Virology 516 (2018) 176–188

Contents lists available at ScienceDirect

Virology

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

Behind the scenes of HIV-1 replication: Alternative splicing as the dependency factor on the quiet

Helene Sertznig[a], Frank Hillebrand[b], Steffen Erkelenz[c], Heiner Schaal[b], Marek Widera[a,*]

[a] Institute for Virology, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
[b] Institute of Virology, Heinrich Heine University, University Hospital, Düsseldorf, Germany
[c] Institute for Genetics, Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Germany

ARTICLE INFO

Keywords:
HIV-1 infection
Alternative pre-mRNA splicing
Splicing regulatory elements
Exon definition
HIV-1 dependency factors (HDF), restriction factors, novel antiretroviral therapy

ABSTRACT

Alternative splicing plays a key role in the HIV-1 life cycle and is essential to maintain an equilibrium of mRNAs that encode viral proteins and polyprotein-isoforms. In particular, since all early HIV-1 proteins are expressed from spliced intronless and late enzymatic and structural proteins from intron containing, i.e. splicing repressed viral mRNAs, cellular splicing factors and splicing regulatory proteins are crucial for the replication capacity. In this review, we will describe the complex network of *cis*-acting splicing regulatory elements (SREs), which are mainly localized in the neighbourhoods of all HIV-1 splice sites and warrant the proper ratio of individual transcript isoforms. Since SREs represent binding sites for *trans*-acting cellular splicing factors interacting with the cellular spliceosomal apparatus we will review the current knowledge of interactions between viral RNA and cellular proteins as well as their impact on viral replication. Finally, we will discuss potential therapeutic approaches targeting HIV-1 alternative splicing.

1. HIV-1 exploits the cellular splicing machinery

About 95% of human genes containing multiple exons are alternatively spliced (Sharp and Hahn, 2011; Wang et al., 2008). Even if not all of the resulting mRNAs may be translated into proteins (Tress et al., 2017), it is not surprising that eukaryotic ssDNA viruses, dsDNA viruses, ss(-)RNA viruses, as well as reverse-transcribing viruses make extensive use of alternative splicing to fully tap the possibilities of regulating their gene expression. In particular, the *human immunodeficiency virus type 1* (HIV-1) although a relatively young human but formerly very successful primate adapted pathogen utilizes the cellular splicing machinery as a central mode of regulating its gene expression (Kim et al., 1989; Klotman et al., 1991; Mbonye and Karn, 2014; Mohammadi et al., 2013; Purcell and Martin, 1993). During the course of infection, the HIV-1 genome is irreversibly integrated into the host's chromosome creating a genetically distinct functional unit under transcriptional control of the HIV-1 promoter localized in the long terminal repeat (LTR). Once integrated the approx. 9.7 kb long proviral DNA is transcribed as a full-length precursor mRNA (pre-mRNA), which contains the densely packed open reading frames (ORFs) for at least eighteen protein and polyprotein-isoforms (Jager et al., 2012; Karn and Stoltzfus, 2012; Leblanc et al., 2013). Subsequently, the HIV-1 primary transcript undergoes extensive alternative splicing, which preserves an equilibrium of more than 50 mRNA isoforms allowing balanced expression of all viral proteins (Ocwieja et al., 2012; Purcell and Martin, 1993; Schwartz et al., 1990). Based on their size HIV-1 mRNAs are classified into the full-length 9 kb, the intron containing 4 kb and the intron-less 2 kb mRNA classes. Hereby, the 4 kb class is mainly spliced at the major splice donor (D1) to one of the acceptors within the central cluster (see below; (Fig. 1) whereas mRNAs of the 2 kb class are additionally spliced from the central splice donor (D4) to the terminal acceptor (A7). Using next generation sequencing (NGS) based transcriptome analysis an additional minor 1 kb mRNA class was identified (Ocwieja et al., 2012).

2. Splicing as the HIV-1 dependency factor on the quiet

In recent years, siRNA-screens have been performed that identified a multitude of HIV-1 dependency factors (HDF) that, by definition, are essential components for HIV-1 replication, but not lethal to the host cell when their expression is silenced (Brass et al., 2008; Konig et al., 2008; Murali et al., 2011; Zhou et al., 2008; Zhu et al., 2014). Based on large-scale evaluation of these orthologous RNAi-screens that generated a quantitatively integrated network of genes, common HDFs were associated with spliceosome and mRNA splicing related proteins. In particular, splicing factors like HNRNPF (Brass et al., 2008; Zhu et al.,

* Corresponding author.
E-mail address: marek.widera@uni-due.de (M. Widera).

https://doi.org/10.1016/j.virol.2018.01.011
Received 29 November 2017; Received in revised form 10 January 2018; Accepted 11 January 2018

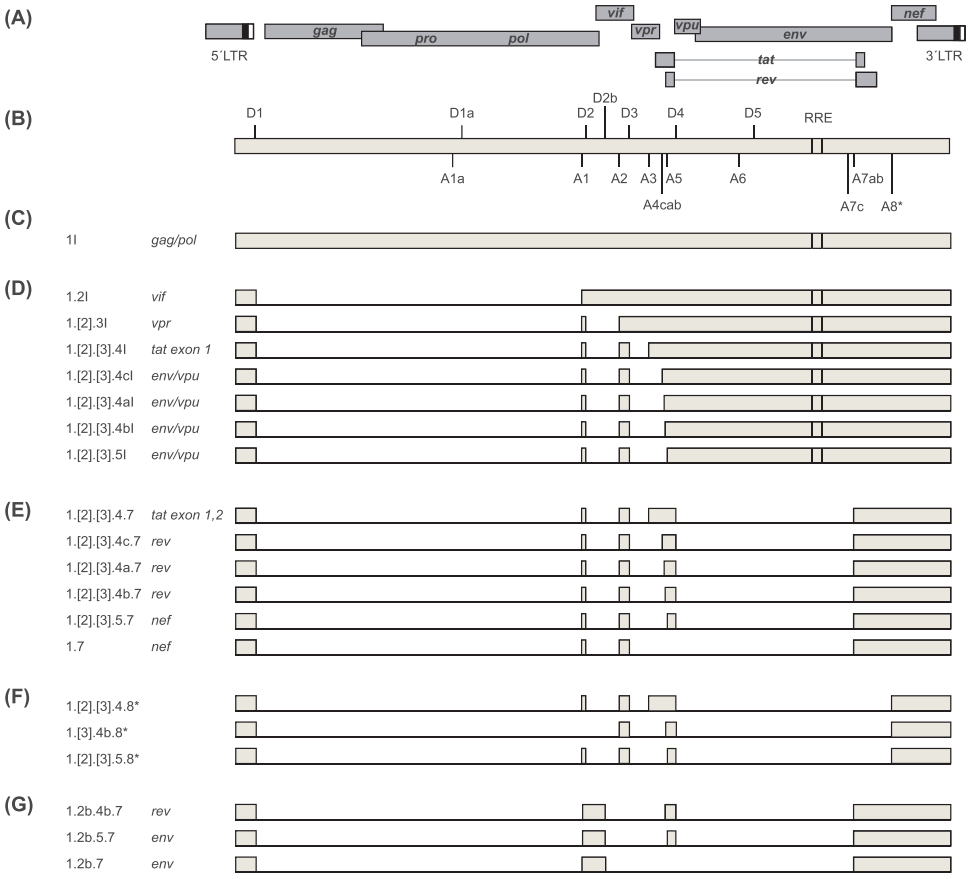

Fig. 1. *Schematic representation of the HIV-1 genome*: (A) The HIV-1 open reading frames (ORF) are shown as grey boxes. The long terminal repeats (LTR) are located at both ends of the HIV-1 genome and are depicted in a three-color-code: U3 (grey), R (black) and U5 (white). (B) Positions of 5′ and 3′ splice sites (ss), their intrinsic strengths are given in Table 1. The Rev-responsive element (RRE) corresponds to the binding site of the viral protein Rev, responsible for the nuclear export of unspliced and intron-containing mRNAs. (C-F) The unspliced 9 kb mRNA is expressed from the 5′ LTR promoter. It serves as genomic RNA or codes for the viral proteins Gag and Pol. Leader exon 1 is non-coding, but exists in all transcript variants. Alternative splicing of non-coding leader exons 2 and 3 and subsequent insertion lead to alternatively spliced mRNAs. The composition of each transcript is explained through the respective nomenclature. Intron-containing transcripts are marked with an "I". Additional splicing occurs for the 2 kb mRNA class (E) between splice sites D4 and A7, which leads to absence of the RRE in these pre-mRNAs. (F) By recognition of the alternative splice acceptor A8, the small 1 kb mRNA class is formed. (G) Transcripts encoding for alternative protein isoforms of Rev4b and Env-gp41 C-terminus (Env-CT) originating from splicing at the alternative D2b.

2014), HNRNPH1 (Konig et al., 2008; Zhu et al., 2014), HNRNPU (Zhou et al., 2008) as well as SRSF2 (Zhou et al., 2008; Zhu et al., 2014) and SRSF6 (Konig et al., 2008; Zhu et al., 2014) were identified as essential HDFs. They interact with pre-mRNA *cis*-regulatory elements and direct the fate of the nascent transcripts and in turn of viral factors essential for replication. Thus, in the following we will refer alternative splicing as the dependency factor on the quiet. The knowledge of HIV-1 alternative splicing and *cis*-regulatory elements has increased massively in recent years and further emphasized its great importance for viral replication determining the fate to fulfil a productive infection (Brillen et al., 2017b; Erkelenz et al., 2015; Widera et al., 2013, 2014). This review summarizes our current knowledge of HIV-1 splicing regulatory elements (SREs), its impact on viral replication, and discusses the possibility to consider these regulators of alternative splicing as a target for anti-retroviral therapy.

3. Splice site bridging as a mechanism of exon and intron definition

The gene architecture in mammals is characterized by multiple short exons of a median size of 123 nt for a middle exon (Scherer, 2008), which are interspersed with long introns. Therefore, the probability of splice site (ss) pairing is more likely to occur over the short exons than over the much longer introns. Hence, exons are expected to represent the first unit of the spliceosomal assembly and initiation-site of the spliceosomal cycle (Berget, 1995; De Conti et al., 2013; Fox-Walsh et al., 2005; Matera and Wang, 2014; Xiao et al., 2007). Exon definition relies on functional cross-exon interactions between the 5′ss (splice donor, D) and the upstream 3′ss (splice acceptor, A) as shown in Fig. 2. They are induced by recognition of the 5′ss by the high molecular complex, the U1 snRNP (Spliceosomal E-complex formation), which in turn activates the upstream 3′ss SA. This occurs by binding of another high molecular complex, the U2 snRNP to the branch point sequence (BPS) resulting in the formation of the exon recognition complex and progression into spliceosomal A-complex formation. Specifically, U1 and U2 snRNP associated components interact with each other and establish a molecular bridge across the exon, whereby the respective splice site pair is connected (De Conti et al., 2013). For intron-excision and exon-joining, the exon spanning molecular bridge must undergo a transition into a state of intron spanning interactions of the splice sites. This implies that the early step of exon-intron definition transits into intron-definition, where an intron-spanning molecular bridge between U1 and U2 snRNP is formed. Consequently, the decision which splice sites are paired together relies on two successive processes, which are recognition and pairing. Here, either step can be modulated by both the intrinsic strength of the 5′ and 3′ss as well as by cellular splicing factors bound to *cis*-acting splicing regulatory elements (SREs) – both together referred to as the "splicing code" (Barash et al., 2010; Wang and Burge, 2008). Since viruses are highly adapted to their host it is not surprising that the gene architecture of HIV-1 has adapted to this cellular regulation mechanism. Thus, HIV-1 splicing architecture and associated SREs will be discussed in the following sections.

4. Splicing of the HIV-1 primary transcript

The HIV-1 genomic 9 kb pre-mRNA is transcribed by the cellular DNA-dependent RNA-Polymerase II resulting in a full-length transcript harbouring eight HIV-1 ORFs, i.e. *gag-pol*, *env*, *vif*, *vpr*, *vpu*, *tat*, *rev*, and *nef* (Karn and Stoltzfus, 2012; Leblanc et al., 2013). Since the initiation of eukaryotic translation begins with the attachment of the 43S ribosomal subunit to the CAP at the 5′ end of the mRNA and then scans the 5′ UTR until an efficient AUG codon is found (Jackson et al., 2010) HIV-1 has to transfer each start codon of its ORFs into close proximity to the 5′ CAP. By hijacking the cellular apparatus, upstream AUGs are spliced out in order to rearrange the downstream ORFs to the 5′ end of the mRNA, thereby extracting the genetic content of the relatively short