerization domain is highly conserved between two morbilliviruses namely RPV and *peste des petits ruminants* virus. These results are discussed in the context of biological functions of P protein among various negative stranded RNA viruses.

*Key words*: Paramyxoviridae, Rinderpest virus, Phosphoprotein, coiled coil, tetramer.

**INTRODUCTION**

Rinderpest virus (RPV) which causes Rinderpest disease in large and small ruminants is an enveloped virus belonging to the morbillivirus genus of the family Paramyxoviridae. The negative sense, single stranded RNA genome codes for six structural proteins, namely nucleocapsid protein (N), phosphoprotein (P), matrix protein (M), fusion protein (F), hemagglutinin protein (H) and large protein (L). The viral N-RNA i.e. genomic RNA encapsidated with N protein is transcribed and replicated by the L (RNA dependent RNA polymerase) and P complex (1). The L protein is associated with N-RNA template through its interaction with P protein to form the transcribing ribonucleoprotein (RNP) complex. In addition to polymerization activity, L exhibits a number of other enzymatic activities
including methyl transferase, 5' cap synthesis of mRNA and poly (A)$^+$ polymerase (2). During transcription, the intergenic start / stop signals are recognized by polymerase complex resulting in the synthesis of monocistronic, capped and polyadenylated mRNAs. Once the intracellular concentration of viral proteins reaches a threshold level, genome replication begins. The intracellular concentration of unassembled N protein ($N_0$) is believed to regulate the switch from transcription to replication (3). During replication, the same polymerase complex ignores stop signals and generates full-length unmodified encapsidated antigenomic RNA to serve as the template for the synthesis of progeny viral genomes.

P proteins of negative-stranded RNA viruses play multiple roles during viral infection. They act as a transcription transactivator and recruit L protein onto viral N-RNA template (1,4). P proteins also bind to N-RNA template independent of being part of L-P polymerase complex and activate transcription (5). In addition to their binding with assembled nucleocapsid structure of N-RNA template, P proteins interact with unassembled N proteins and prevent non-specific aggregation of the latter by forming $N_0$-P complex, a precursor for encapsidating newly synthesized RNA during replication (6). P proteins of mononegavirales undergo phosphorylation in one or more serine residues, which has been shown to be important for its
function (7,8). Although P proteins function as a homooligomer, their oligomerization status as well as the requirement of phosphorylation for oligomerization has been shown to vary among them (1,9-17). P proteins of all the paramyxoviruses harbor a coiled coil region at C-terminus domain and this region has been shown to be important for oligomerization in a number of viruses in the Paramyxoviridae family (13,18-21). The P protein has a modular structure which comprises two major domains: the N-terminus domain is highly variable among various Paramyxoviruses whereas the C-terminus domain though exhibiting low sequence similarity is conserved in terms of secondary structure (13). The C-terminus has been shown to have two sub-domains in Sendai virus P protein: PMD corresponding to the N-terminus region of C-terminus domain which harbors multimerization domain along with L binding domain and Px corresponding to rest of the C-terminus domain involved in nucleocapsid (NC) binding (13). Earlier work on RPV P protein has shown that while the first 59 aa residues at the N-terminus along with the predicted coiled coil region in the C-terminus half are important for interaction with unassembled N protein, the last 17 aa residues along with the predicted coiled coil region are required for interaction with nucleocapsid structure (18). This study also indicated the importance of coiled coil region in P protein self-interaction.
As a first step towards understanding the structure-function relationship of RPV P protein, we have looked at the oligomerization status of RPV P protein and examined the importance of coiled coil region of P protein in oligomerization as well as its function using different biochemical and biophysical approaches. Further, we propose a three-dimensional structure for the multimerization domain of RPV P based on its sequence similarity with that of Sendai virus (SeV) whose crystal structure has recently been solved (21). The importance of phosphorylation and oligomerization of P protein in the transcription and replication processes of negative sense RNA viruses has also been examined.

EXPERIMENTAL PROCEDURES

Materials- E. coli DH5α strain was used for maintenance of plasmids whereas BL21 (DE3) strain for expression of recombinant proteins (Invitrogen, USA). p3e and p4a harboring 1-291 aa (PNT) and 292-508 aa (PCT) regions of P protein respectively were earlier cloned in the laboratory in pRSET vector while the full length P (508 aa) gene isolated from a cDNA library of RPV (RBOK strain) was cloned in the expression vector pRSETB and named as pRP6 (22). Plasmids pKSN-1 (RPV N gene in pBS), pPol10 (RPV L gene in pGEM) and pMDB8A (a RPV minigenome plasmid carrying the 3’ regulatory sequence i.e. leader region, transcription/replication start
regions & 5' trailer sequences flanking CAT reporter gene ORF driven by T7 promoter and T7 terminator and delta ribozyme) were kindly provided by Dr. M.D. Baron, Institute of Animal Health, Pirbright, UK (23). A plasmid harboring the P gene of Peste de Petits Ruminants virus (PPRV P in pGEM vector) was a gift from Dr. T. Barrett, Institute for Animal Health, Pirbright, UK. A549 cells derived from Human Lung Carcinoma cell line were from ATCC (USA). These cells were maintained on HAM12 containing 10% NBCS (GIBCO BRL, USA). VTF7-3: Recombinant vaccinia virus expressing T7 polymerase in mammalian cells was a kind gift from Dr. Bernard Moss, NIH, USA.

Cloning of coiled coil region (RPC), extreme C terminus region (Px) of RPV P and multimerization domain (PPMD) of PPRV P- The nucleotide sequence corresponding to RPC (316-382 aa), a part of RPV P multimerization domain (PMD, 266 -388 aa) was released from pRP6 plasmid DNA by digestion with EcoRV and BamHI and the end-filled insert was sub-cloned into NcoI and XhoI digested pET33b (+) vector after end-filling. The expressed protein from this clone gives nine additional amino acids, one (M) at N-terminus and eight (L, E and H₆) at C-terminus.

The Px (377-508 aa) was cloned by removing NheI and SmaI fragment of pRP6 followed by religation of the backbone. The expressed
protein results in 14 additional amino acids (MRGSH$_6$GMAR) at its N-terminus.

The nucleotide sequence corresponding to PPRV PMD region- PPMD (264-387 aa) was amplified using pTB-P DNA as template and appropriate primers [Forward: 5’ (790) CGA AAT GCG TCT GTG G (805) 3’, Reverse: 5’ TTA (1161) CTC AGA TGT TGG GTC (1147) 3’]. Nucleotide positions of the primers on PPRV P gene are indicated within parentheses. The PCR product was cloned in EcoRV site of pET20b (+) vector. The insert from recombinant was released using NcoI and XhoI and sub-cloned into similar restriction sites of pET33b (+). The expressed protein codes for two additional amino acids (MD) at N-terminus. A stop codon was incorporated in the reverse primer to eliminate any additional amino acid at C-terminus.

**Expression and purification of recombinant proteins-** E.coli BL21 (DE3) strain was transformed with plasmids carrying full-length as well as different parts of RPV P. The transformant was grown in LB containing 100 µg/ml ampicillin (except for RPC) or 50 µg/ml kanamycin (for RPC) and induced with 0.4 mM Isopropyl thio-galactopyranoside (IPTG) at O.D$_{600}$ of 0.6 and grown for another 5h. The cells were harvested and lysed by sonication in MCAC buffer (500 mM NaCl in 20 mM Tris-HCl pH 8) supplemented with 2 mM PMSF and protease inhibitor cocktail. The lysates were centrifuged
and supernatant was mixed with Ni-NTA agarose. The resin was washed with 100 bed volumes of MCAC buffer containing 50 mM imidazole except for RPC in which imidazole was not used. Proteins were eluted with 500 mM imidazole in MCAC supplemented with protease inhibitor cocktail. The RPC so obtained was dialysed against 50mM Tris-Cl, pH8.0 and further purified by passing through a 5 ml Q-sepharose column using 0 -300 mM NaCl in dialysis buffer as gradient. Eluted samples of purified proteins were detected by coomassie-staining of SDS polyacrylamide gels. The protein concentration was measured by taking absorbance at 280 nm except RPC the concentration of which was estimated by Bradford assay. The identity of the proteins was confirmed by western blot analysis using polyclonal antibodies raised in rabbit against bacterially expressed RPV P or PPRV P protein.

Size exclusion chromatography (SEC)- Either sephadex G75 column (45 cm x 2.22 cm², bed volume of 100ml) or sephacryl S300 column (60 cm x 2 cm², 120ml bed volume) was equilibrated with PBS or MCAC buffer respectively and calibrated using standard protein molecular weight markers. One milligram each of P, PNT and PCT in MCAC or RPC and Px in 1 ml PBS were separated on sephacryl S300 or sephadex G75 respectively and the elution profiles were monitored by measuring absorbance at 280 nm except for RPC, which was monitored by protein estimation of the fractions.
by Bradford assay. The proteins were then identified by using SDS polyacrylamide gels and silver staining.

*Dynamic Light Scattering (DLS) Analysis*- About 1 mg/ml of RPC (in 50 mM Tris-Cl, pH8.0 and 50 mM NaCl) or 0.5 mg/ml of P (in 20 mM Tris-Cl, pH8.0 and 500 mM NaCl) was subjected to DLS analysis using DynaPro machine (Proteinsolutions, USA). About 50 observations were made to calculate the hydrodynamic radius ($R_h$) using DynaPro software. The viscosity used for $R_h$ calculation was estimated from the refractive index of the buffer as measured by refractometer.

*Chemical crosslinking*- About 20 µg of purified RPC protein was cross-linked using glutaraldehyde (final concentration of 0.5 mM and 1 mM) for different time intervals from 30 minutes to 4h at 25°C. The reaction was stopped by adding 200 mM glycine and the products were electrophoresed on a 15% SDS polyacrylamide gel and detected by silver staining.

*Circular dichroism (CD) spectroscopy*- Purified RPC at 0.1 mg / ml in phosphate-buffered saline (PBS) or in 50% trifluoro ethanol (TFE) in PBS was analyzed in a spectropolarimeter (JASCO J-715) at room temperature. The CD spectrum was measured in a cuvette of 2-mm path length, with a bandwidth of 0.5 nm and a scan speed of 50 nm/s. The buffer spectrum was subtracted from the protein spectrum. An average of four independent
measurements were used to calculate molar residue ellipticity, $[\theta]_{MRW}$ using the formula, $[\theta]_{MRW} = (\theta \cdot 100 \cdot M_r) / (c \cdot l \cdot N_A)$ where, $[\theta]$ is mean-residue-molar-ellipticity in deg cm$^2$ dmol$^{-1}$; $\theta$ is experimental ellipticity in milli degree; $M_r$ is the molecular weight of protein in Dalton; $c$ is protein concentration in mg/ml; $l$ is cuvette pathlength in cm; $N_A$ is the number of residues of the protein. The percentage helicity was estimated as follows (24,25):

$$\% \text{ helicity} = \frac{([\theta]_{222} - [\theta]_{222}^0)}{(100[\theta]_{222} - [\theta]_{222}^0) \times 100}$$

where $[\theta]_{222}$ is the experimentally observed absolute mean residue ellipticity at 222 nm, and values for $100[\theta]_{222}$ and $0[\theta]_{222}$, corresponding to 100% and 0% helix content at 222 nm, estimated at 32,000 and 2000 deg. cm$^2$ / dmol, respectively (25,26).

**In vivo Replication/Transcription Assay** - To assess the significance of tetramerization on biological function of P protein, an *In vivo* replication-transcription assay using minigenome construct pMDB8A was performed as described earlier (23). The transcript from the minigenome is antigenomic sense, which gets replicated to genomic sense RNA by the virus proteins, L, N and P expressed by cotransfected plasmids in A549 cells infected with recombinant vaccinia virus expressing T7 RNA polymerase. The newly made, genomic RNA is then transcribed into CAT mRNA & the translated protein CAT is measured by ELISA.
A549 cells (1X10^6 cells / 35 mm dish) were infected with recombinant vaccinia virus, VTF7-3 at a multiplicity of infection (MOI) of 10 at 37°C. At 1 hour post-infection, the cells were washed with PBS and transfected using 5 μl of lipofectamine (2 mg/ml) in 1 ml of OPTI-MEM medium (Gibco-BRL, USA) containing 1 μg each of pMDB8A, pKS-N, pRP6 and 100 ng of pGEM-L with or without pRPC / pPPMD. At 48 hour post-transfection, the cells were harvested and CAT (Chloramphenicol acetyltransferase) was assayed by enzyme-linked immunosorbent assay (ELISA; Roche Molecular Biochemicals).

Co-expression of full length RPV P with PPRV P multimerization domain-
The plasmid DNA of pRP6 and pPPMD clones were co-transformed into BL21 (DE3) strains of E.coli and the recombinant cells harboring both the plasmids were selected using two antibiotics i.e. ampicillin (100 μg/ml) and kanamycin (50 μg/ml). Transformed cells were grown in Luria broth supplemented with 100μg/ml of ampicillin and 50μg /ml of kanamycin to OD₆₀₀ 0.6 at 37°C. Expression and the purification of the protein by Ni-NTA agarose affinity chromatography were done as described in the earlier section. The purity of both the purified proteins was tested by electrophoresis on 15% SDS polyacrylamide gel followed by coomassie blue staining and confirmed by western blot analysis using appropriate antibody.
Prediction of secondary structures and coiled Coil Regions- The sequence of the multimerization domain of RPV P-protein (PMD) was subjected to secondary structure prediction using PHD as well as prediction of coiled coil regions if any (27-31). These predictions are employed in order to get views about the potential of this region to adopt α-helical structure as well as to form coiled coil independent of the fact that a distant homologue (Sendai virus phosphoprotein) has these structural features.

Comparative modelling of the coiled coil region of the p protein- The amino acid sequence of RPV PMD protein (266 -388 aa) is aligned with that of Sendai virus phosphoprotein whose crystal structure shows a homotetrameric α-helical coiled coil structure (21). The two proteins are distantly related and the alignment is non-trivial. Hence the structural features (such as solvent accessibility and secondary structure) at every residue were evaluated and relationships such as hydrogen bonding patterns in the crystal structure were assessed. While aligning the sequences the probability of a residue in the RPV PMD protein adopting the structural environment of equivalent residues in the known structure was considered. The match between predicted secondary structures in the P-protein and the observed secondary structures in the crystal structure in arriving at the alignment was also given importance.
The suite of programs encoded in *COMPOSER* and incorporated in *SYBYL* (Tripos Inc., St. Louis) was used to generate a three-dimensional model of the P-protein (32).

The *COMPOSER* generated model was energy minimized in *SYBYL* using the *AMBER* force-field (33). Energy minimized model of a subunit of RPV PMD was superimposed with each one of the four subunits of SeV PMD and the preliminary model for RPV PMD tetramer so obtained was subjected to further energy minimization to optimize the inter-protomer interactions.

**RESULTS**

*Recombinant Proteins*- All the recombinant proteins such as full-length P (1-508 aa), PNT (1-291 aa), PCT (292-508 aa), RPC (316-382 aa) and Px (376-508 aa) were expressed and purified to near homogeneity (Figure 1). The authenticity of the purified proteins was confirmed by western blot analysis using polyclonal antibody made against purified P protein expressed in *E. coli* (data not shown). As shown in Figure 1, the full-length P and PNT migrate at positions corresponding to 80 kDa and 52 kDa respectively which are much higher than their calculated masses (62 kDa and 39 kDa respectively). This anomalous mobility is attributed to the cluster of acidic
residues at the N-terminus domain (18). Mass spectroscopic analysis of full-length P protein further confirmed its authenticity (data not shown).

*P protein exists as a homo-oligomer in solution*- The oligomerization state of bacterially expressed P protein was studied by size exclusion column chromatography. As shown in Figure 2, majority of the P protein elutes at a position that corresponds to a molecular mass more than 300 kDa. This result indicated that P forms a higher order multimer as the monomeric molecular weight is 62 kDa. The hydrodynamic radius of P protein was measured by dynamic light scattering. The $R_h$ of 7 nm for P protein confirms the formation of oligomer in solution. The chemical crosslinking experiment also suggested that P protein exists as a multimer (data not shown).

*Coiled coil region on the C terminal domain is responsible for oligomerization of P protein into a tetramer*- Recombinant proteins corresponding to various parts of P protein were subjected to size exclusion chromatography. Elution profiles are shown in Figure 3. PNT (ca. Mass 39 kDa) elutes at around 100 kDa indicating that it is either an oligomer or is a partially structured monomer. Earlier work had revealed that the C-terminus is involved in P protein self-interaction (18). In measles virus P protein also, the equivalent domain (PNT) has been shown to be a partially structured monomer (34). Further, PCT (ca. Mass 28 kDa) eluted from the gel filtration
column at a position of molecular mass 150 kDa. Taken together, these results suggest that the oligomerization domain lies at the C terminus of P (PCT).

The coiled coil region (RPC) and rest of the C terminus domain (Px) show molecular mass of 35 kDa and 23 kDa respectively in size exclusion chromatography. This suggests that RPC (ca. Mass 8.5 kDa) is an oligomer, possibly a tetramer. The molecular size of Px (ca. Mass 16 kDa) is too small to be a dimer and the increased molecular size of the monomer might result from its elongated shape or partially structured nature. The oligomeric state of RPC was further tested by DLS. The result showed $R_h$ of 2.8 nm (corresponding to approx. 34 kDa) again confirming the oligomeric nature of RPC. To find out the exact stoichiometry of the RPC, chemical crosslinking of RPC was carried out. As shown in Figure 4, in addition to monomers, crosslinked RPC was detected as dimers, trimers and tetramers. With increase in duration of reaction and increase in crosslinker concentration, an increase in the amount of tetramers was observed. Since crosslinked products higher than tetramer were not observed, we conclude that the most common form of RPC is a tetramer. The nature of interaction of such a tetramer was studied by CD spectra analysis in presence and in absence of tri-fluoro ethanol (Figure 5). The result indicated that RPC is rich in $\alpha$-
helical content (approx. 90%) and the ratio of ellipticities at 222 nm/208 nm is greater than 1.0 indicative of presence of interacting helices. Moreover, the ratio of ellipticities at 222 nm/208 nm in 50% TFE decreased to 0.918, a characteristic of non-interacting \(\alpha\)-helices. Since TFE has been shown to disrupt tertiary structure and quaternary structure and to promote secondary structure (35), this result suggests that RPC forms a coiled coil structure. Taken together, these results lead us to conclude that RPV P protein forms a tetramer through coiled coil interaction present in RPC.

\textit{P protein functions as a multimer and the multimerization domain is conserved between two morbilliviruses}- The biological function of the multimerization domain of P protein was assessed employing an \textit{in vivo} replication/transcription system for RPV. As shown in Figure 6a, the CAT protein level is significantly reduced when RPC is coexpressed with wild type P protein as compared to the control where the full length P plasmid alone was used. Earlier studies in our laboratory have revealed that the coiled coil region does not interact with N protein (18) or L protein (36). These results clearly suggest that RPC forms a hetero-oligomer with wild type P protein and thereby inhibits the function of the latter as RPC lacks other functional domains required. A similar result was observed when PPRV P multimerization domain (264-386 aa) was used in the minigenome.
replication/transcription system in place of RPC (Figure 6b). These results indicate that multimerization of P protein is essential for its function and the multimerization domain is highly conserved between these two morbilliviruses. To further confirm this conservation of multimerization domain between the two viruses, wild type RPV P protein with histidine tag at its N terminus was coexpressed with untagged PPRV P multimerization domain in *E. coli* and purified by nickel affinity chromatography. As shown in Figure 7, untagged PPRV P multimerization domain is specifically co-purified with wild type RPV P. Untagged PPRV P multimerization domain expressed alone could not be purified under similar condition (data not shown). These results suggest that multimerization domain of PPRV P can interact with RPV P signifying conserved structure of this domain between the two viruses.

*Prediction of secondary structures and coiled coil regions of RPV PMD-*

The secondary structure prediction analysis using sequences which are closely related to the P-protein showed that the multimerization domain is predominantly helical. In particular, the region from positions 316 to 355 is strongly predicted to be $\alpha$-helical with a break of 4 residues around position 345. A few shorter segments of $\alpha$-helix were also predicted in the beginning
of the multimerization domain. Prediction of long α-helical region is consistent with the fact that it could form a coiled coil structure.

The amino acid sequence of the multimerization domain is also subjected to the prediction of coiled coil regions. The regions from about 315 to 365 and from 310 to 375 are predicted as coiled coil regions. The result is shown in Figure 8. It can be seen that the probability of a coiled coil structure for the region, 340-360 aa is about 0.7 if the window size used in the program is 14. However, the estimated probability of coiled coil formation for other window sizes is suggested to be low (of the order of 0.2-0.4) probably due to the fact that these procedures do not consider the possibility of a four-stranded coiled coil. When the residues of RPV PMD in the heptad repeat positions (a to g) were analysed, most of the nonpolar residues at a and d positions remained conserved. Although there were some drastic substitutions from nonpolar to polar residues at those positions, the interactions were maintained by compensatory changes thereby maintaining the coiled coil structure. Such nonpolar to polar residue substitutions are also seen in many other coiled coil structures (37).

Structural features of RPV PMD- The sequence identity between the multimerization domains of RPV P and SeV P is 14.3%. However if the similarity between the aligned residues in the two proteins are scored, the
percentage sequence similarity is 55.8%. This represents a low level of similarity, but appears convincing based on the compatibility of various residues in RPV P protein with the structural features in SeV P as well as potential evolutionary relationship. Based on the comparative sequence analysis it is very likely that the overall fold of the RPV PMD is similar to that of the crystal structure of SeV P. However, as the sequence identity (14.3%) is low, the details of the structures of these two proteins are likely to differ. The potential roles of various residues in P-protein in rendering stability to the coiled coil structure has also been investigated by generating a low resolution model using the crystal structure of SeV P as the basis.

The crystal structure of SeV PMD shows tetrameric elongated structure. The N-terminal globular structure (approximately first 50 residues from each subunit) is followed by a parallel long coiled coil structure (21). The interaction between the protomers is present both in the small globular region as well as in the coiled coil region.

The three-dimensional model for the amino acid sequence of RPV PMD protein generated using COMPOSER suite of programs based on the alignment with SeV P is shown in Figure 9.

A detailed examination of the model in the interactive graphics suggests that the sequence of the multimerization domain of the RPV P
protein could be comfortably accommodated in the fold of SeV P multimerization domain despite a low sequence similarity between these proteins. Most of the apolar residues are buried in the structure and polar residues are generally exposed. Many of the apparent drastic substitutions in the multimerization domain of RPV P protein compared to SeV PMD are accommodated by stabilizing interactions which are different in nature compared to the observations in the crystal structure of SeV PMD.

Figure 10 shows the amino acid sequences of SeV PMD and the multimerization domain of the RPV P protein with the structural environment at every residue position of SeV PMD shown for both the tetrameric state (cryt) and for a monomeric state (crym). The comparison of these two states show that the four-stranded coiled coil structural model of RPV PMD is largely stabilized by a series of interactions between the subunits involving short apolar side chains and a series of ionic pairs formed by oppositely charged amino acid side chains. This feature is also usually seen in the two and three-stranded coiled coil structures (37). Many of the side chains that are involved in the key interactions across the protomers of the tetramer model are shown in Figure 9. The residues of SeV PMD that are exposed in the monomeric form, but, get buried in the native tetramer form are given with the equivalent residue of the RPV PMD protein in the
brackets (Figure 10): Met 328 (Leu), Leu 332 (Ser), Val 333 (Thr), Ser 351 (Ser), Phe 354 (Ser), Ala 355 (Gln), Ala 358 (Ile), Leu 359 (Glu), Cys 372 (Ile), Gly 373 (Gln), Leu 374 (Asp), Leu 376 (Lys), Ser 377 (Thr), Val 386 (Gln), Leu 393 (Leu), Ile 396 (Leu), Val 400 (Lys), Phe 403 (Ile), Tyr 407 (Lys), Gln 414 (Asn) and Leu 425 (Ser). This shows that in addition to the coiled coil region, the residues in the globular domain participate in the inter-subunit interactions. Some of the residues involved in inter-subunit interactions are either conserved between the two proteins or substituted by another residue of similar chemical characteristics. However in a number of positions the equivalent residues from the two proteins are significantly different.

Val 386 of SeV PMD is replaced by Gln in the RPV PMD. In the modelled structure Gln residues from the adjacent strands are oriented in such a way that a hydrogen bonding network, connecting the side chains, is possible contributing towards the stability of the four-stranded coiled coil structure. Another residue, Val 400, in the coiled coil region of SeV P is replaced by the lysine residue in the RPV P protein. Interestingly, in the model, the side chain of this lysine residue is hydrogen bonded to the side chain of a glutamate residue from an adjacent strand which is equivalent to Ser 402 of SeV P. Thus a cyclic network of salt bridges is predicted to
stabilize the coiled coil structure. Similarly, the lysine residue of the RPV P protein that is equivalent to Tyr 407 of SeV PMD is hydrogen bonded, through its side chain, with the side chain of an Asp in an adjacent strand. This Asp replaces Arg 404 of the crystal structure and the net result of these two residue changes is yet another cyclic network of salt bridges linking adjacent strands. Thus, it is another example of apparently drastic amino acid substitutions (Tyr 407 → Lys and Arg 404 → Asp) resulting in the stabilizing interactions in the coiled coil structure. It is generally known that interactions between oppositely charged residues stabilize coiled coil structures (21). Leu 425 of the crystal structure occurs towards the end of the coiled coil structure and is replaced by Ser in the RPV P protein. In the model, the side chain -OH group of this Ser is hydrogen bonded to the main chain carbonyl in the vicinity.

DISCUSSION

One of the long-standing controversies is the exact oligomerization status of P protein of negative stranded RNA viruses. Studies on a number of viruses ranging from rhabdovirus to paramyxovirus report different oligomeric status of P protein including dimer, trimer and tetramer (9,11-13,17). By gel-filtration, DLS and chemical crosslinking analyses, we have shown that RPV P protein exists as a tetramer in solution and the tetramerization is through
the coiled coil region of the C terminus domain. Although earlier work on P proteins from SeV, NDV and MuV suggested that P could be a trimer, recent structural analyses using various biophysical methods and crystallographic study reveal that SeV P is a tetramer (13,21). Though VSV P protein is reported to be a trimer, it is also suggested to form a tetramer (9,11). It might be noted that the study with rabies virus (a rhabdovirus) P also does not exclude the possibility of tetramer formation (12). All these results lead us to conclude that P proteins from this group of RNA viruses can exist as a tetramer.

P proteins exhibit differences in terms of their requirement of phosphorylation for their multimerization. While P protein of VSV and chandipura virus (ChP) (rhabdovirus) require phosphorylation to facilitate its oligomerization, P proteins from paramyxoviruses such as SeV, NDV, MuV, MV has been shown to oligomerize independent of phosphorylation (7,9,17,19). Moreover unlike VSV P and ChV P, rabies virus (another rhabdovirus) P protein can form an oligomer in the absence of any phosphorylation (12). The RPV P protein used in the present work is unphosphorylated as it was expressed in bacteria and could form a stable tetramer. These results suggest that multimerization of P protein is independent of phosphorylation except in case of VSV and ChP. Notably the
study with VSV P also suggests the presence of a fraction of oligomeric population in absence of any phosphorylation and phosphorylation is only involved in the shift of equilibrium towards the multimer formation. Further at high concentration, the unphosphorylated VSV P exists predominately as a multimer (16). Considering these results, it can be generalized that P proteins of mononegaloviruses exist as a multimer, possibly as a tetramer whose oligomeric form is independent of any phosphorylation.

P proteins of mononegaloviruses also vary with respect to the role of phosphorylation on their biological functions. In SeV, phosphorylation of P has been shown to be dispensable for transcription and replication functions and similar possibility is suggested for human parainfluenza virus 1 (hPIV1) also (1). Although phosphorylation of P has been reported to be essential for its activity in VSV and RSV, later on it has been shown that phosphorylation of HRSV P protein is not essential for transcription and replication (40-42). Also VSV P when used in high concentration is able to bring about transcription activity without requiring any phosphorylation (43). This activity of VSV P could be due to the presence of some oligomers at high concentration (16). From these studies, it appears that oligomerization and not the phosphorylation of P protein of mononegalovirales is essential for its transcription and replication function. Using a minigenome in vivo
replication/transcription assay, we have demonstrated that disruption of RPV P oligomerization inhibits P protein function almost completely. Although a phosphorylation null mutant of RPV P has been shown to be inactive in an in vivo replication-transcription assay (22), this could be either due to an inactive conformation of the mutant (substitution of serine to alanine in three places) or a possible requirement of phosphorylation for an unidentified regulatory role of P protein. We hypothesize that phosphorylation, which is dispensable for transcription and replication, may play an important regulatory role particularly during the initial period of viral infection when the concentration of P protein is very low.

The functional inhibition and interaction studies with PPRV multimerization domain and RPV P protein strongly suggest that this domain is highly conserved between the two morbilliviruses. We extend that this functional conservation may be true for all the members of Paramyxoviridae family, based on in-depth sequence comparison between RPV P and SeV P whose crystal structure is known. Although the sequence identity between these two proteins is insignificant, the structural similarity is likely to be very high. Some of the interacting residues are drastically different (non-polar residues to polar residues substitution and vice-versa) between these two viruses but the four-stranded coiled coil structure is
maintained by compensatory changes elsewhere in the sequence. These suggest an evolutionary significance of this domain to perform a similar function among all the viruses of this family. Thus, the model further substantiates a common mechanism of P protein function among various negative sense RNA viruses. Although there is no evidence about the involvement of coiled coil structure in multimerization of rhabdovirus P proteins, there is a predicted coiled coil motif at its N terminus. This coiled coil motif might be responsible for multimerization involving possible tetramerization of P proteins among rhabdoviruses. Recently, with rabies virus P protein, it has been shown that the coiled coil motif at the N terminus domain is not involved in P protein multimerization and multimerization domain resides at C terminus domain which also harbors predicted coiled coil motif (12). However, in other rhabdoviruses like VSV, there is no predicted coiled coil motif at C terminus domain. This could mean that VSV P either oligomerizes through the N terminal coiled coil motif or through C terminus domain which might harbor a coiled coil motif not detectable by prediction programs. Although further experiments are required to understand the multimerization of VSV P, our results on RPV P along with the results from other investigators discussed above suggest that there is a
common mechanism for P protein function that acts as a multimer, possibly as a tetramer and multimerization is independent of phosphorylation.

Acknowledgement

The infrastructural facilities provided by the Department of Biotechnology, Government of India under the program support is acknowledged. A.R. was a senior research fellow of University Grant Commission, Government of India and N. Srinivasan is supported by the Senior Research Fellowship program for Biomedical Research by the Wellcome Trust, London.
REFERENCES


**Figure Legends**

**Figure 1:** Expression and purification of recombinant deletion proteins of P. Polyacrylamide gels showing clarified cell lysate (1), unbound fraction of Ni-column (2) and purified sample (3) of (a) pRP6, (b) PNT (1-291aa), (c) PCT (292-508aa), (d) RPC (316-382aa) and (e) Px (377-508aa). Molecular weights of the respective proteins were calculated using the standard low molecular weight markers obtained from Amersham Biosciences (UK) and are indicated by arrow.

**Figure 2:** Gel filtration analysis of pRP6. Nickel affinity purified protein was fractionated through sephacryl S300 column (bed vol. 120ml) and protein content in each fraction (1ml each) was monitored by taking absorbance at 280nm. Gelfiltration column was calibrated using the standard molecular weight markers for size exclusion column chromatography obtained from Amersham Pharmacia Biotech (USA) and the positions of different standard molecular weight markers are indicated on the top. The larger peak corresponds to >300kDa whereas small peak corresponds to ~ 68kDa.

**Figure 3:** Elution profile of (a) PCT, (b) PNT, (c) Px and (d) RPC in gel filtration column. Purified proteins at 1mg/ml were fractionated either on a sephacryl S300 column (bed vol. 120ml) or on a sephadex G75 column (bed
vol. 100ml) as mentioned earlier. Protein content in each fraction (1ml each) was monitored by $A_{280}$ except RPC, which was monitored by estimating protein content of the fractions by Bradford assay. PCT, PNT, Px and RPC eluted from the column at the positions corresponding to 150kDa, ~100kDa, 23kDa and 35kDa respectively. The columns were calibrated as mentioned earlier. The positions of different standard molecular weight markers are indicated on the top.

**Figure 4:** Chemical crosslinking of RPC. Purified protein was dialysed in PBS and 10µg protein was incubated with either 0.5mM glutaraldehyde (lane 2,3,4 & 5) or 1mM glutaraldehyde (lane 6,7,8 & 9) for 30 minutes (lanes 2,6), 1h (lanes 3,7), 2h (lanes 4,8) and 4h (lanes 5,9) at 25°C. M, molecular weight marker and lane 1, protein without glutaraldehyde. The cross-linked products were electrophoresed on a 15% SDS polyacrylamide gel and silver stained. The migration of monomer, dimer, trimer and tetramer are indicated on the right.

**Figure 5:** Coiled coil interaction of RPC. CD spectra of 20 µM RPC in PBS (open circle) and 20 µM RPC in presence of 50% TFE in PBS (filled circle). The ratio of ellipticities 222nm/208nm is greater than 1.0 indicating presence of helix-helix interaction i.e. coiled coil interaction. A ratio of less
than 1.0 (0.918) in presence of TFE suggests non-interacting helices due to
disruption of quaternary structure i.e. coiled coil structure of RPC.

**Figure 6:** *In vivo* replication/transcription assay. (a) coiled coil region of
RPV P (RPC) or (b) multimerization domain of PPRV P (PPMD) was co-
expressed with RPV P and the amount of CAT expressed was estimated by
ELISA as described in experimental procedures section. Lane1, A549 cells;
lane2, 1µg pRP6 without any inhibitory plasmid; lane3, 1µg pRP6 with 2µg
pET33b(+) plasmids; lane4, lane5 and lane6 1µg pRP6 with 0.5µg, 1µg and
2µg RPC or PPMD plasmids respectively.

**Figure 7:** Interaction of RPV P with the multimerization domain of PPRV
P. Histidine tagged RPV P was co-expressed with untagged PPRV P
multimerization domain in *E.coli* BL21(DE3) and was purified using Ni-
NTA agarose column. Approx. 50 nanograms pRP6 plasmid DNA and 100
nanograms pRPC plasmid were used for coexpression.

Lane1, co-purification of PPMD (without tag) with pRP6 (histidine tagged)
indicating specific interaction between them. Lane2, molecular weight
markers (kDa) as indicated by arrows.

**Figure 8:** The predicted Coiled coil region of multimerization domain of
RPV P. The probability for different window size is indicated by different
symbols as in the inset. Amino acid position 0-123 aa in the diagram
corresponds to 266-388 aa of P protein. The program by Lupas et al. (1991) has been used (29).

**Figure 9:** Three-dimensional model of multimerization domain of RPV P. The amino acid sequence of multimerization domain of RPV P was aligned with that of sendai virus P and 3D model was generated using comparative modelling. Four different strands of the tetramer are shown in four different colours in which side-chains of residues involved in coiled coil interaction are indicated by balls. The program SETOR has been used to produce this figure (38).

**Figure 10:** Sequence alignment of multimerization domain of RPV P (mod) with that of sendai virus P based on structural features. Structural features at the residue positions are shown for the protomeric state (crym) and in the tetrameric state (cryt) of the SeV PMD crystal structure. The program JOY has been used to produce this figure (39).

Upper case= solvent inaccessible, lower case= solvent accessible, *italic* = positive $\phi$, breve= cis-peptide, tilde= hydrogen bond to other side chain, **bold**= hydrogen bond to main chain amide, **underline** = hydrogen bond to main chain carbonyl, cedilla= disulphide bond
Figure 1
Figure 3 (a)
Figure 3 (b)
Figure 3 (c)
Figure 3 (d)
Figure 5
Figure 7
**Figure 10**

```
(52) NVCLGLAEGKSSARKVQDENKQILKQIGEQVESFRQDIYKRFSEYQKQNS
(52) MVCGGLIWSAEKLSSARKVQDENKQILKQIGEQVESFRQDIYKRFSEYQKQNS
SDIQDIKTALAKLHDDQQIITRLESLLSLKGEIDSIIKKQISKQNSISIST
```

```
(102) LIILMLILAHLHITD
(102) LIILMLILAHLHITD
I EGHLS S S F M A I P G F G K D P N D P T A D
```

The Journal of Biological Chemistry

[Image of the page]

www.jbc.org Downloaded from

[Image of the page]