The structure of the nucleoprotein binding domain of lyssavirus phosphoprotein reveals a structural relationship between the N-RNA binding domains of Rhabdoviridae and Paramyxoviridae.

Olivier Delmas, Rene Assenberg, Jonathan M Grimes, Hervé Bourhy

Point of view:

The Structure of the Nucleoprotein Binding Domain of lyssavirus Phosphoprotein reveals a structural relationship between the N-RNA binding domains of Rhabdoviridae and Paramyxoviridae

Delmas Olivier[1], Rene Assenberg[2,3], Jonathan M Grimes[2] and Hervé Bourhy[1*].

[1]Institut Pasteur, Unit Lyssavirus Dynamics and Host Adaptation, WHO Collaborating Centre for Reference and Research on Rabies, 75724 Paris CEDEX 15, France

[2]Division of Structural Biology and Oxford Protein Production Facility, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom

[3]Current address: Center for Proteomic Chemistry/Expertise Platform Proteases, Novartis Institutes for BioMedical Research, Fabrikstrasse 16, CH-4002 Basel, Switzerland.

* Corresponding author:

Hervé Bourhy. E-mail: hbourhy@pasteur.fr, Phone: + 33 1 45 68 87 85, Fax: +33 1 40 61 30 20

CONFLICT-OF-INTEREST AND FINANCIAL DISCLOSURE:
Authors declare no conflict of interest and financial disclosure.

Abstract

The phosphoprotein P of non-segmented negative-sense RNA viruses is an essential component of the replication and transcription complex and acts as a co-factor for the viral RNA-dependent RNA polymerase. P recruits the viral polymerase to the nucleoprotein-bound viral RNA (N-RNA) *via* an interaction between its C-terminal domain and the N-RNA complex. We have obtained the structure of the C-terminal domain of P of Mokola virus (MOKV), a lyssavirus that belongs to the *Rhabdoviridae* family and mapped at the amino acid level the crucial positions involved in interaction with N and in the formation of the viral replication complex. Comparison of the N-RNA binding domains of P solved to date suggests that the N-RNA binding domains are structurally conserved among paramyxoviruses and rhabdoviruses in spite of low sequence conservation. We also review the numerous other functions of this domain and more generally of the phosphoprotein.

Key words: Lyssavirus, Phosphoprotein, RNA-dependant RNA polymerase, nucleoprotein.

Introduction

Viruses belonging to the Mononegavirales order are enveloped viruses comprising four distinct families, the *Filoviridae*, *Bornaviridae*, *Rhabdoviridae* and *Paramyxoviridae*. The Mononegavirales infect numerous mammal hosts, and exhibit high genetic diversity and rapid evolutionary rates due to the lack of proofreading activity of the replication complex, therefore constituting a large variety of potential and actual human pathogens. Among them the *Paramyxoviridae* and *Rhabdoviridae* will be the main focus of this report as they include several highly pathogenic viruses like measles virus (a paramyxovirus) and rabies virus (a rhabdovirus), respectively. Measles virus kills more than 200,000 children in developing countries in which vaccination provides a significant decrease of paediatric mortality [1]. The measles virus has also been associated with severe respiratory presentations and with rare but extremely serious neurological complications. The agents of rabies, leading invariably to fatal viral encephalitis, all belong to the lyssavirus genus. They are responsible for an estimated 55000 deaths per year worldwide [2]. Seven different species are recognized by the ICTV based on genome sequence homology. New viruses have been recently described, although not yet classified, that could represent new lyssavirus genotypes [3-5].

The genome of Mononegavirales consists of a negative sense non-segmented RNA, which is encapsidated by the nucleoprotein N and encodes at least 5 viral proteins common to all viruses of this family, in the order from 3' to 5' : the nucleoprotein (N), the phosphoprotein (P), the matrix (M) and the large protein (L). The latter contains approximately 2000 amino acids and is the central protein in the replication and transcription of the viral genome, and carries RNA dependent RNA polymerase, RNA polyadenylation, RNA guanylyltransferase and RNA methyltransferase activities. After the first steps of attachment, penetration and uncoating of the virus particle, the nucleocapsid (the N-RNA complex) and all the components necessary for the early transcription are released in the cytoplasm of the infected

cell where the transcription and then the replication steps of the virus life cycle occur. During these processes the genome (and anti-genome) remains bound to the N polymer at all times and recently structures have been solved for N-RNA complexes from Rabies, VSV and respiratory syncytial virus (RSV) [6-8]. Importantly, during transcription the viral mRNAs are not encapsidated by N, probably due to the fact that leader and trailer sequences are absent from mRNA conversely to genomic and antigenomic RNA. Thus, the viral transcripts are, like their cellular counterparts, capped and poly-adenylated at their 5' and 3' ends respectively. One important unresolved question remains how the viral polymerase switches between replication and transcription, which both utilise the same genomic RNA strand. However, in the case of transcription the polymerase recognises specific RNA signals that demarcate the start and end points of the viral genes to produce a gradient of mRNA transcripts, whilst during replication these are ignored. The Mononegavirales have to deal with a limited number of proteins (5-10) to achieve all the necessary functions for their life cycle. As such the P exhibits a modular organization with each of its domains is involved in several aspects of the P "cellular functions" (Fig.1) The P protein is an essential cofactor of the viral RdRP although the precise role of P in facilitating polymerase activity remains to be established. It is thought that one function of P is to deliver, during replication, free N (referred to as "N-degree" or N^o) to the newly synthesised viral RNA. In addition, it is thought that P acts as a bridge between the L polymerase and the N-RNA template to gain access to the genomic RNA. Consistent with this idea is that P has two distinct N-binding sites, one specific for N^o (located at the N-terminus of P) and one for N-RNA (located in the C-terminal domain of P [PCDT]). The general domain organization of P is conserved across the Mononegavirales, although the details differ. In between the two N-binding domains, P contains a multimerisation domain, which in rhabdoviruses mediates P dimerisation [9] whilst in paramyxoviruses it mediates tetramerisation [10]. The multimerisation domain is flanked by

disordered regions of variable length, probably explaining the difficulty of obtaining full length P crystals, as well as its susceptibility to in vitro proteolysis (see below). Interestingly, in paramyxoviruses, and in rhabdoviruses, the P gene is edited during transcription by L, yielding additional proteins termed C and V in paramyxovirus and shorter P in lyssavirus termed P2, P3, P4 and P5 [11, 12]. These P products appear to function primarily to evade the host immune response.

In lyssavirus, the PCDT modulates the sub-cellular localization of P during infection. The PCTD have been shown to contain a nuclear localization signal (NLS) [13] and a nuclear export signal (NES) [14]. The NLS involves the amino acids that also contribute to the interaction with N. The NES involves amino acids positions 227 to 232. These two signals contribute to the cellular localization of P and phosphorylation by protein kinase C may be involved in the switch mechanism to balance the import and export of P to and from the nucleus [14]. The association of P with microtubules also regulates its nuclear trafficking. The CTD of P contains motifs that participate in P association with microtubules but these associations are dependent on oligomerization of P [15]. In the nucleus, the PCTD also interacts with interferon-induced promyelocytic leukaemia (PML) bodies [11]. The exact role of this interaction remains to be fully understood but may be involved in virus pathogenesis.

The PCTD also plays a role to limit the capacity of defense of the hosts. It has been shown that P prevents interferon regulatory factor 3 (IRF3) phosphorylation thus inhibiting the key factor involved in IFN alpha/beta response in infected cells [16]. P is also able to target and interact with STAT1 and STAT 2 [17-19]. This interaction in combination with microtubule association prevents STAT 1 and STAT 2 nuclear localization [15]. The region of P involved in this interaction is also located within the PCTD between positions 267 and 297 [18]

We have recently obtained the structure of the phosphoprotein C-terminal domain (PCTD) of one species of the lyssavirus genus, the Mokola virus (MOKV) and mapped at the amino acid

level the crucial positions involved in interaction with N involved in the viral replication complex formation [20]. In this paper we will focus on the structural aspects of the C-terminal domain in relation to virus replication.

Results

Structure determination of the C-terminal domain of Mokola virus P

Full length MOKV P was recombinantly expressed in *E. coli,* purified and subjected to crystallization experiments. The structure was solved by molecular replacement, using the structure of the RABV P N-RNA binding domain as a model [21] (PDB 1VYI), which revealed that a proteolytic fragment had been crystallized. Mass spectrometry analysis of the protein immediately before the crystallization trials showed the protein to be intact at that point confirming that truncation had occurred during crystallisation. The structure of MOKV P obtained was of the C-terminal domain of the protein, formed by six tightly folded α-helices, a 3_{10}-helix and a small β-sheet consisting of two β-strands (Fig. 2), with a morphology resembling a sliced pear, with a flat and curved face.

Analysis of the crystal structure of MOKV P revealed that one of the contacts points in crystal packing involves the C-terminal 4 residues (E300, E301, A302, E303) of one monomer packing against a predominantly positively charged region consisting of residues located on β1, β2, α4 and α6 (K212, K213, Y214, K215, I223, L225, R261 and Y295) in another monomer. K212, K213, K215, R261 and Y295 engage in both main- and side chain polar contacts with the acidic tail, with residues Y214, K215, I223, L225 and R261 contributing additional hydrophobic interactions.

This positively charged region on P has been implicated previously in the P-N interaction [22]. To confirm and map precisely the relative contribution of each of these amino acids, we tested by yeast two hybrid the interaction of N with single point mutants of the P C-terminal domain. Mutations of Y214, K215 and R261 resulted in the most significant growth defects of the yeast strain whilst for K212 and L225 higher stringency conditions were necessary to reveal an effect on N binding [20]. The degree of association and the reduction in accessible

surface area upon binding the acidic tail by each of the residues shows a good overall correspondence with the yeast two hybrid results.

One of the major roles of P is as part of the replication complex together with N and L [22]. The P is thought to bridge between N (associated with the viral genome) and the L polymerase [23]. This viral replication complex is active in the Negri bodies [24, 25]. Negri bodies are viral cytoplasmic inclusions that contain N, P and L and viral genome whose formation is dependent on Toll like receptor 3. The PCTD interacts with N alone or with N associated with RNA [26-29] and the amino acids K212, Y214, K215, L225 and 261 have been implicated in this interaction. Whether these N-P interactions are involved in Negri bodies formation remains to be explored.

PCTD motif and structure conservation among Lyssaviruses

Lyssaviruses taxonomy is defined on the basis of genetic similarity of the nucleoprotein sequence. Seven species of lyssavirus are recognized by the ICTV (International Committee on Taxonomy of Viruses). Recently, four new lyssaviruses have been described from bats originated from Eurasia (Aravan virus (ARAV), Khujand virus (KHUV), Irkut virus (IRKV)) or from Caucasian region (West Caucasian Bat Lyssavirus (WCBV)) [5]. A virus related to LBV has also been described [3, 4]. Each of these new viruses could represent new species.

The lyssavirus P protein, including those from newly described viruses show a percentage of sequence identity ranging from 42 % (minimum intergenotype identity) to 81% (minimum intragenotype identity). The PCTD is known to be the most conserved part of the protein [30].

The structure of MOKV PCDT is very similar to the PCDT solved previously of RABV P, (another species of lyssavirus) (PDB ID 1VYI) (Fig. 2B) where 101 amino acids can be aligned with a root mean square deviation (rmsd) of 0.66 Å, consistent with the high degree of sequence identity (68%). There are only minor changes in the length of helixes at the N and C

termini of the domain, with a shorter α1 helix but an extended α6 helix in MOKV compared to RABV.

All the amino acids positions which were implicated in P-N association are conserved in all lyssaviruses further reinforcing the probable role of this charged interaction surface of the P in the virus life cycle (Fig3).

A conserved organization of P proteins in the families of *Rhabdoviridae* and *Paramyxoviridae*

The structures of functionally equivalent domains of the P proteins from several other members of the order *Mononegavirales* have been solved [31-34]. The structures of the N-RNA binding domain of Sendai virus (SeV) and Measles virus (MeV) P protein consist of an anti-parallel triple helix bundle and, as for the *Rhabdoviridae*, the N-RNA binding domains among paramyxoviruses are also structurally conserved in spite of low sequence conservation.

Careful analysis of the structures reveals a structural relationship between the N-RNA binding domains of the P proteins of the *Rhabdoviridae* and the *Paramyxoviridae,* where the 3 helix bundle of SeV and MeV P aligns with helices α3, α4 and α6 of MOKV (and RABV) (Fig. 4). Structural superposition using SHP [35], show that 43 and 40 residues of MOKV P (out of 107) can be aligned with SeV P and MeV P, with rmsds of 3.3 Å and 3.2 Å respectively over the aligned residues, despite 14% and 0% sequence identity. The similarity between VSV P and the paramyxovirus P proteins is much poorer due to the shorter C-terminal helix in VSV P (α4 in VSV, α6 in MOKP) and the absence of a helix equivalent to MOKV α3 (Fig 4), although the alignment does suggest that VSV P α3 is conserved (α4 in MOKV P).

The structure based sequence alignment reveals that the highest degree of sequence similarity resides in the region spanning 7 residues in helix α6 in MOKV P (equivalent to helix α3 in

SeV and MeV P) (Fig. 4). Although the alignment indicates that MeV P shows a much poorer sequence conservation in this area, SeV also resembles lyssavirus P in the character of the P-N interactions which are primarily electrostatic (whereas in MeV P with N the interactions are primarily hydrophobic).

This alignment is in agreement with functional data where the structurally equivalent helix in the rhabdo- and paramyxoviruses (helix α3) to MOKV P α6 is involved in binding N. For MeV P a co-crystal structure was solved with the N-RNA binding domain of P bound to a MeV N segment (spanning one α-helix) as shown in Fig. 4. The latter is bound to the topologically equivalent site in MOKV P (the positive cluster) as shown in the superposition of MeV P bound to N with MOKV P in Fig. 4. Clearly however the precise details of the interactions between P and N differ. For instance, none of the MOKV P residues identified in this study as important for the interaction with MOKV N, appear to be conserved in the VSV P structure, suggesting that VSV P interacts differently with N than the lyssavirus P. In addition, biochemical data also indicate that P-N and P-L interactions are likely to be different [36-40]. This has been confirmed by the structure of VSV N-RNA complexed with the C-terminal domain of VSV P which reveals that those residues of VSV-P that are involved in binding lie close to but do not map directly onto the residues that we predict are involved in MOK N-P binding. This lack of agreement is probably due primarily to differences in the size of the P proteins (the CTD of MOK P is 50% larger than the equivalent CTD of VSV).

Thus, although clearly the interactions and mode of binding has diverged between the different members of the Mononegavirales, it seems likely that the P proteins of *Rhabdo- and Paramyxoviridae* originated from a common ancestor in spite of the high degree of amino acid divergence.

A recent analysis of structure and sequence conservation of P of two rhabdoviruses and one paramyxovirus combined the results from multiple prediction methods as well as two-hybrid

and biochemical approaches to locate the boundaries between disordered regions and structured domains and to characterize their potential functions [28]. The proposed modular organization predicted for RABV P confirmed both our and earlier findings that the putative structured domains correspond to autonomous folding units, and that CTD of P is central to the formations of N-P complexes..

Conclusion

The phosphoprotein of the rhabdoviruses has a modular multidomain architecture and is central to the formation of an active RNP. The C-terminal domain of lyssavirus P has been shown to be multifunctional, and structural and biochemical analysis has implicated binding surfaces involved in N-P complex formation. P is found both in the cytoplasm and nucleus of infected cells and this compartmentalization defines to some degree the function of P. Thus P in the cytoplasm either forms part of the active viral polymerase complex or sequesters free N, and it has been shown that P oligomerisation is required for polymerase function. This oligomerisation may be required for the correct presentation of binding surfaces of P to N or L. It is unclear what the oligomerisation state of P is in the nucleus, but its role in inhibiting the IFN response is dependent on its nuclear import and retention.

What role if any structural switching of P has on modulating these different functions remains unclear but further steps in understanding the regulation of these functions will be to obtain the structure of P in complex with its partners; as recently obtained for VSV PCTD in association with N-RNA rings [41]. What is clear from our structure and the analysis of those amino acids crucial to N-P association is that they are different from those observed in the VSV N-P complex structure. Whether this is due to differences in the mode of association or to structural changes to P when in complex with N-RNA remains an open question, but this lack of structural correlation is probably due partly to differences in the size of the P proteins.

Our structural analysis of the N-RNA binding domains of the P proteins from the *Rhabdo-* and *Paramyxoviridae* reveals a low but detectable degree of structural similarity as well as the conservation of the binding surfaces that suggest that these P proteins originated from a common ancestor in spite of the high degree of amino acid divergence.

Obtaining the structure of complexes of P with other proteins involved in the replication complex will help in our understanding of how so many functions could be retained and regulated by this viral protein. It will also allow us to understand the viral RNA replication / transcription complex and model how the viral polymerase accesses the viral RNA genome that is totally covered by the N protein. Understanding this process would help in the design of specific antiviral compounds that could target functional interaction surface on viral proteins and could help in preventing rabies and more widely underpin the treatment of viruses belonging to the Order Mononegavirales.

ACKNOWLEDGEMENTS

This work was supported by the EU-funded FP6 Vizier program (LSHG-CT-2004-511960).

References

1. **Wolfson LJ, Strebel PM, Gacic-Dobo M, Hoekstra EJ, McFarland JW, Hersh BS. Has the 2005 measles mortality reduction goal been achieved? A natural history modelling study. Lancet 2007; 369:191-200.**
2. **Knobel DL, Cleaveland S, Coleman PG, Fevre EM, Meltzer MI, Miranda ME, et al. Re-evaluating the burden of rabies in Africa and Asia. Bull World Health Organ 2005; 83:360-8.**
3. **Delmas O, Holmes EC, Talbi C, Larrous F, Dacheux L, Bouchier C, et al. Genomic diversity and evolution of the lyssaviruses. PLoS ONE 2008; 3:e2057.**
4. **Kuzmin IV, Niezgoda M, Franka R, Agwanda B, Markotter W, Beagley JC, et al. Lagos bat virus in Kenya. J Clin Microbiol 2008; 46:1451-61.**
5. **Kuzmin IV, Wu X, Tordo N, Rupprecht CE. Complete genomes of Aravan, Khujand, Irkut and West Caucasian bat viruses, with special attention to the polymerase gene and non-coding regions. Virus Res 2008; 136:81-90.**
6. **Albertini AA, Wernimont AK, Muziol T, Ravelli RB, Clapier CR, Schoehn G, et al. Crystal structure of the rabies virus nucleoprotein-RNA complex. Science 2006; 313:360-3.**
7. **Green TJ, Zhang X, Wertz GW, Luo M. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. Science 2006; 313:357-60.**
8. **Tawar RG, Duquerroy S, Vonrhein C, Varela PF, Damier-Piolle L, Castagne N, et al. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. Science 2009; 326:1279-83.**
9. **Ivanov I, Crepin T, Jamin M, Ruigrok RW. Structure of the dimerisation domain of the rabies virus phosphoprotein. J Virol.**
10. **Tarbouriech N, Curran J, Ruigrok RW, Burmeister WP. Tetrameric coiled coil domain of Sendai virus phosphoprotein. Nat Struct Biol 2000; 7:777-81.**
11. **Blondel D, Regad T, Poisson N, Pavie B, Harper F, Pandolfi PP, et al. Rabies virus P and small P products interact directly with PML and reorganize PML nuclear bodies. Oncogene 2002; 21:7957-70.**
12. **Devaux P, Hodge G, McChesney MB, Cattaneo R. Attenuation of V- or C-defective measles viruses: infection control by the inflammatory and interferon responses of rhesus monkeys. J Virol 2008; 82:5359-67.**
13. **Pasdeloup D, Poisson N, Raux H, Gaudin Y, Ruigrok RW, Blondel D. Nucleocytoplasmic shuttling of the rabies virus P protein requires a nuclear localization signal and a CRM1-dependent nuclear export signal. Virology 2005; 334:284-93.**
14. **Moseley GW, Filmer RP, DeJesus MA, Jans DA. Nucleocytoplasmic distribution of rabies virus P-protein is regulated by phosphorylation adjacent to C-terminal nuclear import and export signals. Biochemistry 2007; 46:12053-61.**
15. **Moseley GW, Lahaye X, Roth DM, Oksayan S, Filmer RP, Rowe CL, et al. Dual modes of rabies P-protein association with microtubules: a novel strategy to suppress the antiviral response. J Cell Sci 2009; 122:3652-62.**
16. **Brzozka K, Finke S, Conzelmann KK. Identification of the rabies virus alpha/beta interferon antagonist: phosphoprotein P interferes with phosphorylation of interferon regulatory factor 3. J Virol 2005; 79:7673-81.**

17. **Brzozka K, Finke S, Conzelmann KK. Inhibition of interferon signaling by rabies virus phosphoprotein P: activation-dependent binding of STAT1 and STAT2. J Virol 2006; 80:2675-83.**
18. **Vidy A, Chelbi-Alix M, Blondel D. Rabies virus P protein interacts with STAT1 and inhibits interferon signal transduction pathways. J Virol 2005; 79:14411-20.**
19. **Vidy A, El Bougrini J, Chelbi-Alix MK, Blondel D. The nucleocytoplasmic rabies virus P protein counteracts interferon signaling by inhibiting both nuclear accumulation and DNA binding of STAT1. J Virol 2007; 81:4255-63.**
20. **Assenberg R, Delmas O, Ren J, Vidalain PO, Verma A, Larrous F, et al. The Structure of the N-RNA Binding Domain of the Mokola virus Phosphoprotein. J Virol 2010; 84:1089-96.**
21. **Mavrakis M, McCarthy AA, Roche S, Blondel D, Ruigrok RW. Structure and function of the C-terminal domain of the polymerase cofactor of rabies virus. J Mol Biol 2004; 343:819-31.**
22. **Jacob Y, Real E, Tordo N. Functional interaction map of lyssavirus phosphoprotein: identification of the minimal transcription domains. J Virol 2001; 75:9613-22.**
23. **Albertini AA, Schoehn G, Weissenhorn W, Ruigrok RW. Structural aspects of rabies virus replication. Cell Mol Life Sci 2008; 65:282-94.**
24. **Lahaye X, Vidy A, Pomier C, Obiang L, Harper F, Gaudin Y, et al. Functional characterization of Negri bodies (NBs) in rabies virus-infected cells: Evidence that NBs are sites of viral transcription and replication. J Virol 2009; 83:7948-58.**
25. **Menager P, Roux P, Megret F, Bourgeois JP, Le Sourd AM, Danckaert A, et al. Toll-like receptor 3 (TLR3) plays a major role in the formation of rabies virus Negri Bodies. PLoS Pathog 2009; 5:e1000315.**
26. **Chenik M, Chebli K, Gaudin Y, Blondel D. In vivo interaction of rabies virus phosphoprotein (P) and nucleoprotein (N): existence of two N-binding sites on P protein. J Gen Virol 1994; 75 (Pt 11):2889-96.**
27. **Fu ZF, Zheng Y, Wunner WH, Koprowski H, Dietzschold B. Both the N- and the C-terminal domains of the nominal phosphoprotein of rabies virus are involved in binding to the nucleoprotein. Virology 1994; 200:590-7.**
28. **Gerard FC, Ribeiro Ede A, Jr., Leyrat C, Ivanov I, Blondel D, Longhi S, et al. Modular organization of rabies virus phosphoprotein. J Mol Biol 2009; 388:978-96.**
29. **Mavrakis M, Mehouas S, Real E, Iseni F, Blondel D, Tordo N, et al. Rabies virus chaperone: identification of the phosphoprotein peptide that keeps nucleoprotein soluble and free from non-specific RNA. Virology 2006; 349:422-9.**
30. **Nadin-Davis SA, Abdel-Malik M, Armstrong J, Wandeler AI. Lyssavirus P gene characterisation provides insights into the phylogeny of the genus and identifies structural similarities and diversity within the encoded phosphoprotein. Virology 2002; 298:286-305.**
31. **Bernard C, Gely S, Bourhis JM, Morelli X, Longhi S, Darbon H. Interaction between the C-terminal domains of N and P proteins of measles virus investigated by NMR. FEBS Lett 2009; 583:1084-9.**
32. **Houben K, Marion D, Tarbouriech N, Ruigrok RW, Blanchard L. Interaction of the C-terminal domains of sendai virus N and P proteins: comparison of polymerase-nucleocapsid interactions within the paramyxovirus family. J Virol 2007; 81:6807-16.**
33. **Johansson K, Bourhis JM, Campanacci V, Cambillau C, Canard B, Longhi S. Crystal structure of the measles virus phosphoprotein domain responsible for the**

induced folding of the C-terminal domain of the nucleoprotein. J Biol Chem 2003; 278:44567-73.

34. **Kingston RL, Hamel DJ, Gay LS, Dahlquist FW, Matthews BW. Structural basis for the attachment of a paramyxoviral polymerase to its template. Proc Natl Acad Sci U S A 2004; 101:8301-6.**
35. **Stuart DI, Levine M, Muirhead H, Stammers DK. Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 A. J Mol Biol 1979; 134:109-42.**
36. **Chenik M, Schnell M, Conzelmann KK, Blondel D. Mapping the interacting domains between the rabies virus polymerase and phosphoprotein. J Virol 1998; 72:1925-30.**
37. **Das T, Pattnaik AK, Takacs AM, Li T, Hwang LN, Banerjee AK. Basic amino acid residues at the carboxy-terminal eleven amino acid region of the phosphoprotein (P) are required for transcription but not for replication of vesicular stomatitis virus genome RNA. Virology 1997; 238:103-14.**
38. **Ribeiro EA, Jr., Favier A, Gerard FC, Leyrat C, Brutscher B, Blondel D, et al. Solution structure of the C-terminal nucleoprotein-RNA binding domain of the vesicular stomatitis virus phosphoprotein. J Mol Biol 2008; 382:525-38.**
39. **Takacs AM, Banerjee AK. Efficient interaction of the vesicular stomatitis virus P protein with the L protein or the N protein in cells expressing the recombinant proteins. Virology 1995; 208:821-6.**
40. **Takacs AM, Das T, Banerjee AK. Mapping of interacting domains between the nucleocapsid protein and the phosphoprotein of vesicular stomatitis virus by using a two-hybrid system. Proc Natl Acad Sci U S A 1993; 90:10375-9.**
41. **Green TJ, Luo M. Structure of the vesicular stomatitis virus nucleocapsid in complex with the nucleocapsid-binding domain of the small polymerase cofactor, P. Proc Natl Acad Sci U S A 2009.**

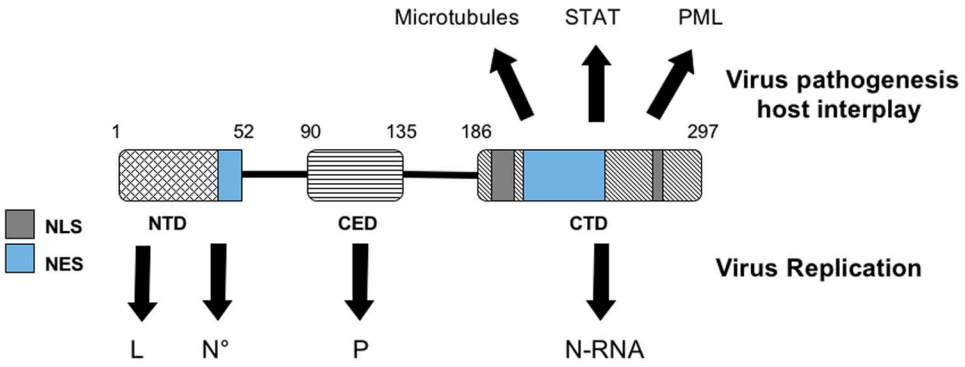
Figure Legend

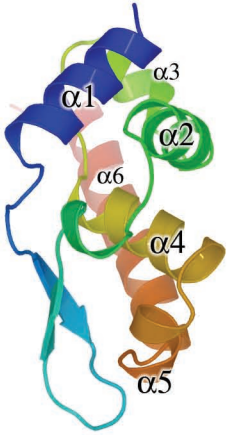
Fig1: Schematic representation of lyssavirus phosphoprotein and its functions in virus pathogenesis/host interplay and virus replication. Abbreviations: PML promyelocytic leukaemia proteins NLS, nuclear localization signal; NES, nuclear export signal, NTD aminoterminal domain, CED central domain, CTD carboxy terminal domain.

Fig2: Structure of MOKV and RABV PCTD. Side-by-side comparison of rhabdovirus P structures: (A) MOKV P; (B) RABV P (PDB accession no. 1VYI); (C) VSV P (PDB accession no. 2K47). The structures are rainbow ramped from blue (N terminus) to red (C terminus).

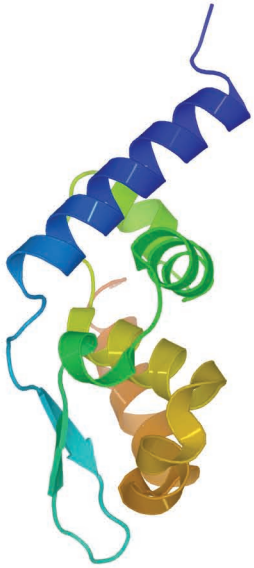
Fig3: PCTD alignment of all lyssavirus genotypes. Arrowheads show conserved aminoacids involved in nucleoprotein binding. Homology plot of each position is indicated in the bottom.

Fig4: Comparison of the structures of *Rhabdovirus* and *Paramyxovirus* N-RNA binding domains of P. (a)) Side-by-side comparison of the superposedconserved elements of MOKV P (orange), VSV P (yellow) (PDB accession no. 2K47), and SeV (magenta) (PDB accession no. 1R4G)., based on SHP alignment of the Cα's (40) (b) Side-by-side comparison of MOKV P (orange) and MeV P bound to MeV N (cyan, PDB Id 1T6O),(based on the superposition from SHP) highlighting the overlap between the N-binding site in the two P proteins.

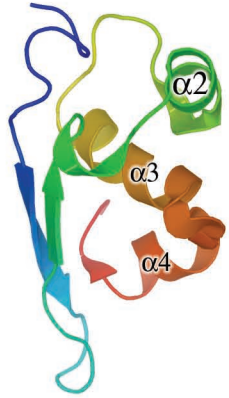

Microtubules
STAT
PML
Virus pathogenesis
host interplay
1
52
90
135
186
297
NLS
NES
NTD
CED
CTD
Virus Replication
L
N°
P
N-RNA



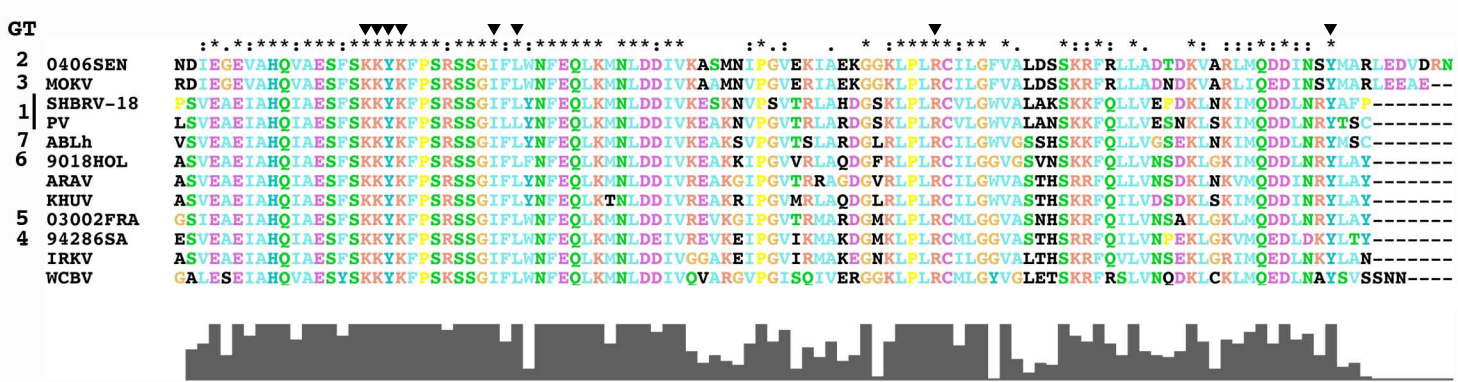
MOK

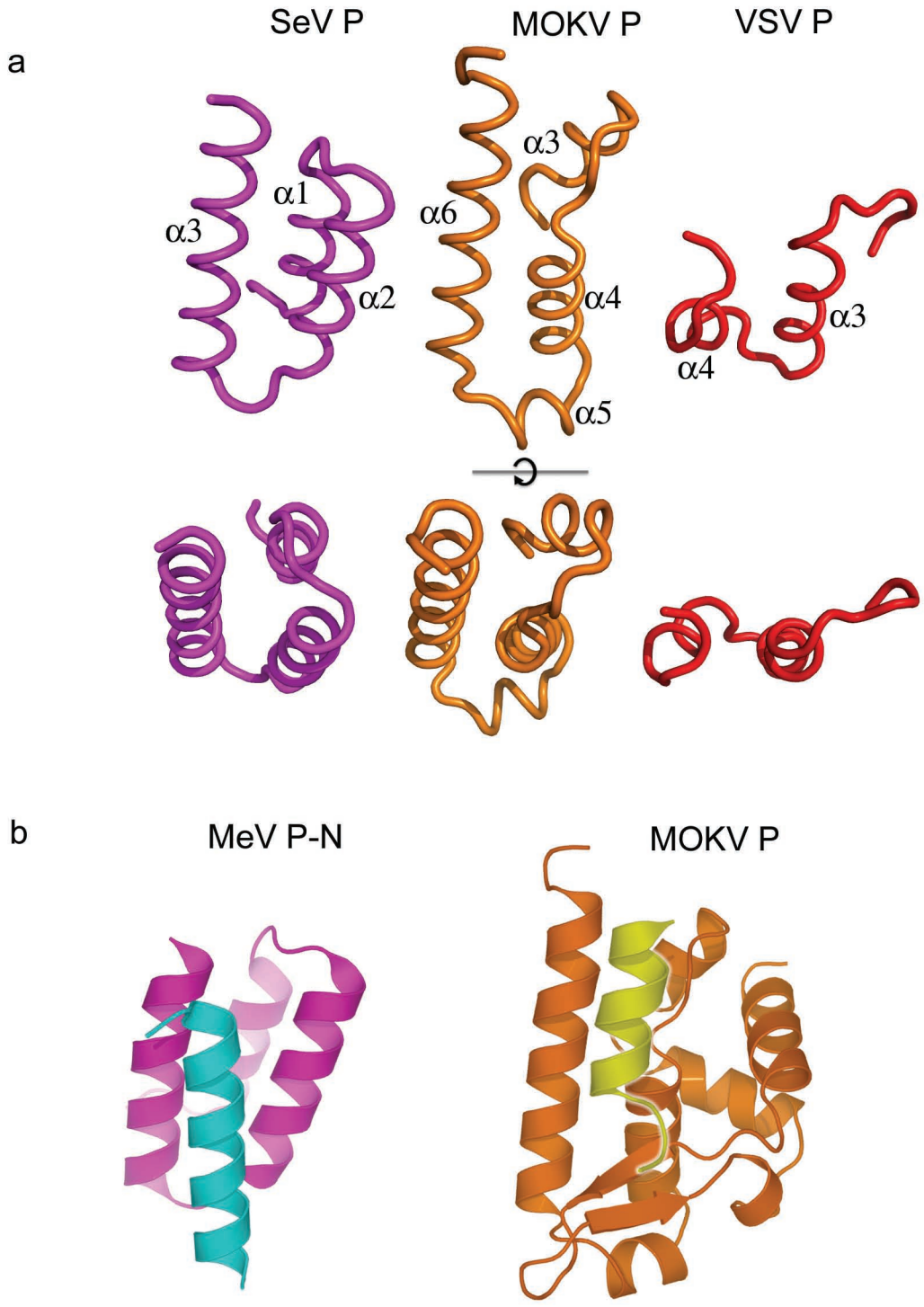


RAB



VSV


GT
2
3
1
7
6
5
4
0406SEN
MOKV
SHBRV-18
PV
ABLh
9018HOL
ARAV
KHUV
03002FRA
94286SA
IRKV
WCBV


a
SeV P
MOKV P
VSV P
α3
α1
α2
α6
α3
α4
α5
α4
α3
b
MeV P-N
MOKV P