

Unusual Alternative Splicing within the Human Kallikrein Genes *KLK2* and *KLK3* Gives Rise to Novel Prostate-specific Proteins*

Received for publication, March 14, 2001, and in revised form, February 4, 2002
Published, JBC Papers in Press, February 7, 2002, DOI 10.1074/jbc.M102285200

Anat David‡§, Nicola Mabjeesh¶, Idit Azar‡§, Sharon Biton‡§, Sharon Engel‡, Jeanne Bernstein‡, Jacob Romano‡, Yoav Avidor¶, Tova Waks¶, Zelig Eshhar¶, Salomon Z. Langer‡, Beatriz Lifschitz-Mercer**, Haim Matzkin¶, Galit Rotman‡ ‡‡, Amir Toporik‡, Kinneret Savitsky‡, and Liat Mintz§§

From ‡Compugen Ltd., 72 Pinchas Rosen St., Tel Aviv 69512, the ¶Department of Immunology, Weizmann Institute of Science, Rehovot 76100, the ¶Department of Urology and the **Department of Pathology, Tel Aviv Sourasky Medical Center, 6 Weizmann St., Tel Aviv 64239, Israel, and §§Compugen Inc., Jamesburg, New Jersey 08831

Prostate-specific antigen (PSA) and human kallikrein 2 are closely related products of the human kallikrein genes *KLK3* and *KLK2*, respectively. Both PSA and human kallikrein 2 are produced and secreted in the prostate and have important applications in the diagnosis of prostate cancer. We report here the identification of unusual mRNA splice variants of the *KLK2* and *KLK3* genes that result from inclusion of intronic sequences adjacent to the first exon. The novel proteins encoded by these transcripts, named PSA-linked molecule (PSA-LM) and hK2-linked molecule (K-LM), share only the signal peptide with the original protein product of the respective gene. The mature proteins are entirely different and bear no similarity to the kallikrein family or to other proteins in the databases. As is the case with PSA, PSA-LM is expressed in the secretory epithelial cells of the prostate and is up-regulated in response to androgenic stimulation. A similar pattern of expression is suggested for K-LM.

Prostate-specific antigen (PSA)¹ is considered the best cancer biomarker currently available and is widely used for screening, diagnosis, and monitoring of prostate cancer (1, 2). PSA (also named human kallikrein 3, hK3) is encoded by the *KLK3* gene (3), a member of the human kallikrein gene family (4). These genes are clustered on chromosome 19q13.3–13.4 and have a highly conserved structural organization (4, 5). Another member of the kallikrein gene family, *KLK2*, displays 80% homology with *KLK3* (3, 6) and is located adjacent to it (4). The *KLK2* protein product, human kallikrein 2 (hK2; previously known as human glandular kallikrein 1, or hGK1), is emerging as an additional prostatic tumor marker (7).

The kallikrein genes code for a group of serine proteases in-

involved in post-translational processing of polypeptide precursors to their bioactive or inactive forms (4). Both PSA and hK2 are regulated by androgens and are synthesized and secreted by the epithelial cells of the prostate tubules (8, 9). hK2 is responsible for the activation of PSA by cleaving its pro-form to the enzymatically active mature form (10, 11). The mature, catalytically active form of PSA, a glycoprotein of 237 amino acids, has a chymotrypsin-like activity. PSA acts on semenogelin I and II, the major gel-forming proteins in seminal plasma, causing lysis of the seminal clot immediately after ejaculation (12).

Alternative splicing is a common mechanism for producing functionally diverse proteins from a single gene (13). In addition to the major 1.6-kb mRNAs, multiple transcripts were found to originate from the *KLK3* and *KLK2* genes (14, 15), suggesting the occurrence of extensive splicing events. Indeed, several alternatively spliced variants have been reported for both genes (15–19). The identity and function of the products encoded by the various *KLK2* and *KLK3* mRNA species are not known. Some of the transcripts formed by alternative splicing events in these genes are predicted to encode different isoforms of the original PSA and hK2 proteins.

In this study, we identified alternative transcripts in each of these genes that arise from unusual splicing events in which sequences from the first intron remain linked to the first exon in the mature mRNA, creating new ORFs that encode for novel polypeptides. These proteins, named PSA-LM and K-LM, share only the signal peptide with PSA and hK2, respectively. In their mature form, their amino acid sequences differ completely from the kallikrein family and other proteins in the databases. Analysis of the pattern of expression of these novel alternative transcripts and their protein products suggests that, like PSA and hK2, they are up-regulated by androgens and are expressed primarily in the secretory cells of the prostate epithelium.

EXPERIMENTAL PROCEDURES

Identification of Alternative Transcripts—The LEADSTM platform for clustering and assembly of genomic sequences, cDNAs and ESTs (www.labonweb.com/sitehtml/leads_overview_toc.html), was used to generate the LEADS Transcriptome Database, which includes predicted transcripts of human genes. The LEADS software models important parts of the transcription process, including alternative splicing. The LEADS databases built from GenBankTM versions 106.0 through 120.0 were searched for prostate-specific novel genes and splice variants. For this purpose, clusters were automatically screened according to the source tissue of the ESTs and cDNAs. Several clusters were composed mostly of transcripts from prostate tissues. The resulting clusters were then subjected to bioinformatic annotation and manually inspected. This analysis led to the identification of several ESTs that support the existence of alternative transcripts of the *KLK3* gene, linking exon 1

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) AF335478, AF335477, and AF336106.

§ These authors contributed equally to this work.

‡‡ To whom correspondence should be addressed: Compugen Ltd., 72 Pinchas Rosen St., Tel Aviv 69512, Israel. E-mail: galitr@compugen.co.il.

¹ The abbreviations used are: PSA, prostate-specific antigen; hK2, human kallikrein 2; *KLK2* and *KLK3*, human kallikrein genes 2 and 3; ESTs, expressed sequence tags; ORF, open reading frame; UTR, untranslated region; RT, reverse transcription; RACE, rapid amplification of cDNA ends; BPH, benign prostatic hyperplasia.



FIG. 1. Schematic representation of the splicing events in the *KLK3* gene that generate the PSA mRNA (top) and the alternative transcripts encoding PSA-LM (bottom). Alternative splicing of intron 1 is shown in dark blue. One of the PSA-LM transcripts includes all of intron 1, whereas the other is derived by alternative splicing within intron 1 (shown in light blue). Coding exons (orange) and untranslated sequences (yellow) are boxed, and their sizes, in base pairs, are indicated. Exons and introns are not drawn to scale. PSA exons 1–5 are marked. The red and black arrowheads indicate initiation and stop codons, respectively, of the PSA and PSA-LM ORFs.

with the adjacent intronic sequences, and coding for a putative novel protein (Fig. 1). Among these ESTs are AA506939, AA524970, AA528287, AI525636, and AI557311.

RNA Isolation—Total RNA was extracted from cells using the Tri-Reagent (Molecular Research Center Inc., Cincinnati, OH). Poly(A) RNA was isolated from total RNA using (dT)₂₅ Dynabeads (Dynal, Oslo, Norway).

RT-PCR—Reverse transcription (RT) was carried out in a final volume of 20 μ l using 2 μ g of total RNA from human tissues or cell lines and 2.5 units of Superscript II Reverse Transcriptase (Invitrogen), in the buffer supplied by the manufacturer and supplemented with 10 pmol of dT₁₅ (Promega, Madison, WI) and 30 units of RNasin (Promega, Madison, WI). PCR was carried out using 1 μ l of the RT reaction, in the presence of 2 mM dNTPs, 25 pmol of primers, and 2.5 units of DNA polymerase mix of the Expand Long Template PCR System (Roche Molecular Biochemicals) in the reaction buffer supplied by the manufacturer. The following primers were used: PSA-LM-specific primers were the forward primer A (5'-CCGGAGAGCTGTGTCACCATGTG-3'), spanning the initiation codon of the PSA-LM ORF, and the reverse primer E (5'-AAGCAGCTGGGAGCAGGGAGCTG-3'), at positions 287–309 within the PSA-LM ORF. K-LM-specific primers were the forward primer H (5'-CTCTCCATCGCCTTGTCTGTGGG-3') at position 16–38 of this ORF, and the reverse primer L (5'-GGACAGGGTTGAATGCTC-3') located at nucleotides 19–36 immediately downstream of the stop codon. PCR products were analyzed on 1.5% agarose gels, and their identity was verified by DNA sequencing.

RACE Analysis of 5' and 3' Ends—Rapid amplification of cDNA ends (RACE) analysis was performed on poly(A) RNA from human prostate tissue using the Marathon cDNA Amplification Kit (CLONTECH, Palo Alto, CA). Adaptor-ligated, double-stranded cDNA libraries were prepared essentially as suggested by the manufacturer. Superscript II Reverse Transcriptase (Invitrogen) was used for the first strand synthesis. First round PCR was performed on these libraries using the Expand Long Template PCR System (Roche Molecular Biochemicals). All PCRs were carried out in the presence of 1.5 μ g of anti-Taq monoclonal antibody (CLONTECH, Palo Alto, CA). PCR products were separated on 1.5% agarose gels, purified, and sequenced. A nested PCR approach was used to isolate 5'- and 3'-RACE products. For 5'-RACE, first round PCR was carried out with an adaptor-specific primer and the PSA-LM-specific reverse primer E (described above). Nested PCR was then performed on these reactions using the reverse primer C (5'-CCCCAACCATGGCCCTCTCAC-3') at positions 47–68 in the PSA-LM ORF. For 3'-RACE, first round PCR was performed with the forward primer A (described above), and nested PCR with the forward primer B (5'-GTGAGAGGGGCCATGGTTGGGG-3') derived from positions 47–68 of the PSA-LM ORF.

Northern Blot Analysis—Multiple Tissue Northern blot (MTNTM, Human II), containing poly(A) RNA samples from a variety of tissues, was purchased from CLONTECH (Palo Alto, CA). Poly(A) RNA (2 μ g) of LNCaP cells was fractionated by electrophoresis on 1.2% agarose gels containing formaldehyde and blotted onto Nytran Super Charge membranes (Schleicher & Schuell). Equal loading was verified by ethidium bromide staining of the gels. Probes were labeled using the Random Primer DNA Labeling Mix (Biological Industries, Beit Haemek, Israel) and [α -³²P]dCTP. Hybridization was carried out in the EZ-Hybridization Solution (Biological Industries, Beit Haemek, Israel) at 68 °C for 18 h. The membranes were rinsed twice with 2 \times SSC, 0.1% SDS at room temperature, followed by two washes with 0.1 \times SSC, 0.1% SDS at 50 °C. Autoradiograms were obtained by exposing the membranes to x-ray films. Quantification was performed by densitometric analysis of hybridizing bands from scanned autoradiograms.

The following probes were obtained by RT-PCR. PSA probe was a product of a forward primer (5'-CTTGCTGGGTCCGACAGCCTG-3') and a reverse primer (5'-GGTCCCCAGTGTGCTGCTCTGG-3'), both derived from exon 3 of the *KLK3* gene. The PSA-LM probe was a product of the forward primer D (5'-CCAACCCAGCACCCAGCCCA-

3') at positions 121–141 within the PSA-LM ORF and the reverse primer E (see above). The K-LM probe was obtained with the forward primer J (5'-CCCAGCCTCGTCCCTTCAGCC-3') and the reverse primer K (5'-GGGACCAGTGGCATGGAGAGAGG-3'), at positions 122–142 and 347–369, respectively, within the K-LM ORF. A 780-bp probe used for the analysis of hormonal regulation of PSA-LM was derived from intron 1 sequences by RT-PCR, using the forward primer D and the reverse primer M (5'-GGTGCTAGGAACCCAGGGA-TAAGTG-3').

In Situ Hybridization—A PSA-LM-specific fragment was obtained by RT-PCR using the forward primer D (see above) and the reverse primer F (5'-GGAGGTGGACCAGTGATACGGTCCAAG-3') derived from the 3'-untranslated region of PSA-LM at positions 81–107 downstream of the stop codon. This fragment was then cloned into the pT-Adv vector (CLONTECH). Clones in opposite orientations were used to prepare sense and antisense riboprobes by digoxigenin RNA labeling kit (Roche Molecular Biochemicals). Prostate sections were prepared as described for immunohistochemistry (see below). *In situ* hybridization was performed essentially according to the procedure described by Roche Molecular Biochemicals (20). After deparaffinization and rehydration, sections were digested with proteinase K (10 μ g/ml) in 10 mM Tris, pH 8.0, 1 mM EDTA, for 20 min at 37 °C. Hybridization with 0.2–1 μ g/ml digoxigenin-labeled riboprobes was carried out in hybridization buffer (2 \times SSC, 50% formamide, 0.02% SDS, 10% dextran sulfate, and 0.25 mg/ml salmon sperm DNA) for 4–5 h at 60 °C. Hybridization was followed by three 20-min washes with 1 \times SSC and 50% formamide at 55 °C, and two 15-min washes with 1 \times SSC at room temperature. Alkaline phosphatase-conjugated anti-digoxigenin antibody was used at 1:500 dilution for 90 min at room temperature, and detection was obtained by the addition of nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate for 15 min at room temperature.

Cell Cultures and Hormonal Stimulation—The human prostatic adenocarcinoma cell line LNCaP (ATCC, Manassas, VA) was propagated in RPMI 1640 containing 10% fetal calf serum. When cells reached 70% confluency, the culture medium was replaced by phenol red-free RPMI 1640 supplemented with 2% charcoal-stripped fetal calf serum for 24 h, followed by addition of 10 nM methyltrienolone (R1881, PerkinElmer Life Sciences) or testosterone (Fluka) for periods of 9–48 h.

Expression Constructs and Cell Transfections—The complete coding sequences of K-LM or PSA-LM were amplified by RT-PCR and cloned into the pcDNA4/Myc-His mammalian expression vector (Invitrogen), enabling the synthesis of proteins fused to the Myc and His epitope tags at the C terminus. LNCaP cells, at about 70% confluency, were transfected with these expression constructs using LipofectAMINE (Invitrogen), according to the manufacturer's instructions. After 48 h, spent media were collected, and cells were washed with phosphate-buffered saline, harvested, and lysed in 50 mM Tris, pH 7.5, 150 mM NaCl, 1% Triton X-100, containing a protease inhibitor mixture (CompleteTM, Roche Molecular Biochemicals). Proteins from the culture medium were immunoprecipitated using c-Myc-conjugated agarose beads (9E10 AC, Santa Cruz Biotechnology Inc., Santa Cruz, CA).

Antibodies—Peptides corresponding to amino acid residues 33–51 of PSA-LM (NH₂-CQAEALSPPTQHPSPDREL-COOH) and residues 85–104 of K-LM (NH₂-PSQIPAPSCFTKEQVPRHLC-COOH) were synthesized. Polyclonal antibodies were raised by immunizing rabbits with the PSA-LM and K-LM peptides conjugated to bovine serum albumin and keyhole limpet hemocyanin, respectively. The anti-PSA monoclonal antibody used for immunoblot analysis was purchased from Santa Cruz Biotechnology, Inc. (A67-B/E3). The anti-PSA monoclonal antibody used for immunohistochemistry was purchased from Zymed Laboratories Inc. (San Francisco, CA). A mouse monoclonal anti-c-Myc antibody (9E10) was purchased from Santa Cruz Biotechnology.

Western Blot Analysis—Tissue extracts, prepared in RIPA buffer, and cell lysates were fractionated by SDS-PAGE on 15% gels, transferred onto polyvinylidene difluoride membranes (Bio-Rad), and subjected to immunodetection using a commercial anti-Myc antibody or the

immunized sera described above. Peroxidase-conjugated donkey anti-rabbit or anti-mouse IgGs (Jackson ImmunoResearch Laboratories, West Grove, PA) were used as secondary antibodies. Proteins were visualized with the SuperSignal West Pico or West Femto ECL systems (Pierce).

Immunohistochemistry—Human prostate samples, taken from patients with benign prostatic hyperplasia (BPH), were fixed in 10% buffered formalin, and 3- μ m sections were prepared and mounted on OptiPlus™ slides (BioGenex, San Ramon, CA). Deparaffinization was performed with xylene for 10 min, and the sections were rehydrated by rinsing three times with 100% ethanol and once in 95% ethanol. Slides were washed in double distilled H₂O and then incubated with 3% H₂O₂ for 5 min. Subsequently, the slides were washed again in double distilled H₂O and twice in Optimax™ Wash Buffer (BioGenex, San Ramon, CA). Immunohistochemical staining was performed using Histostain™ Plus Bulk Kit (Zymed Laboratories Inc., San Francisco, CA). The anti-PSA-LM and K-LM polyclonal antibodies, raised in the course of this study, were used at 1:100 and 1:50 dilutions, respectively. Counterstaining with hematoxylin was employed.

Seminal Fluid Analysis—Seminal plasma, obtained from ejaculates after removal of spermatozoa by centrifugation, was diluted in phosphate-buffered saline containing a protease inhibitor mixture. Proteins were separated by SDS-PAGE on 15% gels, transferred to polyvinylidene difluoride membranes, followed by immunodetection as described above.

GenBank™ Accession Numbers—The nucleotide sequence for the PSA-LM transcript that contains all of intron 1 has been deposited in the GenBank™ database under accession number AF335478. The second PSA-LM-encoding transcript, derived by alternative splicing within intron 1, has been deposited under GenBank™ accession number AF335477. The K-LM encoding transcript was deposited under GenBank™ accession number AF336106.

RESULTS AND DISCUSSION

Novel Proteins Predicted by Inclusion of Intronic Sequences in the *KLK3* and *KLK2* Genes—We found prostate-derived, alternatively spliced variants of the PSA-encoding *KLK3* gene by searching the LEADS Human Transcriptome Database (see “Experimental Procedures”). These variants include all or part of intron 1, thereby linking the first exon to the adjacent intronic sequences. RT-PCR was carried out with primers derived from both sides of this sequence in order to verify the existence of such transcripts in prostate samples. This sequence was then elongated using the RACE methodology. A new open reading frame (ORF) of 312 nucleotides was revealed, beginning with the start codon of PSA and ending in a stop codon within the sequences adjacent to exon 1, which are spliced out as intron 1 in the PSA mRNA (Fig. 1). To characterize these transcripts, RT-PCR was performed with a forward primer from this ORF and reverse primers derived from each of the downstream PSA exons. Sequence analysis of the PCR products indicated that the remaining PSA introns are spliced out and that PSA coding exons 2–5 are located downstream of the stop codon of this ORF, serving therefore as 3′-untranslated sequences (Fig. 1). In addition, this analysis revealed the use of an alternative donor splice site within intron 1 (Fig. 1).

The putative novel protein predicted by this new ORF, which we designated PSA-linked molecule (PSA-LM), is 104 amino acids long and shares amino acid sequence identity with PSA only in the N-terminal 15 amino acids that encompass the signal peptide of the pre-form protein (Fig. 2A). Thus, the mature PSA and PSA-LM proteins are entirely different.

Based on the close similarity between the *KLK3* and *KLK2* genes, and the high level of homology between their non-coding sequences, including introns (3, 6) (80% in intron 1), we looked for evidence of similarly spliced transcripts in the *KLK2* gene. Analysis of the intronic sequences adjacent to exon 1 disclosed an ORF of 423 nucleotides, beginning with the start codon of hK2 within exon 1 and ending in a stop codon within intron 1 of the *KLK2* gene. We verified the existence of alternative transcripts that include intron 1 and contain this ORF by

A

PSA-LM

AGCCCCAAGCTTACCACCTGCACCCGGAGAGCTGTGTACC

ATGTGGGTCCCGTTGTCTTCTCACCCTGTCCGTGACGTGGATTGGTGA
M W V P V V F L T L S V T W I G E

GAGGGGCCATGTTGGGGGATGCAGGAGAGGGAGCCAGCCCTGACTGTC
 R G H G W G D A G E G A S P D C

AAGCTGAGGCTCTTTCCCCCACAACCCAGCAGCCAGCCAGACAGGGAG
 Q A E A L S P P T Q H P S P D R E

CTGGGCTCTTTTCTGTCTCTCCAGCCCCACTCCAAGCCATACCCCCAG
 L G S F L S L P A P L Q A H T P S

CCCCCTCATATTGCAACAGTCTCTACTCCACACCAGGTCCCCGCTCCCT
 P S I L Q Q S S L P H Q V P A P

CCCCTTACCCAGAACTTTCTCCCATTTGCCAGCCAGCTCCCTGCTCC
 S H L P Q N F L P I A Q P A P C S

CAGCTGCTTTACT**TAAGGGGAAGTCTCTGGGCATCTCCGTGTTTCTCTT**
 Q L L Y *

TGTGGGGCTCAAAACCTCCAAGGACCTCTCTCAATGCCATTGGTTCCTTG

GACCGTATCACTGGTCCACCTCCTGAGGCCCTCAATCCTATCACAGTCTA

CTGACTTTTCCCATTCAGCT**GT**GAGTGCCCAACCTATCCAGAGACCTT

B

K-LM

ATGTGGGACCTGGTCTCTCCATCGCCTTGCTGTGGGGTGCACCTGGTGA
M W D L V L S I A L S V G C T G E

GATTGGGGGATAAAGGAAGGGGGGGGGTCTGACTCTTATGCTGAAGC
 I G G I K E G G R V L T L M L K

CCTTTTCTCCACCCAGTGCCCCAGCCTCGTCCCTTCAGCCACAGTTC
 P F S S H P V P Q P R P F S P Q F

AGCCAGACAATGTGCCCTGACTCTTCCACATTGCAATAGTCTCATGC
 S P D N V P L T L P H C N S P H A

CCACACTAGGTCCCGCTCCCTCCCACTTACCTCAGACCTTTCTCTCCAT
 H T R S P L P P T Y L R P F S P

TGCCAGCCAAATCCCTGCTCCAGCTGCTTTACTAAAGACAAAGTTCCT
 L P S Q I P A P S C F T K E Q V P

AGGCATCTCTGTGTTTCTTTATGGGGTTCAAAACCTTTCAAGGACCTC
 R H L C V S L Y G V Q N L S R T S

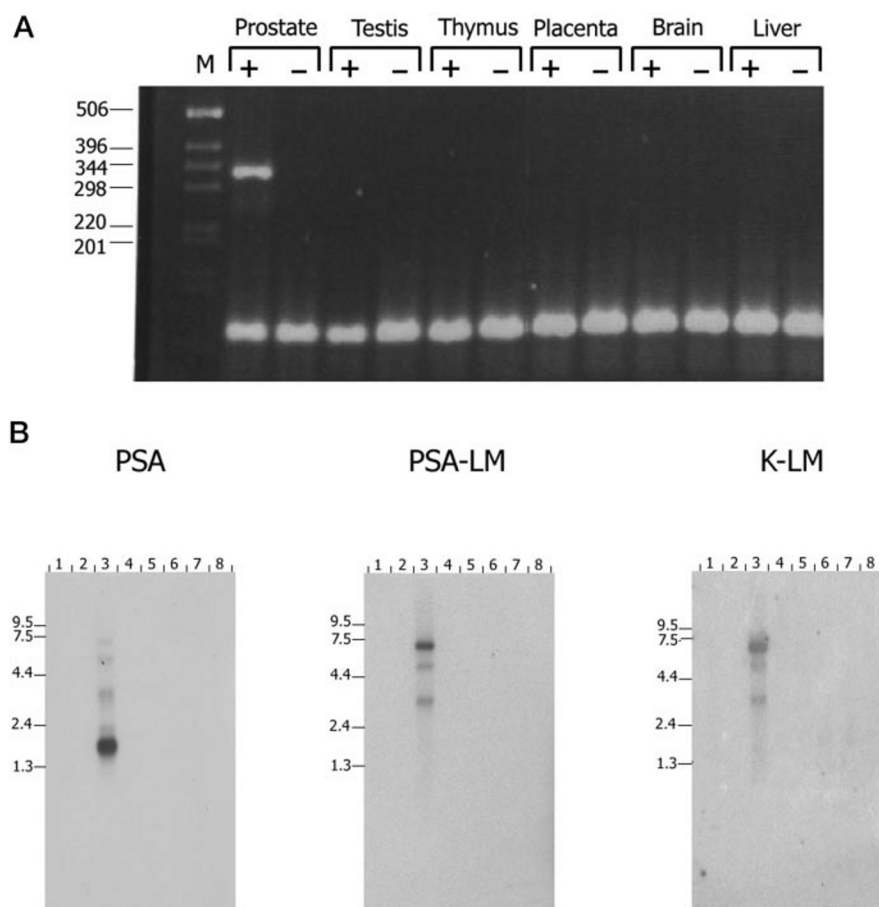
TCTCATGCCACTGGTTCCTTGGACCTATCACTGGGCTGCCTCTGAGC
 L H A T G S L D P I T G L P P E

CCCTCAGTCTACACAGTCTACT**GA**CTTTTCCATTACAGCTGTGAGCAT
 P L S P T T V Y *

TCAACCTGTCCCTGGACCTTGACACCTGGCTCCCCAACCTGTCCAG
 GAAACCCAGATTCCACCAGACACTTCTTCTCCCCCGAGGCTATCTG

FIG. 2. Nucleotide sequence of the PSA-LM (A) and K-LM (B) ORFs and flanking sequences. The start and stop codons are marked in **bold letters**, and the donor site at the beginning of intron 1 is underlined. The alternative donor site within intron 1 in the sequence of PSA-LM is marked in **bold letters** and underlined. The corresponding amino acid sequences are shown below in single letter abbreviations. The signal peptide sequence, common to PSA or hK2, is double underlined.

FIG. 3. Expression of PSA-LM and K-LM. **A**, RT-PCR analysis of PSA-LM expression in several tissues, using a sense primer from exon 1 and an antisense primer derived from the PSA-LM ORF. The expected 326-bp product is obtained only with prostate RNA. +/– indicate the presence or absence of reverse transcriptase in the reaction. Other tissues and cell lines were checked and found negative, including colon, kidney, ovary, B-lymphocytes, and HeLa and T24 cells (not shown). **B**, Northern blot analysis with probes for PSA (left panel), PSA-LM (middle panel), and K-LM (right panel). The Northern blots contain poly(A) RNA from the following human tissues: lanes 1, spleen; 2, thymus; 3, prostate; 4, testis; 5, ovary; 6, small intestine; 7, colon; and 8, leukocytes. Size markers (in kb) are indicated. A different multiple tissue Northern blot (which included prostate, skeletal muscle, uterus, heart, stomach, and bladder) was also used (not shown). In all cases the expression is restricted to prostate. The expected 1.6-kb transcript is the major band detected with the PSA probe. Several transcripts, ranging in size from 3 to 6.5 kb, are detected with the PSA-LM and K-LM probes and weakly with the PSA probe.



performing RT-PCR on prostate samples, using a forward primer at the beginning of this ORF and a reverse primer immediately after the stop codon. In analogy to PSA-LM, the putative 141-amino acid protein specified by this new ORF, K-LM (hK2-linked molecule), shares with hK2 only the N-terminal 15 amino acids, which includes the signal peptide (Fig. 2B), and its mature form displays no similarity with hK2.

The PSA-LM and K-LM proteins exhibit 32% identity and 51% similarity with each other. These novel proteins show no similarity to any other protein in the databases and contain no known protein signatures. Interestingly, however, the PSA-LM and K-LM proteins are rich in proline residues (16 and 17%, respectively) and in serine/threonine residues (15 and 19%). Proline-rich proteins often serve as adaptor molecules, bringing together other proteins (reviewed in Ref. 21). However, it is not clear whether the proline-rich sequences play a role in the biological function of these novel polypeptides.

Transcription of PSA-LM Is Restricted to Prostatic Epithelium and Up-regulated by Androgens—The tissue-specific expression of PSA-LM was analyzed by RT-PCR, using a primer from exon 1 and a reverse primer derived from intron 1 within the new ORF. Out of 12 different tissues and cell lines examined, the expected PCR fragment was detected only in prostate and in the prostate-derived cell line LNCaP (Fig. 3A and data not shown). RT-PCR analysis also showed prostate-specific expression of the K-LM-encoding transcript (data not shown).

The expression pattern of these transcripts was further studied by Northern blotting (Fig. 3B). A probe derived from exon 3 of PSA detected a major transcript of about 1.6 kb only in prostate (Fig. 3B, left panel). Additional minor bands of about 1.9, 3, 5, and 6.5 kb were also observed, in agreement with previous reports (14, 15). A probe specific for PSA-LM, derived from its coding sequences within intron 1, detected RNA tran-

scripts only in prostatic tissue. These were the same size as the larger bands seen with the PSA probe (Fig. 3B, middle panel). This is not surprising because the PSA probe is included in the PSA-LM 3'-UTR and should therefore also detect the PSA-LM-encoding transcripts. The relative intensity of these bands in the left panel of Fig. 3B implies that the PSA-specific transcript of 1.6 kb is considerably more abundant. Prostate-specific transcripts were also detected with a probe derived from the K-LM ORF (Fig. 3B, right panel). The exact nature of the higher bands observed with these probes is not clear. They might represent precursors of mature mRNAs or complex alternative splicing patterns, because several spliced products as well as alternative polyadenylation have been described for both the *KLK2* and *KLK3* genes (15–19). The transcript of about 3 kb might result from inclusion of intron 1, which contains the novel ORF, as described in Fig. 1A. Inclusion of all introns would result in a transcript of about 5 kb, whereas the addition of a 1.5-kb-long 3'-UTR, as reported by Liu *et al.* (17), could explain the 6.5-kb transcript.

The androgenic transcriptional activation of the *KLK2* and *KLK3* genes has been studied extensively (8, 9, 16, 23, 24). We analyzed the effects of androgenic stimulation on the expression of PSA-LM using the androgen-responsive LNCaP cell line. After 24 h of hormonal depletion, cells were fed on charcoal-stripped serum in the absence or presence of either testosterone or a synthetic androgen (R1881). Northern blot analysis with a probe derived from exon 3 of PSA showed a strong induction of the 1.6-kb PSA-encoding transcript after exposure to 10 nM R1881 (Fig. 4, left panel) or testosterone (data not shown) for 9 h, conditions previously shown to yield maximal expression of PSA mRNA (8). Under these conditions, a 2–6-fold up-regulation of the 3-, 5-, and 6.5-kb transcripts was also detected using a 780-bp probe derived from intron 1 that spans

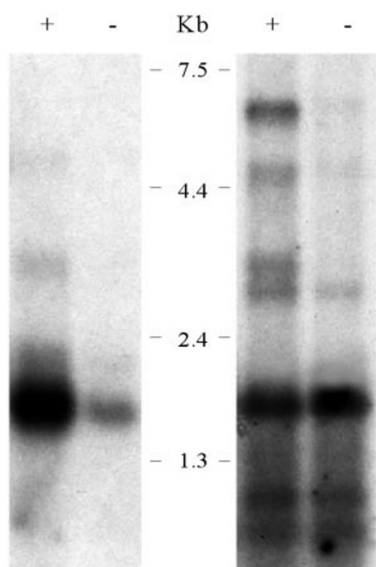


FIG. 4. Hormonal regulation of PSA and PSA-LM mRNA expression. Northern blots were prepared by electrophoresis on 1.2% agarose gels of 2 μ g of poly(A) RNA from LNCaP cells incubated with (+) or without (–) 10 nM of the synthetic androgen R1881. The *left* and *right* panels, respectively, show the hybridization with a PSA probe derived from exon 3, or a PSA-LM probe derived from intron 1 that spans the ORF. Size markers (in kb) are indicated. The major transcript of 1.6 kb and the other minor bands detected with the exon 3-specific probe show a strong up-regulation in the androgen-treated cells. The higher molecular weight bands detected with the intron 1-specific probe also show a significant induction following androgenic stimulation.

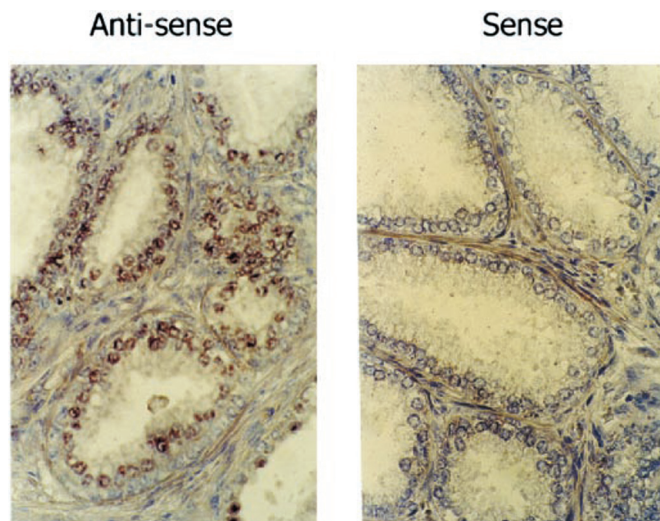


FIG. 5. Expression of PSA-LM in human prostate analyzed by *in situ* hybridization. An antisense RNA probe was used to detect PSA-LM-specific mRNA; the sense probe served as a negative control. The hybridization signal is observed only with the antisense probe and is restricted to prostatic epithelial cells. Original magnification is $\times 40$.

the PSA-LM ORF without the signal peptide common to PSA (Fig. 4, *right* panel). The nature of the other bands observed with this probe is not clear. Higher levels of PSA and PSA-LM mRNAs were also observed up to 48 h (data not shown).

Analysis of prostate sections by *in situ* hybridization indicated specific expression of the PSA-LM-encoding transcripts in the secretory epithelial cells of the prostate tubules (Fig. 5). Expression of PSA in the prostate has also been shown to be restricted to these cells (8, 25). Taken together, these results suggest that expression of PSA and PSA-LM undergo similar modes of transcriptional regulation.

PSA-LM and K-LM Protein Expression and Secretion—Poly-

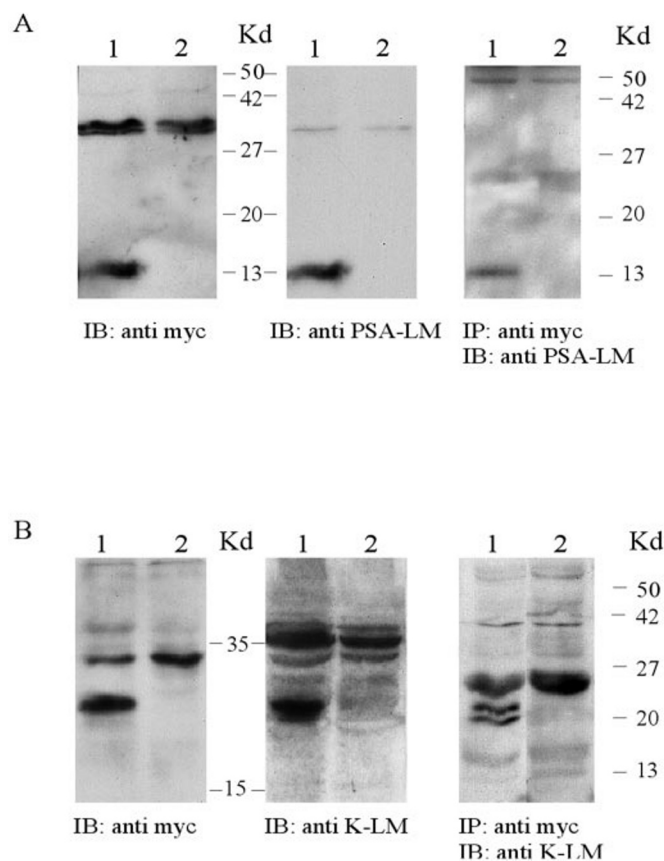


FIG. 6. Detection of Myc-tagged PSA-LM and K-LM proteins. LNCaP cells were transfected with pcDNA4 constructs expressing Myc-tagged PSA-LM (A) and K-LM (B) proteins (lanes 1) or pcDNA4 vector alone (lanes 2). Cell lysates (60 μ g) were fractionated by 15% SDS-PAGE and subjected to immunodetection with a monoclonal anti-Myc antibody (*left* panels), followed by analysis with rabbit sera immunized against PSA-LM (*middle* panel in A) or against K-LM (*middle* panel in B). Secretion of the recombinant proteins was verified by immunoprecipitation (IP) of spent media using Myc-conjugated agarose beads, followed by immunodetection with the respective sera (*right* panels). Size markers (in kilodaltons) are indicated. Equal loading is shown by the intensity of the nonspecific bands. IB, immunoblot.

clonal antibodies were generated against synthetic peptides derived from the amino acid sequences of PSA-LM and K-LM. In order to verify the specificity of these antibodies, the full coding sequences of PSA-LM and K-LM were cloned into the mammalian expression vector pcDNA4, tagged with Myc and His epitopes. LNCaP cells were transfected with these constructs, and after 48 h the culture medium was collected, and cells were harvested and lysed. Immunoblot analysis of the transfected cell lysates, using an anti-Myc antibody or our polyclonal antibodies, showed expression of PSA-LM and K-LM (Fig. 6, A and B, *left* and *middle* panels). The observed sizes of these recombinant proteins (13 and 20 kDa, respectively) are similar to those predicted from their amino acid sequence (11 and 15 kDa), with the addition of about 3 kDa of the epitope tags. The additional bands observed in Fig. 6B most likely represent cross-reacting proteins, as they were not detected with other sera raised against K-LM (not shown). The secretion of these proteins to the growth medium of the transfected cell cultures was confirmed by immunoprecipitation using anti-Myc-coupled beads, followed by immunodetection with the specific antibodies (Fig. 6, A and B, *right* panels). Endogenous PSA-LM and K-LM proteins could not be detected in LNCaP cells or other prostate-derived cells either prior or following androgenic stimulation (data not shown), perhaps due to low protein levels.

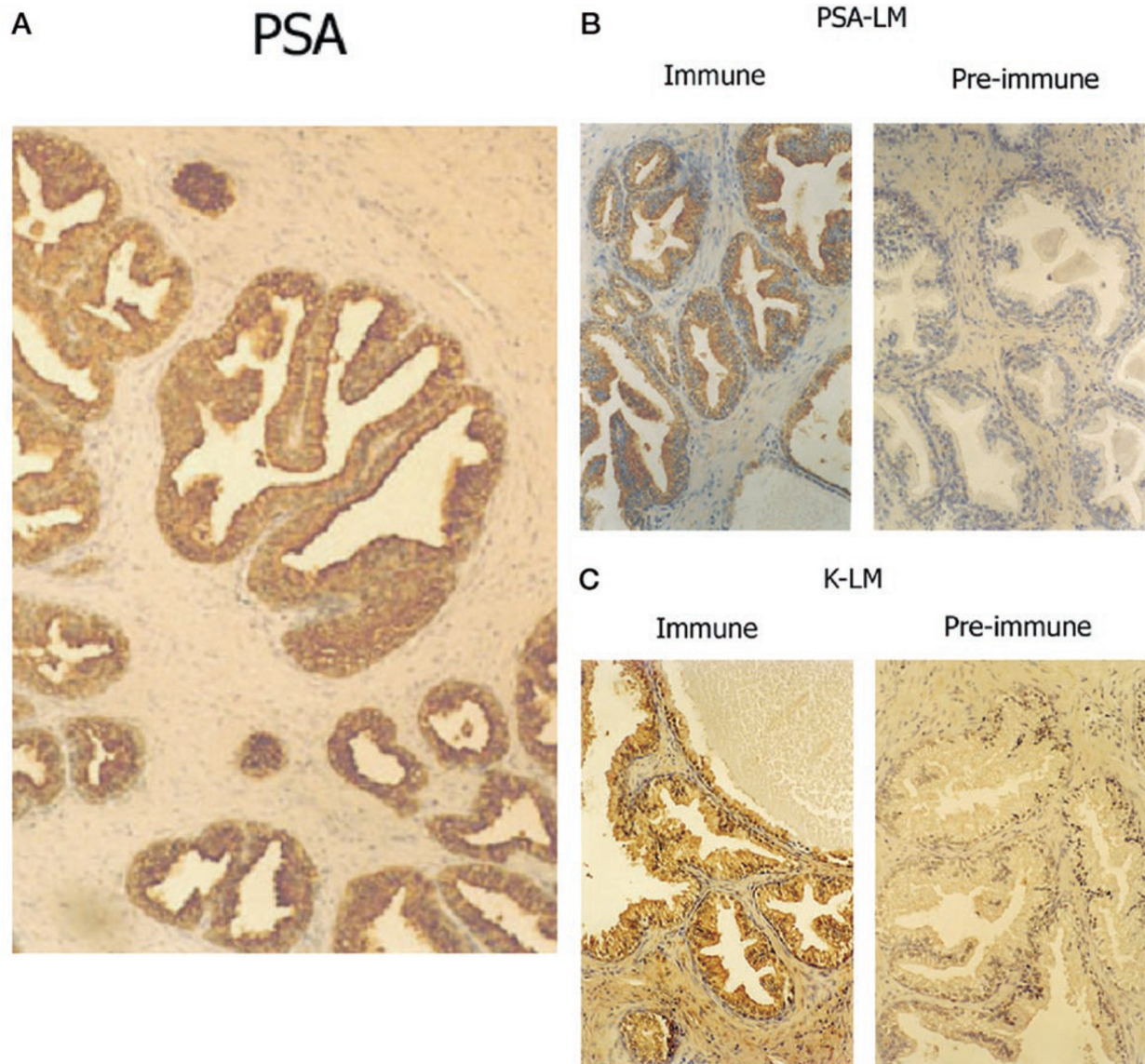


FIG. 7. **Immunohistochemistry of prostate sections, using anti-PSA (A), PSA-LM (B), and K-LM (C)-specific antibodies.** Preimmune sera for the antibodies generated in the course of this study were used as negative control. Marked staining within the epithelial cells surrounding the prostate tubules is seen with all three protein-specific antibodies. We counterstained with hematoxylin. Original magnifications are $\times 20$.

Immunohistochemistry analysis of prostate sections using our PSA-LM and K-LM polyclonal antibodies and a monoclonal antibody against PSA, showed that the three proteins are detected only in the secreting cells of the tubule lumen (Fig. 7). The specific staining observed within the lumen of some of the tubules in Fig. 7 could reflect the secretion of these proteins. It is important to bear in mind, however, that the cross-reacting bands detected with the anti-K-LM antibodies in Fig. 6B raise questions about the specificity of these antibodies and the interpretation of the immunostaining data presented in Fig. 7C.

The presence of a common signal peptide, the secretion of the recombinant proteins upon transfection to LNCaP cells, and the specific expression in the prostatic secretory epithelial cells strongly suggest that, like PSA and hK2 (26), PSA-LM and K-LM are secreted to the seminal fluid. Immunoblot analysis of samples of semen plasma and benign prostatic hyperplasia (BPH) identified a major form of the native K-LM protein, appearing as 27 kDa (Fig. 8). The appearance of this and fainter bands in these samples, which are bigger than the expected size of this protein, could be due to post-translational modifications, such as glycosylation. This assumption gains

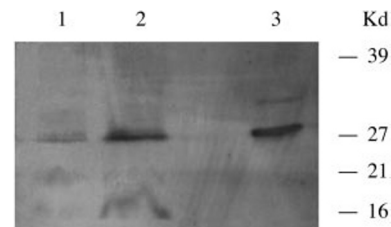


FIG. 8. **Secretion of K-LM to seminal fluid and expression in BPH.** Proteins from seminal plasma (25 μ g, lanes 1 and 2) and a BPH sample (100 μ g, lane 3) were separated by 15% SDS-PAGE and subjected to immunoblot analysis with K-LM-specific polyclonal antibodies.

support from the presence of several putative glycosylation sites in its amino acid sequence (not shown). We failed to identify the PSA-LM protein in these samples, perhaps due to lower expression levels, lower sensitivity of antibodies, protein instability, or masking by other proteins as in the case of complexed PSA (27).

The alternative transcripts of the *KLK2* and *KLK3* genes described here share the same 5'-untranslated sequences and start codon with the major mRNA species of the respective

gene. In addition, the novel alternative protein products have the same signal peptide required for secretion of the original proteins encoded by these genes. Thus, PSA-LM and K-LM are expected to undergo the same mode of hormonal regulation, expression, and secretion as the PSA and hK2 proteins. This assumption is indeed supported by the data presented here. Our observations indicate that expression of PSA-LM is restricted to the glandular epithelium of the prostate and is up-regulated by androgens, implying that the same control elements regulate the transcription of the mRNAs that specify the two distinct protein products of the *KLK3* gene. In addition, similar to PSA and hK2, the K-LM protein appears to be secreted to the seminal fluid.

There are only a few examples of dual utilization of a single promoter to yield two polypeptides that, in their mature form, differ entirely in their amino acid sequences. One extensively studied example is calcitonin and the calcitonin gene-related peptide. In this case, the two transcripts share the first three exons and have the same initiation codon. However, tissue-specific alternative RNA processing associated with differential use of polyadenylation sites, followed by N-terminal proteolytic cleavage, generate two peptide hormones with completely different structures and functions (28). Another example is the leptin receptor and the leptin receptor gene-related protein (OBR/OB-RGRP): alternative splicing yields two transcripts that share only the first two 5'-untranslated exons, but the coding exons are entirely different, resulting in two polypeptides with no sequence similarity (29).

The current scientific literature on the physiology and pathology of PSA is somewhat controversial, and the physiological role of PSA is still not clear (2). It has been suggested that the biological functions of PSA may be influenced by the activity of other kallikreins, such as hK2, the activating enzyme of the pro-form of PSA (7, 10, 11). Efforts are underway to supplement PSA diagnostics with other markers aimed in particular at enhancing the discrimination between patients with benign disease and those with prostate cancers. The hK2 protein is now emerging as an additional prostatic tumor marker that might have clinical applications complementary to those of PSA (7). Additional studies are required to evaluate the involvement of PSA-LM and K-LM in the biology of PSA and hK2, as well as in signaling pathways that lead to the progres-

sion of prostate cancer. Such studies may also clarify their potential clinical value.

Acknowledgments—We thank N. Almogly and N. Kossoy of the Department of Pathology, Tel Aviv Sourasky Medical Center, for their help and professional advice; H. M. Safer, R. Gill-More, and D. Eshel for critical reading of the manuscript; M. Yehezkeili for help with graphic work; and A. Faingold for assistance with the figures.

REFERENCES

1. Diamandis, E. P. (1998) *Trends Endocrinol. Metab.* **9**, 310–316
2. Diamandis, E. P. (2000) *Clin. Chem.* **46**, 896–900
3. Lundwall, A. (1989) *Biochem. Biophys. Res. Commun.* **161**, 1151–1159
4. Diamandis, E. P., Yousef, G. M., Luo, L.-Y., Magklara, A., and Obiezu, C. V. (2000) *Trends Endocrinol. Metab.* **11**, 54–60
5. Harvey, T. J., Hooper, J. D., Myers, S. A., Stephenson, S.-A., Ashworth, L. K., and Clements, J. A. (2000) *J. Biol. Chem.* **275**, 37397–37406
6. Riegman, P. H., Vlietstra, R. J., van der Korput, J. A., Romijn, J. C., and Trapman, J. (1989) *Biochem. Biophys. Res. Commun.* **159**, 95–102
7. Rittenhouse, H. G., Finlay, J. A., Mikolajczyk, S. D., and Partin, A. W. (1998) *Crit. Rev. Clin. Lab. Sci.* **35**, 275–368
8. Young, C. Y., Montgomery, B. J., Andrews, P. E., Qiu, S., Bilhartz, D. L., and Tindall, D. J. (1991) *Cancer Res.* **51**, 3748–3752
9. Young, C. Y., Andrews, P. E., Montgomery, B. T., and Tindall, D. J. (1992) *Biochemistry* **31**, 818–824
10. Kumar, A., Mikolajczyk, S. D., Goel, A. S., Millar, L. S., and Saedi, M. S. (1997) *Cancer Res.* **57**, 3111–3114
11. Lovgren, J., Rajakoski, K., Karp, M., Lundwall, A., and Lilja, H. (1997) *Biochem. Biophys. Res. Commun.* **238**, 549–555
12. Lilja, H. (1993) *World J. Urol.* **11**, 188–191
13. Black, D. L. (2000) *Cell* **103**, 367–370
14. Henttu, P., Lukkariinen, O., and Vihko, P. (1990) *Int. J. Cancer* **45**, 654–660
15. Riegman, P. H., Klaassen, P., van der Korput, J. A., Romijn, J. C., and Trapman, J. (1988) *Biochem. Biophys. Res. Commun.* **155**, 181–188
16. Riegman, P. H., Vlietstra, R. J., van der Korput, J. A., Romijn, J. C., and Trapman, J. (1991) *Mol. Cell. Endocrinol.* **76**, 181–190
17. Liu, X. F., Essan, M., Vasmatazis, G., Lee, B., and Pastan, I. (1999) *Biochem. Biophys. Res. Commun.* **264**, 833–839
18. Heuze, H., Olayat, S., Gutman, N., Zani, M.-L., and Courty, Y. (1999) *Cancer Res.* **59**, 2820–2824
19. Tanaka, T., Isono, T., Yoshiki, T., Yuasa, T., and Okada, Y. (2000) *Cancer Res.* **60**, 56–59
20. Roche Molecular Biochemicals (1996) *Non-radioactive in Situ Hybridization Application Manual*, 2nd Ed., pp. 34–56, Basel, Switzerland
21. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) *FASEB J.* **14**, 231–241
22. Deleted in proof
23. Wolf, D. A., Schulz, P., and Fittler, F. (1992) *Mol. Endocrinol.* **6**, 753–762
24. Sun, Z., Pan, J., and Balk, S. P. (1997) *Nucleic Acids Res.* **25**, 3318–3325
25. Qui, S., Young, C. Y. F., Bilhartz, D. L., Prescott, J. L., He, W. W., and Tindall, D. J. (1990) *J. Urol.* **144**, 1550–1556
26. Lovgren, J., Valtanen-Andre, C., Marsal, K., Lilja, H., and Lundwall, A. (1999) *J. Androl.* **20**, 348–355
27. Carroll, P. R. (2001) *Urology* **57**, 591–593
28. Lou, H., and Gagel, R. F. J. (1998) *J. Endocrinol.* **156**, 401–405
29. Bailleul, B., Akerblom, I., and Strosberg, A. D. (1997) *Nucleic Acids Res.* **25**, 2752–2758