Review article

Chemotherapy induces alternative transcription and splicing: Facts and hopes for cancer treatment

Charles A. Lambert*, Nancy Garbacki, Alain C. Colige

Laboratory of Connective Tissues Biology, GIGA, Quartier Hôpital, University of Liège, Avenue Hippocrate 13, 4000 Liège, Belgium

ABSTRACT

Alternative promoter usage, alternative splicing and alternative cleavage/polyadenylation (referred here as to alternative transcription and splicing) are main instruments to diversify the transcriptome from a limited set of genes. There is a good deal of evidence that chemotherapeutic drugs affect these processes, but the therapeutic incidence of these effects is poorly documented. The scope of this study is to review the impact of chemotherapy on alternative transcription and splicing and to discuss potential implications in cancer therapy. A literature survey identified > 2200 events induced by chemotherapeutic drugs. The molecular pathways involved in these regulations are briefly discussed. The GO terms associated with the alternative transcripts are mainly related to cell cycle/division, mRNA processing, DNA repair, macromolecules catabolism and chromatin. A large fraction (43%) of transcripts are also related to the *new hallmarks of cancer*, mostly genetic instability and replicative immortality. Finally, we ask the question of the impact of alternative transcription and splicing on drug efficacy and of the possible curative benefit of combining chemotherapy and pharmaceutical regulation of this process.

1. Introduction

Chemotherapy is widely used in cancer treatment and is the gold standard for hematologic tumors and for the killing of metastatic cells. In spite of the huge arsenal of chemotherapeutic drugs, resistances and relapses remain very frequent, advocating for the development of new drugs and strategies.

While human cells contain nearly 20,000 protein-coding genes, their transcriptome is tremendously more complex, with > 80,000 distinct mRNA sequences indexed on the latest GENCODE release (release 26, March 2017, http://www.gencodegenes.org/stats/current.html). New original mRNA sequences regularly broaden the list, and the actual number of known transcripts likely is only the tip of the iceberg. This huge number of transcripts is mainly due to three processes: initiation and/or termination of transcription at alternative 5′ and 3′ positions, respectively, and differential inclusion/exclusion of internal sequences into the mature mRNA (alternative splicing, AS). While these three processes are frequently considered together under the term "AS" we deem that they merit to be grouped under a mechanistically more relevant acronym and we will use *alternative transcription and splicing* (ALTS) in the rest of this article. ALTS concerns more than 90% of genes (Table 1 and references therein). While some

variants present in minute amounts in the cells likely belong to a *noise* stemming on weakness of the transcription and splicing machineries, many are the fruit of finely tuned regulations. Read-through and trans-splicing (Jividen and Li, 2014) and non ALTS processes, such as RNA editing, methylation and pseudouridilation (Licht and Jantsch, 2016), also enable transcriptome diversity but are not covered here.

This review summarizes the current knowledge on ALTS modulation by chemotherapy, discusses selected examples and addresses the question "Could combined chemotherapy and pharmacological modulation of ALTS improve cancer treatment?".

2. Expanding the transcriptome

2.1. Alternative promoters

The promoter of a gene is the sequence dictating the location of the transcription start site by recruiting the transcriptional machinery including (for mRNA, the only class of RNA considered here) RNA polymerase II (PolII), transcription factors and associated molecules. Evidence mainly obtained during the two last decades indicates that a large proportion of genes contain two or more alternative promoters, resulting in the expression of transcripts varying in their 5′ termini

Table 1
ALTS events frequency.

ALTS event	% affected genes/transcripts
Alt promoter	18–64
Alt termination	50–82
Cassette exon	~40
Alt 5′splice site	5–8
Alt 3′splice site	8–13
Intron retention	3–7
Mutually exclusive exons	~1
All events	> 90

From Gamazon and Stranger (2014); Gruber et al. (2014); Landry et al. (2003); Pal et al. (2012); Tian and Manley (2013).

(Davuluri et al., 2008; de Klerk and 't Hoen, 2015; Landry et al., 2003; Pal et al., 2012). This process is tightly regulated in a cell/tissue and developmental manner, dependent on three molecular mechanisms: alteration of chromatin state, abundance and activity of specific transcription factors and methylation status of the DNA. Its functional significance is highly variable (Fig. 1). In many cases (60–80%), a common AUG translation initiation site is used so that the alternative sequence does not result in any change in the protein sequence. Alternative 5′- untranslated regions (UTR) may however govern translation efficiencies due to sequence variation and occurrence of secondary structures. A paramount example is provided by the mouse p18 (INK4c) gene whose transcription is controlled by two promoters. During C2C12 myoblast differentiation there is a progressive shift from the upstream to the downstream promoter, resulting in the shortening of the 5′UTR. This is accompanied by a 50-fold increase in the amount of protein in the cells even though the mRNA level is unchanged (Phelps et al., 1998). In other circumstances, alternative promoter usage results in initiation of translation at alternative AUG sites without modification of the reading frame, generating proteins with elongated, truncated or alternative N-termini. Popular examples are the TP63 and TP73 genes. Alternative promoters downstream of the conventional one cause the expression of N-terminally truncated proteins (ΔNp63 and ΔNp73)
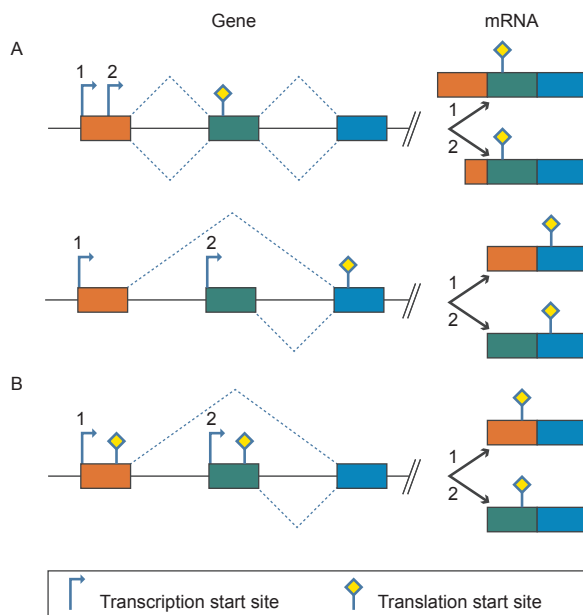
functioning as dominant negative regulators of their classical counterparts and of p53. In other situations, the consequences of these differential N-termini of the proteins are more subtle or not (yet) established. Finally, in rare and extreme situations, use of alternative promoters changes the reading frame of the mRNA by the translation machinery, leading to production of proteins with unrelated sequences and functions. This is exemplified by the CDKN2A gene which contains two alternative first exons followed by a common second exon 2. Since these two mRNA variants are read with different frames, this results in translation of two functionally different proteins: INK4, an inhibitor of CDKs, and ARF, an inhibitor of MDM2 (Quelle et al., 1995; Robertson and Jones, 1999).

2.2. Alternative splicing

Introns (intragenic regions) are sequences that are eliminated (spliced out) by the spliceosome complex during the maturation of most pre-mRNAs into mRNAs. They belong to two classes according to the involvement of the major (accounting > 99% of all introns) or minor spliceosomes in their splicing. Only the major class will be considered here.

Definition of exon/intron junctions involves precise selection of splicing sites on the primary transcripts (De Conti et al., 2013). Introns are characterized by 5′-GU and 3′-AG sequences, a so-called *branch point site* (YNYURAY) with a mandatory A residue, and usually a pyrimidine-rich stretch (Y_n, n between 10 and 40). Other regulatory sequences, the intronic and exonic splicing enhancers and silencers, are located in introns or exons. These sequences are usually highly degenerate, and the exact definition of a splicing code enabling the prediction of exon/intron boundaries from the sequence of primary transcripts remains elusive. After recognition of the splicing sites by the spliceosome, two steps of transesterification eliminate introns and rejoin exons (Fig. 2) (Krämer, 1996 for a review).

The major spliceosome is among the most complex of cell machineries. It is a dynamic assembly implicating hundreds of proteins and five small nuclear ribonucleoproteins (U1, U2, U4, U5 and U6 snRNP) enabling the recognition of complimentary sequences on the primary transcripts. The different catalytic processes require ATP as a source of energy. Precise spliceosome assembly and function have been described
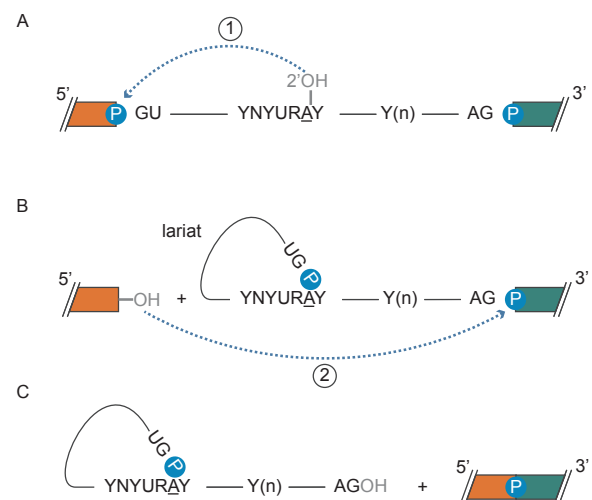

Fig. 1. Alternative promoters expand the transcriptome.A) Alternative transcription start sites 1 and 2 are upstream of the translation start site. The 5′-UTRs of the mRNA variants differ but only a single protein is translated. B) Alternative transcription start sites lead to the use of different translation start sites, and therefore to the synthesis of different proteins. If the translation start sites are in frame, the proteins differ by the N-terminal end only. If not, the proteins produced are highly divergent. Exons appear as boxes and introns as lines.

Fig. 2. Transesterification reactions during splicing.A) The adenine residue of the branch point site engages a nucleophilic attack (1), breaking the 3′-5′ phosphodiester bond between the guanine and the adjacent residue on the upstream exon. B) This creates a 2′-5′ phosphodiester bond forming a lariat within the intron. The hydroxyl group of the upstream exon commits a second nucleophilic attack on the first residue of the downstream exon (2). C) This generates a new 3′-5′ bond between the two exons and the release of the intronic sequence. Exons appear as boxes and introns as lines. The spliceosome is not illustrated. (Adapted from Krämer, 1996).