

## Showcase on Research

# Human Tissue Kallikreins - Is Their Function Solely Proteolytic?

Judith A. Clements

Centre for Molecular Biotechnology, School of Life Sciences, Queensland University of Technology, QLD 4001

### What is a tissue kallikrein?

The serine protease, tissue kallikrein, was first described over 50 years ago, when a kininogenase activity was noted in pig pancreatic extracts - hence its name kallikrein which is derived from *kallikreas* (Greek for pancreas). The proteolytic processes involved in the generation of (lys) bradykinin from low molecular weight (LMW) kininogen by tissue kallikrein and the functional attributes of bradykinin have been extensively studied since that time. The multi-functional peptide, bradykinin, is a powerful vasodilator and pain-producing autacoid that is also involved in vascular permeability, smooth muscle contraction, electrolyte transfer and cell proliferation (1). Thus, tissue kallikrein, through the generation of (lys) bradykinin, and a ubiquitous expression pattern, is involved in the function of most systems in the body. Although LMW kininogen is the designated substrate of tissue kallikrein, many *in vitro* studies have shown that tissue kallikrein can hydrolyse other substrates such as the matrix metalloproteases, pro-insulin and atrial natriuretic factor (1). Indeed, a recent novel finding is that tissue kallikrein can directly activate the G protein-coupled bradykinin B2 receptor, independent of bradykinin generation, suggesting a novel kallikrein effect similar to that of the protease-activated receptors (PARs) (2). Whether this multi-substrate ability is representative of what occurs physiologically or perhaps reflects co-purified contaminants has not yet been proven. This ambiguity is a continuing theme as we biochemically characterise the activity of the newer kallikreins and attempt to understand their physiological roles.

The expansion from one tissue kallikrein to a family of related enzymes is a product of gene cloning. Attempts to clone the tissue kallikrein (*KLK1*) gene identified other highly similar genes, which were eventually recognised to cluster in one locus in both rodents and man and were the result of gene duplication. Although highly conserved with respect to both sequence and

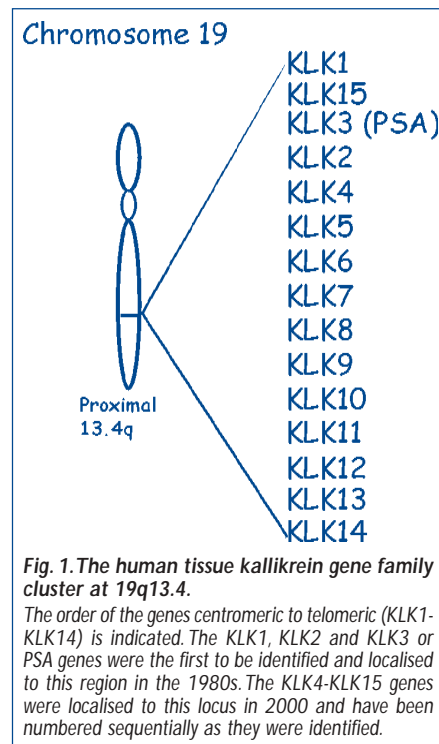


Fig. 1. The human tissue kallikrein gene family cluster at 19q13.4.

The order of the genes centromeric to telomeric (*KLK1*-*KLK14*) is indicated. The *KLK1*, *KLK2* and *KLK3* or *PSA* genes were the first to be identified and localised to this region in the 1980s. The *KLK4*-*KLK15* genes were localised to this locus in 2000 and have been numbered sequentially as they were identified.

structure, their known or predicted enzymic functions are specific and distinct from one another. They also share a wide range of expression patterns that likely indicates involvement in a diverse range of physiological processes. It should be emphasised that there is only *one true tissue kallikrein*, with the substrate specificity noted above, in any one species. Other members of the family are called kallikreins or kallikrein-related enzymes because of their sequence and structural homology to *tissue kallikrein* and clustered position with the *KLK1* gene in the genome. It should also be noted that although *plasma kallikrein* shares a similar substrate specificity with true tissue kallikrein, in the generation of bradykinin from (HMW) kininogen, it is not considered a member of this family since it shares no sequence or structural homology, nor similar genomic location.

### The human tissue kallikrein family

The full extent of the human tissue kallikrein or *KLK* gene family, the focus of this review, has been established only

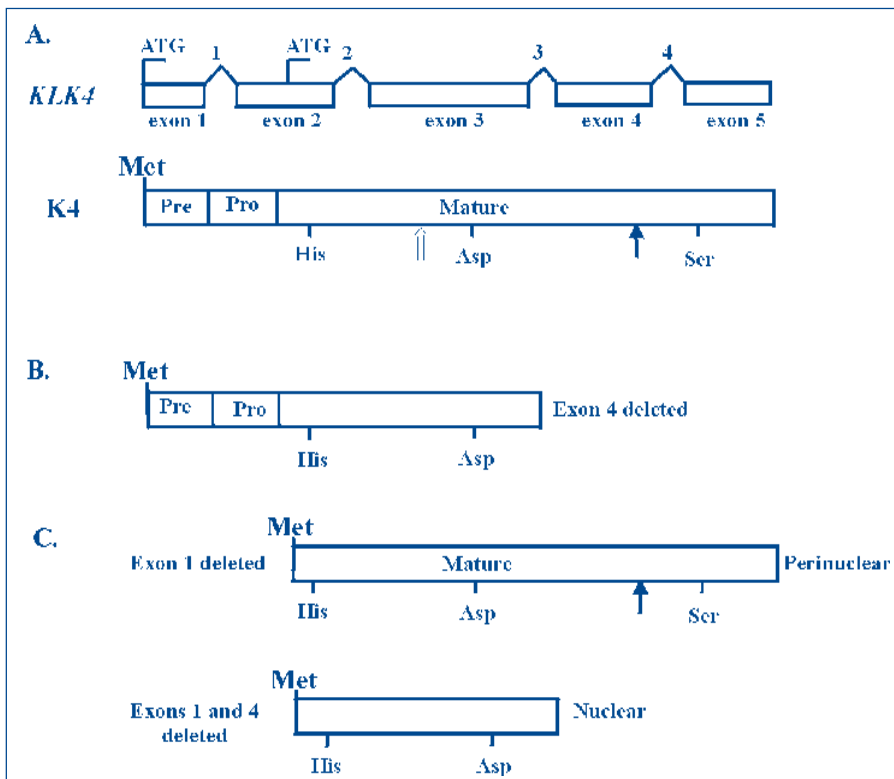
recently. The first three characterised human *KLK* genes - *KLK1*, *KLK2* and *KLK3* - were localised in a cluster of a 60 kilobase region at 19q13.3-13.4. This region is syntenic to the locus where the mouse *Kik* gene family cluster was localised on mouse chromosome 7, further emphasising evolutionary similarities between these families (1). However, the striking difference between the mouse (and rat) and human *KLK* families was their size. The rodent families were large, consisting of 13-26 genes. Thus, by comparison, it appeared logical that the human *KLK* family might also contain more genes specifically within, or adjacent to, the original *KLK* cluster. Recently, both the availability of draft sequence for this region from the Human Genome Sequencing Project (at Lawrence Livermore National Laboratory, California, USA) and physical mapping has allowed the identification of 12 additional kallikrein-related genes (*KLK4*-*15*) within a 320 kilobase region that encompasses the original *KLK* locus (Fig. 1) (3, 4). Given the short space of time since these genes were identified, there is limited biochemical information on their enzymatic activity and function, although inferences can be made from structural and sequence alignments and sites of expression.

### Structural and sequence similarities: conservation yet diversity

Although the size of the individual human *KLK* genes may differ, ranging from 4-10 kilobases, the five coding exons are highly conserved both in size and organisation (Fig. 2A) (3, 4). The first of these has a short 5' untranslated region prior to the initiation of the pre-peptide (signal) coding region. The second exon encodes the remainder of the pre-peptide as well as pro-enzyme sequences. The mature enzyme is encoded on the remainder of the second coding exon through to the fifth coding exon with a variably sized 3' untranslated region completing this last coding exon. In the second, third and fifth coding exons, the positions of the three residues of the catalytic triad (histidine, aspartate, serine)

## Showcase on Research

### Human Tissue Kallikreins - Is Their Function Solely Proteolytic? (contin.)



**Fig. 2. Structural attributes of native and variant forms of kallikrein 4.**

**A.** Structural arrangement of the 5 coding exons of the KLK4 gene (indicating 2 potential ATG start sites in exons 1 and 2) and the K4 pre-pro-enzyme derived from initiation at the first ATG. The positions of the histidine, aspartate and serine of the catalytic triad, the kallikrein loop (open arrow; only present in its full form in KLK1-3) and the residue critical in determining trypsin- or chymotrypsin-like activity (solid arrow) are indicated. This structural organisation is identical for all other kallikreins, except for the second potential translational start site in exon 2, which is peculiar to KLK4.

**B.** Putative K4 protein derived from an mRNA splice variant that lacks exon 4. There is a frame shift and premature stop-codon at the beginning of exon 5 that leads to a truncated protein lacking the serine residue of the catalytic triad.

**C.** Putative K4 proteins that are initiated from the second ATG in exon 2 and are perinuclear or nuclear localised. Although the upper protein retains the residues critical for catalysis, it is not clear that the correct conformation would occur.

critical for protease activity are highly conserved across all 15 genes. This structural organisation is also highly conserved for all rodent KLK genes.

In contrast to the rodent KLK/KLK families and the human KLK1-KLK3 encoded enzymes, which are highly conserved in sequence (rodents: 70-85%; hK1-hK3: 62-77%), the newer hK4-hK15 enzymes are less related (25-66%) (3, 4). This is likely a reason why they were not identified until recently, as most investigators in the field were focussing on a higher order of conservation. This difference may reflect an evolutionary order of gene duplication within this locus, with the KLK1-KLK3 genes perhaps being the result of the most recent duplication. The rodent kallikreins and the hK1-hK3 enzymes also contain an insertion of 11 amino acids prior to the

catalytic aspartate, denoted the "kallikrein loop", that is thought to be important for substrate specificity of these enzymes. The more recently discovered proteases do not have this loop region, although five of them have smaller insertions that also may be important in specifying their enzymatic action. This latter observation has caused some controversy in the field. Some investigators are adamant that this means these enzymes cannot be named kallikreins; others maintain that the common locus, and other commonalities as described above, provides a basis for their being so-named.

#### Proteolytic function

As noted above, key residues integral to the structure and function of these proteases are conserved. There are 12 conserved cysteine residues that are

predicted to form six disulphide bridges common to other serine proteases. The key residues (and much of the adjacent regions) that denote serine protease activity are entirely conserved; any differences in the P and P' residues will likely be shown to be important in substrate specificity. The analysis of these critical residues by *in silico* molecular modelling and design of synthetic peptide substrates is in progress; the determination of the physiological substrates will require more sophisticated approaches such as those based on proteomics and expression libraries. What can be gleaned from simple sequence alignments is that the majority of these enzymes are, or likely to be, trypsin-like in action, with a preference for cleavage at basic residues - a function denoted by an aspartate six residues from the catalytic serine (3, 4). Four enzymes (K3, K7, K9, K15) have alternate residues in this position which likely denotes a chymotrypsin-like action and a preference for hydrophobic residues (3, 4). Some known biochemical, and presumed physiological, functions for some of these enzymes are noted below.

hK3 or prostate-specific antigen (PSA) hydrolyses seminal vesicle proteins in ejaculate, an event that leads to dissolution of the seminal clot and is integral to sperm motility (5). hK2 can activate pro-urinary plasminogen activator and hydrolyse matrix proteins such as fibronectin; it is thus thought to play a role in matrix degradation and cancer cell invasion (5). Like tissue kallikrein, PSA and hK2 are also suggested to act, albeit with different kinetics, on several other substrates, many of which are common (fibronectin, insulin-like growth factor binding proteins, LMW kinogen) (5-7). The biological relevance of these findings is not yet clear and may reflect some redundancy in the actions of this family of enzymes. What it does highlight, however, is the difficulty in determining specific substrates for each individual kallikrein until recombinant forms of all the enzymes are available to test their relative kinetics against a putative substrate.

hK6 or Zyme can hydrolyse Amyloid Precursor Protein and thus is hypothesised to be important in the deposition of amyloid plaques in Alzheimer's disease. Other functions, in tooth development and neural plasticity, have been suggested for the hK4 and hK8 enzymes respectively, by extrapolation of

## Showcase on Research

### Human Tissue Kallikreins - Is Their Function Solely Proteolytic? (contin.)

the known actions of their putative mouse orthologues - enamel matrix serine protease 1 and neuropsin, respectively (3, 4). The enzymatic functions of the remainder of the more recently discovered *KLK*-encoded proteins are not yet known, although their expression patterns would suggest roles in the pancreas, skin, brain and hormone-dependent cancers.

Activation of the zymogens is a critical regulatory process that is still not fully understood. The 7-amino acid pro-peptide must be proteolytically removed for enzyme activation, a process that then allows the N-terminus of the mature enzyme to fold back and initiate a conformational change critical for the formation of the substrate-binding pocket. A new prostatic serine protease, prostin, that is unrelated to the kallikreins, can cleave this pro-peptide and thus may be a physiological activator of PSA (8). hK2 auto-activates by cleaving its own pro-peptide; hK2 can also activate pro-PSA demonstrating the close relationship of these two prostatic kallikreins (5). Similarly, the co-localisation of the tryptic-like hK5 and chymotryptic-like hK7 enzymes in skin, and analysis of the pro-peptide cleavage sites and putative substrate specificities of these enzymes, has led to the proposal that hK5 may be the activator of pro-hK7 in this tissue (9). Indeed, the current speculation is as follows: where several of the kallikreins are expressed in one tissue type, they may activate each other in a cascade-like process. However, as noted above, there may also be other proteases involved, dependent on the tissue type.

The final regulatory event that is common to these and other proteases is the inactivation of the mature enzyme by binding to endogenous inhibitors. Some known kallikrein protease inhibitors are kallistatin (specific for tissue kallikrein/hK1),  $\alpha_1$ -antichymotrypsin (PSA and hK2), protein C inhibitor (hK2, PSA), PI-6/SPI3 (hK2, mouse neuropsin/K8) and  $\alpha_2$ -macroglobulin (5, 10-12). It is apparent that several different serpins can inhibit an individual kallikrein as well as any one serpin being perhaps pan-specific for several kallikreins. The differences appear to be compartmental since hK2, for instance, predominantly forms complexes with  $\alpha_1$ -antichymotrypsin, protein C inhibitor or PI-6, respectively, in serum, seminal fluid and prostate tissue. Putative protease/inhibitor complexes have also been noted for hK4, hK5 and hK7,

although the precise serpin(s) involved in those complexes are yet to be elucidated.

#### Potential novel non-proteolytic functions

There are several recent observations that suggest the kallikreins may have functional roles other than in proteolytic processing. In contrast to the rodent families where only one mRNA transcript has been described for each gene, multiple variant mRNA transcripts - the result of alternative splicing - have been described for many of the human *KLK* genes (4, 13, 14). These splicing events cause a shift in the open reading frame, leading to premature stop codons and truncated proteins if translated. Of note, these variant transcripts would not encode serine proteases as typically one or other of the three residues critical for catalytic activity is missing (**Fig. 2B**). The following possibilities for the significance of these events require further clarification. They may represent: (a) a means of further functional diversification of the human *KLK* gene family; (b) a control mechanism for regulating levels of the native protein; or (c) simply an unrelated genomic processing event. The degree of tissue specificity in expression of these alternatively spliced mRNAs and their protein products, or their relationship to a particular pathophysiological condition, is currently under investigation. Indeed, the protein product of one PSA variant transcript has been detected in biological samples (15).

PSA, and its complexed or free forms, are widely used as diagnostic markers for prostate cancer and monitoring tumour recurrence. hK2 and its different molecular forms are also proving useful diagnostic/prognostic markers for prostate disease. Although no information has yet emerged as to the usefulness of the above PSA variants as bio-markers of disease, several other proteolytic fragments of PSA are proving useful in the clinical setting. For instance, a truncated precursor form of PSA, [-2]pPSA, with only 2 amino acids of the 7-amino acid leader peptide, is more prevalent in serum from men with prostate cancer than benign disease (16). This truncated precursor form is still proteolytically inactive and moreover, cannot be cleaved to give active PSA by any of the known enzymic activators of the PSA zymogen. Conversely, "nicked" or post-translationally cleaved fragments

of mature PSA, are thought to be more representative of benign disease (17). Whether any of these transcriptional or post-translationally modified forms of PSA will have functions in their own right, or are merely useful as indicators of disease, is yet to be established.

PSA appears to exert anti-angiogenic properties by inhibiting vascular endothelial growth factor or basic fibroblast growth factor-stimulated endothelial cell migration, invasion and tube formation (18). Of particular note is that these same effects can be achieved with [-1] PSA (PSA minus the isoleucine at the N-terminus), a PSA isoform that no longer has serine protease activity (19). This finding suggests that the anti-angiogenic effects observed were not the result of a proteolytic action of PSA.

Finally, the latest novel observation in the kallikrein field is the nuclear localisation of one of these purported extracellular proteases. Using an anti-peptide antibody to hK4, our group has noted prominent nuclear staining in several cell lines. Concurrently, another group has shown that green fluorescent protein constructs of 2 different splice variants of *KLK4*, that lack the reported exon 1 and exon 4 or lack exon 1 and retain intronic sequence between exons 3 and 4, are both localised to the nucleus (**Fig. 2C**) (20). An additional form that only lacks exon 1 has a perinuclear localisation (**Fig. 2C**) (20). All three variants lack the signal peptide and pro-sequences of wild type hK4, but retain an identical, albeit shortened, N-terminal sequence, since the coding sequence now starts with an in-frame methionine 20 amino acids C-terminal to the original pro-cleavage site. The perinuclear form would retain the 3 residues critical to the catalytic triad; however, it is not clear that the correct conformation of the substrate-binding pocket would occur. Clearly, all three variants would not be secreted, are not zymogens and likely would have functions unrelated to that of a serine protease.

#### Concluding remarks

The expanded human tissue kallikrein (*KLK*) family of serine proteases clearly represents a diverse group of enzymes that are likely to be involved in the (patho)physiology of a range of different organs as denoted by their tissue-restricted or broader expression patterns. The challenge now is to precisely determine the enzymatic action of these



# Chromosome 19

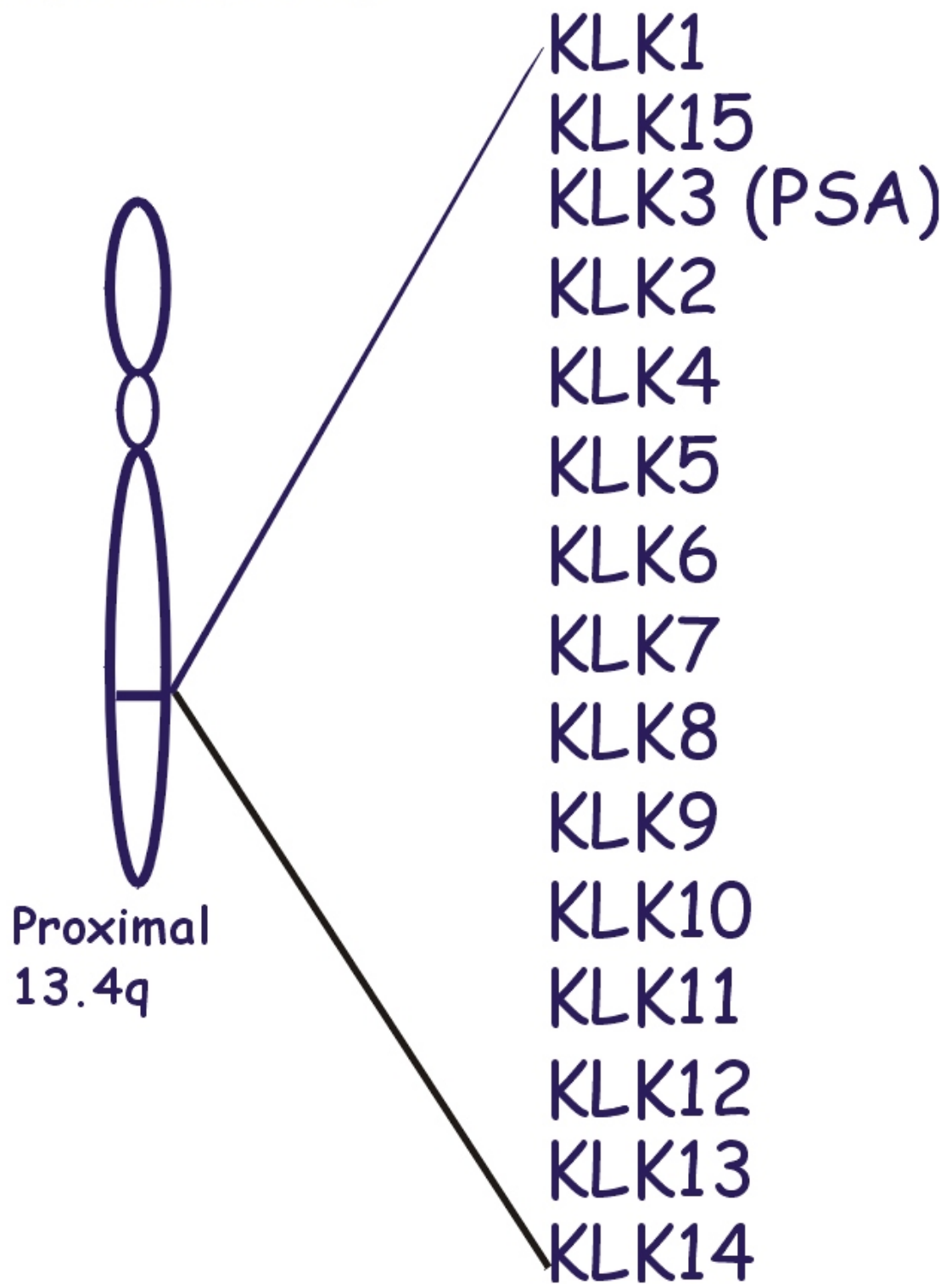


Fig. 1

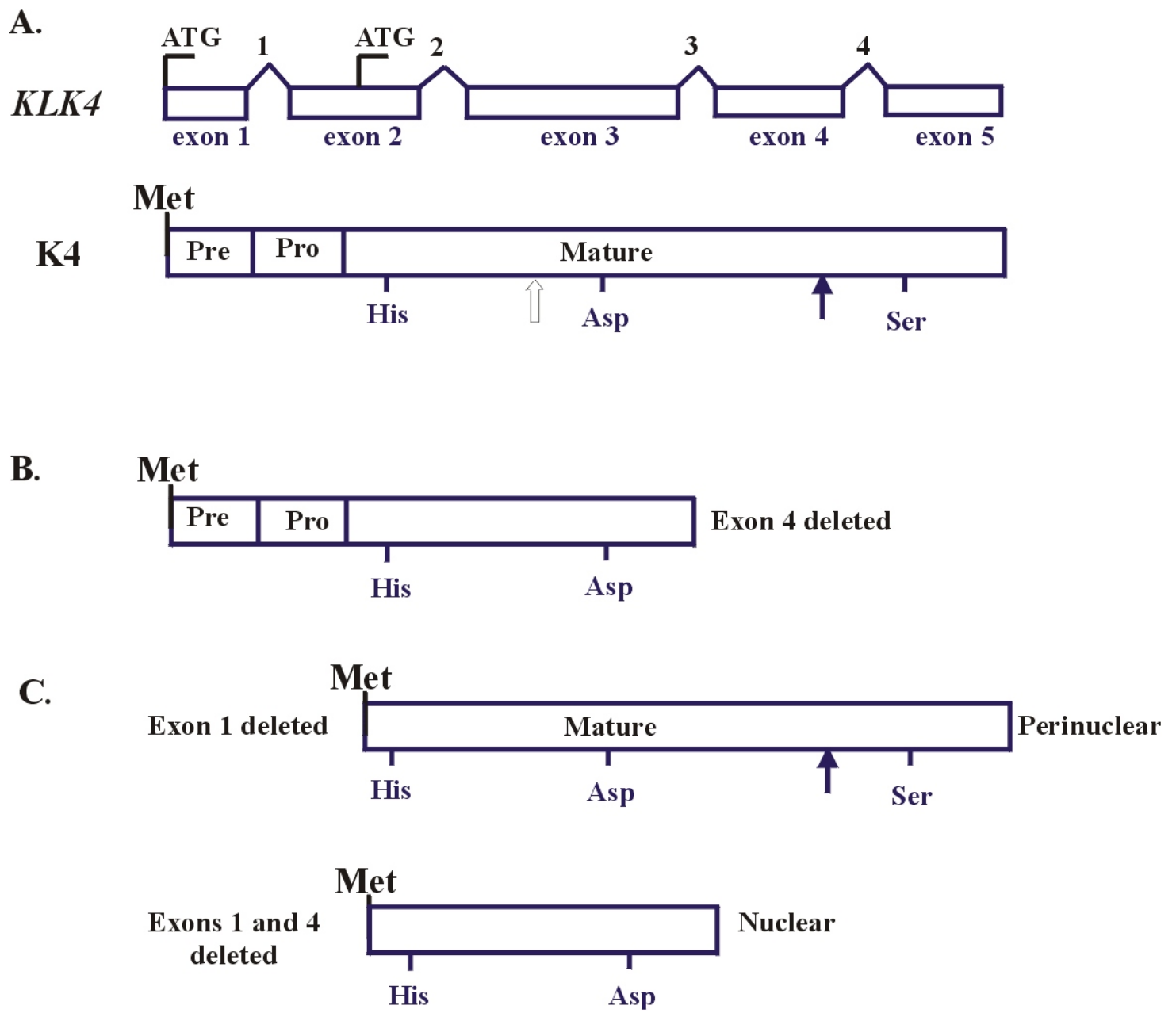


Fig. 2