A paramyxovirus-like model for Ebola virus bipartite promoters

Irina Gutsche, Philippe Le Mercier, Daniel Kolakofsky

HAL Id: hal-03009975
https://hal.uvsq.fr/hal-03009975

Submitted on 24 Nov 2020

PLOS PATHOGENS

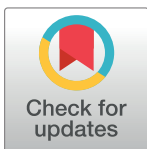# A paramyxovirus-like model for Ebola virus bipartite promoters

Irina Gutsche[ID]¹*, Philippe le Mercier[ID]², Daniel Kolakofsky³*

1 Institut de Biologie Structurale, Univ Grenoble Alpes, CEA, CNRS, IBS, Grenoble, France, 2 Swiss-Prot Group, Swiss Institute of Bioinformatics, Centre Médicale Universitaire, Geneva, Switzerland, 3 Department of Microbiology and Molecular Medicine, University of Geneva Medical School, Geneva, Switzerland

* irina.gutsche@ibs.fr (IG); Daniel.Kolakofsky@unige.ch (DK)

Abstract

Paramyxo- and filovirus nucleocapsids (NCs) have bipartite promoters at their 3′ ends to initiate RNA synthesis. The 2 elements, promoter element 1 (PE1) and promoter element 2 (PE2), are separated by a spacer region that must be exactly a multiple of 6 nucleotides (nt) long. Paramyxovirus NCs have 13 nucleoprotein (NP) subunits/turn, such that PE1 and PE2 are juxtaposed on the same face of the NC helix, for concerted recognition by the viral polymerase. Ebola virus (EBOV) NCs, in contrast, have 25 to 28 subunits/turn, meaning that PE1 and PE2 cannot be juxtaposed. However, there is evidence that the number of subunits/turn at the 3′ end of the EBOV NC is variable. We propose a paramyxovirus-like model for EBOV explaining why there are 8 contiguous copies of the PE2 repeat when 3 are sufficient, why expanding this run to 13 further improves minigenome performance, and why there is a limit to the number of hexa-nt that can be inserted in the spacer region.

Non-segmented negative-strand RNA virus (nsNSV) genomes are 12 to 20 kB in length and contain 5 to 10 tandemly arranged genes. There are 4 main families: Rhabdoviridae (vesicular stomatitis virus (VSV) and rabies), Paramyxoviridae (Sendai virus (SeV) and measles), Pneumoviridae (respiratory syncytial virus (RSV)), and Filoviridae (Ebola virus (EBOV) and Marburg virus (MARV)). nsNSV genome RNAs are enclosed within a non-covalent polymer of nucleoprotein (NP) to form a helical nucleocapsid (NC), the substrate for all viral RNA synthesis. The NP subunits of paramyxo- and filovirus NCs bind precisely 6 nucleotides (nt) each, with 3 contiguous bases pointing toward the protein core and 3 pointing away, whereas those of pneumo- and rhabdoviruses bind 7 and 9 nt/subunit, respectively. Paramyxo- and filoviruses also differ from the other 2 families in that their mRNAs can be co-transcriptionally edited to express alternate open reading frames (ORFs). EBOV edits its glycoprotein mRNA to make the essential full-length glycoprotein, and MARV (a member of another filovirus genus), which does not need to edit its glycoprotein mRNA, may edit its NP and polymerase (L) mRNAs [1]. Furthermore, paramyxo- and filovirus genomes have bipartite promoters at their 3′ ends to initiate RNA synthesis. There is both a 3′ promoter element (PE1) within the short leader region, and an internal element (promoter element 2 (PE2)) within the 5′ UTR of the first gene. PE1 and PE2 are separated by a spacer region that includes the first mRNA start site near position 56 from the genome 3′ end. Initiation of RNA synthesis from this end not only

requires both elements but also that they be separated by a spacer region that is a multiple of 6 nt long. The insertion or deletion of even a single nt here (or a stretch of RNA that is not a multiple of 6 nt long) will inactivate these promoters [2–5].

Paramyxovirus genome replication is governed by the "rule of six" [6], i.e., all paramyxovirus genomes found in nature are a multiple of 6 nt long, and only minigenomes of hexamer length replicate well in cell culture. This rule imposes a hexamer phase on the entire genome, which is composed of a series of hexa-nt bound to each NP subunit. Paramyxovirus PE2 simply consists of 3 contiguous hexamers where only 1 or 2 of these nt are at all important, e.g., $3'$-CN_5 in NPs number 14–16 (NP#14–16, counting from the genome $3'$ end) for SeV, and $3'$-N_4GC in NP#13–15 for parainfluenza virus type 5 (PIV5) [2,7]. Because paramyxovirus NCs contain approximately 13 NP subunits/turn [8–12], these tripartite PE2s are juxtaposed on the same axial face of the NC helix at the $3'$ end of the genome, presumably for concerted recognition by the viral polymerase [13,14]. The manner in which this bipartite promoter operates was recently examined for PIV2, a close relative of PIV5. Ten residues of their NP RNA-binding groove make contact with the RNA, 9 with the ribose-PO_4 backbone, and 1, gln202, with an nt base modeled as uracil [15]. Minigenomes containing wild-type (wt) NP-Q202 require both PE1 and PE2, as well as hexamer-length genomes for activity (i.e., wt hexamer phasing). However, when Q202 is mutated to 1 of several other residues, PE2 is no longer required at all, and non-hexamer-length minigenomes remain active [16]. PE1 was proposed to contain a negative element, namely Q202 making base-specific contact with the genome $3'$ uridine, where all RNA synthesis likely begins. This interaction presumably prevents promoter activity unless polymerase can also make contact with the correctly phased PE2 tripartite repeat [17]. In this manner, paramyxovirus bipartite promoters ensure the hexamer phase of the entire genome, including that of the *cis*-acting mRNA editing signal where hexamer phase participates in regulating this process [18].

Paramyxo- and filoviruses share many properties, but there are also important differences. EBOV actually acts like a pair of viruses, whose overall genome lengths differ by a single nt because their mRNA editing site can contain a run of either 7 or 8 uridines (EBOV/7U and EBOV/8U); both genomes are present during infection, regardless of which EBOV has initiated the infection [19]. Moreover, EBOV and MARV minigenomes are indifferent to their overall length, even though their bipartite promoters must obey the same rules of PE2 hexamer phasing as those of paramyxoviruses [5,20,21]. Remarkably, EBOV Zaire (or ZEBOV) contain 8 contiguous copies of its PE2 hexamer repeat ($3'$-NNURRN for EBOV/8U that is of hexamer length and $3'$ NNNURR for EBOV/7U) [22]. A minimum of 3 are still active, and the promoter gains strength when more hexamer repeats are present [5]. MARV in contrast contains the apparent minimum of 3 contiguous hexamer repeats [23]. Notably, the run of 8 hexamer repeats is not conserved in the 5 other species of the EBOV genus; EBOV Bundibugyo virus (BDBV) and Tai Forest virus (TAFV) have 6, Sudan virus (SUDV) and Reston virus (RESTV) have 5, and Bombali virus (BOMV) has 4 (Fig 1). When ZEBOV's run of 8 hexamer repeats is extended into the leader region (via a single base change) to 13 contiguous repeats, this further improves minigenome reporter gene expression. Most remarkably, this spacer region can be expanded to accommodate as many as 8 or 9 additional hexa-nt, as long as the hexamer phase of the PE2 repeats remains the same [20], and BOMV, in fact, appears to have included 4 additional hexa-nt between the almost perfectly conserved leader plus gene start site and the hexamer repeats. This is in strong contrast to paramyxovirus bipartite promoters where displacement of the 3 PE2 hexamers by deletion of 6 nt in the spacer region between PE1 and PE2 abolished SeV and PIV5 minigenome replication [2,4]. For SeV, 6 nt insertions in this region of the promoter were tolerated, but insertion of 12 nt largely ablated promoter activity [24]. Thus, paramyxovirus bipartite promoters have a more limited tolerance for alteration of
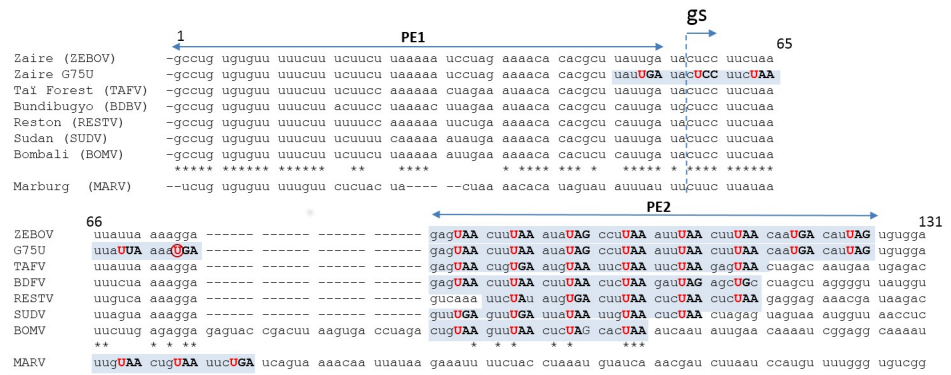
Fig 1. Filovirus bipartite promoters. This alignment of the EBOV/U7 genome 3′ ends is based on the premise that genomes are assembled directly from their 5′ ends, 6 nt/NP protomer, generating genome hexamer phase where only nt in hexamer positions 3, 4, and 5 point away from the protein core and can readily interact with the polymerase (underlined below). At least for ZEBOV, sequences upstream of the editing site will have different hexamer phases, e.g., 3′ NNN**U**RR for EBOV/U7 and 3′ NN**U**RRN for EBOV/U8 for the PE2 repeats. The conserved 3′ URR of contiguous PE2 repeats (shaded in blue) is capitalized. The conserved U of the repeat is highlighted in red and that of G75U is circled, and asterisks below the EBOV sequences indicate their overall sequence conservation. MARV, aligned below, does not appear to have any spacer region between gs (indicated by bent arrow) and the 3 PE2 repeats. Only the ZEBOV and MARV PE2 repeats have been experimentally verified [23]. BDBV, Bundibugyo virus; BOMV, Bombali virus; EBOV, Ebola virus; gs, gene start; MARV, Marburg virus; NP, nucleoprotein; nt, nucleotides; PE1, promoter element 1; PE2, promoter element 2; RESTV, Reston virus; SUDV, Sudan virus; TAFV, Tai Forest virus; ZEBOV, EBOV Zaire.

the spacing of the 2 promoter elements. The superabundance of EBOV PE2 hexamer repeats, and the tolerance of the hexa-nt length of the spacer region suggests that the NC structure recognized by EBOV polymerase for initiation of RNA synthesis is unlikely to be unique or fixed, as suggested for paramyxoviruses.

If paramyxo- and filoviruses evolved from a common ancestor and have adopted bipartite promoters to assure genome hexamer phasing (in large part to regulate mRNA editing), we would expect their bipartite promoters to operate similarly to the extent that their biology permits. The EBOV model we recently presented based on a 13 NP/turn helix, for which there is no evidence of its existence, is more than questionable and does not explain the filovirus data [20,22], while remaining valid for paramyxoviruses. Recent cryo-electron tomography structures of EBOV NCs within intact viruses and recombinant NC-like assemblies allow estimating their approximate helical symmetries from subtomogram averaging and show a variable number of subunits per turn ranging between roughly 25 and 28 [25]. This assembly, however flexible, does not lend itself to a paramyxovirus-like model (polymerase's simultaneous interaction of PE1 and PE2), as the 8 PE2 hexamer repeats are in NP#14–21 from the 3′ end. A paramyxovirus-like model of interaction would require abandoning the notion of a uniform EBOV NC structure that includes the genome 3′ end. If PE1 and PE2 need to be juxtaposed on the first 2 turns of the helix so that the polymerase can interact with both (and judge the intervening sequence length to within a single nt), this requires much more helix plasticity for the 3′ end of the EBOV NC, and here, another nsNSV may be instructive.

Rhabdo- and filoviruses form rod-like virions whose length is determined by that of their NC, whereas paramyxo- and pneumovirus NC are contained within spherical envelopes. The 3′ end of the VSV NC, at the pointed end of the bullet-like virion, is itself also pointed or conical. Its tip (the 3′-most turn) likely consists of around 10 subunits/turn and increases to the uniform 33 subunits/turn of the NC trunk after 5 to 6 turns [26]. Remarkably, the formation of this conical tip of the VSV NC turns out to be an inherent property of its NP [27]. Electron

microscopy images of the EBOV and MARV particles show that the tip of the virus and the visible NC filament within is actually also conical [28,29], but until now, all structural studies were focused on the straightest and most ordered helical portion of the EBOV NC [25,28,30]. While the structure at the tip of the NC has not as yet been resolved, the variation in the number of subunits per turn of VSV NC also occurs at the 3′ end of EBOV NCs, at least inside the virus particle. VSV NCs within virions are of course fixed structures, whereas EBOV NCs within the cell cytoplasm, perhaps with polymerase in place to initiate RNA synthesis, need not be fixed. In such a scheme for EBOV, the diameter of the first turn of the helix may be variable and dynamic. The G75U mutant of Bach and colleagues [20] offers some insights into the structural plasticity of the EBOV tip enabling alignment of PE1 and PE2 for promoter activation. Due to this single base change in this mutant, the contiguous stretch of 8 3′ NNURRN repeats (NP#14–21) is extended to 13 (NP#9–21) (Fig 1). If the lower limit of the first turn of the helix were 10 subunits/turn similarly to VSV NC, the tripartite repeats of NP#11–13 would align with NP#1–3 on the first 2 turns of the helix for activity (Fig 2A), which may explain why the G75U mutant outperforms the wt minigenome [20]. When there are 13 subunits/first turn, the first of the wt tripartite repeats (NP#14–16) would align (Fig 2B). When there are 14 subunits/first turn, the second of the wt tripartite repeats (NP#15–17) would align, etc., and when there are 18 subunits/first turn, the last of the wt tripartite repeats (NP#19–21) would align (Fig 2C). Notably, in this last case of 18 subunits/first turn, the insertion of additional hexa-nt to the spacer region would simply displace NP#19–21 downstream, to be replaced by the more 3′ tripartite repeats, until NP#14–16 are aligned with the 3′ end (after the addition of 5 hexa-nt
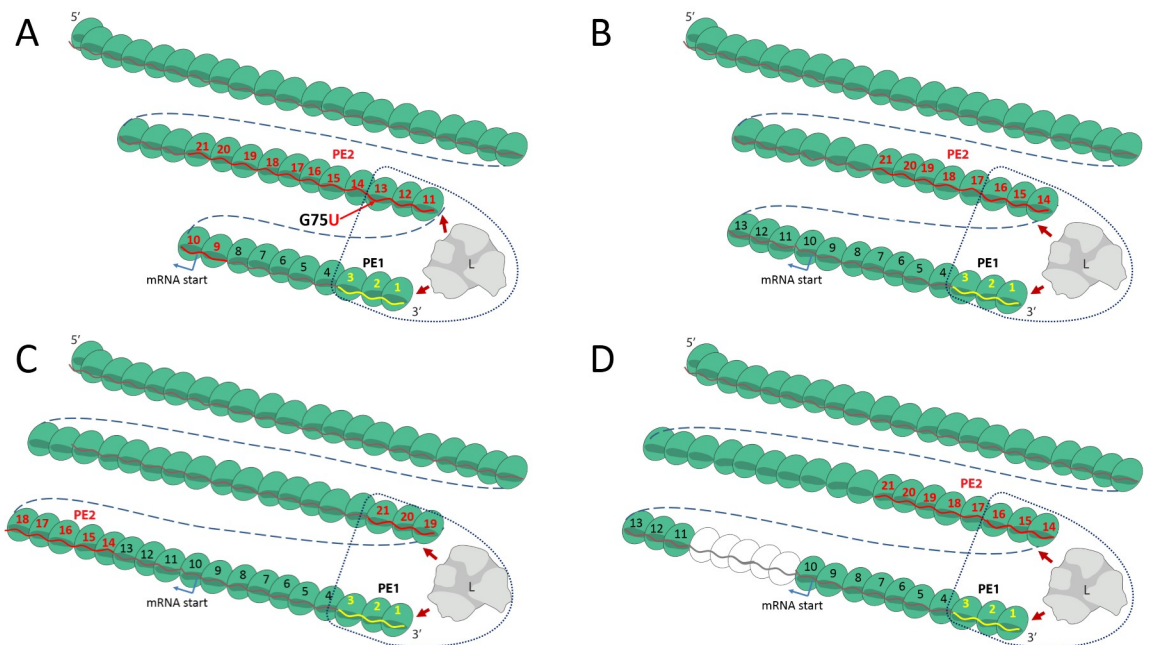
Fig 2. A paramyxovirus-like model for EBOV bipartite promoters, based on a variable and dynamic 3′ end of the NC helix. The G75U and wt NC helices are schematically shown as 2-dimensional chains of NP subunits (green spheres), numbered from their 3′ ends, whose first turn is either 10 (panel A), 13 (panel B), or 18 subunits long (panels C and D). The open spheres in panel D represent hexa-nt insertions in the wt minigenome. Their genome RNAs are shown as a wavy line within their NP RNA-binding grooves (darker green). PE1, modeled as the first 3 subunits, is highlighted in yellow. Subunits containing contiguous PE2 hexamer repeats (3′ NNURRN for EBOV/8U and 3′ NNNURR for EBOV/7U) are highlighted in red. The polymerase (L) is shown in gray, interacting with PE1 and the minimum PE2 tripartite repeat needed for the initiation of RNA synthesis. For further detail, see text. EBOV, Ebola virus; NC, nucleocapsid; NP, nucleoprotein; nt, nucleotides; PE1, promoter element 1; PE2, promoter element 2; wt, wild-type.

to the spacer region (Fig 2D)). The addition of more than 5 hexa-nt to the spacer in this case leads to a loss of promoter activity because PE2 can no longer align with PE1, and this places a limit on the length of RNA that can be added to the spacer [20]. The variable length of the first turn allows all 8 of the wt repeats and 3 of the additional repeats of the G75U mutant to participate in the alignment of PE1 and PE2. If the first 2 turns of the helix are dynamic, with the PE1-contacting region of the L polymerase fixed near the 3′ end so that it can negatively control the promoter, the PE2-contacting region of L can scan for an NC region with the correctly phased PE2 tripartite repeat. Only upon this concerted interaction can the negative effect on PE1 be overcome and RNA synthesis initiate. In this way, paramyxo- and filovirus bipartite promoters would operate similarly, even though EBOV genomes need not, nor cannot, given their biology, always be of hexamer length (e.g., EBOV/7U is 6n+5 nt long, (Fig 1)). As for MARV that contains the apparent minimum of 3 contiguous PE2 hexamer repeats, it may not require this degree of promoter plasticity [21,23].

Acknowledgments

References

1. Shabman RS, Jabado OJ, Mire CE, Stockwell TB, Edwards M, Mahajan M, et al. Deep sequencing identifies noncanonical editing of Ebola and Marburg virus RNAs in infected cells. MBio. 2014; 5: e02011. https://doi.org/10.1128/mBio.02011-14 PMID: 25370495

2. Tapparel C, Maurice D, Roux L. The activity of Sendai virus genomic and antigenomic promoters requires a second element past the leader template regions: a motif (GNNNNN)3 is essential for replication. J Virol. 1998; 72:3117–3128. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9525637 https://doi.org/10.1128/JVI.72.4.3117-3128.1998 PMID: 9525637

3. Murphy SK, Parks GD. RNA replication for the paramyxovirus simian virus 5 requires an internal repeated (CGNNNN) sequence motif. J Virol. 1999; 73:805–809. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9847393 https://doi.org/10.1128/JVI.73.1.805-809.1999 PMID: 9847393

4. Murphy SK, Ito Y, Parks GD. A functional antigenomic promoter for the paramyxovirus simian virus 5 requires proper spacing between an essential internal segment and the 3' terminus. J Virol. 1998; 72:10–19. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9420195 https://doi.org/10.1128/JVI.72.1.10-19.1998 PMID: 9420195

5. Weik M, Enterlein S, Schlenz K, Mühlberger E. The Ebola virus genomic replication promoter is bipartite and follows the rule of six. J Virol. 2005; 79:10660–10671. https://doi.org/10.1128/JVI.79.16.10660-10671.2005 PMID: 16051858

6. Calain P, Roux L. The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. J Virol. 1993; 67:4822–4830. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8392616 https://doi.org/10.1128/JVI.67.8.4822-4830.1993 PMID: 8392616

7. Murphy SK, Parks GD. RNA replication for the paramyxovirus simian virus 5 requires an internal repeated (CGNNNN) sequence motif. J Virol. 1999; 73:805–809. https://doi.org/10.1128/JVI.73.1.805-809.1999 PMID: 9847393

8. Bhella D, Ralph A, Yeo RP. Conformational flexibility in recombinant measles virus nucleocapsids visualised by cryo-negative stain electron microscopy and real-space helical reconstruction. J Mol Biol. 2004; 340:319–331. https://doi.org/10.1016/j.jmb.2004.05.015 PMID: 15201055

9. Schoehn G, Mavrakis M, Albertini A, Wade R, Hoenger A, Ruigrok RWH. The 12 A structure of trypsin-treated measles virus N-RNA. J Mol Biol. 2004; 339:301–312. https://doi.org/10.1016/j.jmb.2004.03.073 PMID: 15136034

10. Desfosses A, Goret G, Farias Estrozi L, Ruigrok RWH, Gutsche I. Nucleoprotein-RNA orientation in the measles virus nucleocapsid by three-dimensional electron microscopy. J Virol. 2011; 85:1391–1395. https://doi.org/10.1128/JVI.01459-10 PMID: 21106738

11. Gutsche I, Desfosses A, Effantin G, Ling WL, Haupt M, Ruigrok RWH, et al. Structural virology. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. Science. 2015; 348:704–707. https://doi.org/10.1126/science.aaa5137 PMID: 25883315

12. Desfosses A, Milles S, Jensen MR, Guseva S, Colletier J-P, Maurin D, et al. Assembly and cryo-EM structures of RNA-specific measles virus nucleocapsids provide mechanistic insight into paramyxoviral

replication. Proc Natl Acad Sci U S A. 2019; 116:4256–4264. https://doi.org/10.1073/pnas.1816417116 PMID: 30787192

13. Kolakofsky D, Roux L, Garcin D, Ruigrok RWH. Paramyxovirus mRNA editing, the "rule of six" and error catastrophe: a hypothesis. J Gen Virol. 2005; 86:1869–1877. https://doi.org/10.1099/vir.0.80986-0 PMID: 15958664

14. Kolakofsky D. Paramyxovirus RNA synthesis, mRNA editing, and genome hexamer phase: a review. Virology. 2016; 498:94–98. https://doi.org/10.1016/j.virol.2016.08.018 PMID: 27567257

15. Alayyoubi M, Leser GP, Kors CA, Lamb RA. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex. Proc Natl Acad Sci U S A. 2015; 112:E1792–E1799. https://doi.org/10.1073/pnas.1503941112 PMID: 25831513

16. Matsumoto Y, Ohta K, Kolakofsky D, Nishio M. A point mutation in the RNA-binding domain of human parainfluenza virus type 2 nucleoprotein elicits abnormally enhanced polymerase activity. J Virol. 2017; 91. https://doi.org/10.1128/JVI.02203-16 PMID: 28179533

17. Matsumoto Y, Ohta K, Kolakofsky D, Nishio M. The control of paramyxovirus genome hexamer length and mRNA editing. RNA. 2018; 24:461–467. https://doi.org/10.1261/rna.065243.117 PMID: 29358233

18. Iseni F, Baudin F, Garcin D, Marq J-B, Ruigrok RWH, Kolakofsky D. Chemical modification of nucleotide bases and mRNA editing depend on hexamer or nucleoprotein phase in Sendai virus nucleocapsids. RNA. 2002; 8:1056–1067. https://doi.org/10.1017/s1355838202029977 PMID: 12212849

19. Volchkova VA, Dolnik O, Martinez MJ, Reynard O, Volchkov VE. Genomic RNA editing and its impact on Ebola virus adaptation during serial passages in cell culture and infection of guinea pigs. J Infect Dis. 2011; 204(Suppl 3):S941–S946. https://doi.org/10.1093/infdis/jir321 PMID: 21987773

20. Bach S, Biedenkopf N, Grünweller A, Becker S, Hartmann RK. Hexamer phasing governs transcription initiation in the 3'-leader of Ebola virus. RNA. 2020; 26:439–453. https://doi.org/10.1261/rna.073718.119 PMID: 31924730

21. Enterlein S, Schmidt KM, Schümann M, Conrad D, Krähling V, Olejnik J, et al. The marburg virus 3' non-coding region structurally and functionally differs from that of ebola virus. J Virol. 2009; 83:4508–4519. https://doi.org/10.1128/JVI.02429-08 PMID: 19225002

22. le Mercier P, Kolakofsky D. Bipartite promoters and RNA editing of paramyxoviruses and filoviruses. RNA. 2019; 25:279–285. https://doi.org/10.1261/rna.068825.118 PMID: 30587495

23. Hume AJ, Mühlberger E. Distinct genome replication and transcription strategies within the growing filovirus family. J Mol Biol. 2019; 431:4290–4320. https://doi.org/10.1016/j.jmb.2019.06.029 PMID: 31260690

24. Pelet T, Delenda C, Gubbay O, Garcin D, Kolakofsky D. Partial characterization of a Sendai virus replication promoter and the rule of six. Virology. 1996; 224:405–414. https://doi.org/10.1006/viro.1996.0547 PMID: 8874501

25. Wan W, Kolesnikova L, Clarke M, Koehler A, Noda T, Becker S, et al. Structure and assembly of the Ebola virus nucleocapsid. Nature. 2017; 551:394–397. https://doi.org/10.1038/nature24490 PMID: 29144446

26. Ge P, Tsao J, Schein S, Green TJ, Luo M, Zhou ZH. Cryo-EM model of the bullet-shaped vesicular stomatitis virus. Science. 2010; 327:689–693. https://doi.org/10.1126/science.1181766 PMID: 20133572

27. Desfosses A, Ribeiro EA, Schoehn G, Blondel D, Guilligay D, Jamin M, et al. Self-organization of the vesicular stomatitis virus nucleocapsid into a bullet shape. Nat Commun. 2013; 4:1429. https://doi.org/10.1038/ncomms2435 PMID: 23385574

28. Bharat TAM, Noda T, Riches JD, Kraehling V, Kolesnikova L, Becker S, et al. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. Proc Natl Acad Sci U S A. 2012; 109:4275–4280. https://doi.org/10.1073/pnas.1120453109 PMID: 22371572

29. Bharat TAM, Riches JD, Kolesnikova L, Welsch S, Krähling V, Davey N, et al. Cryo-electron tomography of Marburg virus particles and their morphogenesis within infected cells. PLoS Biol. 2011; 9:e1001196. https://doi.org/10.1371/journal.pbio.1001196 PMID: 22110401

30. Sugita Y, Matsunami H, Kawaoka Y, Noda T, Wolf M. Cryo-EM structure of the Ebola virus nucleoprotein-RNA complex at 3.6 Å resolution. Nature. 2018; 563:137–140. https://doi.org/10.1038/s41586-018-0630-0 PMID: 30333622