Gene 587 (2016) 107–119

Contents lists available at ScienceDirect

Gene

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

Review

Towards understanding pre-mRNA splicing mechanisms and the role of SR proteins

Mahbod Sahebi [a,b], Mohamed M. Hanafi [a,b], Andre J. van Wijnen [c], Parisa Azizi [d], Rambod Abiri [e], Sadegh Ashkani [f,g], Sima Taheri [h]

[a] Laboratory of Plantation Crops, Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[b] Department of Land Management, Faculty of Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[c] Mayo Clinic, Rochester, MN, USA
[d] Laboratory of Food Crops, Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[e] Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, UPM, 43400 Serdang, Selangor, Malaysia
[f] Laboratory of Food Crops, Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[g] Department of Agronomy and Plant Breeding, Shahr-e-Rey Branch, Islamic Azad University, Tehran, Iran
[h] Department of Crop Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 25 April 2016
Accepted 30 April 2016
Available online 3 May 2016

Keywords:
Pre-mRNA splicing
Post-transcriptional splicing
Co-transcriptional splicing
Post-transcriptional splicing
SR proteins

ABSTRACT

Alternative pre-mRNA splicing provides a source of vast protein diversity by removing non-coding sequences (introns) and accurately linking different exonic regions in the correct reading frame. The regulation of alternative splicing is essential for various cellular functions in both pathological and physiological conditions. In eukaryotic cells, this process is commonly used to increase proteomic diversity and to control gene expression either co- or post-transcriptionally. Alternative splicing occurs within a megadalton-sized, multi-component machine consisting of RNA and proteins; during the splicing process, this complex undergoes dynamic changes *via* RNA–RNA, protein–protein and RNA–protein interactions. Co-transcriptional splicing functionally integrates the transcriptional machinery, thereby enabling the two processes to influence one another, whereas post-transcriptional splicing facilitates the coupling of RNA splicing with post-splicing events. This review addresses the structural aspects of spliceosomes and the mechanistic implications of their stepwise assembly on the regulation of pre-mRNA splicing. Moreover, the role of phosphorylation-based, signal-induced changes in the regulation of the splicing process is demonstrated.

Contents

Abbreviations: snRNA, small nuclear RNA; snRNPs, small nuclear ribonucleic proteins; hnRNA, heterogeneous nuclear RNA; PPT, polypyrimidine tract; BPS, branch point sequence; DYRKs, dual-specificity tyrosine-phosphorylation-regulated kinases; CRKRS, CDC2-related kinase with RS domain-7 (also known as CRK7); PTB, polypyrimidine tract-binding protein; CTD, carboxy-terminal domain; ISEs, intronic splicing enhancers; ESEs, exonic splicing enhancers; ISSs, intronic splicing silencers; ESSs, exonic splicing silencers; Pol II, RNA polymerase II; CLIP, crosslinking immunoprecipitation; HTS, high-throughput sequencing; EJC, exon junction complex; NMD, nonsense-mediated decay; PKCI-1, protein kinase C-interacting protein; UTR, untranslated region; aORF, alternative open reading frame; pORF, downstream primary open reading frame; uORF, upstream open reading frame; TIS, translation initiation site; TSSs, transcription start sites; SAGE, serial analysis of gene expression.

E-mail addresses: mahbod_sahebi@yahoo.com (M. Sahebi), mmhanafi@agri.upm.edu.my (M.M. Hanafi).

http://dx.doi.org/10.1016/j.gene.2016.04.057
0378-1119/

1. Introduction

Pre-mRNA splicing plays a crucial role in gene expression in the eukaryotic cell nucleus. For instance, it has been reported recently that pre-mRNA splicing is responsible for regulation of gene expression in temperature signaling (Capovilla et al., 2015) and response to abiotic stress in plants (Cui and Xiong, 2015). Protein-coding genes are disrupted by intervening intronic sequences. The intervening sequences (introns) in most pre-mRNAs are removed from the snRNAs, thereby linking all exon sequences within the same reading frame.

RNA polymerase II transcribes eukaryotic pre-mRNAs and is also responsible for multiple post-transcriptional modifications. Although the majority of constitutive events that occur in the nuclei of higher eukaryotes are precise and efficient, different mRNA isoforms can be produced through alternative mRNA splicing. Hence, the precise removal of introns *via* pre-mRNA splicing is an indispensable step in the regulation of gene expression. The spliceosome is a macromolecular machine consisting of five snRNAs – U1, U2, U4, U5, and U6 – and many different proteins such as small nuclear ribonucleic proteins (snRNPs). This machine catalyzes mRNA splicing in eukaryotic cells by removing the introns and ligating the exons (Wahl et al., 2009).

The canonical stepwise assembly of the spliceosome occurs twice on each molecule of heterogeneous nuclear RNA (hnRNA). Specific sequence elements contained within the 5′ and 3′ splice sites include the polypyrimidine tract (PPT) and the branch point sequence (BPS), which are both involved throughout spliceosome assembly. The 3′ splice site of hnRNA is rich in AG nucleotides and called the polypyrimidine tract (PPT); it contains a variable number of polypyrimidines. This polypyrimidine tract provides a suitable substrate for the recruitment of different factors to the 3′ splice sites and potentially facilitates their recruitment to the branch point sequence (BPS) of the hnRNA. The branch point sequence includes the conserved adenosine that is required to initiate the first splicing step. The canonical mRNA splicing pathway starts with the U1 snRNP binding the 5′ splice site, followed by the U2 snRNP binding the branch point sequence. Subsequently, they bind the U4/U5/U6 tri-snRNP to produce the mature spliceosome, which is required for catalyzing RNA-based sequences (Matlin and Moore, 2008; Smith et al., 2008).

Briefly, two main steps known as SN2-type transesterification reactions are involved in occurring pre-mRNA splicing (Fica et al., 2013). The 2′-OH group of "conserved RNA adenine nucleotide" in BPS of an individual intron attacks at the 5′ end of the intron to the phosphorous of guanine nucleotide (first nucleophilic attack) in the first step of pre-mRNA splicing. This step of pre-mRNA splicing causes to release 5′-exon and form an "intronlariat-3′-exon intermediate". Following the first step, the 3′-OH group of RNA nucleotide on 3′ end of the 5′-exon attacks at the 5′ end of the 3′-exon on the phosphorous of the RNA guanine nucleotide (second nucleophilic attack), resulting in joining two exons and releasing the intron lariat.

The important features of pre-mRNA splicing are related to: i) the enhanced functional diversity of the proteome and ii) the regulation of gene expression by introducing additional regulatory layers; different mRNA isoforms have distinct coding capacities and stabilities (Bindereif, 2015).

The nature of spliceosome assembly steps has been distinctly explained for decades, but it is still enigmatic that how the spliceosome assists these reactions? How is a reacting piece placed into close of one another in a stringent temporal arrangement? With the different lengths of the introns and exons, how the spliceosome hold the 5′-exon in two reaction steps and accommodate pre-mRNA?

Several *in vitro* studies have been performed to understand the chemical mechanisms of pre-mRNA splicing, primarily that of pre-mRNAs poised at two exons and their intervening intron (Brow, 2002; Will and Lührmann, 2011). But the function of the protein components involved in the spliceosome have not understood yet. A growing number of studies have explored the structures of the components and domains of the spliceosome (Stefl et al., 2005; Auweter et al., 2006; Cléry et al., 2008), and the literature describes how alternative mRNA splicing is influenced by extracellular signals through different signal transduction pathways (Stamm, 2002; Lynch, 2004; Shin and Manley, 2004). This review briefly addresses the splicing process, discusses the structure of the pre-mRNA splicing machine, and reviews the effect of extracellular signal transduction pathways on the regulation of alternative mRNA splicing.

2. Transcription initiation *via* alternative promoters

The first layer of control in gene expression is related to transcription initiation during mRNA biogenesis (Sanyal et al., 2012; Consortium, 2014). Different transcripts result from the initiation of transcription at the first exon or within the upstream 5′-untranslated-region (UTR). Alternative first exons result in different transcripts with different open reading frames (ORFs), giving rise to diverse protein isoforms with alternative N-termini (Goossens et al., 2007). Transcripts derived from the same coding region but with different 5′-UTRs may experience differential translational regulation (Barbosa et al., 2013) through translational control of the short upstream ORFs (Calvo et al., 2009; Fritsch et al., 2012), leading to the production of biologically relevant peptides (Magny et al., 2013; Slavoff et al., 2013). Prior to the development of broad transcriptomic approaches, transcription start sites (TSSs) and alternative promoters were identified within protein-coding transcripts using "cap analysis of gene expression" (CAGE) (Auboeuf et al., 2005). Two molecular mechanisms, the chromatin state and the regulation of tissue- and cell-specific transcription factors, are responsible for the selection of TSSs and alternative promoters (de Klerk and AC't Hoen, 2015). A biological understanding of the importance of tissue-specific TSSs requires understanding how a given TSS is selected, as well as which transcription factors and regulatory networks are involved.

Inferred in transcriptional networks facilitate selection of transcription start sites. For instance, an investigation over human monocytic leukemia cells differentiation (Deep CAGE study) (Suzuki et al., 2009), the transcription factor binding sites were predicted around the TSSs in different condition. This survey, subsequently analyzed the motif activity responses and provided a network model of gene expression as well as the main insights of the key regulators for controlling transcription at separate differentiation phases. Similarity, results of another survey over investigation effects of disordering in particular transcription factors (IRF8, SP1, MYB and PU.1) led to finding out target genes related to each transcription factor, and identifying different de novo binding site motifs (Vitezic et al., 2010). It has been shown that the selection of a particular TSS plays important roles during cell differentiation (Pozner et al., 2007) and development (Levanon and Groner, 2004). Aberrations in use of alternative promoter and TSS leads to different diseases such as neuropsychiatric disorders (Tan et al., 2007), developmental disorders, and cancer (Pedersen et al., 2002). Although the epigenetic changes in the promoter region cause some of disorders, others emerged through genetic changes in distal elements affective