Virus Research 144 (2009) 1–7

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Characterization of *Ebolavirus* regulatory genomic regions

Gabriele Neumann[a], Shinji Watanabe[a], Yoshihiro Kawaoka[a,b,c,d,*]

[a] Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin-Madison, Madison, WI 53706, USA
[b] Division of Virology, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
[c] Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Saitama 332-0012, Japan
[d] International Research Center for Infectious Diseases, Institute of Medical Science, University of Tokyo, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

ARTICLE INFO

Article history:
Received 7 January 2009
Received in revised form 9 February 2009
Accepted 11 February 2009
Available online 21 February 2009

Keywords:
Ebolavirus
Transcription
Minireplicon
Nontranslated region (NTR)
Intergenic region (IGR)

ABSTRACT

For filoviruses, such as *Ebolavirus* and the closely related *Marburgvirus*, transcriptional regulation is poorly understood. The open reading frames (ORFs) that encode the viral proteins are separated by regulatory regions composed of the 3′ nontranslated region (NTR) of the upstream gene, highly conserved transcription stop and start signals, and the 5′NTR of the downstream gene. The conserved transcription stop and start signals either overlap, or they are separated by intergenic regions (IGRs) of different lengths. To assess the contribution of the regulatory regions to transcription, we established bicistronic minireplicons in which these regions were flanked by upstream and downstream ORFs, the *Ebolavirus* leader and trailer regions, and by T7 RNA polymerase promoter and ribozyme sequences. We found that the individual viral regulatory regions differ in their ability to direct protein synthesis from the upstream or downstream ORFs. Deletion or modification of the NTRs, IGRs, or transcription stop and start signals affected protein expression levels to various extents; for example, 5′NTRs appear to affect efficient protein expression from the downstream ORF, whereas 3′NTRs seem to attenuate protein expression from the upstream ORF. Overall, our data suggest that the regulation of *Ebolavirus* protein levels is complex.

1. Introduction

Ebolavirus and the closely related *Marburgvirus* form the family *Filoviridae* in the order *Mononegavirales* (reviewed in Sanchez et al., 2007). Within the genus *Ebolavirus*, there are four species, *Zaire ebolavirus* (*ZEBOV*), *Sudan ebolavirus*, *Ivory Coast ebolavirus*, and *Reston ebolavirus*. *Ebola-* and *Marburgviruses* cause hemorrhagic fever with mortality rates as high as 80%.

The genome organization and the replication strategy of filoviruses closely resemble those of other nonsegmented, negative-sense RNA viruses (Feldmann et al., 1992, 1993). The genomes of filoviruses consist of a nonsegmented, single-stranded, negative-sense RNA of about 19 kb in length that encodes seven structural proteins in the following order (in the positive-sense orientation): NP (nucleoprotein); VP35 (a component of the replication/transcription complex); VP40 (a matrix protein); GP (glycoprotein; *Ebolaviruses* also encode a secreted glycoprotein from the GP gene); VP30 (a component of the replication/transcription complex); VP24 (a minor matrix protein); L (an RNA-dependent RNA polymerase) (Fig. 1a). Each gene is flanked by conserved transcription start and stop signals (Fig. 1) (Feldmann et al., 1992; Sanchez et al., 1993). At the ends of the genome, there are extragenic regions, referred to as *leader* and *trailer*, which contain promoters for replication and transcription.

Very little is known about the regulation of filoviral gene transcription. Genes located closer to the 3′-end of the genome are transcribed at higher levels than those located further downstream (Feldmann et al., 1992; Muhlberger et al., 1996). This 'transcriptional gradient' likely results from polymerase complexes that, after transcription terminates, fail to reinitiate at the downstream start signal. However, other factors may also contribute to the regulation of transcription, such as the organization of the transcription stop and start signals, or the composition of noncoding and/or intergenic regions.

For ZEBOV, the transcription start signals comprise 12 nucleotides that differ in only one position among the ZEBOV genes (Fig. 1b). Likewise, the transcription stop signals (11–12 nucleotides in length) are highly conserved among ZEBOV genes; the only difference is the length of the uridine stretch (adenosine stretch in the positive-sense orientation), which functions as a polyadenylation signal, and the transcription stop signal of the L gene, which deviates from the consensus sequences at two positions (Fig. 1b).

The transcription stop and start signals of nonsegmented, negative-sense RNA viruses are separated by intergenic regions (IGRs). For *Ebolavirus*, the IGRs differ in length; for three gene

* Corresponding author at: Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin-Madison, 2015 Linden Drive, Madison, WI 53706, USA. Tel.: +1 608 265 4925; fax: +1 608 265 5622.
E-mail addresses: kawaokay@svm.vetmed.wisc.edu, VirusResearch@svm.vetmed.wisc.edu (Y. Kawaoka).

0168-1702/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.virusres.2009.02.005

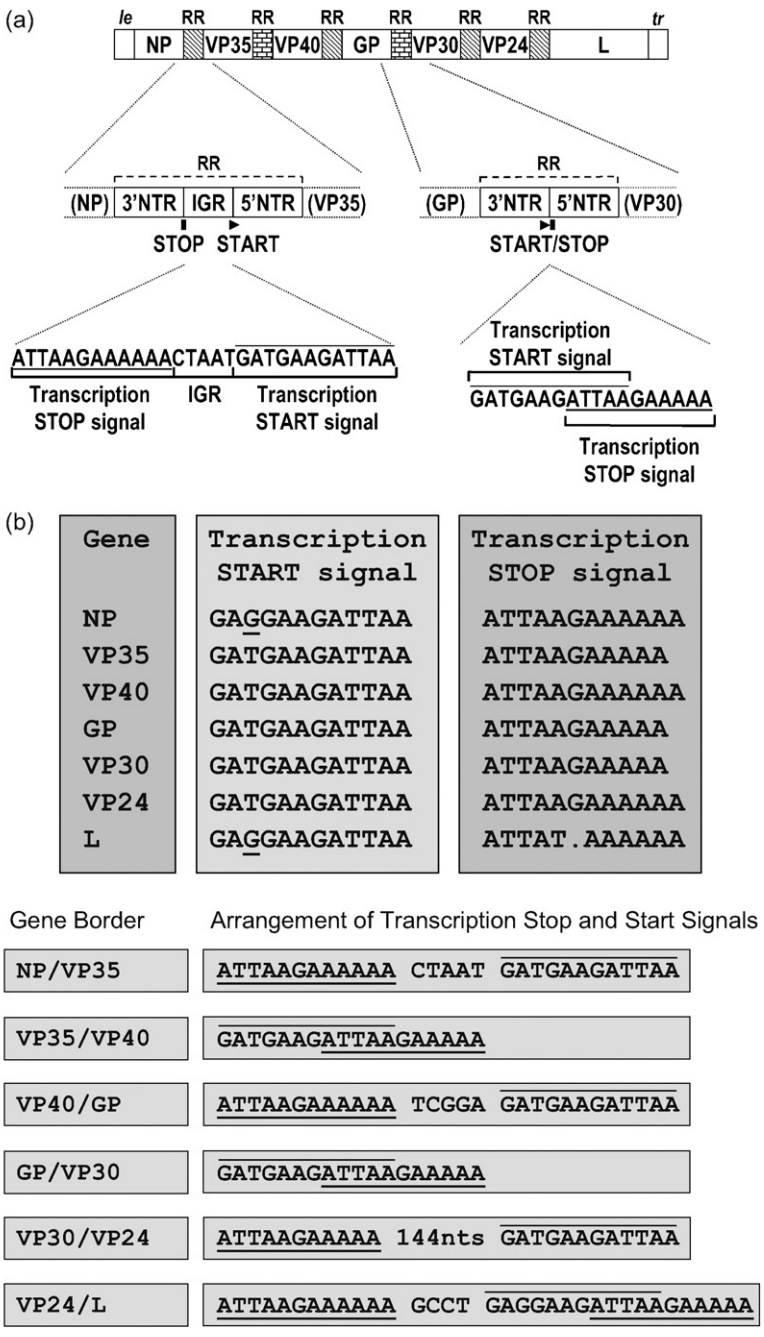

Fig. 1. *Ebolavirus* genome organization. (a) *Ebolavirus* genome organization in the positive-sense orientation. The open reading frames (ORFs) are flanked by leader (*le*) and trailer (*tr*) sequences and are separated by regulatory regions. The RRs are composed of 3′ nontranslated region (NTRs), highly conserved transcription stop and start signals, and 5′NTRs. The transcription stop and start signals overlap or are separated by intergenic regions (IGRs). RRs with nonoverlapping or overlapping transcription start/stop signals are shown in diagonal bars or brick bars, respectively. Transcription start and stop signals are indicated by lines above or below the nucleotide sequence, respectively. (b) Comparison of transcription stop and start signals. Nucleotides that deviate from the consensus sequences are underlined (for transcription start signals) or nonunderlined (for transcription stop signals). (c) Arrangements of transcription stop and start signals for the *Ebolavirus* gene borders. For the NP/VP35, VP40/GP, VP30/VP24, and VP24/L regulatory regions, the transcription stop and start signals are separated by IGRs. By contrast, the VP35/VP40 and GP/VP30 transcription stop and start signals overlap. The VP24/L regulatory region contains two transcription stop signals, the second of which overlaps with the transcription start signal. Transcription start and stop signals are indicated by lines above or below the nucleotide sequence. All sequences shown are in the positive-sense orientation.

boundaries, VP35-VP40, GP-VP30, and VP24-L, the transcription stop and start signals overlap (Fig. 1a and c) (Sanchez et al., 1993). In this case, the transcription start signal of the downstream gene precedes the transcription stop signal of the upstream gene; i.e., after terminating transcription, the polymerase has to track back to initiate transcription of the downstream gene. The overlap of the two signals is brought about by a conserved pentanucleotide ATTAA that is shared by transcription stop and start signals (Fig. 1). Overlapping transcription stop and start signals also exist for *Marburgviruses*, albeit for different genes. The biological significance of filoviral IGRs and these gene overlaps is not known.

The filoviral genes contain gene-specific 5′- and 3′ nontranslated regions (NTRs; Fig. 1a) that are relatively long compared to those of other nonsegmented, negative-sense RNA viruses (Feldmann et al., 1992; Muhlberger et al., 1992; Sanchez et al., 1989, 1992, 1993). Computer analysis has revealed potential stem–loop structures in