Available online at www.sciencedirect.com

SCIENCE DIRECT®

Biochimica et Biophysica Acta 1755 (2005) 1–14

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

http://www.elsevier.com/locate/bba

Review

A survey of alternative transcripts of human tissue kallikrein genes

Lisa Kurlender[a,b], Carla Borgono[a,b], Iacovos P. Michael[a,b], Christina Obiezu[a,b], Marc B. Elliott[a,b], George M. Yousef[c], Eleftherios P. Diamandis[a,b,*]

[a]*Department of Laboratory Medicine and Pathobiology, University of Toronto, Ontario, Canada M5G 1L5*
[b]*Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada M5G 1X5*
[c]*Discipline of Laboratory Medicine, Memorial University, St. John's, Newfoundland, Canada A1B 3V6*

Received 30 August 2004; received in revised form 10 February 2005; accepted 11 February 2005
Available online 10 March 2005

Abstract

Alternative splicing is prevalent within the human tissue kallikrein gene locus. Aside from being the most important source of protein diversity in eukaryotes, this process plays a significant role in development, physiology and disease. A better understanding of alternative splicing could lead to the use of gene variants as drug targets, therapeutic agents or diagnostic markers. With the rapidly rising number of alternative kallikrein transcripts, classifying new transcripts and piecing together the significance of existing data are becoming increasingly challenging. In this review, we present a systematic analysis of all currently known kallikrein alternative transcripts. By defining a reference form for each of the 15 kallikrein genes (*KLK1* to *KLK15*), we were able to classify alternative splicing patterns. We identified 82 different kallikrein gene transcript forms, including reference forms. Alternative splicing may lead to the synthesis of 56 different protein forms for *KLK1-15*. In the kallikrein locus, the majority of alternative splicing events occur within the protein-coding region, and to a lesser extent in the 5′ untranslated regions (UTRs). The most common alternative splicing event is exon skipping (35%) and the least common events are cryptic exons (3%) and internal exon deletion (3%). Seventy-six percent of kallikrein splice variants that are predicted to encode truncated proteins are the result of frameshifts. Eighty-nine percent of putative proteins encoded by splice variants are predicted to be secreted. Although several reports describe the identification of kallikrein splice variants and their potential clinical utility, this is the first extensive review on this subject. Accumulating evidence suggests that alternative kallikrein forms could be involved in many pathologic conditions or could have practical applications as biomarkers. The organization and analysis of the kallikrein transcripts will facilitate future work in this area and may lead to novel clinical and diagnostic applications.

Keywords: Kallikrein; Alternative splicing; Differential expression; Serine protease; Splice variant; Cancer biomarker

Contents

1. Introduction . . . 2
2. Methodology . . . 3
   2.1. Alternative splicing in the kallikrein locus . . . 4
   2.2. The coding exons . . . 5
   2.3. The non-coding exons . . . 5
   2.4. Encoded proteins . . . 5

Abbreviations: KLK, hK Kallikrein gene, kallikrein protein; UTR, untranslated region; RACE, rapid amplification of cDNA ends; EST, expressed sequence tag; ORF, open reading frame

* Corresponding author. Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario, Canada M5G 1X5. Fax: +1 416 586 8628.
E-mail address: ediamandis@mtsinai.on.ca (E.P. Diamandis).

0304-419X/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbcan.2005.02.001

3. Conclusion . . . 10
Acknowledgements . . . 11
References . . . 11

1. Introduction

Human tissue kallikreins (*KLK*, hKs) are a group of serine proteases encoded by 15 structurally similar, hormonally-regulated genes that tandemly localize on chromosome 19q13.4 [1,2] and found to have clinical utility in various diseases, including cancer [3]. The presence of more than one mRNA form for the same gene is common among kallikreins. In the kallikrein locus, alternative transcripts are the result of splicing in the coding or non-coding regions, use of alternative transcription or translation start and stop sites and combinations thereof. Here, we review the numerous alternative transcripts of the kallikrein genes in order to facilitate the identification of novel kallikrein transcripts and future investigations towards their diagnostic and therapeutic applicability.

Alternative splicing, first proposed by Gilbert in 1978 [4], generates multiple mRNA forms from one gene and may yield several different proteins. Splice variants were found for 35–74% of all human genes [5–12]. Of all mechanisms that increase protein diversity, alternative pre mRNA splicing is considered to be the most significant source in vertebrates, as approximately 70–88% of alternative splicing events result in changes in the encoded protein [5,8,11]. Processes such as alternative use of promoters, splice sites, translational start sites and translational termination codons can serve as mechanisms for regulating alternative splicing. The use of alternative transcriptional start sites and poly A sites can also generate a variety of mRNAs.

Alternative transcripts can be detected using a combination of RT-PCR techniques, cloning and sequencing or by using fiber-optic arrays and exon junction microarrays [10,13–18]. Although they have several limitations [7], expressed sequence tags (ESTs) are also very useful in identifying putative alternative transcripts.

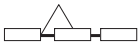
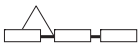
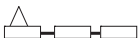
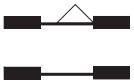
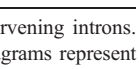
Alternative splicing may generate segments of mRNA variability that can insert or remove amino acids, shift the reading frame, or introduce a termination codon. Alternative splicing can also remove or insert regulatory elements controlling translation, mRNA stability or localization [19]. The use of alternative translation initiation sites is key in generating a versatile repertoire of functionally different proteins within individual cells. Diseases associated with mutations may disturb the initiation step of translation by changing the context around the AUG start codon or by introducing upstream AUG codons [20–25]. See reference [26] for more details.

The mechanism that allows transcription to initiate at an alternative start site or terminate at a different poly A site is associated with the use of alternative promoters, a less well-characterized phenomenon than alternative splicing that still contributes to genome complexity [27]. Although 60–80% of genes with alternative promoters produce transcripts with identical open reading frames (ORFs) [27], this mechanism can also lead to different proteins through the use of alternative ORF or creation of novel ORFs [28,29]. A recent review describes the interrelationships between splicing and promoter regions [27].

Before being transported to the cytoplasm for translation, the mRNA is spliced and polyadenylated [30]. The poly A tail enhances translation and mRNA stability [31,32]. Aberrant polyadenylation may alter cell viability, growth and development and may ultimately lead to disease [33].

Alternative splicing has been implicated in many physiological and pathophysiological processes and 15% of mutations in the mammalian genome that cause disease are associated with an affected RNA splicing signal [34].

Table 1
Nomenclature and schematic representation of splicing events

Our definition	Definition by Wang et al. [17]	Schematic representation
Skipped exon	Type I deletion	
5′ Truncated exon	Type II deletion	
3′ Truncated exon	Type III deletion	
Internal exon deletion	Type IVdeletion	
Cryptic exon	Type I insertions	
5′ Exon extension	Type II insertions	
3′ Exon extension	Type III insertions	
Complete intron retention	Type IV insertions	
Classical splicing	Regular splicing	

Boxes represent exons and horizontal lines represent intervening introns. Diagonal lines show splicing arrangement. The first 4 diagrams represent splicing events involving only exons, while the last 5 diagrams represent splicing events involving introns and exons.