Molecular Cloning of the Human *Kallikrein 15* Gene (*KLK15*)

UP-REGULATION IN PROSTATE CANCER*

George M. Yousef‡§, Andreas Scorilas‡§, Klaus Jung¶, Linda K. Ashworth‖, and Eleftherios P. Diamandis‡§

From the ‡Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario M5G 1X5, Canada, the §Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, M5G 1L5 Canada, the ¶Department of Urology, University Hospital Charite, Humboldt University, Berlin, 10048, Germany, and the ‖Biology and Biotechnology Program, Lawrence Livermore National Laboratory, Livermore, California 94551

Kallikreins are a subgroup of serine proteases with diverse physiological functions. Growing evidence suggests that many kallikreins are implicated in carcinogenesis. By using molecular cloning techniques, we identified a new human kallikrein gene, tentatively named *KLK15* (for *kallikrein 15* gene). This new gene maps to chromosome 19q13.4 and is located between the *KLK1* and *KLK3* genes. *KLK15* is formed of five coding exons and four introns, and shows structural similarity to other kallikreins and kallikrein-like genes. *KLK15* has three alternatively spliced forms and is primarily expressed in the thyroid gland and to a lower extent in the prostate, salivary, and adrenal glands and in the colon testis and kidney. Our preliminary results indicate that the expression of *KLK15* is up-regulated by steroid hormones in the LNCaP prostate cancer cell line. The *KLK15* gene is also up-regulated, at the mRNA level, in prostate cancer in comparison to normal prostatic tissue. *KLK15* up-regulation was found to be associated with more aggressive forms of prostate cancer. This newly discovered gene has the potential of being used as a diagnostic and/or prognostic marker for prostate cancer.

Kallikreins are a group of serine proteases that are found in diverse tissues and biological fluids. The term "kallikrein" was first introduced by Werle and colleagues (1, 2) who found high levels of their original isolates in the pancreas (in Greek, the "kallikreas"). Kallikreins are divided into two main groups; the plasma kallikrein, which is a single gene (3), and the tissue kallikreins, which are encoded by a large multigene family in rodents (4, 5). Until recently, the human kallikrein gene family was thought to consist of only three members (6). However, we and others have identified 11 new members of the kallikrein gene family (7–18). The progress in this area of investigation has recently been reviewed (7).

Prostate-specific antigen (PSA),[1] currently the most useful tumor marker for prostate cancer diagnosis and monitoring, is a member of the human kallikrein gene family of serine proteases (19, 20). In addition to PSA, human glandular kallikrein 2 (hK2, encoded by the *KLK2* gene) has been proposed as an adjuvant diagnostic marker for prostate cancer (21, 22). Moreover, accumulating evidence indicates that other members of the expanded kallikrein gene family may be associated with malignancy (7). The normal epithelial cell-specific 1 gene (*NES1*; *KLK10*, according to the approved human tissue kallikrein gene nomenclature) was found to be a novel tumor suppressor, which is down-regulated during breast cancer progression (23). Other gene family members, including zyme (*KLK6*), neuropsin (*KLK8*), and human stratum corneum chymotyrptic enzyme (HSCCE; *KLK7*) were also found to be differentially expressed in certain types of malignancies (24–26). The diagnostic usefulness of PSA in prostate cancer led us to speculate that other, related molecules might be valuable biomarkers for prostate cancer and other malignancies.

During our efforts to identify new kallikrein genes that might be involved in malignancy, we cloned a new gene, tentatively named *kallikrein 15* gene (*KLK15*). Here, we describe the identification of the new gene, its genomic and mRNA structure, its location in relation to other known kallikreins, and its tissue expression pattern. The *KLK15* gene has another three alternatively spliced forms, expressed in many tissues. Our preliminary data suggest that *KLK15* is up-regulated in prostate cancer and that it is under steroid hormone regulation in the LNCaP prostate cancer cell line. Higher expression of *KLK15* is associated with more aggressive (higher stage and higher grade) prostate tumors.

MATERIALS AND METHODS

Identification of the New Gene—We previously constructed the first contiguous map for the human kallikrein gene locus extending from the *KLK1* gene (centromere) to the *KLK14* gene (telomere) (7, 8, 11, 12, 27). Overlapping bacterial artificial chromosome (BAC) clones spanning this area were identified by screening of a human BAC library using different radiolabeled gene-specific probes. An area of ~300 kb of genomic sequence was established using different techniques, as described previously (11, 27). By performing an *Eco*RI restriction analysis, we were able to orient the kallikrein locus along the *Eco*RI restriction map of chromosome 19q13 available from the Lawrence Livermore National Laboratory (LLNL). A BAC clone that extends more centromerically (BC 781134) was then identified. Contigs of linear genomic sequences from this clone are available from the LLNL. Initially, we have used these contig sequences to predict the presence of novel genes, using

[1] The abbreviations used are: PSA, prostate-specific antigen; KLK, kallikrein; KLK-L, kallikrein-like; PCR, polymerase chain reaction; hK2, human glandular kallikrein 2; dNTPs, deoxynucleoside triphosphates; LLNL, Lawrence Livermore National Laboratory; RT, reverse transcription; EST, expressed sequence tag; TLSP, trypsin-like serine protease; HSCCE, human stratum corneum chymotryptic enzyme; *NES1*, normal epithelial cell-specific 1 gene; kb, kilobase(s); BAC, bacterial artificial chromosome; contig, group of overlapping clones.

TABLE I
Primers used for genomic PCR amplification

Gene	Primer name and sequence[a]		GenBank™ number
KLK1	KLK1-A	ATC CCT CCA TTC CCA TCT TT	L10038
	KLK1-B	CAC ATA CAA TTC TCT GGT TC	
KLK2	KLK2-A	AGT GAC ACT GTC TCA GAA TT	M18157
	KLK2-B	CCC CAA TCT CAC GAG TGC AC	
PSA	E5-A	GTC GGC TCT GGA GAC ATT TC	M27274
	E5-B	AAC TGG GGA GGC TTG AGT C	
KLK15	KLK15-F1	CTC CTT CCT GCT GGC ATC CA	AF242195
	KLK15-R1	ATC ACA CGG GTG GTC ATG TG	
	KLK15-F2	CAA GTG GCT CTC TAC GAG CG	
	KLK15-R2	GAC ACC AGG CTT GGT GGT GT	

[a] All primers are presented in $5' \rightarrow 3'$ direction.

bioinformatic approaches, as described previously (8, 12), and a putative new serine protease was identified. The sequence of the putative gene was then verified by different approaches, including sequencing, EST data base search, PCR screening of tissues, as described below.

Expressed Sequence Tag Searching—The predicted exons of the putative new gene were subjected to homology search using the BLASTN algorithm (28) from the National Center for Biotechnology Information web server against the human EST data base (dbEST). Clones with >95% homology were obtained from the IMAGE consortium (29) through Research Genetics Inc. (Huntsville, AL). The clones were propagated, purified as described elsewhere (30), and sequenced from both directions with an automated sequencer, using insert-flanking vector primers.

Prostate Cancer Cell Line and Hormonal Stimulation Experiments—The LNCaP prostate cancer cell line was purchased from the American Type Culture Collection (ATCC), Manassas, VA. Cells were cultured in RPMI media (Life Technologies, Inc., Gaithersburg, MD) supplemented with glutamine (200 mM), bovine insulin (10 mg/liter), fetal bovine serum (10%), antibiotics, and antimycotics, in plastic flasks, to near confluency. The cells were then aliquoted into 24-well tissue culture plates and cultured to 50% confluency. 24 h before the experiments, the culture media were changed into phenol red-free media containing 10% charcoal-stripped fetal bovine serum. For stimulation experiments, various steroid hormones dissolved in 100% ethanol were added into the culture media at a final concentration of 10^{-8} M. Cells stimulated with 100% ethanol were included as controls. The cells were cultured for 24 h, then harvested for mRNA extraction.

RT-PCR for the KLK15 Gene—Total RNA was extracted from the LNCaP cell line or from prostate tissues using Trizol reagent (Life Technologies, Inc.) following the manufacturer's instructions. RNA concentration was determined spectrophotometrically. 2 μg of total RNA was reverse-transcribed into first-strand cDNA using the Superscript preamplification system (Life Technologies, Inc.). The final volume was 20 μl. Based on the combined information obtained from the predicted genomic structure of the new gene and the EST sequences (see below), two gene-specific primers were designed (KLK15-F1 and KLK15-R1) (Table I), and PCR was carried out in a reaction mixture containing 1 μl of cDNA, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl$_2$, 200 μM dNTPs (deoxynucleoside triphosphates), 150 ng of primers, and 2.5 units of HotStar DNA polymerase (Qiagen Inc., Valencia, CA) on a PerkinElmer Life Sciences 9600 thermal cycler. The cycling conditions were 95 °C for 15 min to activate the *Taq* DNA polymerase, followed by 35 cycles of 94 °C for 30 s, 64 °C for 30 s, 72 °C for 1 min, and a final extension step at 72 °C for 10 min. Equal amounts of PCR products were electrophoresed on 2% agarose gels and visualized by ethidium bromide staining. All primers for RT-PCR spanned at least two exons to avoid contamination by genomic DNA. To verify the identity of the PCR products, they were cloned into the pCR 2.1-TOPO vector (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. The inserts were sequenced from both directions using vector-specific primers, with an automated DNA sequencer.

Tissue Expression—Total RNA isolated from 26 different human tissues was purchased from CLONTECH (Palo Alto, CA). We prepared cDNA as described above for the tissue culture experiments and used it for PCR reactions. Tissue cDNAs were amplified at various dilutions using two gene-specific primers (KLK15-F2 and KLK15-R1) (Table I). Due to the high degree of homology between kallikreins, and to exclude nonspecific amplification, PCR products were cloned and sequenced.

Prostate Cancer Tissues—Prostate tissue samples were obtained from 29 patients who had undergone radical retropubic prostatectomy for prostatic adenocarcinoma at the Charite University Hospital, Berlin, Germany. The patients did not receive any hormonal therapy before

surgery. The use of these tissues for research purposes was approved by the Ethics Committee of the Charite Hospital. Fresh prostate tissue samples were obtained from the cancerous and non-cancerous parts of the same prostates that had been removed. Small pieces of tissue were dissected immediately after removal of the prostate and stored in liquid nitrogen until analysis. Histological analysis from all the tissue pieces was performed as described previously (31) to ensure that the tissue was either malignant or benign. The tissues were pulverized with a hammer under liquid nitrogen, and RNA was extracted as described above, using Trizol reagent.

Statistical Analysis—Statistical analysis was performed with SAS software (SAS Institute, Cary, NC). The analysis of differences between *KLK15* expression in non-cancerous *versus* cancerous tissues from the same patient was performed with the non-parametric McNemar test. The binomial distribution was used to compute the significance level. Prostate tumor *KLK15* mRNA levels were qualitatively classified into two categories (KLK15-low and KLK15-high groups), and associations between KLK15 status and other variables were analyzed using the Fisher's exact test.

Structure Analysis—Multiple alignment was performed using the ClustalX software package and the multiple alignment program available from the Baylor College of Medicine (Houston, TX). Phylogenetic studies were performed using the Phylip software package. Distance matrix analysis was performed using the Neighbor-Joining/UPGMA program, and parsimony analysis was done using the Protpars program. Hydrophobicity study was performed using the Baylor College of Medicine search launcher. Signal peptide was predicted using the SignalP server. Protein structure analysis was performed by the SAPS (structural analysis of protein sequence) program.

RESULTS

Cloning of the KLK15 Gene—A contiguous map for the human kallikrein gene locus extending from the *KLK1* gene (centromere) to the *KLK14* gene (telomere) was previously established (7, 8, 11, 12, 27). To investigate the presence of other kallikrein-like genes centromeric to *KLK1*, a BAC clone (BC 781134) was obtained as described under "Materials and Methods." According to the published genomic sequence of prostate specific antigen (*PSA*) and human renal kallikrein (*KLK1*) genes, we designed gene-specific primers for each of these genes (Table I) and developed polymerase chain reaction (PCR)-based amplification protocols, which allowed us to generate specific PCR products with genomic DNA as a template. PCR screening of the BAC clone by these gene-specific primers indicated that this clone is positive for *KLK1* but negative to PSA, thus, confirming its location to be centromeric to PSA.

A putative new serine protease was predicted from the sequence of this clone by computer programs as described previously (12). This clone was digested, blotted on a membrane, and hybridized with gene-specific primers for the putative *KLK15* gene (according to the predicted sequence), and positive fragments were subcloned and sequenced to verify the structure of the putative gene. This putative gene sequence was then compared using the BLASTN program against the human EST data base, and two EST clones were identified (GenBank™ accession numbers AW274270 and AW205420). These two clones were 99% identical to the last exon and the 3'-untrans-

(ATG) TGG CTT CTC CTC ACT CTC TCC TTC CTG CTG GCA TCC ACA
  M   W   L   L   L   T   L   S   F   L   L   A   S   T
  G gtgaggtggccccaggaggggggccaggtctgtgggagcaggtg............................
  ...Intron1............ ....gcatcctctaccccttctcttag CA GCC CAG
                                                                      A   A   Q
GAT GGT GAC AAG TTG CTG GAA GGT GAC GAG TGT GCA CCC CAC
 D   G   D   K   L   L   E   G   D   E   C   A   P   H
TCC CAG CCA TGG CAA GTG GCT CTC TAC GAG CGT GGA CGC TTT
 S   Q   P   W   Q   V   A   L   Y   E   R   G   R   F
AAC TGT GGC GCT TCC CTC ATC TCC CCA CAC TGG GTG CTG TCT
 N   C   G   A   S   L   I   S   P   H   W   V   L   S
GCG GCC CAC TGC CAA AGC CG gtatgaaggcaggggctcagggtcctga
 A   A [H] C   Q   S   R
ggga...................... .. ....Intron 2 ....cgcactccactggcgggaaa
accactcgcccgcacag C TTC ATG AGA GTG CGC CTG GGA GAG CAC
                    F   M   R   V   R   L   G   E   H
AAC CTG CGC AAG CGC GAT GGC CCA GAG CAA CTA CGG ACC ACG
 N   L   R   K   R   D   G   P   E   Q   L   R   T   T
TCT CGG GTC ATT CCA CAC CCG CGC TAC GAA GCG CGC AGC CAC
 S   R   V   I   P   H   P   R   Y   E   A   R   S   H
CGC AAC GAC ATC ATG TTG CTG CGC CTA GTC CAG CCC GCA CGC
 R   N [D] I   M   L   L   R   L   V   Q   P   A   R
CTG AAC CCC CAG GTG CGC CCC GCG GTG CTA CCC ACG CGT TGC
 L   N   P   Q   V   R   P   A   V   L   P   T   R   C
CCC CAC CCG GGG GAG GCC TGT GTG GTG TCT GGC TGG GGC CTG
 P   H   P   G   E   A   C   V   V   S   G   W   G   L
GTG TCC CAC AAC GAG CCT GGG ACC GCT GGG AGC CCC CGG TCA
 V   S   H   N   E   P   G   T   A   G   S   P   R   S
CAA G gtgcgtgaaaggatggagctggat...........................Intron 3.................
 Q
ctccaagtccactgtcttccccag TG AGT CTC CCA GAT ACG TTG CAT
                          V   S   L   P   D   T   L   H
TGT GCC AAC ATC AGC ATT ATC TCG GAC ACA TCT TGT GAC AAG
 C   A   N   I   S   I   I   S   D   T   S   C   D   K
AGC TAC CCA GGG CGC CTG ACA AAC ACC ATG GTG TGT GCA GGC
 S   Y   P   G   R   L   T   N   T   M   V   C   A   G
GCG GAG GGC AGA GGC GCA GAA TCC TGT GAG gtcagagcctagagg
 A   E   G   R   G   A   E   S   C   E
ggccatcaggcggaagaagaggg................................Intron 4................ .. ..cct
gagacccccтctттtccccacag GGT GAC TCT GGG GGA CCC CTG GTC
                        G   D [S] G   G   P   L   V
TGT GGG GGC ATC CTG CAG GGC ATT GTG TCC TGG GGT GAC GTC
 C   G   G   I   L   Q   G   I   V   S   W   G   D   V
CCT TGT GAC AAC ACC ACC AAG CCT GGT GTC TAT ACC AAA GTC
 P   C   D   N   T   T   K   P   G   V   Y   T   K   V
TGC CAC TAC TTG GAG TGG ATC AGG GAA ACC ATG AAG AGG AAC
 C   H   Y   L   E   W   I   R   E   T   M   K   R   N
(TGA)CTATTCTAGCCTATCTCCTGTGCCCCTGACTGAGCAGAAGCCCCCACAGCTGGCCAGCAGCCC
CGCCTGACATGGAACAGAACGGAGCCATCCCCCAAGACCCTGTCCAAGGCCCAGATGTTAGCCAAGG
ACTTGTCCCACCTGAGGACAAAGCTGGCGCTCAAGGTCACCTGTTTAATGCCAAGATAACAAAGCGC
TGATCCAAGTTGCTCTGTAGGAATTTCTGTGACTTTTTTCTGGGGTCAAAGAGAAACCCCGAGACAC
TGTACACTGTTCCTTTTCACCCACCACCCCGATCCCTAGGTGAGGAGAAGCGGCTTGAAGCAGGGCT
CCATTCATTCAACACACATGACCACCCGTGTGATCTTGAACAAGAGGCCCAATCTCACTTCGCCTTG
GTTTCCTTATCTGTAAAATGAGACCATCTTATTGCTGACTTCAAAGGGCTGTTGTGAGGATTAAATG
AGATGATTCGTCTGAACTGATTAAAATCGTGTCTGGCACTGA

FIG. 1. **Genomic organization and partial genomic sequence of the KLK15 gene.** Intronic sequences are not shown except for the splice junction areas. Introns are shown with *lowercase letters* and exons with *capital letters*. The coding nucleotides are shown in *boldface* and the 3′-untranslated region follows the TGA stop codon (*encircled*). The translated amino acids of the coding region are shown underneath by a single letter abbreviation. For the full sequence, see GenBank® accession number AF242195. The start and stop codons are *encircled*, and the exon-intron junctions are *underlined*. The catalytic residues are *boxed*. The putative polyadenylation signal is *underlined*. The exact start of the first coding exon was not determined.

lated region of the gene, and the second EST ends with a stretch of 17 adenine (A) nucleotides that were not found in the genomic sequence, thus verifying the 3′-end of the gene and the position of the poly(A) tail.

To identify the full mRNA structure of the gene and to determine the exon/intron boundaries, PCR reactions were performed using primers located in different computer-predicted exons, using a panel of 26 human tissue cDNAs as templates. PCR products were sequenced. Two of these primers (KLK15-F1 and KLK15-R1) (Table I) were able to amplify the full coding region of the gene from different tissues. Comparing the mRNA with the genomic structure indicated the presence of a gene formed from five coding exons with four intervening introns. Translation of the mRNA sequence in all possible

reading frames revealed the presence of only one frame that gives an uninterrupted polypeptide chain, which also contains the highly conserved structural motifs of the kallikreins, as discussed below.

Structural Characterization of the KLK15 Gene—As shown in Fig. 1, the *KLK15* gene is formed of five coding exons and four intervening introns, although, as with other kallikrein genes, the presence of further upstream untranslated exon(s) could not be ruled out (17, 32, 33). All of the exon/intron splice sites conform to the consensus sequence for eukaryotic splice sites (34). The gene further follows strictly the common structural features of other members of the human kallikrein multigene family, as described below. The predicted protein-coding region of the gene is formed of 771 bp, encoding a deduced

```
                                                                              ↓
zyme       1   --------------------MKKLMV--------------VLSLIAAAWA----EEQ---------------
KLK-L4     1   -------------------MWPIAL--------------VIASLTLALSGGVSQES----------------
KLK-L6     1   -------------------MFLLLT--------------ALQVLAIAMTQS-QEDE----------------
TLSP       1   --MQRLRWLRDWKSSGRGLTAAKEPGARSSPLQAMRILQLILLALATGL-VGGE-----------------
KLK-L3     1   -------MK--LGLLCA---------------------LLSLLAGHG----WAD------------------
KLK15      1   -------------------MWLLLT--------------LSFLLASTAAQ---DG-----------------
NES1       1   MRAPHLHLSAASGARALARTLP----------------LLMAQLWAAEEAAL-LPQN---------------
KLK-L5     1   -------------------MGLSIFLL-----------LCVLGLSQAATPK-IFN-----------------
neuropsin  1   ------------MGRPRPRAAKTWMFL-----------LLLGGGAWAGHSRA-QEDK---------------
PSA        1   -------------------MWVPVV--------------FLTLSVTWIGAAPLIL-----------------
hK2        1   -------------------MWDLVL--------------SIALSVGCTGAVPLIQ-----------------
hK1        1   -------------------MWFLVL--------------CLALSLGGTGAAPPIQ-----------------
KLK-L2     1   --------MATARPPWMWVLCA----------------LITALLLGVTEHVLANNDVSCDHPSNTVPSGSNQDL
prostase   1   --------MATAGNPWGWFLG-----------------YLILGVAGSLVSG--------------------
HSCCE      1   -------MARSLLLLPLQILLL----------------SLALETAGEEAQG-------------------

                                    +                      +                    +
zyme      20   -------------NKLVHGG---------PCDKTSHPYQAALYTSG--HLLCGGVLIHPLWVLTAAHCKKP
KLK-L4    24   -------------SKVLNTNGTSGFLPGGYTCFPHSQPWQAALVQG--RLLCGGVLVHPKWVLTAAHCLKE
KLK-L6    23   -------------NKIIGGH---------TCTRSSQPWQAAILAGPRRRFLCGGALLSGQWVITAAHCGRP
TLSP      52   -------------TRIIK----------GFECKPHSQPWQAALFEKT--RLLCGATLIAPRWLLTAAHCLKP
KLK-L3    21   -------------TRAIG----------AEECRPNSQPWQAGLFHLT--RLFCGATLISDRWLLTAAHCRKP
KLK15     20   -------------DKLLE----------CBCGDKHSQPWQVALYERG--RFNCGASLISPHWVLSAAHCQSR
NES1      40   -------------DTRLDPEAY------GAPCARGSQPWQVVSLFNGL--SFHCAGVLVDQSWVLTAAHCGNK
KLK-L5    25   -------------GTECGRNSQPWQVGIFEGT--SLRCGGVLIDHRWVLTAAHCSGS
neuropsin 33   ----------------VLG---------GHECQPHSQPWQAALFQGQ--QLLCGGVLVGGNWVLTAAHCKKP
PSA       23   -------------SRIVG----------GWECEKHSQPWQVLVASRG--RAVCGGVLVHPQWVLTAAHCIRN
hK2       23   -------------SRIVG----------GWECEQHSQPWQAAVYSHG--WAHCGGVLVHPQWVLTAAHCLKK
hK1       23   -------------SRIVG----------GWECEQHSQPWQAALYHFS--TFQCGGILVHRQWVLTAAHCISD
KLK-L2    51   GAGAGEDARSDDSSSRIIN---------GSDCDMHTQPWQAALLLRP-NQLYCGAVLVHPQWLLTAAHCRKK
prostase  27   -------------SCSQIIN--------GEDCSPHSQPWQAALVME--NELFCSGVLVHPQWVLSAAHCFQN
HSCCE     28   -------------DKLIID---------GAPCARGSHPWQVAILSG--NQLHCGGVLVNERWVLTAAHCKMN
                                   ↑                                •          •      ••   ••••

zyme      67   NLQVFLGKHNLRQSSQEQSSVVRAVIHPDYDAAS-----------HDQDIMLLRLARPAK-LSELIQPLP
KLK-L4    81   GLKVYLCKHALG-RVEAGEQVREVVHSIPHPEYRRSPT----HL---NHDHIMLLELQSPVQ-LTGYIQTLP
KLK-L6    72   ILQVALGKHNLR-RWEATQQVLRVVRQVTHPNYNSRT---------HDNDLMLLQLQQPAR-IGRAVRPIE
TLSP      99   RYIVHLGQHNLQ-KEEGCEQTRTATESFPHPGFNNSLP----NK---DHRNLIMLVKMASPVS-ITWAVRPLT
KLK-L3    68   YLWVRLGEHHLW-KWEGPEQLFRVTDFFPHPGFNKDLS----AN---DHNDIMLIRIPQAR-LSPAVQPLN
KLK15     67   FMRVRLGEHNLR-KRDGPEQLRTTSRVIPHPRY----EA----R---SHRNDLMLLRLQPAR-LNPQVRPAV
NES1      91   PLWARVCDDHLL-LLQG-EQLRRTTRSVVHPKYHQGSGPILPRR---TDEHDLMLLRLARPVV-PGPRVRALQ
KLK-L5    67   RYWVRLGEHSLS-QLDWTEQIRHSGFSVTHPGYLGAS------T---SHEHDLRLIRLRLPVR-VTSSVQPLP
neuropsin 78   KYTVRLGDHSLQ-NKDGPEQEIPVVQSIPHPCYNSSD-----VE---DHNHDLMLLQLRDQAS-LGSKVKPIS
PSA       70   KSVILLGRHSLF-HPEDTGQVFQVSHSFPHPLYDMSLLKNRFLRPGDDSSHDLMLLRLSEPAE-LTDAVKVMD
hK2       70   NSQVWLGRHNLF-EPEDTGQRVPVSHSFPHPLYNMSLLKHQSLRPDEDSSHDLMLLRLSEPAK-ITDVVKVLG
hK1       70   NYQLWLGRHNLF-DDENTAQFVHVSESFPHPGFNMSLLENHTRQADEDYSHDLMLLRLTEPADTITDAVKVVE
KLK-L2   113   VFRVRLGHYSLSPVYESGQMFQGVKSIPHEGYSHPG----------HSNDLMLIKINRRIR-PTKDVRPIN
prostase  76   SYTIGLGLHSLEADQEPGSQMVEASLSVRHPEYNRPL----------LANDLMLIKLDESVS-ESDTIRSIS
HSCCE     75   EYTVHLGSDTLG---DRRAQRIKASKSFRHPGYSTQT----------HVNDLMLVKLNSQAR-LSSMVKKVR

                                              ..........
                          +            +                           +        +            +
zyme     127   LERDCSANT-TSCHILGWGKTA--------DG----DFPDTIQCAYIHLVSREECEHAYPGQITQNMLCAGDE
KLK-L4   145   LSHNNRLTPGTTCRVSGWGTTT---------SP--QVNYPKTLQCANIQLRSDEECRQVYPGKITDNMLCAGTK
KLK-L6   132   VTQACASPG-TSCRVSGWGTIS---------SP--IARYPASLQCVNINISPDEVCQKAYPRTITTPGMVCAGVP
TLSP     163   ISSRCVTAG-TSCLISGWGTS----------SP--QLRPHTLRCANITIIEHQKCENAYPGQITDNMVCASVQ
KLK-L3   132   ISSQTCVSPG-MQCLISGWGCAVS--------SP--KALFPVTLQCANISILENKLCHWAYPGHISDSMLCAGLW
KLK15    127   LPTRCPHPG-EACVVSGWGLVSHNEPGTAGSPRSQVSLPDTLHCANISIISDTSCDKSYPGRLTNTMVCAGAE
NES1     158   LPYRCAQPG-DQCQVAGWGTTA---------AR--RVKYNKGLTCSSITHLSPKECEVFYPGVVTNNMICAGL-
KLK-L5   129   LPNDCATAG-TECHVSGWGLTN----------HP--RNPFPDLLQCINLSIVSHATCHGVYPGRITSNMVCAGS
neuropsin141  LADHCTQPG-QKCTVSGWGTVT----------SP--RENFPDTLNCAEVKIFPQKKCEDAYPGQITDGMVCAGS
PSA      141   LPTQEPALG-TTCYASGWGSIE-----------P--EEFLTPKKLQCVDLHVISNDVCAQVHPQKVTKFMLCAGRW
hK2      141   LPTQEPALG-TTCYASGWGSIE-----------P--EEFLRPRSLQCVSLHLLSNDMCARAYSEKVTEFMLCAGLW
hK1      142   LPTEEPEVG-STCLASGWGSIE-----------P--ENFSFPDDLQCVDLKILPNDECKKAHVQKVTDFMLCVGHL
KLK-L2   174   VSSHCPSAG-TKCLVSGWGTTK----------SP--QVHPKVLQCINISVLSQKRCEDAYPRQIDDTMFCAG-D
prostase 137   IASQCPTAG-NSCLVSGWGLLA--------NG----RMPTVLQCVNVSVVSEEVCSKLYDPLYHPSMFCAGGG
HSCCE    133   LPSRCEPPG-TTCTVSGWGTTT----------SP--DVTFPSDLMCVDVKLISPQDCTKVVKDLLENSMLCAGIP
               •                                                       ••        •        ••

                       +          +                    +          +
zyme     187   KYGKDSCQGDSGGPLVCGDHLRGLVSWG-NIPCGSKEKPGVYTNVCRYTNWIQKTIQAK------------
KLK-L4   208   EGGKDSCEGDSGGPLVCDGVLRGITSWG-DFPCGQPDRPGVYTNVWIRETIRKYETQQQKWLKGPQ
KLK-L6   194   QGGKDSCQGDSGGPLVCRGQLQGLVSWG-MERCALPGYPGVYTNLCKYRSWIEETMRDK--------
TLSP     225   EGGKDSCQGDSGGPLVCNQSLQGIISWG-QDPCAITRKPGVYTKVCKYVDWIQEIMKNN--------
KLK-L3   194   EGGRGSCQGDSGGPLVCNGTLAGVVSGG-AEPCSRPRPAVYTSVCHYLDWIQEIMEN---------
KLK15    199   GRGAESCEGDSGGPLVCGGILQGLVSWG-DVPCDNTTKPGVYTKVCHYLEWIRETMKRN------
NES1     219   DRGQDPCQSDSGGPLVCDETLQGILSWG-VYPCGSAQHPAVVTQICKYMSWINKVIRSN------
KLK-L5   190   VPGQDACQGDSGGPLVCGGVLQGLVSWG-SVGPCGQDGIPGVYTYICKYVDWIRMIMRNN----
neuropsin202  SKGADTCQGDSGGPLVCDGALQGITSWG-SDPCGRSDKPGVYTNICRYLDWIKKIIGSKG----
PSA      203   TGGKSTCGGDSGGPLVCNGVLQGITSWG-SEPCALPERPSLYTKVVHYRKWIKDTIVANP----
hK2      203   TGGKDTCGGDSGGPLVCNGVLQGITSWG-PEPCALPERPAVYTKVVHYRKWIKDTIAANP----
hK1      204   EGGKDSCQGDSGGPLMCDGVLQGITSWG-YVPCGTPNKPSVAVRVLSYVKWIEDTIAENS----
KLK-L2   235   KAGRDSCQGDSGGPVVCNGSLQGIVSWG-DYPCARPNRPGVYTNLCKFTKWIQETIQANS----
prostase 197   HDQKDSCNGDSGGPLICNGYLQGLVSFG-KAPCGQVGVPGVYTNLCKFTEWIEKTVQAS------
HSCCE    195   DSKKNACNGDSGGPLVCRGTLQGLVSWG-TFPCGQPNDPGVYTQVCKFTKWINDTMKKHR----
          *    •  ••••••    •          •      •            •
```

FIG. 3. **Plot of hydrophobicity and hydrophilicity of the KLK15 protein, as compared with the prostate specific antigen (PSA).** Note the hydrophobic region at the amino terminus, suggesting presence of a signal peptide. For details, see text.
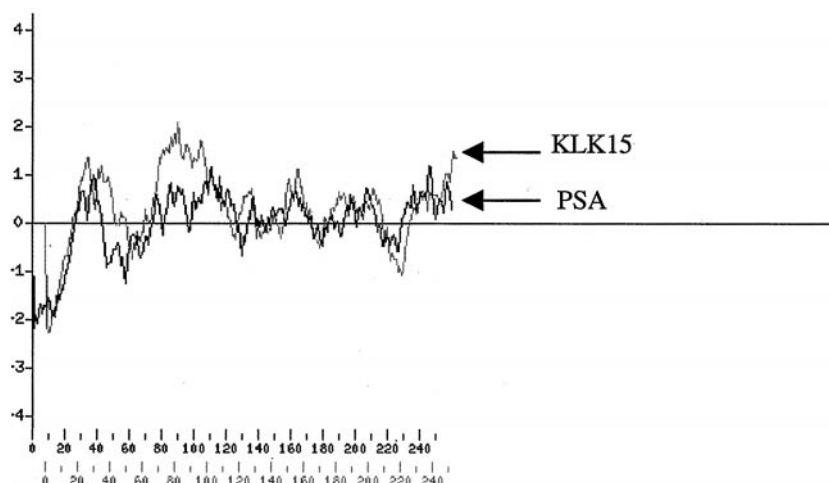
256-amino acid polypeptide with a predicted molecular mass of 28.1 kDa. The potential translation initiation codon matches the consensus Kozak sequence (35), moreover, there is a purine at position −3, which occurs in 97% of vertebrate mRNAs (36). It should also be noted that, like most other kallikrein-like genes, *KLK15* does not have the consensus G nucleotide at position +4.

Nucleotides 7764–7769 (ATTAAA) (the numbers refer to our GenBank™ submission number AF242195) closely resemble a consensus polyadenylation signal (37) and are followed, after 17 nucleotides, by the poly(A) tail. No other potential polyadenylation signals were discernible in the 3′-untranslated region, suggesting that the above sequence is the actual polyadenylation signal. Although AATAAA is highly conserved, natural variants do occur, and the ATTAAA sequence is reported to occur as a natural polyadenylation variant in 12% of vertebrate mRNA sequences (38). The presence of glutamic acid (E) at position 203 suggests that *KLK15* will likely possess a unique substrate specificity. PSA has a serine (S) residue in the corresponding position and has chymotryptic-like activity. Many other kallikreins usually have aspartate (D) in this position, indicating a trypsin-like activity (Fig. 2) (6).

Although the KLK15 protein sequence is unique, comparative analysis revealed that it has a considerable degree of homology with other members of the kallikrein multigene family. KLK15 shows 51% protein identity and 66% similarity with the trypsin-like serine protease (TLSP) and 49% and 48% identity with the neuropsin and KLK-L3 proteins, respectively. Hydrophobicity analysis revealed that the amino-terminal region is quite hydrophobic (Fig. 3), consistent with the possibility that this region harbors a signal sequence, analogous to other serine proteases. Computer analysis of the KLK15 protein sequence predicted a cleavage site between amino acids 16 and 17 (TAA-QD). Sequence alignment (Fig. 2) also revealed another potential cleavage site (Lys[21]), at a site homologous to the activation site of other serine proteases (lysine (K) or arginine (R) is present in most cases) (39). Several evenly distributed hydrophobic regions throughout the KLK15 polypeptide are consistent with a globular protein, similar to other kallikreins and serine proteases. Thus, as is the case with other kallikreins, and although direct experimental evidence is lack-
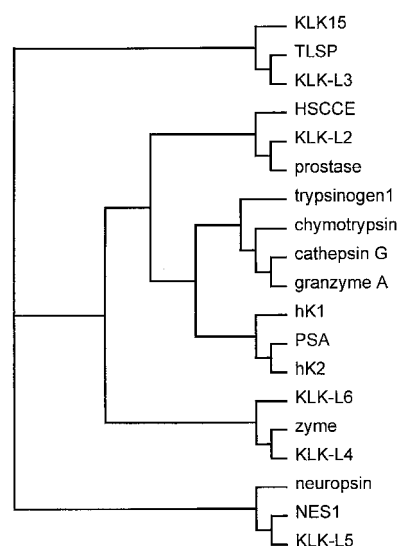

FIG. 4. **Dendrogram of the predicted phylogenetic tree for 15 kallikreins and a few other serine proteases.** The Neighbor-Joining method was used to align *KLK15* with other serine proteases and members of the kallikrein gene family. The tree grouped the classical kallikreins (hK1, hK2, and PSA) together and aligned *KLK15* in one group with *TLSP* and *KLK-L3* genes. Other serine proteases were aligned in different groups, as shown. For full gene names, see the abbreviations footnote.

ing, KLK15 is presumably translated as an inactive 256-amino acid preproenzyme precursor. Prepro-KLK15 has 21 additional residues, which constitute the preregion (the signal peptide formed of 16 residues), and the propeptide (5 residues).

The *dotted area* in Fig. 2 indicates an 11-amino acid loop characteristic of the classical kallikreins (PSA, KLK1, and KLK2) but not found in KLK15 or other members of the kallikrein multigene family (10, 11, 13, 14). However, KLK15 has a unique 8-amino acid loop (HNEPGTAG) at positions 148–155, not found in any other kallikrein (Fig. 2). 29 "invariant" amino acids surrounding the active site of serine proteases have been described (40). Of these, 28 are conserved in KLK15. One of the unconserved amino acids (Ser[173] instead of Pro) is also found in prostase, KLK-L2, and KLK-L5 proteins, and

Fig. 2. **Alignment of the deduced amino acid sequence of KLK15 with members of the kallikrein multigene family.** *Dashes* represent gaps to bring the sequences to better alignment. The residues of the catalytic triad (*H*, *D*, and *S*) are shown in *italics*. Identical amino acids are highlighted in *black* and similar residues in are *gray*. The 29 invariant serine protease residues are marked by the symbol ● on the *bottom* and the cysteine residues by the + symbol on *top* of each block. The predicted cleavage sites of the signal and activation peptides are indicated by *arrows*. The *dotted area* represents the kallikrein loop sequence. The trypsin-like cleavage pattern predicted by the presence of the Asp residue is indicated by an *asterisk*. *KLK15* has a Glu in this position. A unique 8-amino acid sequence, HNEPGTAG, is present at positions 148–155 of the *KLK15* gene.
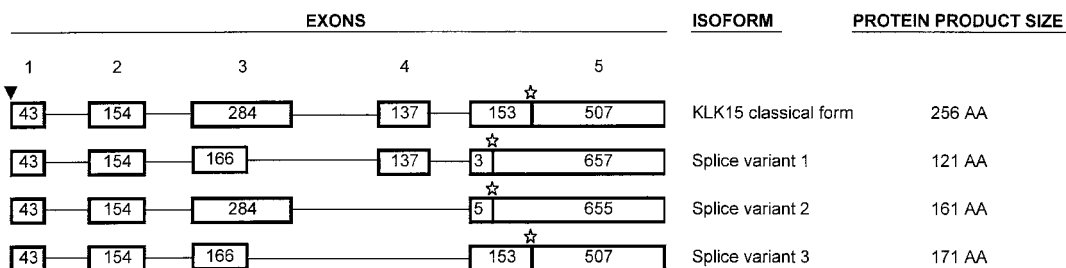
FIG. 5. **Schematic presentation of the different splice variants of the *KLK15* gene.** Exons are shown by *boxes* and introns by the *connecting lines*. *Numbers* inside *boxes* represent the exons lengths in base pairs. The *arrowhead* points to the common start codon and *stars* to the stop codon positions. The length of the predicted polypeptide product is indicated beside each variant in amino acids (*AA*). For full sequence information, see our GenBank® submission number AF242195. The alternative splicing and/or exon skips create a frameshift, which leads to a premature termination.
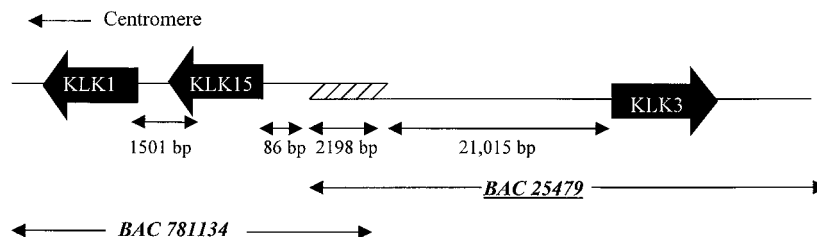
FIG. 6. **The relative locations of *KLK1*, *KLK15*, and *KLK3* genes on chromosome 19q13.4.** The two overlapping BAC clones are identified, and the overlap region is *hatched*. Genes are represented by *horizontal arrows* denoting the direction of the coding sequence. Distances between genes are given in base pairs. The figure is not drawn to scale.

represents a conserved evolutionary change to a protein of the same group, according to protein evolution studies (41). Twelve cysteine residues are present in the putative mature KLK15 protein; 10 of them are conserved in all kallikreins and would be expected to form disulfide bridges. The other two (Cys[131] and Cys[243]) are not found in *PSA*, *KLK1*, *KLK2*, or *KLK-L4*, however, they are found in similar positions in all other kallikrein genes and are expected to form an additional disulfide bond.

To predict the phylogenetic relatedness of the KLK15 protein with other serine proteases, the amino acid sequences were aligned together using the Unweighted Pair Group Method with Arithmetic mean (UPGMA) and the Neighbor-Joining distance matrix methods, and the Protpars parsimony method. All phylogenetic trees obtained agreed that other serine proteases (non-kallikreins) can be grouped together as a separate group, indicating that kallikreins represent a separate step in the evolution of serine proteases. *KLK15* was grouped with the *KLK-L3* and TLSP (Fig. 4) and the classical kallikreins (hK1, hK2, and PSA) are grouped together in all trees, suggesting that the separation between classical kallikreins and the kallikrein-like genes occurred early during evolution, consistent with suggestions of previous studies (13).

Splice Variants of the KLK15 Gene—PCR screening for *KLK15* transcripts using gene-specific primers (KLK15-F2 and KLK15-R2) (Table I) revealed the presence of three bands in most of the tissue cDNAs examined (Fig. 7). These bands were gel-purified, cloned, and sequenced. The *upper band* represents the classical form of the gene, and the *lower band* is splice variant 3 (Fig. 5). The *middle band* represents other two-splice variants. Restriction digestion of the PCR product of the *middle band* with *Stu*I, followed by gel separation, purification, and sequencing, revealed that it is composed of splice variants 1 and 2, which have approximately the same length (splice variant 1 has exon 4 (137 bp) but is missing 118 bp from exon 3, whereas splice variant 2 has an additional 118 bp of exon 3 but missing exon 4. All splice variants are expected to encode for truncated protein products (Fig. 5).

Chromosomal Localization of the KLK15 Gene—Restriction

analysis study of a number of overlapping BAC clones spanning the human kallikrein locus followed by comparison with the *Eco*RI restriction map of the area (available from the LLNL web site) enabled us to identify a BAC clone (BC 25479) that is telomerically adjacent to BC 781134 (which harbors the *KLK15* gene). A BLASTN search of the sequences of the two clones showed that the ends of these clones are overlapping. By identifying the position of the *KLK1*, *KLK3*, and *KLK15* genes along these clones, we were able to precisely define the relative location and the direction of transcription of these three genes. *KLK1* is the most centromeric, and its direction of transcription is from telomere to centromere, followed by *KLK15*, which is more telomeric and transcribes in the same direction. The distance between the two genes is a length of 1501 bp. The *KLK3* gene is more telomeric, located at a distance of 23,335 bp from the position of *KLK15*, and is transcribed in the opposite direction (Fig. 6). These results are consistent with previous reports where the distance between *KLK3* and *KLK1* was roughly estimated to be ~31 kb (6, 27).

Tissue Expression and Hormonal Regulation of the KLK15 Gene—As shown in Fig. 7, the *KLK15* gene is expressed at highest levels in the thyroid gland. Lower levels of expression are also seen in the prostate, salivary, and adrenal glands and in the colon, testis, and kidney. To verify the RT-PCR specificity, representative PCR products were cloned and sequenced. Fig. 8 shows that the *KLK15* gene is up-regulated by steroid hormones in the human LNCaP prostate cancer cell line.

KLK15 Expression in Prostate Cancer—The expression of the *KLK15* gene in normal and cancerous prostatic tissues was examined by RT-PCR. Actin was included as a control gene to ensure the quality and amount of the cDNA used. To examine the relative expression of the *KLK15* gene in normal compared with malignant tissues, we analyzed 29 pairs of prostatic tissues. Each pair represented normal and cancerous tissue obtained from the same patient. The intensity of the PCR band was compared between each pair of normal/cancerous tissue to establish if the expression was higher or lower in cancer, compared with normal tissue. The results are summarized in Table II. Thirteen out of 29 patients had significantly higher *KLK15*
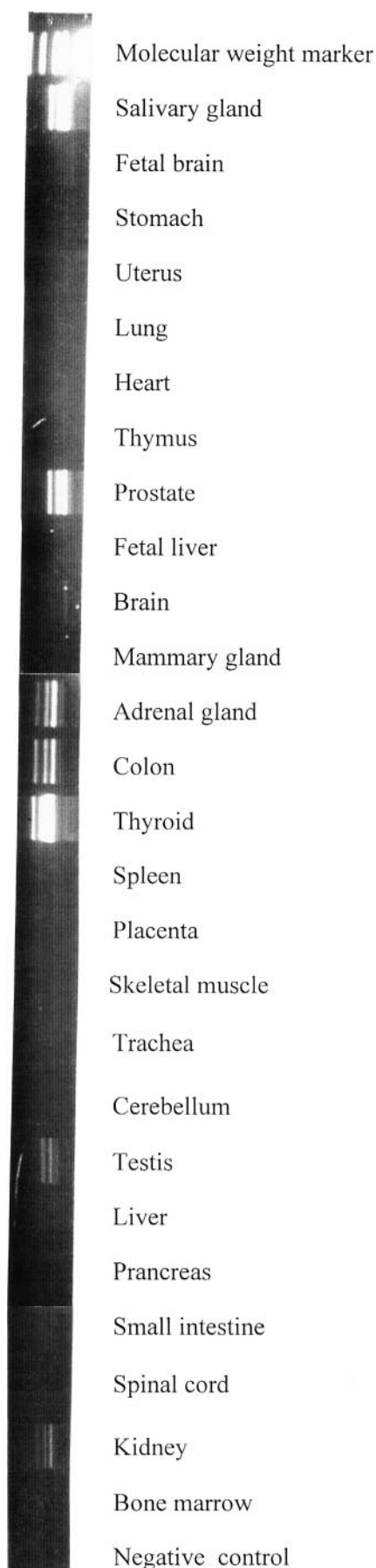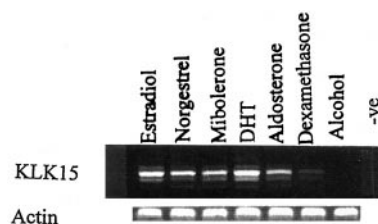
FIG. 8. **Hormonal regulation of the *KLK15* gene in the LNCaP prostate cancer cell line.** *DHT*, dihydrotestosterone. Steroids were added at 10^{-8} M final concentrations. *-ve*, negative control. Actin was used as a control gene. For more details, see text.

TABLE II

KLK15 expression in 29 pairs of cancerous and non-cancerous prostatic tissues

KLK15 expression	Number of patients	p value[a]
Higher in cancer *versus* normal	13	
Lower in cancer *versus* normal	3	
High expression but approximately equal in both tissues	8	
Low (or no) expression but approximately equal in both tissues	5	0.021

[a] The *p* value was calculated by the McNemar test using binomial distribution.

TABLE III

Relationship between KLK15 expression and other clinicopathological variables in 29 patients with primary prostate cancer

Variable	Patients	No. of patients (%)		p value[a]
		KLK15-negative	KLK15-positive	
		%		
Stage				
I/II	20	8 (40)	12 (60)	
III	9	0 (0)	9 (100)	0.033
Grade				
G1/2	23	8 (34.8)	15 (65.2)	
G3	6	0 (0)	6 (100)	0.15
Gleason score				
≤6	22	7 (31.8)	15 (62.2)	
>6	6	0 (0)	6 (100)	0.14
Unknown	1			

[a] Fisher's exact test.
[b] Numbers in parentheses are percentages.

expression in the cancer tissue and only three had the expression of KLK15 higher in non-cancer than to cancer tissues. Analysis by the McNemar test indicated that the differences between normal and cancerous tissues are statistically significant ($p = 0.021$). Because of the small number of cases, the binomial distribution was used to compute the significance level. We have further qualitatively classified the prostate cancer patients in two groups: (*a*) *KLK15* expression-positive ($n = 21$) and (*b*) *KLK15* expression-negative (or very low) ($n = 8$). When we compared the association of *KLK15* expression with clinicopathological prognostic variables, we found that higher *KLK15* expression was more frequent in patients with late stage disease and tumors of higher grade (Table III).

DISCUSSION

Kallikreins are a subgroup of serine proteases. The term "kallikrein" is usually utilized to describe an enzyme that acts upon a precursor molecule (kininogen) for release of a bioactive

FIG. 7. **Tissue expression of the *KLK15* gene, as determined by RT-PCR.** KLK15 is primarily expressed in the thyroid gland, and to a

lower extent in the prostate, salivary, and adrenal glands and in the colon, testis, and kidney. For an explanation of the multiple PCR bands (alternatively spliced forms) see text. PCR was performed with primers KLK15-F2 and KLK15-R1.
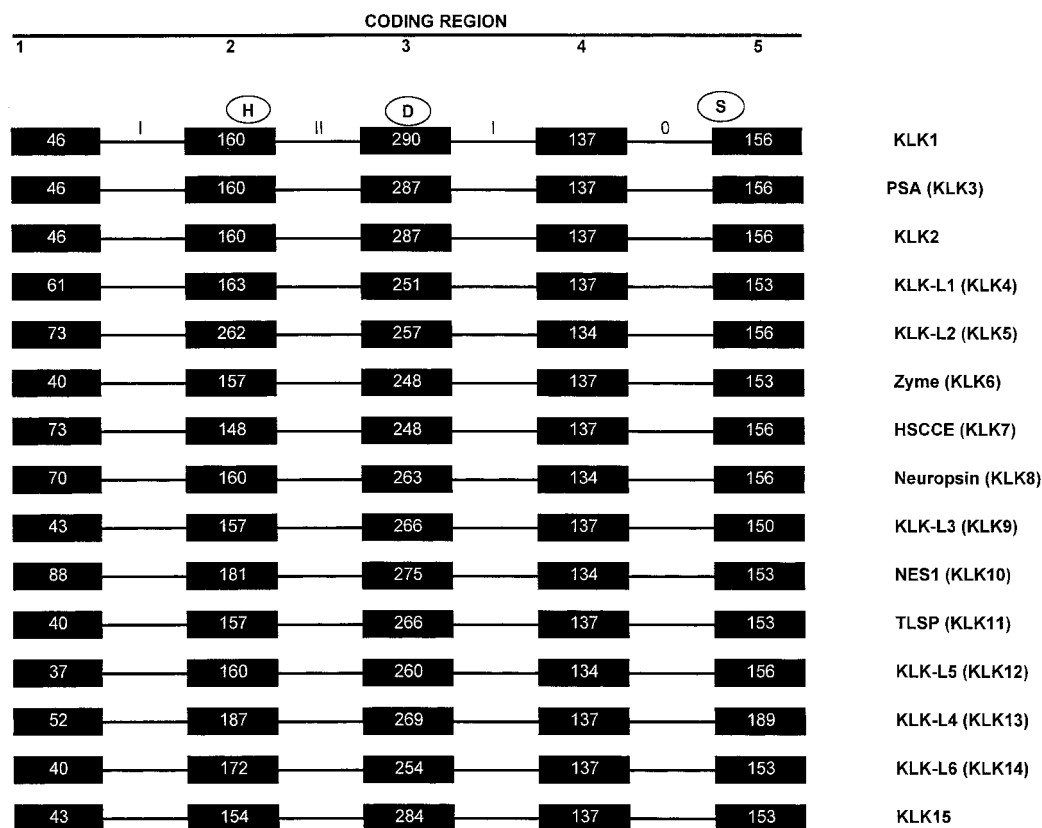
FIG. 9. **Schematic diagram showing the comparison of the coding regions of the 15 kallikrein genes.** Exons are shown by *solid bars* and introns by the *connecting lines*. *Letters* above *boxes* indicate relative positions of the catalytic triad that was found to be conserved in all genes; *H* denotes histidine, *D* aspartic acid, and *S* serine. *Roman numerals* indicate intron phases. The intron phase refers to the location of the intron within the codon; *I* denotes that the intron occurs after the first nucleotide of the codon, *II* the intron occurs after the second nucleotide, and *0* the intron occurs between codons. The intron phases are conserved in all genes. *Numbers* inside *boxes* indicate exon lengths in base pairs. *Names* inside *brackets* represent the official nomenclature approved by the Human Gene Nomenclature Committee. Untranslated 3′- and 5′-regions and 5′-untranslated exons are not shown.

peptide (kinin) (3, 42). However, the generic term "tissue kallikrein" is not restricted to the functional definition of the enzyme. This term is now used to describe a group of enzymes with a highly conserved gene and protein structure, which also co-localizes in the same chromosomal locus. Among the three classical human kallikrein genes, only *KLK1* encodes for a protein with potent kininogenase activity. The enzymes encoded by *KLK2* and *KLK3* genes have very weak kininogenase activity. The already cloned 14 members of the human kallikrein gene family have a number of similarities (7, 11) as follows: 1) All genes localize to the same chromosomal region (proximal19q13.4). 2) All genes encode for putative serine proteases with a conserved catalytic triad in the appropriate positions, *i.e.* histidine near the end of the second coding exon, aspartic acid in the middle of the third exon, and serine at the beginning of the fifth (last) exon. 3) All genes have five coding exons (some members contain one or more 5′-untranslated exons). 4) Coding exon sizes are similar or identical. 5) Intron phases are fully conserved. 6) All genes have significant sequence homologies at the DNA and amino acid levels (30–80%). 7) Many of these genes are regulated by steroid hormones.

Figs. 2 and 9 show that the newly identified *KLK15* gene shares all the above similarities and is thus suggested to be a new member of the human kallikrein multigene family. We named this gene *KLK15* based on the recommendations for the new kallikrein gene nomenclature, as approved by the Human Gene Nomenclature Committee and available on the Web (43).

Many kallikrein genes are related to the pathogenesis of human diseases, depending on the tissue of their primary ex-

pression. The *KLK1* gene is involved in many disease processes, including inflammation (3), hypertension (44), renal nephritis and diabetic renal disease (45, 46), The connections of HSCCE (*KLK7*) with skin diseases, including pathological keratinization and psoriasis, have already been reported (47, 48). Little *et al.* suggested that zyme (*KLK6*) is amyloidogenic and plays a role in the development of Alzheimer's disease (14). There are other reports describing connection of neuropsin (*KLK8*) expression with diseases of the central nervous system, including epilepsy (49, 50). Being primarily expressed in the thyroid, *KLK15* may play an important role in the normal physiology and pathophysiology of this gland. Among all other discovered kallikreins, many are expressed in the thyroid but none at the highest levels in this tissue (7, 11).

Our preliminary findings indicate that the *KLK15* gene is up-regulated, at the mRNA level, in a subset of prostate cancers. The distributions of *KLK15* qualitative expression status (high or low) between subgroups of patients differing by disease stage, tumor grade, and Gleason score indicated that high *KLK15* expression was found more frequently in grade 3 tumors as well as in stage III and Gleason score >6 patients. These findings suggest that overexpression of *KLK15* is associated with more aggressive forms of the disease and may be an indicator of poor prognosis. These findings require verification, because our patient population was relatively small (Table III).

There is now growing evidence that many kallikreins and kallikrein-like genes are related to malignancy. PSA is the best marker for prostate cancer so far (20). Recent reports suggest that hK2 (encoded by the *KLK2* gene) could be another useful diagnostic marker for prostate cancer (21, 51). *NES1* (*KLK10*)

appears to be a novel tumor suppressor gene (23). The zyme (*KLK6*) gene was shown to be differentially expressed in primary breast and ovarian tumors (24), and the human stratum corneum chymotryptic enzyme (HSCCE, *KLK7*) has been shown to be expressed at abnormally high levels in ovarian cancer (25). Another recently identified kallikrein-like gene, tentatively named the tumor-associated differentially expressed gene-14 (*TADG-14*/neuropsin) (*KLK8*) was found to be overexpressed in about 60% of ovarian cancer tissues (26). *prostase/KLK-L1/(KLK4)*, another newly discovered kallikrein-like gene, is speculated to be linked to prostate cancer (13). Two newly discovered kallikreins, *KLK-L4* (*KLK13*) and *KLK-L5* (*KLK12*), were also found to be down-regulated in breast cancer (10). Thus, extensive new literature suggests multiple connections of various kallikrein genes to many forms of human cancer.

The existence of multiple alternatively spliced mRNA forms is frequent among the kallikreins. Distinct RNA species are transcribed from the *PSA* gene, in addition to the major 1.6-kb transcript (19, 52, 53). Also, Reigman *et al.* (54) reported the identification of two alternatively spliced forms of the human glandular kallikrein 2 (*KLK2*) gene. A novel transcript of the tissue kallikrein gene (*KLK1*) was also isolated from the colon (55). Neuropsin, a recently identified kallikrein-like gene, was found to have two alternatively spliced forms, in addition to the major form (26, 56). *KLK-L4* was also found to have different alternatively spliced forms (10). Because the splice variants of *KLK15* have an identical 5′-sequence required for translation, secretion, and activation, it is possible to assume that they encode for a secreted protein (53). These proposals need experimental verification, and the role, if any, of these isoforms in tissue physiology and pathophysiology are worth examining in the future.

In conclusion, we characterized a new member of the human kallikrein gene family, *KLK15*, which maps to the human kallikrein locus (chromosome 19q13.4). This gene has three related splice forms in addition to the classical form. *KLK15* is expressed in a variety of tissues but predominantly in the thyroid, it appears to be up-regulated in more aggressive forms of prostate cancer, and its expression is influenced by steroid hormones. Because a few other kallikreins are already used as valuable tumor markers, we speculate that *KLK15* may also find similar clinical applications. This possibility, as well as the physiological function of this protein in various tissues need further investigation.

REFERENCES

1. Kraut, H., K., F. E., and Werle, E. (1930) *Physiol. Chem.* **189,** 97–106
2. Werle, E. (1934) *Biochem. Z.* **269,** 415–434
3. Clements, J. (1997) in *The Kinin System* (Farmer, S., ed) pp. 71–97, Academic Press, New York
4. Evans, B. A., Drinkwater, C. C., and Richards, R. I. (1987) *J. Biol. Chem.* **262,** 8027–8034
5. Ashley, P. L., and MacDonald, R. J. (1985) *Biochemistry* **24,** 4520–5427
6. Riegman, P. H., Vlietstra, R. J., Suurmeijer, L., Cleutjens, C. B., and Trapman, J. (1992) *Genomics* **14,** 6–11
7. Diamandis, E. P., Yousef, G. M., Luo, L. Y., Magklara, A., and Obiezu, C. V. (2000) *Trends Endocrinol. Metab.* **11,** 54–60
8. Yousef, G. M., Obiezu, C. V., Luo, L. Y., Black, M. H., and Diamandis, E. P. (1999) *Cancer Res.* **59,** 4252–4256
9. Yousef, G. M., and Diamandis, E. P. (1999) *J. Biol. Chem.* **274,** 37511–37516
10. Yousef, G. M., Chang, A., and Diamandis, E. P. (2000) *J. Biol. Chem.* **275,** 11891–11898
11. Yousef, G. M., and Diamandis, E. P. (2000) *Genomics* **65,** 184–194
12. Yousef, G. M., Luo, L. Y., and Diamandis, E. P. (1999) *Anticancer Res.* **19,** 2843–2852
13. Nelson, P. S., Gan, L., Ferguson, C., Moss, P., Gelinas, R., Hood, L., and Wang, K. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 3114–3119
14. Little, S. P., Dixon, E. P., Norris, F., Buckley, W., Becker, G. W., Johnson, M., Dobbins, J. R., Wyrick, T., Miller, J. R., MacKellar, W., Hepburn, D., Corvalan, J., McClure, D., Liu, X., Stephenson, D., Clemens, J., and Johnstone, E. M. (1997) *J. Biol. Chem.* **272,** 25135–25142
15. Liu, X. L., Wazer, D. E., Watanabe, K., and Band, V. (1996) *Cancer Res.* **56,** 3371–3379
16. Hansson, A., Stromqvist, M., Backman, A., Wallbrandt, P., Carlstein, A., and Egelrud, T. (1994) *J. Biol. Chem.* **269,** 19420–19426
17. Yoshida, S., Taniguchi, M., Hirata, A., and Shiosaka, S. (1998) *Gene* **213,** 9–16
18. Stephenson, S. A., Verity, K., Ashworth, L. K., and Clements, J. A. (1999) *J. Biol. Chem.* **274,** 23210–23214
19. Riegman, P. H., Vlietstra, R. J., van der Korput, J. A., Romijn, J. C., and Trapman, J. (1989) *Biochem. Cell Biol. Commun.* **159,** 95–102
20. Diamandis, E. P. (1998) *Trends Endocrinol. Metab.* **9,** 310–316
21. Stenman, U. H. (1999) *Clin. Chem.* **45,** 753–754
22. Partin, A. W., Catalona, W. J., Finlay, J. A., Darte, C., Tindall, D. J., Young, C. Y., Klee, G. G., Chan, D. W., Rittenhouse, H. G., Wolfert, R. L., and Woodrum, D. L. (1999) *Urology* **54,** 839–845
23. Goyal, J., Smith, K. M., Cowan, J. M., Wazer, D. E., Lee, S. W., and Band, V. (1998) *Cancer Res.* **58,** 4782–4786
24. Anisowicz, A., Sotiropoulou, G., Stenman, G., Mok, S. C., and Sager, R. (1996) *Mol. Med.* **2,** 624–636
25. Tanimoto, H., Underwood, L. J., Shigemasa, K., Yan Yan, M. S., Clarke, J., Parmley, T. H., and O'Brien, T. J. (1999) *Cancer* **86,** 2074–2082
26. Underwood, L. J., Tanimoto, H., Wang, Y., Shigemasa, K., Parmley, T. H., and O'Brien, T. J. (1999) *Cancer Res.* **59,** 4435–4439
27. Yousef, G. M., Chang, A., Scorilas, A., and Diamandis, E. P. (2000) *Biochem. Cell Biol. Commun.* **276,** 125–133
28. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
29. Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M. B. (1996) *Genomics* **33,** 151–152
30. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
31. Meyer, A., Jung, K., Lein, M., Rudolph, B., Schnorr, D., and Loening, S. A. (1997) *Int. J. Cancer* **74,** 630–636
32. Luo, L., Herbrick, J. A., Scherer, S. W., Beatty, B., Squire, J., and Diamandis, E. P. (1998) *Biochem. Cell Biol. Commun.* **247,** 580–586
33. Yousef, G. M., Luo, L. Y., Scherer, S. W., Sotiropoulou, G., and Diamandis, E. P. (1999) *Genomics* **62,** 251–259
34. Iida, Y. (1990) *J. Theor. Biol.* **145,** 523–533
35. Kozak, M. (1991) *J. Cell Biol.* **115,** 887–903
36. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8148
37. Proudfoot, N. J., and Brownlee, G. G. (1976) *Nature* **263,** 211–214
38. Sheets, M. D., Ogg, S. C., and Wickens, M. P. (1990) *Nucleic Acids Res.* **18,** 5799–5805
39. Keil, B. (1971) in *The Enzymes* (Boyer, P. D., ed) Vol. 3, 3rd Ed., pp. 249–275, Academic Press, New York
40. Dayhoff, M. O. (1978) *Natl. Biomed. Res. Found.* **5,** 79–81
41. Miyata, T., Miyazawa, S., and Yasunaga, T. (1979) *J. Mol. Evol.* **12,** 219–236
42. Bhoola, K. D., Figueroa, C. D., and Worthy, K. (1992) *Pharmacol. Rev.* **44,** 1–80
43. Diamandis, E. P., Yousef, G., Clements, J., et al., (2000) *Clin. Chem.* **46,** 1855–1858
44. Margolius, H. S., Horwitz, D., Pisano, J. J., and Keiser, H. R. (1974) *Circ. Res.* **35,** 820–5
45. Jaffa, A. A., Chai, K. X., Chao, J., Chao, L., and Mayfield, R. K. (1992) *Kidney Int.* **41,** 789–795
46. Cumming, A. D., Walsh, T., Wojtacha, D., Fleming, S., Thomson, D., and Jenkins, D. A. (1994) *Clin. Sci.* **87,** 5–11
47. Sondell, B., Dyberg, P., Anneroth, G. K., Ostman, P. O., and Egelrud, T. (1996) *Acta Dermato-Venereol.* **76,** 177–181
48. Ekholm, E., and Egelrud, T. (1999) *Arch. Dermatol. Res.* **291,** 195–200
49. Momota, Y., Yoshida, S., Ito, J., Shibata, M., Kato, K., Sakurai, K., Matsumoto, K., and Shiosaka, S. (1998) *Eur. J. Neurosci.* **10,** 760–764
50. Kishi, T., Kato, M., Shimizu, T., Kato, K., Matsumoto, K., Yoshida, S., Shiosaka, S., and Hakoshima, T. (1999) *J. Biol. Chem.* **274,** 4220–4224
51. Black, M. H., Magklara, A., Obiezu, C. V., Melegos, D. N., and Diamandis, E. P. (1999) *Clin. Chem.* **45,** 790–799
52. Riegman, P. H., Klaassen, P., van der Korput, J. A., Romijn, J. C., and Trapman, J. (1988) *Biochem. Cell Biol. Commun.* **155,** 181–188
53. Heuze, N., Olayat, S., Gutman, N., Zani, M. L., and Courty, Y. (1999) *Cancer Res.* **59,** 2820–2824
54. Riegman, P. H., Vlietstra, R. J., van der Korput, H. A., Romijn, J. C., and Trapman, J. (1991) *Mol. Cell. Endocrinol.* **76,** 181–190
55. Chen, L. M., Murray, S. R., Chai, K. X., Chao, L., and Chao, J. (1994) *Braz. J. Med. Biol. Res.* **27,** 1829–1838
56. Mitsui, S., Tsuruoka, N., Yamashiro, K., Nakazato, H., and Yamaguchi, N. (1999) *Eur. J. Biochem.* **260,** 627–634