Ebola Virus GP Gene Polyadenylation Versus RNA Editing

Valentina A. Volchkova, Jaroslav Vorac, Laurie Repiquet-Paire, Philip Lawrence, and Viktor E. Volchkov

Molecular Basis of Viral Pathogenicity, CIRI, INSERM, U1111-CNRS UMR5308, Université de Lyon, Université Claude Bernard Lyon 1, Ecole Normale Supérieure de Lyon, France

Synthesis of Ebola virus (EBOV) surface glycoprotein (GP) is dependent on transcriptional RNA editing. Northern blot analysis of EBOV-infected cells using GP-gene-specific probes reveals that, in addition to full-length GP messenger RNAs (mRNAs), a shorter RNA is also synthesized, representing >40% of the total amount of GP mRNA. Sequence analysis demonstrates that this RNA is a truncated version of the full-length GP mRNA that is polyadenylated at the editing site and thus lacks a stop codon. An absence of detectable levels of protein synthesis in cellulo is consistent with the existence of tight regulation of the translation of such mRNA. However, nonstop GP mRNA was shown to be only slightly less stable than the same mRNA containing a stop codon, against the general belief in nonstop decay mechanisms aimed at detecting and destroying mRNAs lacking a stop codon. In conclusion, we demonstrate that the editing site indeed serves as a cryptic transcription termination/polyadenylation site, which rarely also functions to edit GP mRNA for expression of surface GP. This new data suggest that the downregulation of surface GP expression is even more dramatic than previously thought, reinforcing the importance of the GP gene editing site for EBOV replication and pathogenicity.

Keywords. Ebola virus; transcriptional RNA editing; polyadenylation; GP gene; nonstop mRNA.

Zaire ebolavirus (EBOV) is 1 of 5 distinct species within the genus Ebolavirus, Filoviridae family, order Mononegavirales. The virus was first discovered in 1976 in Zaire during an outbreak of a previously unknown disease [1]. This virus is characterized by high mortality rates, reaching up to 90% in both humans and nonhuman primates. The viral genome encodes 7 structural and a number of nonstructural proteins [2–4]. The glycoprotein (GP) is the single surface spike protein that enables viral entry into susceptible target cells. During intracellular processing and maturation, GP is oligomerized, highly N–and O-glycosylated, and cleaved by the cellular protease furin, resulting in 2 cysteine linked subunits, GP1 and GP2 [5–7]. Significant amounts of surface GP are released from cells in both soluble and membrane-associated forms [8, 9]. It has been demonstrated that when overexpressed, GP causes strong cytopathic effects, such as cell rounding and detachment [10]. This phenomenon was explained by masking of cellular surface proteins, including integrins, preventing the function of cellular proteins and eventually leading to the death of detached cells [11, 12].

All EBOV genes are transcribed into individual messenger RNAs (mRNAs) with the exception of the GP gene. In virus-infected cells, several GP-specific mRNAs are synthesized due to transcriptional RNA editing phenomenon [4, 13, 14]. The majority of GP gene transcripts are unedited and encode nonstructural secreted GP (sGP). Surface GP is synthesized from edited mRNAs containing an extra adenosine inserted by the viral polymerase at the editing site that consists of 7 consecutive uridine residues (genomic sense). Insertion of adenosine at the editing site is explained by slippage or stuttering of the viral polymerase [4, 13], caused by a temporal loss of interaction and then realigning of the genome template and newly synthesized RNAs, and is due to weak bonding between adenine and uridine residues. It had earlier been demonstrated that the ratio between edited versus unedited GP mRNAs is about 1:4 [4, 13]. A similar stuttering mechanism has been described for paramyxoviruses where the P gene is transcribed into several P-gene-specific mRNAs [15, 16].

Correspondence: Viktor E. Volchkov, PhD, Human Virology, Laboratory Molecular Basis of Viral Pathogenicity, International Centre for Research in Infectiology (CIRI), INSERM U1111, UCBL1, 21 Ave Tony Garnier, 69365 LYON Cedex 07, France (viktor.volchkov@inserm.fr).

The Journal of Infectious Diseases® 2015:212:S191–8
© The Author 2015. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.
DOI: 10.1093/infdis/jiv150
An apparently similar event is also occurring during replication of genomic RNA of EBOV, resulting in the appearance of an editing site variant, EBOV/8U, containing 8 uridine residues at the editing site \([2, \ 4, \ 17]\). In the case of EBOV/8U, surface GP is synthesized from nonedited mRNA transcripts, whereas synthesis of sGP is dependent on editing. Notably, the accuracy of editing with EBOV/8U appears to be altered. Multiple adenosine insertions or deletions at the editing site were observed \([4]\). The addition of 1 residue in this case gives 9A at the editing site, and such mRNA codes for a third GP gene product known as second secreted GP (ssGP). In this way, ssGP mRNA contains a stop codon 2 amino acids downstream of the editing site and codes for a protein that is 26 amino acids shorter than sGP. In this case, insertion of 2 residues at the editing site or the deletion of 1 again gives rise to sGP.

In this study, we demonstrate that the editing site indeed serves as a cryptic transcriptional termination/polyadenylation signal. Rarely it also functions for editing, giving rise to mRNA coding for surface GP indispensable for EBOV replication.

**MATERIALS AND METHODS**

**Viruses**

The recombinant wild-type EBOV (EBOV/7U), and mutants EBOV/8U and EBOV/NE were generated as described elsewhere \([17, \ 18]\) (see also Supplementary Experimental Procedures).

**Cell Lines, Infection, Transfection, and Plasmids**

293T and Vero E6 cells were grown in Dulbecco modified Eagle medium with the addition of 10% fetal calf-serum. BSR T7/5 (baby hamster kidney [BHK]) cells were maintained in Glasgow medium (Gibco) supplemented with 10% newborn calf serum.

Transfection of 293T or BSR T7/5 cells with capped RNA was performed using TransIT-mRNA transfection kit (Mirus) or the Neon electroporation system (Invitrogen) according to manufacturer’s instruction. Details of plasmid constructions are provided in the Supplementary Experimental Procedures. Western blot analysis of cell lysates and culture supernatants was performed with corresponding antibodies (see Supplementary Experimental Procedures). For infection, subconfluent VeroE6 cells were infected with EBOV at a multiplicity of infection (MOI) 5. Cells were harvested 48 hours postinfection in RLT buffer and used for total RNA or polyA RNA extraction and further northern blot analysis (see Supplementary Experimental Procedures).

**Biotinylated Probes**

Biotinylated probes were amplified in the presence of 40% Biotin-16-UTP by standard polymerase chain reaction (PCR) protocol (Roche). Probes 1, 2, and 3 were generated using plasmid containing GP and the following primer pairs, respectively: 5’ CTTTTCAAAGAACATTTCGCC3’ and 5’ ATACCTTGTGCATACCGG3’, 5’ GACTTCTTCACTCACACC3’ and 5’ GCTCTTTCTCCACTTTGTA3’, 5’ CAAAGACATCAGTGGTCAGA3’ and 5’ TGTGTCTGGTTCTGGGA3’. Probes were purified by Qiaquick PCR Purification Kit (Qiagen).

**Analysis of Viral mRNA PolyA Tail Length**

RNA isolated from approximately 10^6 VeroE6 cells infected with different EBOVs as described above was ligated with Universal anchor primer (5’p-CAUGGCCAUCUCUGACC GGAUCUUA) as described elsewhere \([19]\). The entire product was reverse transcribed by Superscript III (Invitrogen) using of P1 primer (antisense to anchor primer: 5’-TATAGGATCC GACTGAGAGATG). Resulting complementary DNA (cDNA) was amplified by Pfu Ultra (Invitrogen) by annealing P1 and GP gene-specific primers (see Supplementary Experimental Procedures). The PCR products were then cloned within the BamHI and HindIII sites of the pGem3zf(+) vector and sequenced.

**In Vitro Transcription/Translation**

T7 TNT Quick coupled transcription/translation system (Promega) was used according to manufacturer’s instructions (see Supplementary Experimental Procedures).

**RESULTS AND DISCUSSION**

The GP gene editing site (UUUUUUU) somewhat resembles the polyadenylation/transcription termination site present at the end of every viral gene, including that of GP gene (UAAAUU CUUUUU). All of these sites contain a stretch of uridine residues (genomic sense) that are the basis for insertion of nontemplate adenosomes during transcription \([2]\). In this respect, it is plausible to suggest that due to stuttering at this stretch of U residues, the viral polymerase could form a polyA sequence and potentially give rise to a prematurely truncated GP (GPTr)-gene-specific mRNA if polyadenylation results in transcription termination. Indeed, such short, GP-gene-specific mRNA was detected in the work of Sanchez et al \([2]\) when a plaque-purified variant EBOV/8U was used instead of wild-type virus. Expression of surface GP by this variant of EBOV principally occurs without the need for editing because the 2 overlapping reading frames present in the GP gene of wild-type EBOV are linked together through an additional uridine residue at the editing site. As a result, editing phenomenon was not noted in this study and the role of the short GP mRNA was not defined. After the publication of RNA editing as the principal mechanism for surface GP expression in wild-type EBOV \([4, \ 13]\), the existence of shorter GP-specific mRNA was overlooked and all following publications have concentrated only on the 2 full-length mRNA transcripts coding for sGP and surface GP. In these following studies, reverse-transcriptase-PCR techniques and primers flanking each side of the editing site were used to address...
editing frequency. This technique provided an estimation of the proportion of the edited form versus nonedited mRNAs. However, this approach clearly makes it impossible to detect and appreciate truncated GP mRNAs (GPTr mRNA). In this study to look for the existence of GPTr mRNA in wild-type EBOV, we designed 3 cDNA probes to target specific sequences, 2 upstream and 1 downstream of the editing site (Figure 1A). Moreover, we used a series of recombinant EBOV constructs differing in their GP gene editing site; wild-type virus carrying 7 uridine residues at the editing site (EBOV/7U), a mutant bearing 8 Us (EBOV/8U), and also a “no-editing” mutant (EBOV/NE) containing a silently mutated editing site (UU→CU) that prevents a basis for stuttering and codes only for surface GP while preserving conservation of both the reading frame and amino acid sequence of the protein [18].

VeroE6 cells were infected with these viruses at an MOI of 5 and cells were harvested for RNA extraction 24 hours postinfection. Samples of RNA extracted from the cells were quantified for total RNA levels and equal amounts of RNA were subjected to Northern blot analysis using the 3 biotinylated probes generated by PCR (Figure 1B). When probe #3 was used for detection, a single, high-molecular-weight RNA band was observed for all viruses. In contrast, with probes #1 and #2, 2 distinct bands were found for the 7U and 8U variants, whereas the lower band was not seen for EBOV/NE. The lower band of around 1000 nucleotides corresponds to the expected size of the mRNA that would be produced if the editing site is indeed able to function as a transcription termination signal. To confirm that this smaller band represents a polyadenylated GP-specific mRNA, we next performed purification of RNA using oligo-dT columns (Qiagen) followed by Northern blot analysis (Figure 1C). Two bands were detected, supporting the notion that both are polyadenylated GP gene-specific mRNA species. Levels of both mRNAs were found to increase over the course of infection (Figure 1D, left and right). Notably, quantification of GP mRNAs revealed that GPTr mRNA amounts for over 40% of total GP gene-specific mRNAs synthesized in EBOV-infected cells.

To further explore the nature of this truncated mRNA species, we designed a strategy based on the ligation of an RNA primer to the polyA tail of mRNAs. This RNA primer is then targeted with a specific antisense primer in order to generate cDNA fragments that are amplified by PCR using the same universal antisense primer and gene-specific forward primers designed to detect sequences downstream of either the editing site or GP gene transcription termination site. Analysis of cDNA clones allowed us to quantify the variation in length of the polyA tail for both mRNA species. As seen in Figure 1E, polyA tails for both species were found to be rather short (<48 nt), whereas the existence of longer polyA tails was suggested for some other nonsegmented negative-polarity RNA viruses [19, 20]. While the selected cDNAs represent the majority of mRNA detected and amplified with this method, some mRNA species with longer tails may have been excluded during purification of PCR products. Of note, at least some polyA sequences may represent mRNAs with incomplete polyA tails due to interruption of the polyadenylation process at the moment of cell lysis.

Further sequence analysis revealed that short GP-specific mRNAs do not contain a translational stop codon. Such mRNA would result in translation of a polyA sequence and therefore in synthesis of a protein with a poly-lysine tail at the carboxy-terminal end. Importantly, this nonstop mRNA would represent a potential problem for the release of ribosomes reaching the 3′-end of the mRNA. The vast majority of research on nonstop mRNA has been performed in yeast or bacteria where such transcripts occur often due to the more frequent presence of cryptic sequences recognized as polyadenylation/transcription termination sites [21]. To resolve the problem of stalled ribosomes, yeast for example possess several mechanisms. It has been shown that nonstop mRNA is rapidly degraded by a nonstop mRNA decay pathway (NSD) [22, 23]. In this case, Ski7p recognizes stalled ribosome at the 3′ end of mRNA and recruits both the exosome and Ski complex, resulting in release of ribosomes, rapid degradation of nonstop mRNA, and thus in low levels of protein expression from the offending transcript. Furthermore, translation of nonstop mRNA is also repressed after initiation, most likely again due to stalling of ribosomes at the polyA end, thus limiting the following round of translation and working as a kind of “roadblock.” Stalled ribosomes prevent interaction of polyA-binding protein with the translation initiation factor eIF4G, a subunit of the cap-binding complex eIF4F [24, 25], and in this way interrupt the synergistic effects of the cap and the polyA tail in translation [26]. In addition, translation of the polyA into poly-lysine may lead to protein destabilization and degradation by the proteasome that would also result in low protein levels [27].

To address the capacity of the viral nonstop mRNA to produce a protein product, we first used a T7 in vitro transcription/translation system. For this, and in order to simulate polyadenylation at the editing site, the GP open reading frame ending at the editing site (8A) and containing an additional 18 adenosines was cloned into pGEM3Zf+ plasmid (pGEM3Zf-GPTr). To ensure the absence of any additional coding sequence after the polyA tail, a HindIII restriction site (AAGCTTT)
Figure 1. EBOV infection results in synthesis of full-length and truncated GP (GPTr) gene specific mRNAs. A, Schematic representation of GP gene editing site and probes used for Northern blot. B–D, Northern blot analysis of RNA isolated from EBOV-infected cells. GP-gene specific RNAs were isolated from VeroE6 infected either with recombinant wild-type EBOV (EBOV/7U), variant EBOV/8U, or EBOV/NE. C, Northern blot analysis of total and oligo-dT-purified RNA from EBOV-infected cells. D, Northern blot analysis of GP-specific mRNAs over the course of infection. Histogram shows quantification of full-length and truncated mRNA normalized to ribosomal 28S RNA. E, Sequence analysis of 3′-end of full-length and GPTr mRNAs. Histogram shows distribution of cDNA clones according to the number of adenosine residues at the 3′ end of GP mRNAs. Abbreviations: cDNA, complementary DNA; EBOV, Ebola virus; GP, glycoprotein; GPTr, truncated GP mRNA; mRNAs, messenger RNAs; sGP, secreted GP.
was placed immediately downstream of the 26 adenosine residues. In this way, HindIII-mediated linearization of the plasmid prior to transcription/translation is certain to give nonstop mRNA with a polyA tail of 27 adenosines. As a control, a construct was made in which a stop codon was introduced upstream of the editing site (pGEM3Zf-GPTr/stop). The protein produced from this plasmid serves as a control of translation from an mRNA containing a stop codon. The resulting biotinylated GP-specific proteins were visualized either with the aid of streptavidin–horseradish peroxidase (HRP) detection or with specific anti-GP antibody. As seen in Figure 2A, both plasmids clearly give proteins that are detected with both systems. As expected, the pGEM3Zf-GPTr construct results in a slightly higher molecular weight protein compared with pGEM3Zf-GPTr/stop. As the molecular weight of 36 kDa for this protein corresponds to the size predicted for GPTr, it would suggest that the entirety of the mRNA coding sequence, including the polyA tail, is being successfully translated by the T7 TNT system. Equal synthesis efficiency for both proteins suggests that the reticulocyte translation system is not affected by the lack of a stop codon and that ribosomes can presumably freely run off the template in the case of GPTr. Why nonstop mRNA does not pose a problem in this artificial system and does not result in ribosome stalling is unclear. Notably, when streptavidin–HRP detection was used for pGEM3Zf-GP, we observed the presence of 3 distinct protein bands, corresponding to full-length GP, sGP, and presumably ssGP. The presence of multiple GP-specific bands with a T7 TNT transcription/translation system was seen previously [4] and was explained by the ability of the T7 polymerase to also perform insertion of nontemplate adenosine residues at the editing site. The presence of the lower band product, corresponding to the molecular weight of ssGP (9A in mRNA), would suggest that the T7 polymerase is able to add at least 1 adenosine at this site. The presence of the SGP band would again suggest that the T7 polymerase is able to add at least 2 adenosines, or to delete 1, at this site. Of note and as observed in Figure 2A, in this assay the short polyA tail generated with the pGEM3Zf-GPTr leads to the addition of 6 lysine residues and would thus actually produce a protein that is very similar in length to ssGP containing a natural stop codon 2 amino acids downstream of the editing site.

Considering these results, we next investigated whether nonstop GP mRNAs can be translated in cellulo. For this, the same plasmid constructs described above were used with a T7 RiboMax system. RNA transcripts were capped, quantified (Figure 2B), and used to transfect 293T cells. As a control we used a similar RNA construct coding for green fluorescent protein (GFP) (Figure 2C). Western blot analysis of cells using anti-GP antibody allowed detection of protein translated from the GPTr/STOP encoding RNA (Figure 2D). Surprisingly, and in contrast to the results obtained above with the in vitro T7 TNT system, we failed to detect GP-specific protein in cells transfected with GPTr transcripts (Figure 2D). No release of GP-specific protein was detected in the medium for any construct. This result would imply that peptide synthesis from nonstop mRNA is in some way prohibited in mammalian cells. To ensure that the in vitro transcribed RNA encoding GPTr is functional, the same RNAs were tested in parallel in an in vitro rabbit reticulocyte lysate system. Both GPTr-STOP and GPTr encoding synthetic RNAs give high levels of protein in this system, confirming that the latter RNA samples were indeed translatable and give rise to a protein of the expected size.

Indeed, the fate of nonstop mRNA and its protein product in mammalian cells is less understood and sometimes controversial. Akimitsu and coauthors [28] postulate that nonstop mRNAs are less prone to fast degradation, whereas another group [29] identified an analog of Skip7 and showed that Hbs1-Dom34 protein complex is involved in degradation of nonstop RNA. Both groups agree that stalled ribosomes would result in translation repression of the protein coded by such mRNA, most likely through an interruption in synergy between cap and polyA. Of note, the majority of studies on nonstop mRNA in mammalian cells were performed using synthetic constructs based on GFP or luciferase sequences attached to polyA and lacking a stop codon. To date, and to the best of our knowledge, the existence of nonstop mRNA has only been described in the case of a human genetic disorder affecting fertility and suggested in the case of a mutation in 1 of the human Ski complex genes that leads to a rare congenital syndrome (Tricho-Hepato-Enteric Syndrome) [21], although the precise molecular mechanisms behind these afflictions are unknown.

Considering all that is mentioned above, it seems reasonable to assume that nonstop GP mRNA could be a perfect target for RNA decay that presumably would result in RNA degradation and thus low levels of protein expression. In yeast, for example, nonstop mRNA was shown to have a very short half-life of 0.8 minutes ± 0.1 minute [21]. Importantly, if such degradation of nonstop RNA is indeed occurring, then it would suggest that in reality even more truncated mRNA is synthesized in EBOV-infected cells than we actually detect in our assays. This hypothesis is in line with a recent publication where using deep-sequencing of RNA from EBOV-infected patients, researchers revealed that only 1% of GP gene transcripts apparently coded for surface GP [30]. This led us to test the stability of the nonstop GPTr mRNA. We compared the stability of the same GPTr/STOP versus GPTr-encoding RNAs described above following electroporation of BHK cells with equal amounts of each RNA species. As seen in Figure 2E, the nonstop GPTr RNA displays increased degradation kinetics early posttransfection compared with GPTr/STOP RNA. Interestingly, after around 30 minutes, the degradation of the nonstop mRNA appears to slow, and indeed this RNA remains rather stable for the remainder of the assay. This is in contrast to the mRNA coding for GPTr/STOP where degradation continues
Figure 2. Expression and stability of truncated nonstop GP mRNA. A, Western blot analysis of GP gene-specific proteins. In vitro transcription/translation was performed using coupled TNT T7 kit (Promega). Nonradioactive biotin-labeled proteins were revealed either with streptavidin-coupled HRP (left) or GP-specific antibodies (right). Introduction of a stop codon upstream of the editing site resulted in synthesis of protein with low molecular weight compared to protein expressed from nonstop GP mRNA. B, Analysis of capped RNA transcribed from templates using T7 RiboMAX in the presence of Ribo m7G Cap analog on agarose gel. C, Cells transfected with capped RNA encoding GFP. D, Western blot analysis of 293T cells transfected with capped RNA samples as in (B) using anti-GP antibodies. Truncated GP (GPTr) produced in vitro serves as a control. Protein produced in cells shows an increase in molecular mass due to glycosylation. E, Analysis of RNA stability. Samples of RNA were transfected into BHK cells as indicated and then analyzed over time by Northern blot. Histogram (right) shows relative quantification of RNA normalized to 28S RNA. Amounts of RNA detected immediately posttransfection were adjusted to 100%. F, Schematic representation of the EBOV GP gene expression strategy. Editing site is recognized by viral polymerase as a cryptic polyadenylation site, giving rise to several GP gene specific mRNA species as indicated. Abbreviations: BHK, baby hamster kidney; EBOV, Ebola virus; HRP, horseradish peroxidase; GFP, green fluorescent protein; GP, glycoprotein; mRNA, messenger RNA; ORF, open reading frame; sGP, secreted GP; ssGP, second secreted GP [31].
over the course of the assay at about the same speed (Figure 2E, left and right). Overall, this observation would suggest that BHK cells have a somewhat limited capacity for degradation of such nonstop mRNA, considering the general belief in nonstop decay mechanisms aimed at detecting and destroying mRNAs lacking a stop codon, as seen in bacteria and yeast. In this regard, it is interesting to speculate that while the regulation machinery responsible for the recognition of nonstop mRNA is likely to be present in mammalian cells, it would appear that this machinery can be saturated when this RNA is present in a high enough amount. This is particularly important to consider in the context of an EBOV-infected cell in which this nonstop mRNA would be readily present, and indeed its continuing expression would be expected to quickly saturate the cellular mRNA decay machinery. Such a depletion of the RNA decay machinery by EBOV could represent a novel pathogenicity mechanism, which allows infected cells to replicate high amounts of viral particles and at the same time to survive for a longer period of time postinfection. Such a hypothesis will need to be tested in future experiments.

Given the high cytotoxicity of GP, it is now certain that GP gene editing phenomenon simultaneously provides the basis for synthesis of EBOV surface GP, indispensable for virus replication, and at the same time constitutes the means to downregulate and to control GP synthesis. This control dramatically reduces the levels of full-length GP mRNA synthesis via massive expression of nonstop GPTr mRNA (Figure 2F). The ongoing study of the effect of this GPTr mRNA will continue to shed light on the extreme pathogenicity of EBOV and help in finding novel ways to combat this pathogen.

### Supplementary Data

Supplementary materials are available at The Journal of Infectious Diseases online (http://jid.oxfordjournals.org). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

### Notes

**Acknowledgments.** All experiments involving live Ebola virus were carried out in the L’Institut national de la santé et de la recherche médicale (INSERM) Biosafety Level 4 (BSL-4) laboratory Jean Merieux in Lyon, France.

**Financial support.** This work was supported by INSERM, Agence Nationale de la Recherche (ANR-07-MIME-006-01), and Fondation pour la Recherche Médicale (DFI20091117323), the European Union FP7 project ANTIGONE (278976).

**Potential conflicts of interest.** All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

### References


