Open access • Journal Article • DOI:10.1515/BC.2006.082

A comprehensive nomenclature for serine proteases with homology to tissue kallikreins — Source link

Åke Lundwall, Vimla Band, Michael Blaber, Judith A. Clements ...+14 more authors

Institutions: Florida State University, Queensland University of Technology, François Rabelais University, University of Toronto ...+7 more institutions

Published on: 01 Jun 2006 - Biological Chemistry (De Gruyter)

Topics: Tissue kallikrein, Kallikrein, Tissue Kallikreins and Proteases

Biol. Chem., Vol. 387, pp. 637–641, June 2006 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/BC.2006.082

Nomenclature Paper

A comprehensive nomenclature for serine proteases with homology to tissue kallikreins

Åke Lundwall[1,*], Vimla Band[2], Michael Blaber[3], Judith A. Clements[4], Yves Courty[5], Eleftherios P. Diamandis[6], Hans Fritz[7], Hans Lilja[8], Johan Malm[1], Lois J. Maltais[9], A. Yvonne Olsson[10], Constantina Petraki[11], Andreas Scorilas[12], Georgia Sotiropoulou[13], Ulf-Håkan Stenman[14], Carsten Stephan[15], Maroulio Talieri[16] and George M. Yousef[17]

Published with the consent of: Mireille Ainciburu[5], Maria Brattsand[18], Charlotte Becker[1], Adam Clauss[1], Mekdes Debela[19], Ying Dong[4], Nathalie Heuzé-Vourc'h[5], John Hooper[4], Mary-Anne Kedda[4], Tadaaki Kishi[20], Maciej Kwiatkowski[21], Georgios Pampalakis[13], Chris Planque[5], Dan Sexton[22], Thomas Takayama[23], Olivia Tan[4], Antonia Vlahou[24], Astrid Whitbread[4]

* Corresponding author
e-mail: ake.lundwall@med.lu.se

Abstract

The human kallikrein locus on chromosome 19q13.3–13.4 contains kallikrein 1 – the tissue kallikrein – and 14 related serine proteases. Recent investigations into their function and evolution have indicated that the present nomenclature for these proteins is inadequate or insufficient. Here we present a new nomenclature in which proteins without proven kininogenase activity are denoted kallikrein-related peptidase. Names are also given to the unique rodent proteins that are closely related to kallikrein 1.

Keywords: cancer; evolution; hormone; inflammation; kininogen.

Introduction

The term kallikrein (derived from Greek, *kallikreas*, for pancreas) was coined by Kraut and colleagues in 1930, when they demonstrated that an earlier described hypotensive substance in urine is present at high concentration in the pancreas (Frey and Kraut, 1926; Kraut et al., 1930). Today, the substance is known as kallikrein 1 or tissue kallikrein (EC 3.4.21.35), an enzyme that generates Lys-bradykinin by specific proteolysis of kininogen 1. There is also a proteolytic enzyme in blood plasma that gives rise to bradykinin that is known as plasma kallikrein (EC 3.4.21.34). Several other proteases also exhibit kallikrein activity, albeit usually less efficiently than the tissue and plasma kallikreins. A recent review of the kallikrein-kinin system is provided by Moreau et al. (2005).

Some 25 years ago, it was shown that mouse and rat salivary glands secrete proteins with homology to tissue kallikrein – at that time known as glandular kallikrein (Bothwell et al., 1979). Owing to their close relationship, glandular kallikrein and its homologs were assigned to a subfamily of serine proteinases, which was named the glandular kallikrein family. Although some of the novel glandular kallikreins displayed potent kallikrein – i.e., kininogenase – activity, they were primarily considered to be involved in prohormone processing, as some of them formed complexes and cleaved precursor proteins of epidermal and nerve growth factors (Thomas et al., 1981; Blaber et al., 1987). A comprehensive analysis showed that there were 24 or 25 glandular kallikrein genes in the mouse genome, designated *mGK-1* to *mGK-25* (Evans et al., 1987). Similar analysis of the rat genome identified 10 or 11 glandular kallikrein genes, denoted *rGK-1* to *rGK-10* (Wines et al., 1989). In a revision of the kallikrein nomenclature, the gene family was renamed the tissue kallikrein gene family and the symbol *GK* was replaced by *KLK*, e.g., the new designation of *rGK-4* was *rKLK4*

Affiliations: [1]Lund University, Clinical Chemistry, Department of Laboratory Medicine, University Hospital MAS, S-205 02 Malmö, Sweden; [2]Department of Medicine, Evanston Northwestern Healthcare Research Institute, Evanston, IL 60201, USA; [3]Department of Biomedical Sciences, College of Medicine, Florida State University, Tallahassee, FL 32306-4300, USA; [4]School of Life Sciences and Science Research Center, Queensland University of Technology, Brisbane, QLD 4001, Australia; [5]INSERM, U618, Protéases et Vectorisation Pulmonaires, Universite Francois Rabelais, F-37000 Tours, France; [6]Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, ON M5G1X5, Canada; [7]Division of Clinical Biochemistry at the Surgical Department City of the Ludwigs Maximilians University, D-80336 Munich, Germany; [8]Departments of Clinical Laboratories, Urology, and Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA; [9]Mouse Genomic Nomenclature Committee (MGNC), Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, ME 04609, USA; [10]Molecular Carcinogenesis, Institute of Cancer Research, Sutton SM2 5NG, UK; [11]Department of Pathology, Evangelismos Hospital, GR-11364 Athens, Greece; [12]Department of Biochemistry and Molecular Biology, University of Athens, GR-15701 Athens, Greece; [13]Department of Pharmacy, School of Health Sciences, University of Patras, GR-26500 Rion-Patras, Greece; [14]Department of Clinical Chemistry, Helsinki University Central Hospital, FIN-00029 Helsinki, Finland; [15]Department of Urology, Universitätsmedizin Charité, Campus Mitte, D-10098 Berlin, Germany; [16]G. Papanicolaou Research Center of Oncology, Saint Savas Hospital, GR-11522 Athens, Greece; [17]Discipline of Pathology, Health Science Corporation of St. John's, St. John's, NF A1B 3V6, Canada; [18]Department of Public Health and Clinical Medicine, Umeå University, S-901 87 Umeå, Sweden; [19]Max Planck Institute of Biochemistry, D-82152 Martinsried, Germany; [20]Med Discovery S.A., Chemin des Aulx 16, CH-1228 Plan-les-Ouates, Switzerland; [21]Urological Clinic, Kantonsspital Aarau, CH-5001 Aarau, Switzerland; [22]Dyax Corp., 300 Technology Square, Cambridge, MA 02139, USA; [23]Department of Urology, University of Washington, Seattle, WA 98195-6510, USA; [24]Department of Biotechnology, Foundation for Biomedical Research of the Academy of Athens, GR-11527 Athens, Greece

(Berg et al., 1992). In contrast to the large number of murine genes, the human tissue kallikrein family seemed to consist of only three genes, which coded for tissue kallikrein, prostate-specific antigen (PSA) and human glandular kallikrein 1 (hGK-1) – later renamed human kallikrein 2 (hK2) (Fukushima et al., 1985; Lundwall and Lilja, 1987; Schedlich et al., 1987).

The discrepancy in the number of genes was recently explained by comparative studies on the kallikrein locus in mammals (Olsson and Lundwall, 2002; Olsson et al., 2004a,b). These investigations showed that several duplications of the tissue kallikrein gene (*KLK1*) occurred very late in phylogeny and created 23 *KLK1* paralogs that seem to be unique to the mouse and nine *KLK1* paralogs that seem to be unique to the rat. Late duplication of *KLK1* was also observed in the horse, but not in artiodactyls, carnivores, cavian rodents and primates. Another, presumably primate-specific, duplication yielded the hK2 (*KLK2*) and PSA (*KLK3*) genes. A functional gene related to the progenitor of this duplication is present in the dog, whereas in the mouse and rat there is a nonfunctional pseudogene.

Around the turn of the millennium, investigators identified several genes of simple serine proteases adjacent to the human tissue kallikrein locus on chromosome 19q13.3–13.4 (Gan et al., 2000; Harvey et al., 2000; Yousef et al., 2000). The mutual sequence agreement of the tissue kallikrein family members was higher than the similarity between any of the novel serine protease genes. However, the chromosomal location in combination with overlapping expression in hormone-dependent tissues suggested both a common ancestry and overlapping functionality. Therefore, the old tissue kallikrein family was expanded with the adjacently located serine proteinases into what has become known as the extended kallikrein family (Yousef and Diamandis, 2001; Borgoňo and Diamandis, 2004; Borgoňo et al., 2004).

Soon after its discovery, a rational nomenclature was adopted for members of the extended kallikrein family (Diamandis et al., 2000). The human kallikrein family, as we now know it, consists of 15 genes, designated kallikrein 1–15 and denote by the gene symbols *KLK1*–*KLK15*. The nomenclature has served its purpose and is widely accepted by scientists from many different disciplines. However, some shortcomings of the nomenclature have been recognized, as follows:

- The nomenclature was developed for human genes and does not provide names for unique animal genes;
- The term kallikrein was introduced and has been used for decades to identify enzymes with kininogenase activity. Most of the enzymes in the extended kallikrein family are presumed to not display kininogenase activity and thus use of the term kallikrein might be misleading and confusing;
- At present, the nomenclature does not fully comply with the guidelines provided by the human (http://www.gene.ucl.ac.uk/nomenclature/guidelines.html) and mouse (http://www.informatics.jax.org/mgihome/nomen/gene.shtml) gene nomenclature committees (HGNC and MGNC), e.g., there should not be two separate symbols, such as hK1 and *KLK1*, to depict the protein and the gene.

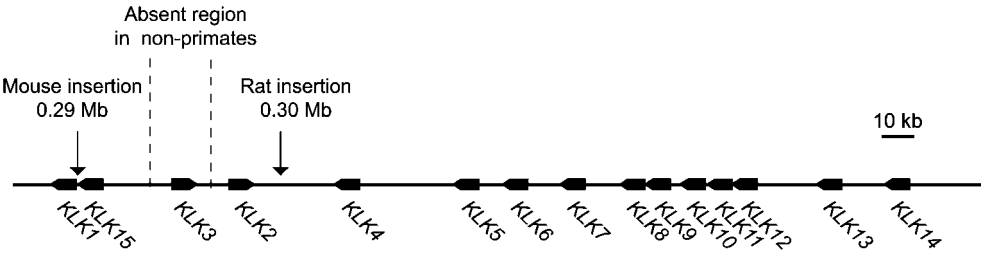

Figure 1 Schematic drawing of the human kallikrein locus on chromosome 19q, 56.0–56.3 Mb.
Approximate locations of genes are indicated by their symbols and arrowheads to mark the direction of transcription. The dashed lines surrounding *KLK3* depict the approximate location of the duplicated region that so far has only been detected in primate species. The arrows show the location of expanded regions in murine species that contain closely related *Klk1* paralogs.

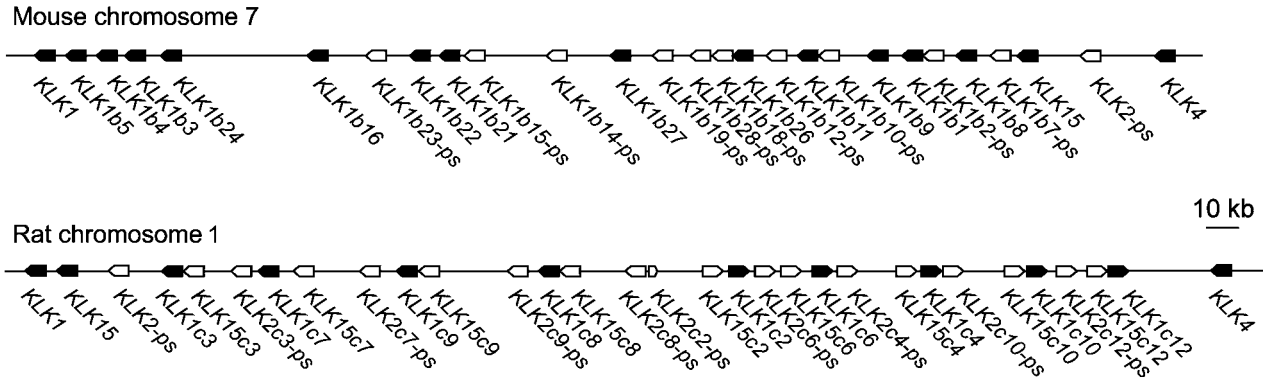

Figure 2 Illustration of the genomic region between *Klk1* and *Klk4*, encompassing taxon-specific genes in the mouse and the rat. Expressed genes are indicated by filled and pseudogenes by empty arrowheads.

Table 1 Proposed new nomenclature for mouse kallikrein 1-related peptidases – family b.

New gene symbol	Old symbol	New gene name	Alternative and old gene names	GenBank accession no.
Klk1	*mGK-6*	Kallikrein 1	Tissue kallikrein	NM_010639
Klk1b1	*mGK-1*	Kallikrein 1-related peptidase b1		NM_010645
Klk1b2-ps	*mGK-2*			AY152419
Klk1b3	*Ngfg*, *mGK-3*	Kallikrein 1-related peptidase b3	γ subunit of the 7S NGF complex	NM_008693
Klk1b4	*Ngfa*, *mGK-4*	Kallikrein 1-related peptidase b4	α subunit of the 7S NGF complex	NM_010915
Klk1b5	*mGK-5*	Kallikrein 1-related peptidase b5		NM_008456
Klk1b7-ps	*mGK-7*			AY152420
Klk1b8	*mGK-8*	Kallikrein 1-related peptidase b8		NM_008457
Klk1b9	*mGK-9*, *Egfbp3*	Kallikrein 1-related peptidase b9	EGF-BP type C, true EGF-BP	NM_010116
Klk1b10-ps	*mGK-10*			AY152421
Klk1b11	*mGK-11*	Kallikrein 1-related peptidase b11		NM_010640
Klk1b12-ps	*mGK-12*			AY152422
Klk1b14-ps	*mGK-14*			AY152423
Klk1b15-ps	*mGK-15*			AY152424
Klk1b16	*mGK-16*	Kallikrein 1-related peptidase b16	γ-renin	NM_008454
Klk1b18-ps	*mGK-18*			AY152426
Klk1b19-ps	*mGK-19*			AY152427
				AY152428
Klk1b21	*mGK-21*	Kallikrein 1-related peptidase b21		NM_010642
Klk1b22	*mGk-22*, *Egfbp1*	Kallikrein 1-related peptidase b22	β-NGF endopeptidase, EGF-BP type A	NM_010114
Klk1b23-ps	*mGK-23*			AY152429
Klk1b24	*mGK-24*	Kallikrein 1-related peptidase b24		NM_010643
Klk1b26	*mGK-26*, *mGK-13*, *Egfbp2*	Kallikrein 1-related peptidase b26	Prorenin-converting enzyme, EGF-BP type B	NM_010644
Klk1b27	*mGK-27*	Kallikrein 1-related peptidase b27		NM_020268
Klk1b28-ps	*mGK-28*			AY152425
Klk2-ps	*mGK-25*			AY152430

Proposed new nomenclature

The organization of the human kallikrein locus is schematically illustrated, with major discrepancies in mouse, rat and dog indicated (Figure 1). The human genes are depicted by the symbols that were introduced in a previous nomenclature paper (Diamandis et al., 2000). The same symbols are used in the new nomenclature, despite the fact that they do not acknowledge the close relationship for *KLK1–KLK3*. However, the gene names are changed for all but *KLK1*, which is still called kallikrein 1. The new names of *KLK2–KLK15* are kallikrein-related peptidase, followed by the number of the gene symbol, e.g., *KLK2* is kallikrein-related peptidase 2. The symbols such as hK1, hK2, etc. previously used to depict the protein should be avoided. To distinguish between the protein and the gene, the former is written in standard font (e.g., KLK2) and the latter in italics (e.g., *KLK2*), as recommended by HGNC. To distinguish a transcript of a gene, the relevant abbreviation is written as a prefix within parentheses, e.g., (mRNA)*KLK2* and (cDNA)*KLK2* to emphasize the message and complementary DNA of the gene for kallikrein-related peptidase 2. If the species needs to be specified, the codes established by SWISS-PROT should be used (http://www.expasy.ch/cgi-bin/speclist). The codes are written as a prefix within parentheses, e.g., (HUMAN)*KLK4* and (MOUSE)*Klk4* to distinguish between human and mouse *KLK4*; note that the gene symbols are written in capital letters, with the exception of the mouse and rat symbols, which are written with an initial capital letter followed by lower case letters. No species-specific prefix is therefore needed in articles only relating to the human and mouse genes.

The proposed new nomenclature for unique kallikrein 1-related peptidases in rodents is according to the proposal by Olsson et al. (2004a). They should be named kallikrein 1-related peptidase followed by a letter depicting the subfamily and the number from the old *GK* nomenclature, e.g., the gene cloned with the designation *mGK-5* has the new gene symbol *Klk1b5* and is called kallikrein 1-related peptidase b5. The proposed new nomenclature for murine kallikrein 1-related peptidases is displayed in Tables 1 and 2, with their location on the chromosome illustrated in Figure 2. The gene subfamilies seem to overlap with single or very closely related animal species, so that the b-family might be confined to *Mus musculus* and the c-family to *Rattus norwegicus* and perhaps also *Rattus rattus*. In non-rodent species, *KLK1* expansion is only known to occur in the horse, where the subfamily is designated by the letter d.

The canine gene with homology to the progenitor of *KLK2* and *KLK3* gives rise to the dog prostate arginine esterase. The proteolytic specificity of this enzyme is similar to that of KLK2, but not to that of PSA, which displays an expanded chymotrypsin-like activity (Chappelaine et al., 1984; Lazure et al., 1984; Malm et al., 2000). Thus, it is proper to assign the symbol *KLK2* to the gene for dog arginine esterase and, as a conse-

Table 2 Proposed new nomenclature of rat kallikrein 1-related peptidases – family c.

New gene symbol	Old symbols	New gene name	Alternative and old gene names	Associated pseudogenes	Accession no.
Klk1	*rGK-1*, *PS*	Kallikrein 1	Tissue kallikrein		M11563
Klk1c2	*rGK-2*, *RSKG-5*, *S2*, *rKLK2*	Kallikrein 1-related peptidase c2	Tonin		M11565
				Klk15c2-ps	BK001365[a]
				Klk2c2-ps	
Klk1c3	*rGK-3*, *RSKG-50*, *S1*, *rKLK3*	Kallikrein 1-related peptidase c3			M11564
				Klk15c3-ps	BK001366[a]
				Klk2c3-ps	BK001376[a]
Klk1c4	*rGK-4*, *rKLK4*	Kallikrein 1-related peptidase c4			L33839
				Klk15c4-ps	BK001367[a]
				Klk2c4-ps	BK001377[a]
Klk1c6	*rGK-6*, *rKLK6*	Kallikrein 1-related peptidase c6			BK001361[a]
				Klk15c6-p	BK001368[a]
				Klk2c6-ps	BK001378[a]
Klk1c7	*RSKG-7*, *K1*, *rK7*, *rKLK7*	Kallikrein 1-related peptidase c7	Esterase B		M19647
				Klk15c7-ps	BK001369[a]
				Klk2c7-ps	BK001379[a]
Klk1c8	*rGK-8*, *P1*, *rK8*, *rKLK8*	Kallikrein 1-related peptidase c8			M27215
				Klk15c8-ps	BK001370[a]
				Klk2c8-ps	BK001375[a]
Klk1c9	*S3*, *rK9*, *SEV*, *rKLK9*	Kallikrein 1-related peptidase c9			M11566
				Klk15c9-ps	BK001371[a]
				Klk2c9-ps	BK001380[a]
Klk1c10	*rK10*, *rKLK10*	Kallikrein 1-related peptidase c10	Endopeptidase k, T-kininogenase, proteinase B, antigen D3b region		S48142
				Klk15c10-ps	BK001372[a]
				Klk2c10-ps	BK001381[a]
Klk1c12	*RSKG-3*, *rKLK12*	Kallikrein 1-related peptidase c12			M19648
				Klk15c12-ps	BK001373[a]
				Klk2c12-ps	BK001382[a]
Klk2-ps					BK001374[a]

Each functional *Klk1* paralog on the chromosome is followed by associated pseudogenes that are paralogous with *Klk15* and *Klk2*, as illustrated in Figure 2.

[a]Third party annotation that has been removed from GenBank, but is still retrievable.

quence, also to the homologous gene in other species, such as the rodent pseudogenes that show equally strong similarity to *KLK2* and *KLK3*.

The nomenclature suggested here is based on our current understanding of genes at the kallikrein locus and may need to be updated in the future as our knowledge widens. If novel genes are discovered, they should have the same stem symbol, but with a novel number (e.g., *KLK16*). Genes created by duplication after the divergence of murine rodents from the lineage leading to primates are exemptions to the rule and should have a name based on the founder gene.

References

Berg, T., Bradshaw, R.A., Carretero, O.A., Chao, J., Chao, L., Clements, J.A., Fahnestock, M., Fritz, H., Gauthier, F., MacDonald, R.J., et al. (1992). A common nomenclature for members of the tissue (glandular) kallikrein gene families. Agents Actions *38* (Suppl. 1), 19–25.

Blaber, M., Isackson, P.J., and Bradshaw, R.A. (1987). A complete cDNA sequence for the major epidermal growth factor binding protein in the male mouse submandibular gland. Biochemistry *26*, 6742–6749.

Borgoňo, C.A., and Diamandis, E.P. (2004). The emerging roles of human tissue kallikreins in cancer. Nat. Rev. Cancer *4*, 876–890.

Borgoňo, C.A., Michael, I.P., and Diamandis, E.P. (2004). Human tissue kallikreins: physiologic roles and applications in cancer. Mol. Cancer Res. *2*, 257–280.

Bothwell, M.A., Wilson, W.H., and Shooter, E.M. (1979). The relationship between glandular kallikrein and growth factor-processing proteases of mouse submaxillary gland. J. Biol. Chem. *254*, 7287–7294.

Chapdelaine, P., Dube, J.Y., Frenette, G., and Tremblay, R.R. (1984). Identification of arginine esterase as the major androgen-dependent protein secreted by dog prostate and preliminary molecular characterization in seminal plasma. J. Androl. *5*, 206–210.

Diamandis, E.P., Yousef, G.M., Clements, J., Ashworth, L.K., Yoshida, S., Egelrud, T., Nelson, P.S., Shiosaka, S., Little, S., Lilja, H., et al. (2000). New nomenclature for the human tissue kallikrein gene family. Clin. Chem. *46*, 1855–1858.

Evans, B.A., Drinkwater, C.C., and Richards, R.I. (1987). Mouse glandular kallikrein genes. Structure and partial sequence analysis of the kallikrein gene locus. J. Biol. Chem. *262*, 8027–8034.

Frey, E.K., and Kraut, H. (1926). Über einen von der Niere ausgeschiedenen die Herztätigkeit anregenden Stoff. Hoppe-Seyler's Z. Physiol. Chem. *157*, 32–61.

Fukushima, D., Kitamura, N., and Nakanishi, S. (1985). Nucleotide sequence of cloned cDNA for human pancreatic kallikrein. Biochemistry *24*, 8037–8043.

Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J., Moss, P., Paeper, B., and Wang, K. (2000). Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region. Gene *257*, 119–130.

Harvey, T.J., Hooper, J.D., Myers, S.A., Stephenson, S.A., Ashworth, L.K., and Clements, J.A. (2000). Tissue-specific expression patterns and fine mapping of the human kallik-

rein (*KLK*) locus on proximal 19q13.4. J. Biol. Chem. *275*, 37397–37406.

Kraut, H., Frey, E.K., and Werle, E. (1930). Der Nachweis eines Kreislaufhormons in der Pankreasdrüse. Hoppe-Seyler's Z. Physiol. Chem. *192*, 1–21.

Lazure, C., Leduc, R., Seidah, N.G., Chretien, M., Dube, J.Y., Chapdelaine, P., Frenette, G., Paquin, R., and Tremblay, R.R. (1984). The major androgen-dependent protease in dog prostate belongs to the kallikrein family: confirmation by partial amino acid sequencing. FEBS Lett. *175*, 1–7.

Lundwall, A. and Lilja, H. (1987). Molecular cloning of human prostate specific antigen cDNA. FEBS Lett. *214*, 317–322.

Malm, J., Hellman, J., Hogg, P., and Lilja, H. (2000). Enzymatic action of prostate-specific antigen (PSA or hK3): substrate specificity and regulation by Zn^{2+}, a tight-binding inhibitor. Prostate *45*, 132–139.

Moreau, M.E., Garbacki, N., Molinaro, G., Brown, N.J., Marceau, F., and Adam, A. (2005). The kallikrein-kinin system: current and future pharmacological targets. J. Pharmacol. Sci. *99*, 6–38.

Olsson, A.Y. and Lundwall, A. (2002). Organization and evolution of the glandular kallikrein locus in Mus musculus. Biochem. Biophys. Res. Commun. *299*, 305–311.

Olsson, A.Y., Lilja, H., and Lundwall, A. (2004a). Taxon-specific evolution of glandular kallikrein genes and identification of a progenitor of prostate-specific antigen. Genomics *84*, 147–156.

Olsson, A.Y., Valtonen-Andre, C., Lilja, H., and Lundwall, A. (2004b). The evolution of the glandular kallikrein locus: identification of orthologs and pseudogenes in the cotton-top tamarin. Gene *343*, 347–355.

Schedlich, L.J., Bennetts, B.H., and Morris, B.J. (1987). Primary structure of a human glandular kallikrein gene. DNA *6*, 429–437.

Thomas, K.A., Baglan, N.C., and Bradshaw, R.A. (1981). The amino acid sequence of the γ-subunit of mouse submaxillary gland 7 S nerve growth factor. J. Biol. Chem. *256*, 9156–9166.

Wines, D.R., Brady, J.M., Pritchett, D.B., Roberts, J.L., and MacDonald, R.J. (1989). Organization and expression of the rat kallikrein gene family. J. Biol. Chem. *264*, 7653–7662.

Yousef, G.M., Chang, A., Scorilas, A., and Diamandis, E.P. (2000). Genomic organization of the human kallikrein gene family on chromosome 19q13.3–q13.4. Biochem. Biophys. Res. Commun. *276*, 125–133.

Yousef, G.M., and Diamandis, E.P. (2001). The new human tissue kallikrein gene family: structure, function, and association to disease. Endocr. Rev. *22*, 184–204.