FEBS Letters 584 (2010) 4401–4407

journal homepage: www.FEBSLetters.org

Regulation of splicing enhancer activities by RNA secondary structures

Wei Liu [a], Yu Zhou [a,b,c], Zexi Hu [a], Tao Sun [a], Alain Denise [c,d,e], Xiang-Dong Fu [a,b,*], Yi Zhang [a,f,*]

[a] State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan, Hubei 430072, China
[b] Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093-0651, USA
[c] Univ. Paris-Sud, LRI UMR8623 and IGM UMR8621, Orsay F-91405, France
[d] CNRS, Orsay F-91405, France
[e] INRIA, Saclay, F-91400, France
[f] ABLife Inc, Building E-1301, 115 Guangba Road, Wuhan, Hubei 430072, China

ARTICLE INFO

Article history:
Received 28 August 2010
Accepted 28 August 2010
Available online 1 October 2010

Edited by Michael Ibba

Keywords:
Alternative splicing
Secondary structure
ESE

ABSTRACT

In this report, we studied the effect of RNA structures on the activity of exonic splicing enhancers on the SMN1 minigene model by engineering known ESEs into different positions of stable hairpins. We found that as short as 7-bp stem is sufficient to abolish the enhancer activity. When placing ESEs in the loop region, AG-rich ESEs are fully active, but a UCG-rich ESE is not because of additional structural constraints. ESEs placed adjacent to the 3′ end of the hairpin structure display high enhancer activity, regardless of their sequence identities. These rules explain the suppression of multiple ESEs by point mutations that result in a stable RNA structure, and provide an additional mechanism for the C6T mutation in SMN2.

1. Introduction

Mammalian genes are commonly interrupted by multiple introns, and the spliceosome-catalyzed intron removal and exon ligation, termed pre-mRNA splicing, represents an important layer of gene regulation [1]. Deep sequencing analysis of the transcriptomes in 15 diverse human tissues and cell lines reveals that 92–94% of human genes are alternatively spliced [2]. Alternative splicing of many mammalian genes are regulated during cell differentiation, proliferation and apoptosis, development, or in response to environmental stimuli [3–6].

Alternative splicing of pre-mRNAs is controlled by precise and highly dynamic interactions between trans-acting regulatory proteins and their corresponding cis-acting splicing regulatory elements (SREs) residing in exons and/or introns. Analysis of such interactions has recently been extended to the genomic level by coupling immunoprecipitation of splicing factor-bound RNA with deep sequencing, termed CLIP-seq or HITS-seq [7].

Although the importance of RNA structures in splicing regulation has been increasingly appreciated [8,11], the prevalence of this layer of regulation is further underscored by a recent genome-wide study suggesting that about 15% of alternative splicing events in the human genome may subject to influence by RNA secondary structures [10]. Most of previous studies focused on the impact of RNA secondary structures on modulating the accessibility of the 5′ and 3′ splice sites to U1 and U2 snRNPs, respectively, or the recognition of the polypyrimidine tract and 3′ splice site by U2AF65-U2AF35 heterodimers. RNA structures that cause sequestration of these essential splicing signals inhibit splicing of associated exons, while other RNA structures that bridging long-distance RNA–RNA interactions in the intron promote splicing by enhancing splice site pairing (lately reviewed by Warf and Berglund [11]).

Besides essential splicing signals (5′ and 3′ splice sites, the pytract and A branch point), exons and introns that are subjected to alternative choices are often enriched with positive and negative SREs, which can be experimentally identified by mutational analysis or predicted by computational algorithms based on existing experimental evidence [7,12,13]. These SREs can be divided into exonic and intronic enhancers (ESE and ISE) and exonic and intronic silencers (ESS and ISS). These regulatory sequences are central to decipher the "splicing code" governing the splicing outcome of most higher eukaryotic genes, which have received considerable attention in the last decade [14–20]. However, although it is conceivable that RNA secondary structures may modulate the accessibility of regulatory proteins to their corresponding exonic and intronic SREs, the progress has been slow. One of the best-understood cases is the regulated accessibility of ESE for SF2/ASF in the EDA exon of the fibronectin pre-mRNA by a fold-back RNA structure, which enhances splicing by presenting the ESE in a

* Corresponding authors. Addresses: ABLife Inc, Building E-1301, 115 Guangba Road, Wuhan, Hubei 430072, China. Fax: +86 27 68754945 (Y. Zhang); Department of Cellular and Molecular Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-5004, USA. Fax: +1 858 822 6920 (X.-D. Fu).
E-mail addresses: xdfu@ucsd.edu (X.-D. Fu), yizhang101@hotmail.com (Y. Zhang).

0014-5793/$36.00
doi:10.1016/j.febslet.2010.09.039

single-stranded loop [21,22]. Computational analysis indicates that some known SREs appear to be more single-stranded than control sequences, and consistently, exonic SREs engineered in the loop of a stable hairpin structure display much higher splicing enhancing activities than those in the stem [23].

Here, we further characterized how local RNA structures regulate the activity of ESEs by engineering ESEs into the basepaired stems with varying thermostability, into the loops of varying sizes, or into the single-stranded region associated with a stem to different degrees. Interestingly, in contrast to GA-rich ESEs, an UCG-rich ESE compromises most of its enhancer activity in the loop region. ESEs inserted immediately downstream of the hairpin stem are more active in enhancing the inclusion of the alternative exon. We also present evidence that the enhancer strength of multiple ESEs in an alternative exon is predictable when the effect of local RNA structures is taken into account.

2. Materials and methods

2.1. Plasmid constructs and mutagenesis

The minigene *SMN1* and *SMN2* containing exon 7 (54-bp) and portions of the adjacent intron 6 and 7 from PCI-*SMN1* and PCI-*SMN2* (gifts from Dr. J. Zhou, Arizona State University) was cloned into the GFP coding sequence of pZW8 (a gift from Dr. C. B. Burge, Massachusetts Institute of Technology) to replace the original *SIRT1* minigene, and therefore yielded pZW9-*SMN1* and pZW9-*SMN2*. To insert the desired ESE sequences and hairpin structures, two restriction enzyme sites *Xho*I and *Eco*RI were cloned into exon 7 of pZW9-*SMN1* to replace the Tra2β-ESE sequence (5′ AAAAGAAGGAAGGTG) to create pZW9-*SMN1*-m used in this study.

To make 5D, 5D-L, 5D-a, ASF-L and Tra2β-L series of constructs, sense and antisense oligos containing a full or truncated *Xho*I and *Eco*RI restriction sites were synthesized (SBS Genetech, Beijing) and diluted in the linker buffer (50 mM Tris–HCl pH 8.0, 100 mM NaCl, 1 mM EDTA) at a final concentration of 100 nM. Some oligos contained truncated restriction sites to avoid the splicing-induced frameshift and the consequent non-sense mediated mRNA decay. Each pair of sense and antisense oligos were mixed for annealing: denaturation at 95 °C for 2 min, annealing at 52 °C for 10 min, then chilled on ice. The product diluted to a concentration of 10 nM was then ligated into the *Xho*I and *Eco*RI-digested pZW9-*SMN1*-m.

ASF-a and Tra2β-a constructs were made according to one-step site-directed mutagenesis as previously described [24], using pZW9-*SMN1* as the template and high-fidelity DNA polymerase KOD PLUS (TOYOBO).

For all the construct designs, the unintended ESEs were avoided according to a computational method (manuscript in preparation). The corresponding oligo sequences are listed in the Supplementary data.

2.2. Cell culture and transfection

HeLa (human cervical carcinoma) cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum. Cells were planted on the 24-well plate with cell concentration of 8×10^4 per well. Cells were transfected when at 60% confluence. The transfection followed standard protocol of Lipofectamine-2000, and 500 ng plasmid was used per well. The cell were kept 24 h after transfected, then harvested for total RNA preparation.

2.3. RNA preparation and RT-PCR analysis

Total RNA was extracted from transfected cells by a one-step extraction method with trizol (Invitrogen). A standard reverse transcription (RT) protocol supplied by Promega (whose MMLV was utilized) was used with oligo$(T)_{18}$, followed by a PCR amplification with the products being separated on 2% agarose gels. The gel images were captured by GeneGenius Gel Imaging System (SynGene) and quantified by the associated GeneTools analysis software.

3. Results

Pre-mRNA folding is limited to a region of about 50-nt downstream of the transcribing polymerase [25], and thus secondary structures of pre-mRNA tend to form locally whereas those require long-range interactions is disfavored [26]. Simple fold-back hairpin structures involving only short-range local basepairing are prevalent in pre-mRNAs, which can be predicted by a number of available algorithms [27]. Although accurate prediction of complex RNA structures remains a main challenge in the field, the local hairpin structures formed within 50-nt of a pre-mRNA can be reliably predicted by minimal free energy algorithms.

3.1. Stems as short as 7-bp block the ESE function

In order to assess the effect of this simple class of RNA secondary structures on the enhancer activity of ESEs, we constructed a *SMN1* (survival of motor neuron 1, telomeric copy) based splicing reporter in which two restriction sites were introduced adjacent to the ESE sequence for the splicing activator Tra2β (Tra2β-ESE) located in the alternative exon 7 (Fig. 1A). This allows insertion of various ESE-containing structures to replace the Tra2β-ESE (Fig. 1A and B). As for formation of the local alternative secondary structures is known to compete with each other and influence their functions [28], we specifically designed various hairpin structures to exclude the possibility of forming competitive alternative secondary structures other than the desired structure within the 50-nt folding window.

Most well-characterized ESEs are AG-rich. However, an unbiased computational analysis revealed that some of the predicted ESEs are CT-rich, as exemplified by Group 5D with a TCGTCG consensus [12]. Fig. 1 shows that a heptamer sequence containing the consensus has a strong ESE activity comparable to that of the Tra2β-ESE, validating the biological importance of this group of ESEs (Fig. 1B and C, Construct 5D). However, the enhancer activity is completely abolished when placing the ESE in the stem region of hairpin structures containing 13–7 contiguous basepairs (Fig. 1B and D, all other constructs except for 5D). This result supports the previous observation that the single-strandedness of SREs is important for efficient recognition by trans-acting splicing factors [23].

3.2. Structural constraints of the hairpin loop interfere with ESE activity

The lost splicing enhancer activity of the ESE-5D series in Fig. 1 is likely due to the hairpin structure that blocks the ESE function by base-pairing. In order to provide further support to this conclusion, we placed two known ESEs, a strong Tra2β-ESE and a weaker ASF-ESE from EPB41 [29], in the loop of a stable hairpin structure (Fig. 2A). We observed either comparable (all Tra2β-ESE constructs and ASF-ESE construct L-3), or higher (ASF-ESE constructs L-0 and L-7) enhancer activities than the parental construct containing a structure-free single-stranded ESE (Lane "SMN1" for Tra2β-ESE and lane "ASF linear" for ASF-ESE) (Fig. 2D and C). This result demonstrated that a simple stable hairpin alone in the exon is not sufficient to have a negative impact on splicing.