Review Article
Structure and function of LIM domain proteins are regulated by alternative splicing

Quan-Hui Zheng¹, Ai-Hong Zhang², Yong Zhao³

¹Hebei Key Laboratory for Chronic Diseases, Tangshan Key Laboratory for Preclinical and Basic Research on Chronic Diseases, School of Basic Medical Sciences, North China University of Science and Technology, Tangshan, China; ²Department of Intensive Care Unit, ³Institute of Zoology, Chinese Academy of Sciences, Beijing, China

Received February 2, 2016; Accepted May 2, 2016; Epub July 15, 2016; Published July 30, 2016

Abstract: LIM domain proteins display various functions including gene regulation, cell fate determination, tumour formation and cytoskeleton organization. A diverse group of LIM domain proteins have been identified, and alternative RNA splicing combined with alternative transcription/translation initiation or termination sites is one of the key molecular mechanisms contributing to the proteome and functional diversity of LIM domain proteins. However, to date, few of the alternative spliced LIM domain proteins has been fully clarified. Further studies on functional significance of the alternative spliced LIM domain proteins and their relevance to the full length isoforms should be done in a cell and tissue specific manner.

Keywords: LIM domain protein, alternative splicing, isoforms

Introduction

The Lin-11, Isl-1 and Mec-3 (LIM) domain is one of the protein-protein interaction motifs that are 50-60 amino acids in size and share two characteristic zinc fingers. Based on the arrangement of LIM domains and the overall structure of proteins, LIM domain proteins can be roughly classified into three groups. The first group consists of LIM homeodomain proteins (LHXs), which feature two tandem LIM domains in their N-terminal and a DNA binding homeodomain in the C-terminal. The second group is LIM domain only protein (LMO), which consists of two or more LIM domains clustered at their N- or C-terminal. The third group includes most of the LIM domain proteins that contain various other protein-protein interaction motifs such as postsynaptic density-95, discs large, zona occludens-1 (PDZ), leucine-aspartate repeat (LD), prickle, espinas and testin (PET), and actin binding domain (ATD), besides the characterized LIM domains [1, 2]. The conditional expression of LIM domain proteins are controlled by different regulatory mechanisms. However, one of the most remarkable observations stemming from the sequencing of LIM domain proteins is that most of these genes are regulated by alternative splicing (AS) and encode different protein isoforms (Table 1). In the present review, we focus on the extensive AS events of LIM domain proteins and their effects on the structural and functional diversity of LIM domain proteins. In addition, as alternative transcription/translation initiation or termination sites are of equal importance to the diversity of LIM domain proteins, which were also included in the manuscript.

Alternative splicing of LIM homeodomain proteins (LHXs)

Isl-1 is one of the originally isolated LHX proteins by virtue of its ability to regulate insulin gene expression and/or islet cell development. The primary mRNA of Isl-1 contains six exons and generates two transcripts due to the alternative 3’splicing in exon5 of Isl-1 [3]. The longer transcript of Isl-1 (isl1-α) encodes a protein of 349 amino acids (aa) and expresses in normal islet cells. Mice deficient in isl1-α fail to form the dorsal exocrine pancreas [4, 5], although
Alternative splicing of LIM domain proteins

<table>
<thead>
<tr>
<th>Group/subgroups</th>
<th>Genes</th>
<th>Alternative splicing patterns</th>
<th>Number of mRNAs</th>
<th>Number of protein isoforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. LIM homeodomain proteins</td>
<td>ISL1</td>
<td>Alternative exon splicing</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>LHX3</td>
<td>Alternative transcription/translation initiation sites/exon splicing</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>LHX4</td>
<td>Alternative translation/initiation sites</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>LHX6</td>
<td>Alternative transcription/termination sites</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>LHX7</td>
<td>Alternative transcription/termination sites</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>LHX9</td>
<td>Alternative transcription/termination sites</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2. LIM domain only proteins</td>
<td>LMO4</td>
<td>Alternative exon splicing</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>TLP</td>
<td>Alternative exon splicing</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>FHL1</td>
<td>Alternative exon splicing</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>KyoT</td>
<td>Alternative exon splicing</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3. LIM domain proteins with other protein-protein interaction motifs,</td>
<td>3.1 PDZ-LIM domain proteins</td>
<td>ALP</td>
<td>Alternative exon splicing</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>ENH</td>
<td>Alternative exon splicing</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>ZASP</td>
<td>Alternative exon splicing</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>CYPHER</td>
<td>Alternative exon splicing</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Lmo7</td>
<td>Alternative exon splicing/exon splicing/intron retention</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>LIMK1</td>
<td>Alternative transcription initiation sites/exon splicing</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>LIMK2</td>
<td>Alternative transcription initiation sites/exon splicing</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>3.2 PET-LIM domain proteins</td>
<td>Testin</td>
<td>Alternative transcription initiation/termination sites/exon splicing</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Prickle</td>
<td>Alternative transcription initiation site</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3.3 Zyxin and Paxillin subgroup</td>
<td>Hic-5</td>
<td>Alternative transcription initiation sites/exon splicing/intron retention</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Trip6</td>
<td>Alternative exon splicing</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3.4 ATD-LIM domain proteins</td>
<td>ABLIM</td>
<td>Alternative transcription initiation sites/exon splicing</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>EPLIN</td>
<td>Alternative transcription initiation sites/exon splicing</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>ZNF185</td>
<td>Alternative translation initiation sites/exon splicing</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

the shorter *Isl-1* encodes *isl-1* isoforms (*isl1-β*), lacking the C-terminal 23aa of *isl1-α*. Unlike *Isl-1*, the LHX3 LIM homeodomain transcription factor is essential for the expression of pituitary hormone genes [6, 7]. Both mouse and the human LHX3 gene contain seven exons including two alternatively spliced first exons (Figure 1). This structure encodes two distinct protein isoforms, LHX3a and LHX3b, which differ exclusively in their N-terminus [8]. In addition, another protein isoform of LHX3, M2-LHX3, is specifically expressed from the *LHX3a* transcript using an alternate in-frame translation initiation site, and translates a LHX3 protein lacking the amino terminus and both LIM domains (Figure 1).

Biochemical and gene regulatory analysis revealed that LHX3a can effectively interact with several protein partners such as pituitary POU domain factor, Pit-1 and CREB binding protein (CBP) to exhibit stronger gene activation properties than LHX3b and M2-LHX3. LHX3a and LHX3b overexpression impact fertility in both female and male mice. By contrast, mouse with overexpressed M2-LHX3 is reproductively
Alternative splicing of LIM domain proteins

Figure 1. The alternative splicing of human LHX3 gene and its protein isoforms. LIM, LIM domains, HD, homeodomain ▼, Transcriptional start sites, ▽, Translational start sites, ▼, Translational end sites.

unaffected, and its specific functions need to be further studied [9, 10].

LIM domain only proteins (LMOs)

LMO4, a member of the nuclear LIM domain only protein, is overexpressed in more than 50% of primary breast cancers and cell lines [11]. Two non-coding 5'UTR exons (exon1a and exon1b) in the human LMO4 gene are differentially utilized in a mutually exclusive manner, which result in the expression of two LMO4 mRNAs among breast epithelial cells. Furthermore, two promoter regions are defined: upstream of exon1a, and upstream of exon1b. Northern blot analysis revealed that the longer transcript of LMO4 is expressed in both normal breast epithelial cells and cancer cell lines, while the shorter transcript is primarily expressed in transformed breast epithelial cells. Therefore, the overexpression of LMO4 in breast cancer reflects increased promoter activity and appears to involve the aberrant activation of the second promoter [12, 13]. Similarly, three alternatively spliced isoforms of human four-and-a-half LIM protein 1 (FHL1): FHL1, FHL1B and FHL1C; and two of its mouse homolog, KyoT1 and KyoT2, are generated by alternative exon splicing. Genomic organization reveals that the FHL1 gene has six exons (1, 2, 3, 4, 4b and 5). FHL1 transcript contains exons 1, 2, 3, 4 and 5; while FHL1B transcript contains exons 1, 2, 3, 4b and 5. In contrast, FHL1C is different from both FHL1 and FHL1B transcripts, which only contains exons 1, 2, 3 and 5 (Figure 2).

Amino-acid sequence analysis reveal that both the additional alternative exon 4b in FHL1B and the absence of exon 4 in FHL1C cause a different open reading frame in exon 5 compared with that of FHL1, and the frame-shifted exon 5 produces a putative RBP-J (J kappa recombination signal binding protein) binding region in the C-terminus of FHL1B and FHL1C [14, 15]. On the other hand, FHL1C is different from FHL1B, in which FHL1C lacks the third LIM domain encoded by exon 4 and three potential bipartite nuclear localization signals (NLS) and a putative nuclear export sequence (NES) encoded by exon 4b (Figure 2).

FHL1 is widely expressed in human tissues including the skeletal muscle and heart at a high level, while FHL1C is expressed in the testis, skeletal muscle and heart at a relatively low level, compared with FHL1. In contrast, FHL1B is specifically expressed in the brain. Functional analysis revealed that FHL1 binds to myosin-binding protein C (MyBP-C) with the fourth LIM domain and plays an essential role in myosin thick filament formation; while FHL1B and FHL1C cannot bind to MyBP-C, but can negatively regulate transcription by association with RBP-J DNA-binding proteins [16-18].
Alternative splicing of LIM domain proteins

The third group of LIM domain proteins

PDZ-LIM domain proteins

PDZ-LIM domain proteins including actinin-associated LIM protein (ALP), Enigma, LIM kinases (LIMKs) subfamily and LIM domain only7 (LMO7) are characterized by a single N-terminal or centrally positioned PDZ domain, and single or multiple LIM domains positioned either in the N-terminal or C-terminal from the PDZ domain. All these proteins are involved in cytoskeleton-associated activities through interaction with other protein partners [19, 20].

Enigma homolog (ENH) and Z-band alternatively spliced PDZ motif protein (ZASP/Cypher) belong to the Enigma family of PDZ-LIM domain proteins, which are characterized by an N-terminal PDZ domain and three C-terminal LIM domains. To date, four ENH isoforms have been identified in both human and mouse tissues. The larger isoform of mice enh, enh1, contains 16 exons; whereas the shorter isoforms named enh2, enh3 and enh4 lack the last six exons of the enh gene. Moreover, exon 5 is only found in the enh4 isoform; and exons 6 and 7 belong to the enh3 and enh4 isoforms. Interestingly, exon 4, found in enh1 and enh2, is internally spliced and gives rise to a truncated exon in enh3 (Figure 3). Consistent with nucleotide sequences of mouse enh isoforms, enh1 encodes an N-terminal PDZ domain and three C-terminal LIM motifs; while the three other isoforms do not contain the three LIM domains in the C-terminal (Figure 3). ENH1 is expressed in various tissues such as the heart, brain, spleen, liver and kidney. In contrast, ENH2, ENH3 and ENH4 are primarily expressed in cardiac and skeletal muscles [21, 22].

LIMKs composed of LIMK1 and LIMK2 are a group of unique PDZ-LIM domain proteins, the full length of LIMK1 and LIMK2 that consist of a C-terminal serine/threonine kinase domain, an internal PDZ domain, and two tandem N-terminal LIM domains. However, due to alternative transcription initiation sites and alternative exon splicing, LIMK1 and LIMK2 produce several tissue specific isoforms that differ mostly in the LIM and PDZ domains, but with some variants in the kinase domain. Two isoforms of LIMK1, the LIM and PDZ domain variants, have been found. The former contains the LIM domain but lacks the PDZ domain, and the latter contains the PDZ domain but lacks the LIM domain. Although these two isoforms of LIMK1 are both found in most tissues, the LIM domain variant of LIMK1 predominantly express in the gut, while the PDZ domain variant is preferably expressed in the middle brain [23, 24].

Figure 2. The alternative splicing of human FHL gene and its protein isoforms. Z, Zinc finger; NLS, Nuclear Location Signal; RBP-J: J kappa recombination signal binding protein.
At least five LIMK2 protein isoforms have been identified. One of them is denoted as LIMK2t, which lacks both LIM domains and a small part of the PDZ domain due to alternative transcription initiation site usage; and is specifically expressed in intact adult mouse testis [25] (Figure 4). In addition to LIMK2t, the full length one is usually denoted as LIMK2a, and a shorter version, LIMK2b, contains the kinase domain and PDZ domain; but only one and a half LIM domain have been described in humans, mice and rats. Genomic analysis revealed that the mouse limk2a gene consists of 16 exons, while limk2b lacks the exon 2 that encodes the first half of the LIM domain (Figure 4). LIMK2a and LIMK2b are expressed at different ratios in a variety of mouse tissues. LIMK2a is preferentially expressed in digestive organs of fetal and adult tissues, while LIMK2b is expressed more abundantly than LIMK2a in the placenta [23, 26]. Another transcript of limk2, termed limk2c, codes for a protein with a 6-amino-acid insert within the protein kinase domain, which has been found in the mouse brains [27]. In addition, a kinase-deleted alternative splice product of LIMK2 was also identified and characterized (Figure 4). In vivo analysis revealed that this alternate splice product inhibited LIM kinase activity, indicating that the amino terminus of LIMK2 may suppress the activity of the kinase domain [28]; as LIMK controls actin organization through phosphorylation, and actin binding/depolymerizing factor cofilin is inactivated. Thus, the kinase-deleted protein of LIMK2 may function as a negative regulator of LIMK2 in actin cytoskeleton organization.

Zyxin and paxillin subfamily

Zyxin and paxillin are prototypes of two related subfamilies of LIM domain proteins that are localized primarily at focal adhesion plaques. Zyxin/paxillin proteins have two major domains, N-terminal proline-rich sequences and the C-terminal LIM domain region. In addition, paxillin proteins have additional protein interaction motifs such as multiple LD repeats [29]. Zyxin/paxillin proteins have been shown to play key roles in various processes, and distinct isoforms of this group of proteins have been described [30-32]. For example, two primary isoforms of hydrogen peroxide inducible clone-5...
Alternative splicing of LIM domain proteins

Figure 4. The alternative splicing of mouse limk2 gene and its protein isoforms.

(Hic-5), a member of the paxillin subfamily, are produced by using alternative transcription initiation sites; and they are described as Hic-5α and Hic-5β, respectively [33]. Hic-5α has an exon 1' and translates four LD motifs, while Hic-5β lacks the exon 1' and only translates the last three LD motifs (Figure 5). Further studies using RT-PCR with Hic-5 isoform-specific primers and sequencing reveal a total of 12 transcripts of the Hic-5 gene, in which five of them fall into Hic-5α and others associate with Hic-5β. Isoforms in the Hic-5α group are mainly generated by exon skipping, while isoforms in the Hic-5β group use intron retention. All Hic-5 protein isoforms are identical within the LIM domains, which serve to target Hic-5 and paxillin to focal adhesions, as well as mediate hormone co-activation [31, 34, 35]. Both the Hic-5α and Hic-5β transcripts are extensively expressed with the exception of the nervous system such as the brain. However, Hic-5β transcripts are much less abundant than that of Hic-5α. Conflicting reports have suggested that Hic-5 can both repress and enhance a number of biological processes including apoptosis, tumorigenesis, myogenesis, hormone action and integrin signaling [36-38]. The identification of numerous Hic-5 isoforms may help to explain the diversity of biological functions attributed to Hic-5.

Actin-binding LIM domain proteins

Actin-binding LIM domain protein (ABLIM) was initially isolated from the human retinal cDNA library, and three ABLIM transcripts designated as ABLIM-L, ABLIM-M and ABLIM-S were gen-
Alternative splicing of LIM domain proteins

Figure 5. The alternative splicing of mouse Hic-5 gene and its protein isoforms. LD: leucine-aspartate repeat.

Figure 6. The alternative splicing of the human ABLIM gene and its protein isoforms. VHP: villin head piece.

erated from alternative transcription initiation sites and alternative exon splicing (Figure 6). The full length of ABLIM, ABLIM-L, contains 26 exons; while exons 9, 10 and 16 are completely skipped in ABLIM-M and ABLIM-S. Conceptual translation of the coding sequences of ABLIM-L, ABLIM-M and ABLIM-S shows that, except the common C-terminal VHP domain, ABLIM-L, ABLIM-M contain four and three N-terminal LIM domains, respectively. ABLIM-S, in contrast, does not contain any LIM domain [39, 40] (Figure 6).

ABLIM-L is found in the retina only, whereas ABLIM-M and ABLIM-S are ubiquitously expressed. Further studies have shown that
ABLIM-L is highly enriched in the retina ganglion cell layer (RGCs). During embryonic development, ABLIM-L expression is at a basal level at embryonic E12, but it is steadily upregulated postnatally. In contrast, the expression of ABLIM-M and ABLIM-S remains constant during this period. Thus, the expression of ABLIM-L is both retina-specific and temporally regulated. However, ABLIM-L knockout mice display no obvious defects in the development and maturation of retinofugal projections and in vitro growth of RGCs. It is known that ABLIM-L and ABLIM-M are structurally extremely similar, and both are expressed in the retina; while ABLIM-L may only play a minor role in regulating axon path finding/remodeling compared with ABLIM-M, or is functionally redundant in the presence of ABLIM-M. ABLIM-S, on the other hand, is unlikely to compensate for the loss of ABLIM-L function, since it does not have any LIM domain [41, 42].

Zinc finger protein 185 (ZNF185) is a novel actin-cytoskeleton-associated LIM domain protein that localizes to F-actin structures, and is enriched at focal adhesions. At least four ZNF185 mRNA isoforms have been identified. They have the same nucleotide sequence encoding a C-terminal LIM domain and a N-terminal ATD domain, but differ from exon 10 to 16 due to alternative exon splicing. In addition, a ZNF185 cDNA encoding only a single C-terminal LIM domain was assembled based on the sequence of an expressed sequence tagged clone and a genomic DNA sequence. The highest expression of various ZNF185 mRNA isoforms is in normal prostate, but decreases dramatically in metastatic prostate cancer tissues and several PCa cell lines. Although the functional difference of these ZNF185 isoforms remain to be established, in vitro transfection analysis with different domains of ZNF185 reveal that the N-terminal ATD is necessary and sufficient to suppress proliferation and anchorage-independent growth of PCa cells; indicating that ZNF185 isoforms with the ATD domain may function as tumor-suppressors by associating with the actin-cytoskeleton, although the biological function of this ATD deleted isoform of ZNF185 and its pathological relevance to PCa is unknown [43, 44].

Closing remarks

Alternative splicing (AS) is one of the most extensively used mechanisms that act to regulate and diversify gene functions. A large number of LIM domain proteins have been discovered to be regulated by AS, and produce even larger numbers of tissue specific proteins. Although the issue on how the AS of genes coding for LIM domain proteins may be regulated in different tissues remain diffused, some clues may be obtained from other AS studies [45, 46]. On the other hand, alternative splicing is likely to be important for LIM domain proteins to bring into full play in many, if not all, developmental pathways. This has also been observed in some diseases such as cancer. To date, very few alternatively spliced LIM domain proteins have been characterized at a functional level. Further studies emphasized on the splicing mechanism and functional difference of the respective isoform of LIM domain proteins should be performed in a cell and tissue specific manner.

Acknowledgements

This work was supported by National Natural Science Foundation of China (81373111) and Natural Science Foundation of Hebei Province in China (H2013209019).

Disclosure of conflict of interest

None.

Address correspondence to: Quan-Hui Zheng, Hebei Key Laboratory for Chronic Diseases, Tangshan Key Laboratory for Preclinical and Basic Research on Chronic Diseases, School of Basic Medical Sciences, North China University of Science and Technology, No. 57 of Jianshe South Road, Lubei District, Tangshan 063000, China. Tel: +8603153724598; Fax: +8603153725427; E-mail: zhqhdip@sohu.com

References

Alternative splicing of LIM domain proteins


Alternative splicing of LIM domain proteins


