International Journal of Biochemistry and Cell Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Biochemistry and Cell Biology

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

A matter of maturity: The impact of pre-mRNA processing in gene expression and antigen presentation

Rodrigo Prado Martins[a,*], Robin Fåhraeus[a,b,c]

[a] Équipe Labellisée Ligue Contre le Cancer, Université Paris 7, INSERM UMR 1162, 27 rue Juliette Dodu, 75010 Paris, France
[b] Department of Medical Biosciences, Umeå University, Umeå, Sweden
[c] RECAMO, Masaryk Memorial Cancer Institute, Brno, Czech Republic

ARTICLE INFO

Keywords:
pre-mRNA processing
Intronless genes
Nuclear translation
Antigen presentation
Viral infections

ABSTRACT

RNA processing plays a pivotal role in the diversification of high eukaryotes transcriptome and proteome. The expression of gene products controlling a variety of cellular and physiological processes depends largely on a complex maturation process undergone by pre-mRNAs to become translation-competent mRNAs. Here we review the different mechanisms involved in the pre-mRNA processing and disclose their impact in the gene regulation process in eukaryotic cells. We describe some viral strategies targeting pre-mRNA processing to control gene expression and host immune response and discuss their relevance as tools for a better understanding of cell biology. Finally, we highlight accumulating evidences toward the occurrence of a translation event coupled to mRNA biogenesis in the nuclear compartment and argue how this is relevant for the production of antigenic peptide substrates for the major histocompatibility complex class I pathway.

1. Introduction

In human, less than 20,000 genes encodes a highly diversified proteome comprising 250,000 to 1 million proteins (de Klerk and t Hoen, 2015). These numbers suggest that the huge diversity of protein and transcripts observed in high eukaryotes rely extensively on gene regulation at the transcriptional, post-transcriptional and translational levels. In fact, pre-mRNA processing plays a relevant role in the diversification of cellular and organismal functions. All pre-mRNAs of protein-coding genes undergo a maturation process, which is regulated in a temporal and cell-specific manner or in response to physiological cues. Diverse cell processing machineries recognize *cis*-acting regulatory elements in pre-mRNA sequences and work in an interdependent way to generate the mature mRNA ready for translation.

In this review, we focus on the different mechanisms involved in the pre-mRNA processing and discuss how they contribute to the modulation of gene expression in eukaryotic cells. We also describe viral strategies targeting the pre-mRNA processing and discuss how these provide insights into processes guiding mRNA translation of full-length proteins in the cytoplasm and the production of antigenic peptide substrates in the nuclear compartment.

2. Pre-mRNA processing

Synthesis and processing of mRNAs are demanding tasks subjected to tight regulation in eukaryotic cells. Signalling pathways and feedback controls ensure a well-balanced expression of proteins under changing cellular conditions. Although transcription is an essential first

Abbreviations: APA, alternative polyadenylation; BiFC, bimolecular fluorescence complementation; CF, cleavage factor; CJE, c-Jun gene's enhancer; CPSF, cleavage and polyadenylation specificity factor; CRM1, chromosome region maintenance 1 protein homolog; CstF, cleavage stimulation factor; CTD, RNA pol II c-terminal domain; CTE, constitutive transport element; DENV, dengue virus; DR, direct repeats; DSE, downstream sequence element; EBV, Epstein-Barr virus; EJC, exon junction complex; EDA, exon-defined A complex; EDE, exon-defined early complex; ESE, exonic splicing enhancer; ESS, exonic splicing silencer; FG-nucleoporins, nucleoporins rich in phenylalanine-glycine repeats; FIV, feline immunodeficiency virus; GANP, germinal centre-associated protein; HBSP, hepatitis B splice-generated protein; HBV, hepatitis B virus; HCMV, human cytomegalovirus; HIV, human immunodeficiency virus; hnRNP, heterogeneous nuclear ribonucleoprotein; HPV, human papilloma virus; HSV, herpes simplex virus; HTLV-1, human T-cell leukemia virus 1; ISE, intronic splicing enhancer; ISS, intronic splicing silencer; KSHV, Kaposi's sarcoma-associated herpesvirus; MHC, major histocompatibility complex; MLV, murine leukemia virus; MPMV, Mason-Pfizer monkey virus; mRNP, messenger ribonucleoprotein; NES, nuclear export signal; NMD, nonsense mediated decay; NPC, nuclear pore complex; ORF, open reading frame; p15/NXT1, NTF2-related export protein 1; PABP, poly(A) binding protein; PAP, poly(a)polymerase; PAS, poly(A) signal; PLA, proximity ligation assay; pRNA, 5′-monophosphorylated RNA; PTC, premature termination codon; RBP, RNA binding protein; RNP, ribonucleoprotein; RPM, ribopuromycylation method; RRE, Rev-responsive elements; RSV, Rous sarcoma vírus; SF1, splicing factor 1; snRNP, small ribonucleoprotein particle; SR, serine-arginine-rich; SRE, splicing regulatory element; SRV, simian retrovírus; SS, splice site; Tap/NXF1, nuclear RNA export factor 1; TREX, transcription-export complex; UTR, untranslated region; VZV, varicella-zoster virus

* Corresponding author.
E-mail address: rodrigo.prado-martins@inserm.fr (R.P. Martins).

http://dx.doi.org/10.1016/j.biocel.2017.05.023
Received 10 March 2017; Received in revised form 15 May 2017; Accepted 18 May 2017

step, proper gene regulation also requires a complex pre-mRNA maturation process including mRNA capping, splicing and polyadenylation. Finally, mature mRNA is exported to cytoplasm, where it is committed to translation. The different steps in this process and how they are regulated on the level of individual mRNAs is becoming better understood, bringing new opportunities for therapeutic intervention aimed at targeting the translation of specific mRNAs or groups of specific mRNAs.

2.1. Pre-mRNA 5′ end capping

The addition of a cap structure (m7GpppN; cap 0) is the first modification made to RNA polymerase II-transcribed RNA. This takes place co-transcriptionally in the nucleus as soon as the nascent transcript is 25–30 nucleotides long (Ramanathan et al., 2016). In addition to its essential role for cap-dependent initiation of protein synthesis, the mRNA cap also functions as a protective group from 5′ to 3′ exonuclease cleavage. Incompletely capped RNAs that lack the N7 methyl moiety, as well as defective mRNAs with premature termination codons, are decapped by a cellular quality-control machinery. In the 5′-to-3′ exonucleolytic pathway, the RNA decapping enzyme Dcp2 removes the cap as m7Gpp, leaving a 5′-monophosphorylated RNA (pRNA) that is progressively degraded by a 5′-to-3′ exonuclease (Xrn1/Rat1) (Schoenberg, 2011).

However, Mukherjee et al. (2012) provided evidence that non-translating mRNAs can accumulate in the cytoplasm in an uncapped state. Contrary to the premise that capping occurs only in the nucleus, they asserted that cytoplasmic capping affects the stability and translation of some mRNAs. Similarly, Ignatochkina et al. (2015) described a decapping/recapping pathway in the cytoplasm of *Trypanosoma* that might function to regulate gene expression in response to stress or environmental changes. These reports raised the hypothesis that the cytoplasmic re-capping system may represent another mechanism employed by the cell for translation regulation.

The mRNA cap is also a unique identifier for recruiting the pre-mRNA splicing, polyadenylation and nuclear export machineries. A heterodimer of CBP80-CBP20 binds the cap in the nucleus and interacts with cell factors coordinating many of the subsequent steps in pre-mRNA processing and mRNA surveillance. In mRNAs exported to the cytoplasm, the CBP80-CBP20 heterodimer is replaced by eIF4E for the translation initiation through the eIF4F complex (Ramanathan et al., 2016).

2.2. Constitutive and alternative splicing

Splicing is required for expression of most of high eukaryotes genes. During constitutive splicing, all introns (non-coding regions) are removed from the pre-mRNAs and all the protein-coding exons of a gene are spliced together to obtain a mature mRNA. The major spliceosome is responsible for the splicing of the vast majority of pre-mRNAs. It consists of a macromolecular complex comprising the five small ribonucleoprotein particles (snRNPs), U1, U2, U4, U5 and U6, in addition to many non-snRNP proteins. Some animal species and plants contain a second minor spliceosome, which is composed of distinct but functionally analogous snRNPs (Turunen et al., 2013).

The information needed for the recruitment of the splicing machinery is contained in the intronic sequences themselves. Essential sequence components are the conserved intron/exon junctions both at the 5′ and 3′ splice sites (SSs), the branch site and the polypyrimidine tract. Nuclear pre-mRNA introns are removed by two consecutive transesterification reactions. First, the 2′ hydroxyl of the branch site performs a nucleophilic attack on the 5′ phosphate of the first intronic nucleotide. This results in cleavage at this site and ligation of the 5′ end of the intron to the branch adenosine, forming a lariat structure. In the second transesterification, the 3′ SS is attacked by the 3′ OH group of the 5′ exon, thereby ligating the exons together and releasing the lariat intron (Will and Luhrmann, 2011). The majority of mRNA splicing events is cotranscriptional. However, individual introns within a pre-mRNA can be excised posttranscriptionally while neighbouring introns are excised cotranscriptionally (Han et al., 2011).

In human pre-mRNAs, spliceosome assembly starts with the ATP-independent base-pairing binding of the U1 snRNP to the intron 5′ SS. This interaction is stabilized by members of the serine-arginine-rich (SR) protein family and proteins of the U1 snRNP. In addition, splicing factor 1 (SF1) associates with the mRNA's branch region and promotes the recruitment of U2AF to the polypyrimidine tract and 3′ SS, resulting in the formation of the E complex. Subsequently, the U2 snRNA engages with the branch region in an ATP-dependent way and replaces SF1, leading to the formation of the A complex. Association of the preassembled U4–U5–U6 tri-snRNP with this complex yields the B complex, which upon release of U1 and U4 become the catalytically active B* complex. The active spliceosome then promote the first transesterification step of splicing, generating the C complex. After additional rearrangements in the spliceosomal ribonucleoprotein (RNP), the second catalytic step takes place and the spliceosome dissociates, releasing the mRNA in the form of a messenger ribonucleoprotein (mRNP) and snRNP (U2, U5 and U6) to be recycled for additional rounds of splicing (for details refer to Wahl et al., 2009). The exon junction complex (EJC) is a major component of the mRNP. The EJC subunits (eIF4AIII, Y14 and Magoh) are recruited co-transcriptionally to the activated spliceosome and upon exon ligation, this complex is stably bound to the mRNA 24 nt upstream of exon-exon junctions. EJC acts as a molecular guide, coupling mRNA splicing to the downstream posttranscriptional processes (Woodward et al., 2016). Following splicing, nuclear–cytoplasmic shuttling proteins also integrate the mRNP, remaining bound to mature mRNA as it transits through nuclear pores (Wickramasinghe and Laskey, 2015).

The initial assembly of the spliceosome across an intron appears to be limited to pre-mRNAs containing single or very short introns. In the case of long introns, which are the most common in metazoans, spliceosomal components first assemble across the exon in an exon-defined early (EDE) complex. Following formation of an exon-defined A (EDA) complex, assembly across an intron takes place for the generation of the A complex and subsequent assembly of the mature spliceosome as previously described (Schellenberg et al., 2008).

Whereas some mRNAs are constitutively spliced, over 90% of human transcripts are alternatively spliced in a tissue or developmental-specific manner. Alternative splicing results in differential intron or exon retention or skipping and choice of alternative splice sites to alter exon size. This generates variable forms of mRNA from a single pre-mRNA species, representing an essential gene-regulatory mechanism. Besides, alternative splicing corresponds to a major contributor to the diverse repertoire of transcriptomes and proteomes found in higher eukaryotes (Hallegger et al., 2010).

The rate at which the gene is transcribed as well as the presence of strong exon–intron consensus sequences determine whether a target exon is included or skipped from the mature mRNA (Black, 2003). Additionally, alternative splicing is regulated via the action of numerous RNA binding proteins (RBPs) that mediate their effects by binding splicing regulatory elements (SREs). SREs include exonic and intronic splicing enhancers (ESE and ISE) and silencers (ESS and ISS), which are located within or surrounding alternatively spliced exons. The interplay between antagonistic splicing factors determine whether a target exon is included or skipped from the mature mRNA (Will and Luhrmann, 2011). SR proteins generally enhance splicing reactions by binding to ESEs or ISEs. In contrast to the robust effect in constitutive splicing, SR-mediated alternative splicing regulation is more complex and subtle. Alternative exons generally hold shorter length and weaker 5′ SSs. Therefore, SR proteins induce the inclusion of the alternative exon through increasing the recognition of weak SSs by the splicing machinery (Jeong, 2017). On the other hand, heterogeneous nuclear ribonucleoproteins (hnRNPs) generally bind ESSs or ISSs and display an