A Novel Protease Homolog Differentially Expressed in Breast and Ovarian Cancer

Anthony Anisowicz,* Georgia Sotiropoulou,*† Goran Stenman,‡ Samuel C. Mok,§ and Ruth Sager*

*Division of Cancer Genetics, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, U.S.A.

†Department of Pharmacy, School of Health Sciences, University of Patras, Patras, Greece

‡Department of Pathology, University of Goteborg, Goteborg, Sweden

§Laboratory of Gynecologic Oncology, Department of Obstetrics, Gynecology, and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, U.S.A.

ABSTRACT

Background: Using differential display (DD), we discovered a new member of the serine protease family of protein-cleaving enzymes, named protease M. The gene is most closely related by sequence to the kallikreins, to prostate-specific antigen (PSA), and to trypsin. The diagnostic use of PSA in prostate cancer suggested that a related molecule might be a predictor for breast or ovarian cancer. This, in turn, led to studies designed to characterize the protein and to screen for its expression in cancer.

Materials and Methods: The isolation of protease M by DD, the cloning and sequencing of the cDNA, and the comparison of the predicted protein structure with related proteins are described, as are methods to produce recombinant proteins and polyclonal antibody preparations. Protease M expression was examined in mammary, prostate, and ovarian cancer, as well as normal, cells and tissues. Stable transfectants expressing the protease M gene were produced in mammary carcinoma cells.

Results: Protease M was localized by fluorescent in situ hybridization analysis to chromosome 19q13.3, in a region to which other kallikreins and PSA also map. The gene is expressed in the primary mammary carcinoma lines tested but not in the corresponding cell lines of metastatic origin. It is strongly expressed in ovarian cancer tissues and cell lines. The enzyme activity could not be established, because of difficulties in producing sufficient recombinant protein, a common problem with proteases. Transfectants were selected that overexpress the mRNA, but the protein levels remained very low.

Conclusions: Protease M expression (mRNA) may be a useful marker in the detection of primary mammary carcinomas, as well as primary ovarian cancers. Other medical applications are also likely, based on sequence relatedness to trypsin and PSA.

INTRODUCTION

Serine proteases are protein-cleaving enzymes that contain a serine residue in their active sites and which play important roles in diverse physiological processes, including digestion (e.g., trypsin, chymotrypsin) and blood clotting (e.g., plasminogen activator, thrombin). Serine proteases also act as regulators of a variety of processes by proteolytic activation of precursor proteins.

The kallikreins are a subfamily of serine proteases originally defined as those cleaving vasoactive peptides (kinins) from kininogen (1). Currently, the kallikreins comprise a large, mul-

Address correspondence and reprint requests to: Ruth Sager, Dana-Farber Cancer Institute Division of Cancer Genetics, 44 Binney Street, Boston, MA 02115, U.S.A.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank/EMBL Data Bank with accession number U62801.

Copyright © 1996, *Molecular Medicine*, 1076-1551/96/$10.50/0

Molecular Medicine, Volume 2, Number 5, September 1996 624-636

tigene family in rodents, although only three members of this family, hKLK1, hKLK2, and hKLK3, are known in humans. These three genes encode the proteins pancreatic/renal kallikrein (hK1), glandular kallikrein (hK2), and prostate-specific antigen (PSA; hK3), respectively (2).

The hK1 protein is secreted from pancreas, kidney, and salivary glands (3), and is the only member of the family having true kallikrein activity. Its major function is the generation of kinins from kininogens and the regulation of blood pressure (1). The hK2 protein has yet to be detected in human tissue or fluids, but its sequence has been inferred from a genomic clone (4), as well as cDNA clones isolated from prostate libraries (5). hK2 expression is specific for prostate and is regulated by androgens (5). PSA is produced predominantly in males by prostate epithelial cells and is secreted into the seminal fluid, where it serves to degrade the gel-like seminogelin protein and increase sperm motility (6,7). Although PSA is produced at higher levels in normal than in malignant prostate tissue, a defect in the malignant tissues ultimately results in the leakage of PSA into the bloodstream (8). This is the basis of the use of PSA as a circulating tumor marker for prostate cancer.

Here, we describe the isolation by differential display (9–11) of a novel member of the serine protease family which is most homologous to trypsin and members of the kallikrein family. This novel protein, which we have named protease M, is down-regulated in metastatic breast cancer lines but strongly expressed at the mRNA level in some primary breast cancer cell lines and in ovarian cancer tissues and tumor cell lines.

MATERIALS AND METHODS

Mammary Cell Strains and Lines

Normal human mammary epithelial cell strains (70N and 76N) were derived from reduction mammoplasties in this laboratory as described (12). Primary (21PT, 21NT) and metastatic (21MT-1, 21MT-2) tumor lines were established in this lab from a single patient as described (12,13). Human mammary epithelial tumor cell lines MCF-7, T47D, ZR75-1, BT549, MDA-MB-157, MDA-MB-231, MDA-MB-435, MDA-MB436, MDA-MB-361, and BT-474 were obtained from American Tissue Culture Collection (ATCC, Rockville, MD, U.S.A.). Cells were grown in DFCI-I media (12) and harvested at approximately 70% confluence for RNA isolation.

Prostate Cell Lines

Normal, immortalized prostate epithelial cell lines: CF3 (HPV immortalized), CF91 (SV40 immortalized), and MLC (SV40 immortalized) were provided by Dr. Johng Rhim and were cultured in KGM medium (DIFCO, Detroit, MI, U.S.A.). The tumor cell lines DU145, LNCaP, and PC3 (ATCC) were grown in Dulbecco's modified Eagle's medium (DMEM) plus 10% fetal calf serum (FCS, Hyclone, Logan, UT, U.S.A.).

Ovarian Cell Cultures and Tissues

The primary human ovarian surface epithelial cell cultures (HOSE 10/11, 16, and 21) were established from the ovarian surface epithelium as described (14). Immortalized ovarian surface epithelial cells (HOSE6.3E6E7) were obtained by infecting the HOSE cells with a replication-defective retrovirus construct, LXSN16E6E7, as described (14). The eight ovarian carcinoma cell lines used for this comparative study include DOV13, OVCA420, OVCA429, OVCA432, and OVCA433, which were established in the laboratory of Gynecologic Oncology; CAOV3 and SKOV3, which were purchased from ATCC; and OVCA3, which was obtained from the National Cancer Institute (Frederick, MD, U.S.A.).

Ovarian tumors obtained (15) include 6 borderline ovarian tumors (354A, 373A, 395A, 405A, 466A, and 469A); 20 stage III/IV high grade invasive ovarian adenocarcinomas from the primary ovarian site; 2 metastatic adenocarcinoma from colon primary tumors (327A, 339A); and 3 normal ovaries (366N, 379N, and 465N).

Differential Display of mRNA

Total cell RNAs (50 μg) from 21PT and 21MT-1 were treated with DNAaseI (Worthington DPRF, Freehold, NJ, U.S.A.) in the presence of RNAasin ribonuclease inhibitor (Promega, Madison, WI, U.S.A.) to remove residual DNA contamination, as described elsewhere (11). Differential display of the mRNA was performed as described (9,10). Basically, the RNAs were reverse transcribed using the 3′-anchored primer $T_{12}MG$ (where M is a mixture of A, G, or C). The resultant cDNAs were then polymerase chain reaction (PCR) amplified in the presence of ^{35}S-dATP using $T_{12}MG$ and the

arbitrary primer OPA1 (CAGGCCCTTC) and run in adjacent lanes on a 6% sequencing gel. Differentially displayed bands were recovered from the dried gel, reamplified by PCR, ^{32}P-labeled by the oligo method (16) and used as a probe on Northern blots prepared with 21PT and 21MT-1 total.

Cloning, Sequencing, and Analysis of cDNAs

The reamplified band from differential display was cloned into the TA cloning vector PCRII (Invitrogen, San Diego, CA, U.S.A.) and sequenced on both strands using T7 and SP6 primers. cDNA libraries from 21PT and 76N cells constructed in Lambda Zap II (Stratagene, San Diego, CA, U.S.A.), were screened using the cloned PCR product as a probe and several cDNA clones were isolated and sequenced on both strands. The longest cDNA clone (from the 76N library) was sequenced on both strands using an ABI automated sequencer (Model 373A) by the Dana-Farber Molecular Biology Core Facility. Oligonucleotides used for sequencing were synthesized by the Dana-Farber Molecular Biology Core facility or by Amitof, Inc. (Cambridge, MA, U.S.A.). The predicted protein coding region and non-translated regions were determined and formatted using the GCG Publish program. The predicted protein sequence was compared to protein databases using the Blast algorithm (17). Protein alignment with related proteins was performed on GCG using the Pileup, Distances, and Prettyplot programs.

Northern Analysis

Total cell RNA was isolated by the guanidinium isothiocyanate/cesium chloride method and analyzed on Northern blots as described (18). 36B4 (19), a ribosomal protein whose message is constant under a variety of conditions, was used to normalize the blots. Densitometric analysis of autoradiographs was performed with an imaging densitomer (Biorad GS-700) using the Molecular Analyst software.

Mapping of the Protease M Gene

A panel of 24 human-rodent somatic cell hybrids (Mapping panel 2 from NIGMS, Coriell Institute for Medical Research, Camden, NJ, U.S.A.) was used for mapping the protease M gene. DNAs from hybrid and parental cell lines were restricted with *Eco*RI, electrophoretically separated in 0.8% agarose gels, and transferred to nylon filters. Blots were hybridized with a ^{32}P-labeled protease M cDNA probe (IG3-8) corresponding to nucleotides 566 to 1526. For fine mapping of the protease M gene, fluorescent in situ hybridization (FISH) was performed on normal human lymphocyte spreads as previously described (20). A λ clone (λIg3-1) containing the 5′ portion of the protease M gene (from position 1 to 1016) was labeled with biotin by nick translation and co-hybridized with an α-satellite probe specific for chromosomes 1, 5, and 19 (D1Z7/D5Z2/D19Z3; Oncor, Gaithersburg, MD, U.S.A.). Chromosomes were counterstained with 4.6-diamino-2-phenylindole (DAPI). Slides were examined in a Zeiss Axiophot epifluorescence microscope using the appropriate filter combinations. Fluorescence signals were digitalized, enhanced, and analyzed using the ProbeMaster FISH image analysis system (Perceptive Scientific Instruments, Houston, TX, U.S.A.).

Production of Polyclonal Antibody and Western Blotting

The multiple antigen peptide (MAP) (21) ^{73}GKNNLRQRESSQEQS^{87} (0.5 mg) was emulsified with an equal volume of Freund's adjuvant and injected into 3- to 9-month-old New Zealand white rabbits. Boosts were done 2 and 6 weeks later. The animals were bled and serum was collected and stored at −20°C. Peptide and antibody production was done at Research Genetics (Huntsville, AL, U.S.A.).

Whole cell lysates were prepared by sonicating 10^7 cells/ml for twenty 30-sec pulses in a Sonicator Ultrasonic Processor in mammalian lysis buffer (4 mM $NaHCO_3$, 100 mM NaF, 20 mM KH_2PO_4, 2 mM sodium orthovanadate, 5 mM EDTA, 5 mM disfluorophosphate, 2 mM PMSF, 2 μg/ml leupeptin, 2 μg/ml aprotinin, pH 7.2). Lysates were clarified by spinning at 14,000 $\times$ g for 30 min in a microfuge.

Fifty to 100 μg of cell lysate was denatured by heating in SDS-PAGE sample buffer (50 mM Tris-HCl, pH 6.8, 0.1 mM DTT, 2% SDS, 0.1% bromphenol blue, 10% glycerol) at 90°C for 5 min and run on a 12% acrylamide/SDS minigel (Biorad), electroblotted onto a PDVF membrane (0.2 μ, Biorad), and reacted with immune serum (1:1000). Anti-rabbit IgG horseradish peroxidase-linked whole antibody (Amersham) (1:2000) was used as secondary antibody, and immunoreactive bands were detected with en-

hanced chemiluminescence (ECL; Amersham, Arlington Heights, IL, U.S.A.).

Expression of GST Fusion Protein

The full-length cDNA clone was PCR amplified using the sense 5′ 26-mer oligonucleotide 5′-GGAATTCCGTTGGTGCATGGCGGACC-3′ and the antisense 3′ oligonucleotide 5′-GTCGGAATT CAGGGTCACTTGGCCTG-3′ at 95°C, 1 min, 60°C, 1 min, 72°C, 1 min for 30 cycles to yield a 0.7-kb product which contained the open reading frame without the hydrophobic N-terminal amino acids. The resultant PCR product encoding for leu^{22} to lys^{244} was digested with *Eco*RI and ligated to alkaline phosphatase-treated *Eco*RI linearized pGEX-2T vector (Pharmacia, Piscataway, NJ, U.S.A.) to produce plasmids encoding a GST-protease M fusion protein. *Escherichia coli* strains XL-1 blue or DH5λ transformed with this construct were grown and induced with 0.2 mM IPTG at 37°C for 1 hr to produce GST fusion protein which was solubilized from bacteria and purified on glutathionine agarose beads by standard methods (22).

Expression of Baculovirus Recombinant Protein

A full-length cDNA clone was cut with *Eco*N1 and *Bst*X1 to give a fragment (nucleotides 233 to 1019) which was incorporated into the baculovirus transfer vector pVL1392 (Pharmingen, San Diego, CA, U.S.A.). Generation and amplification of recombinant baculovirus was as described (23,24). For production of protease M, *Spodoptera frugiperda* (cell line SF9) was infected with amplified recombinant virus to obtain nearly 100% infection as gauged by enlarged cells. Ninety-six hours postinfection, cells were harvested and lysed by sonication in mammalian lysis buffer, adjusted to 500 mM NaCl and rocked for 1 hr at 4°C. All subsequent purifications were done at 4°C.

The lysate was adjusted to 125 mM NaCl, loaded onto *p*-aminobenzamidine agarose (Sigma A7155, St. Louis, MO, U.S.A.), washed with loading buffer, and eluted with 25 mM $NaPO_4$, 0.02% NaN_3, 500 mM NaCl, 10 mM benzamidine, pH 6.0. The eluted fractions were loaded onto concanavalin A agarose (Sigma C8402) by rocking for 1 hr, washed with 25 mM $NaPO_4$, 0.02% NaN_3, 500 mM NaCl, pH 6.0, and eluted in wash buffer containing 10% methyl-α-*D*-mannopyranoside (Sigma M6882).

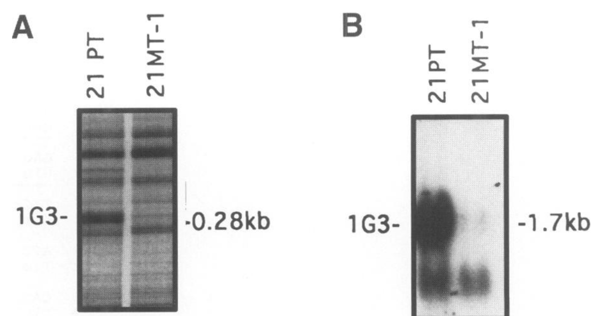

FIG. 1. Identification of protease M (1G3) of DD gel and on Northern blot

(A) DD gel: 21 PT and 21 MT-1 RNA was reverse transcribed with $T_{12}MG$ primer and PCR-amplified with $T_{12}MG$ and OPA1 primers in the presence of ^{35}SdATP, run on a 6% acrylamide/urea sequencing gel, and exposed to X-ray film for 18 hr. The portion of the gel surrounding the differentially displayed 0.28-kb band is shown. (B) Northern blot: 10 μg of total cell RNA was Northern blotted and probed with ^{32}P-labeled PCR-amplified 0.28-kb band from the DD gel shown in Panel A.

Expression Vector Construct and Transfection

A full-length cDNA clone was cut with *Eco*N1 and *Bst*X1 to give a fragment that spanned nucleotides 233 to 1019. This fragment was incorporated into pCMVneo plasmids (25) and checked for correct orientation of the insert. MDA-MB435C cells (5×10^6) were electroporated at 220 V with 10 μg of this construct in the presence of 10 μg/ml DEAE dextran. Vector alone was used as a negative control. Cells (10^6) were plated in five P100 dishes in Alpha + 5% FCS. After 14 days of selection in media containing 1 mg/ml G418, the transfected clones were refed with media containing 0.5 mg/ml G418 for an additional week. Clones were picked in cloning cylinders, expanded, and maintained in Alpha + 5% FCS containing 0.5 mg/ml G418.

RESULTS

Differential Display

Total RNA from a primary breast cancer cell line (21 PT) was compared with that from a metastatic breast cancer cell line from the same patient (21 MT-1) by differential display (DD). Approximately 100 bands appeared for each primer pair tested, and on average two to three bands were differentially expressed. One of the bands

```
   1  AGGCGGACAAAGCCCGATTGTTCCTGGGCCCTTTCCCCATCGCGCCTGGGCCTGCTCCCCAGCCCGG   67
  68  GGCAGGGGCGGGGGCCAGTGTGGTGACACACGCTGTAGCTGTCTCCCCGGCTGGCTGGCTCGCTCTC  134
 135  TCCTGGGGACACAGAGGTCGGCAGGCAGCACACAGAGGGACCTACGGGCAGCTGTTCCTTCCCCCGA  201
 202  CTCAAGAATCCCCGGAGGCCCGGAGGCCTGCAGCAGGAGCGGCC  245

 246  ATG AAG AAG CTG ATG GTG GTG CTG AGT CTG ATT GCT GCA GCC TGG GCA↓GAG  296
   1  Met Lys Lys Leu Met Val Val Leu Ser Leu Ile Ala Ala Ala Trp Ala Glu   17

 297  GAG CAG AAT AAG↓TTG GTG CAT GGC GGA CCC TGC GAC AAG ACA TCT CAC CCC  347
  18  Glu Gln Asn Lys Leu Val His Gly Gly Pro Cys Asp Lys Thr Ser His Pro   34

 348  TAC CAA GCT GCC CTC TAC ACC TCG GGC CAC TTG CTC TGT GGT GGG GTC CTT  398
  35  Tyr Gln Ala Ala Leu Tyr Thr Ser Gly His Leu Leu Cys Gly Gly Val Leu   51

 399  ATC CAT CCA CTG TGG GTC CTC ACA GCT GCC CAC TGC AAA AAA CCG AAT CTT  449
  52  Ile His Pro Leu Trp Val Leu Thr Ala Ala His Cys Lys Lys Pro Asn Leu   68

 450  CAG GTC TTC CTG GGG AAG CAT AAC CTT CGG CAA AGG GAG AGT TCC CAG GAG  500
  69  Gln Val Phe Leu Gly Lys His Asn Leu Arg Gln Arg Glu Ser Ser Gln Glu   85

 501  CAG AGT TCT GTT GTC CGG GCT GTG ATC CAC CCT GAC TAT GAT GCC GCC AGC  551
  86  Gln Ser Ser Val Val Arg Ala Val Ile His Pro Asp Tyr Asp Ala Ala Ser  102

 552  CAT GAC CAG GAC ATC ATG CTG TTG CGC CTG GCA CGC CCA GCC AAA CTC TCT  602
 103  His Asp Gln Asp Ile Met Leu Leu Arg Leu Ala Arg Pro Ala Lys Leu Ser  119

 603  GAA CTC ATC CAG CCC CTT CCC CTG GAG AGG GAC TGC TCA GCC AAC ACC ACC  653
 120  Glu Leu Ile Gln Pro Leu Pro Leu Glu Arg Asp Cys Ser Ala Asn Thr Thr  136

 654  AGC TGC CAC ATC CTG GGC TGG GGC AAG ACA GCA GAT GGT GAT TTC CCT GAC  704
 137  Ser Cys His Ile Leu Gly Trp Gly Lys Thr Ala Asp Gly Asp Phe Pro Asp  153

 705  ACC ATC CAG TGT GCA TAC ATC CAC CTG GTG TCC CGT GAG GAG TGT GAG CAT  755
 154  Thr Ile Gln Cys Ala Tyr Ile His Leu Val Ser Arg Glu Glu Cys Glu His  170

 756  GCC TAC CCT GGC CAG ATC ACC CAG AAC ATG TTG TGT GCT GGG GAT GAG AAG  806
 171  Ala Tyr Pro Gly Gln Ile Thr Gln Asn Met Leu Cys Ala Gly Asp Glu Lys  187

 807  TAC GGG AAG GAT TCC TGC CAG GGT GAT TCT GGG GGT CCG CTG GTA TGT GGA  857
 188  Tyr Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gly  204

 858  GAC CAC CTC CGA GGC CTT GTG TCA TGG GGT AAC ATC CCC TGT GGA TCA AAG  908
 205  Asp His Leu Arg Gly Leu Val Ser Trp Gly Asn Ile Pro Cys Gly Ser Lys  221

 909  GAG AAG CCA GGA GTC TAC ACC AAC GTC TGC AGA TAC ACG AAC TGG ATC CAA  959
 222  Glu Lys Pro Gly Val Tyr Thr Asn Val Cys Arg Tyr Thr Asn Trp Ile Gln  238

 960  AAA ACC ATT CAG GCC AAG   977
 239  Lys Thr Ile Gln Ala Lys   244

 978  TGACCCTGACATGTGACATCTACCTCCCGACCTACCACCCCACTGGCTGGTTCCAGAACGTCTCTCA  1044
1045  CCTAGACCTTGCCTCCCCTCCTCTCCTGCCCAGCTCTGACCCTGATGCTTAATAAACGCAGCGACGT  1111
1112  GAGGGTCCTGATTCTCCCTGGTTTTACCCCAGCTCCATCCTTGCATCACTGGGGAGGACGTGATGAG  1178
1179  TGAGGACTTGGGTCCTCGGTCTTACCCCCACCACTAAGAGAATACAGGAAAATCCCTTCTAGGCATC  1245
1246  TCCTCTCCCCAACCCTTCCACACGTTTGATTTCTTCCTGCAGAGGCCCAGCCACGTGTCTGGAATCC  1312
1313  CAGCTCCGCTGCTTACTGTCGGTGTCCCCTTGGGATGTACCTTTCTTCACTGCAGATTTCTCACCTG  1379
1380  TAAGATGAAGATAAGGATGATACAGTCTCCATCAGGCAGTGGCTGTTGGAAAGATTTAAGATTTCAC  1446
1447  ACCTATGACATACATGGGATAGCACCTGGGCCGCCATGCACTCAATAAAGAATGTATTTTAAAAAAA  1513
1514  AAAAAAAAAAAAA  1526
```

FIG. 2. Protease M cDNA

The cDNA sequence and putative protein coding sequence of the longest clone from the 76N library is shown. The postulated *pre* and *pro* N-terminal amino acids are underlined. The predicted cleavage sites of *pre* and *pro* amino acids after Ala^{16} and Lys^{21}, respectively, are indicated by arrows. The potential N-linked glycosylation site at amino acids 134–136 and Asp^{191} at the bottom of the binding cleft are boxed. The residues of the catalytic triad (His^{62}, Asp^{106}, and Ser^{197}) are circled. The actual polyadenylation signal at nucleotide 1490 and an alternative polyadenylation signal at nucleotide 1095 are underlined.

that was overexpressed in the 21 PT lane (with primer pair OPA 1/T12MG; 280 bp in Fig. 1A) was excised from the gel and PCR amplified. The resulting 280-bp PCR product was used to probe a Northern blot (Fig. 1B). Two bands were detected: a band of 1.7 kb, which was very abundant in 21 PT and barely detectable in 21 MT-1, and a band of approximately 1 kb, which was equal in both cells lines. The mixture was purified, and the differentially expressed clone of 1.7 kb was recovered.

Protease M: Sequence Identification

The 0.28-kb insert was used to screen a cDNA library constructed in λZapII from a normal human mammary epithelial cell line (76N). The longest clone isolated was nearly full-length. This clone of 1526 nt contains 245 bp of 5′-untranslated sequences, 732 bp of coding sequences (coding for a postulated protein of 244 amino acids), and 549 bp of 3′-untranslated sequences (Fig. 2). The presumptive protein coding region begins with an ATG codon, which lies in a good Kozak consensus sequence (26), CGGCCATGA, and ends with a TGA translation stop codon. The amino terminal portion has 13 consecutive hydrophobic residues (Leu^4 to Ala^{16}) which is characteristic of a signal peptide followed by Glu^{17}-Glu-Gln-Asn-Lys^{21}, which resembles a propeptide with a potential trypsin susceptible cleavage site after Lys^{21}. A potential N-linked glycosylation site is found at Asn^{134}-Thr-Thr^{136}. The expected polyadenylation signal AATAAA was found 11 bp upstream of the poly A tail at 1490 bp. Another polyadenylation signal AATAAA was found at 1095 bp.

The postulated protein sequence was compared with the four most closely related proteins using the Pileup and Distances programs, and the comparison was displayed by the Prettyplot program (Fig. 3). Glandular kallikrein 2 (4,5) has 44% exact matches and 48% matches with conservative changes. Trypsin I (27) has 43% exact

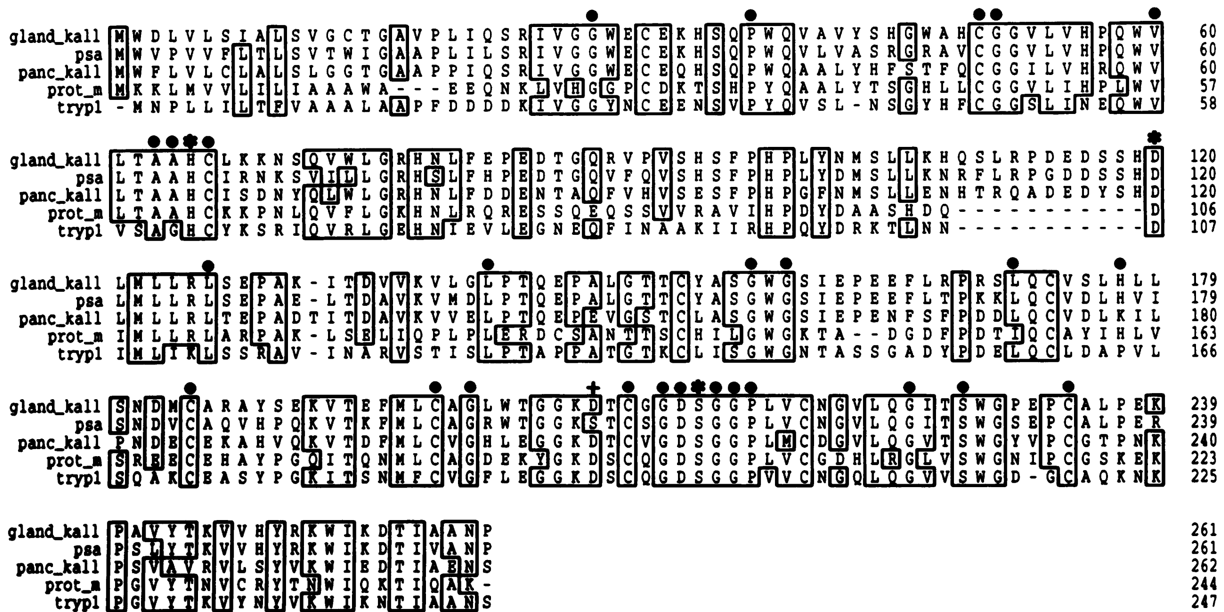

FIG. 3. Alignment of protease M with closely related members of the serine protease family

The GCG Pileup and Pretty plot programs were used to align protease M with closely related human serine proteases. These are (from top to bottom): glandular kallikrein-hk2 (accession no. SP|P06870|), PSA-hk3 (accession no. SP|P07288|), pancreatic kallikrein-hk-1 (accession no. SP|P20511|), and trypsinogen 1 (accession no. SP|P07477|). Amino acids comprising the catalytic triad are marked with an asterisk. The 29 "invariant" amino acids (Dayhoff) are marked with a dot or an asterisk.

matches and 49% match with conservative changes. Both glandular kallikrein 1 (3,28–31) and prostate-specific antigen (32–36) contain 39% exact matches and 44% match with conservative changes. The catalytic triad of serine proteases is conserved in the new protease (i.e., histidine^{62}, aspartate^{106}, and serine^{197}). The presence of aspartate at position 191 predicts that this protein will produce trypsin-like cleavage, unlike PSA, which has a serine at the corresponding position and produces chymotrypsin-like cleavage.

Protease M contains 12 cysteine residues. Ten of these are conserved in the two kallikreins, PSA, and human trypsin, and would be expected to form the following disulfide bridges: Cys^{28}-Cys^{157}, Cys^{47}-Cys^{63}, Cys^{138}-Cys^{203}, Cys^{168}-Cys^{182}, and Cys^{193}-Cys^{218}. The other two cysteines (Cys^{131} and Cys^{231}) are not found in the kallikreins, PSA, and human trypsin, but are found in similar positions in bovine trypsin and would be expected to form a disulfide bond.

Twenty-seven of the 29 "invariant" amino acids surrounding the active site of serine proteases (37) are conserved in protease M. One of the two nonconserved amino acids in protease M, Ileu^{155} in place of Leu, is a conservative change. The other nonconserved amino acid, His^{161} instead of Pro, is also found in glandular kallikrein and PSA. The kallikreins and PSA have 11 amino acid residues, 109–119, which are not found in protease M or trypsin. The function of these amino acids is not clear, but they would be expected to form the so-called kallikrein loop which would determine substrate specificity (38).

Chromosomal Localization

In human DNA, the protease M probe detected a major *Eco*RI fragment of about 18 kb, while in mouse and hamster DNAs several smaller fragments were detected. The human protease M fragment in the hybrid clones segregated with human chromosome 19 (data not shown). There were no discordancies for localization to chromosome 19. To sublocalize the protease M locus, two-color FISH was carried out, using the genomic clone λIG3-1 as a probe. A total of 22 cells were analyzed. Fluorescent signals on one or both chromatids were found in the telomeric region of 19q in 15 metaphases spreads (Fig. 4). Twin-spot signals were not observed on any

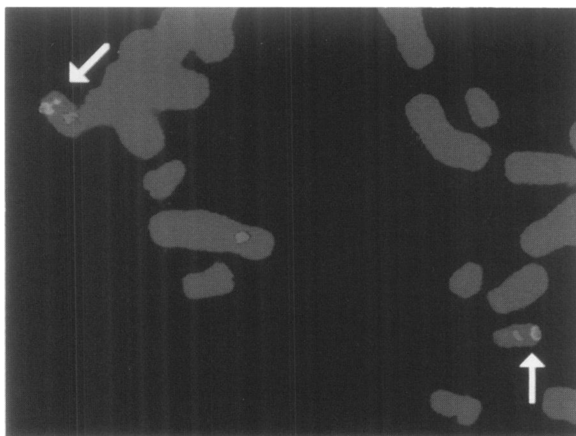

FIG. 4. Chromosomal location of Protease M

The protease M locus was mapped to 19q13.3 by FISH. A genomic protease M probe was co-hybridized with an α-satellite probe specific for chromosomes 1, 5, and 19. Arrows point to protease M–specific hybridization signals in the telomeric region of the long arm of both chromosomes 19 at band q13.3.

other chromosome. Comparison of the banding pattern of chromosome 19 following DAPI-staining allowed us to assign the protease M locus to 19q13.3.

Expression of mRNA in Mammary and Prostate Cells

Figure 5A shows the results of Northern blots of mammary cell lines and strains. The two normal cells strains shown (76N and 70N) and another normal cell strain (81N) not shown expressed the 1.7-kb protease M message at low levels. Two primary tumor lines (21 PT and 21 NT) as well as one metastatic line from the same patient (21 MT-2) expressed high levels of message (approximately 20- to 100-fold higher than the normal strains). The most metastatic cell line from the same patient (21 MT-1), however, expressed low levels of RNA (Fig. 1A). One other primary tumor cell line (BT474) and nine other metastatic cell lines (MCF-7, T47D, ZR-75-1, MDA-MB-157, MDA-MB-231, MDA-MB-361, MDA-MB-435, MDA-MB-436, and BT549) did not have detectable message. Figure 5B shows Northern blots of prostate cell lines. The normal, immortalized cell strains CF3 and CF91 express moderate levels of protease M mRNA, while another normal immortalized strain, MLC, expresses just trace amounts. In contrast, all three of the tumor cell lines examined (DU145, LNCaP, and PC3) failed to express any protease M message.

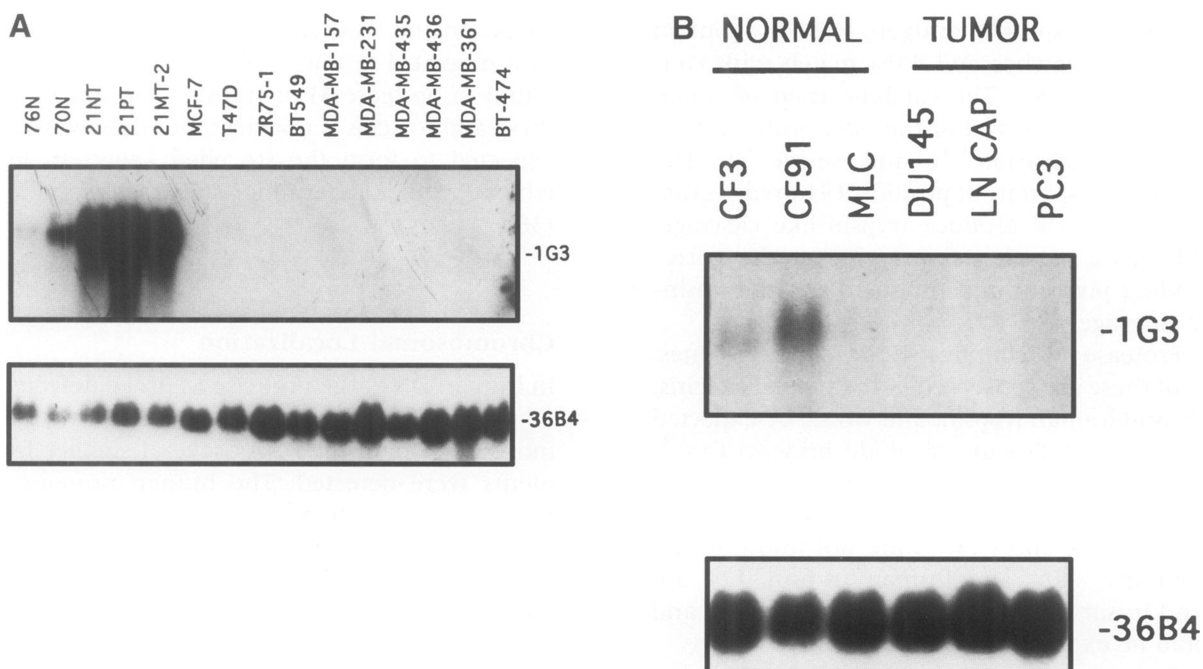

FIG. 5. Protease M mRNA expression in mammary and prostate cell lines

(A) Ten micrograms of total mammary cell RNA was run on an agarose/formaldehyde gel, blotted, hybridized to ^{32}P-labeled protease M probe, and exposed to X-ray film for 20 hr. (B) Ten micrograms of total prostate cell RNA was blotted and hybridized (as in Panel A) and exposed to X-ray film for 20 hr.

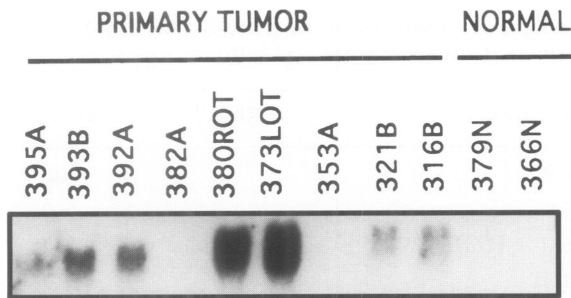

FIG. 6. Protease M mRNA expression in ovarian tissue

Ten micrograms of total cell RNA isolated from ovarian tissue was blotted, hybridized to protease M probe (as in Fig. 4), and exposed to X-ray film for 5 days.

Expression of mRNA in Ovarian Cell Lines and Tissue

A series of normal immortalized and primary tumor derived ovarian cell lines were examined for expression of protease M mRNA on Northern blots. The message was not expressed in any of the five normal immortalized cell lines, but was detected in five of the eight primary tumor cell lines examined (not shown). We also examined the RNA from a series of normal ovarian tissue and biopsies from primary tumors (one of the two Northern blots is shown in Fig. 6). While mRNA was not expressed in the three normal tissues examined, the six borderline ovarian tumor tissues, or the two metastatic tumors from colon primaries, it was expressed in 16 of the 20 primary ovarian tumor tissue specimens examined.

Expression of Protease M mRNA in Normal Human Tissue

A Northern blot containing 2 μg of polyA$^+$ RNA from eight normal human tissues (Clontech, Palo Alto, CA, U.S.A.) was examined for expression of protease M (Fig. 7). While the message was not detected in heart, placenta, lung, liver, or skeletal muscle, high levels of message were detected in brain, kidney, and pancreas. The message detected in brain and kidney was 1.7 to 1.8 kb, but the message detected in pancreas was only about 1.2 kb. A probable explanation for the smaller message in pancreatic RNA would be the use of the alternative polyadenylation signal at 1090 bp noted in Fig. 1.

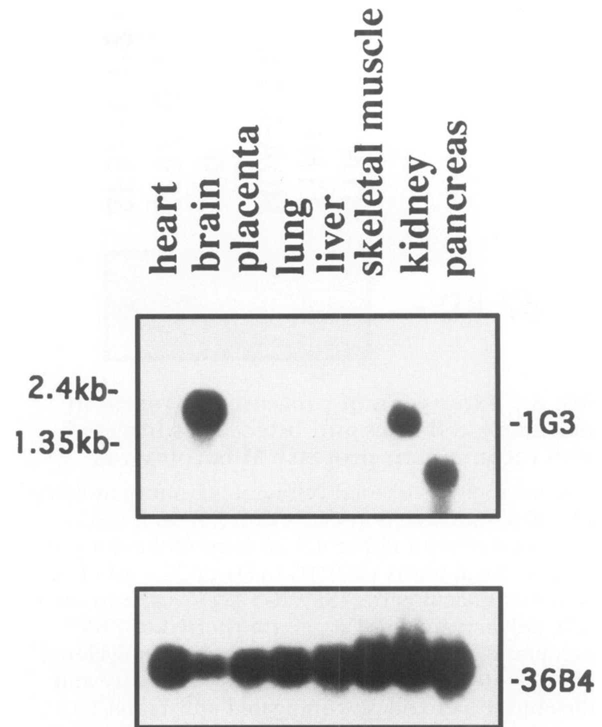

FIG. 7. Protease M mRNA expression in human tissue

A Northern blot containing 2 μg of polyA$^+$ RNA from normal human tissue (Clontech) was hybridized to protease M probes (as in Fig. 4). The blot was exposed to X-ray film for 2 days.

Production of Polyclonal Antibody and Its Use to Study Expression of Protein in Mammary Cell Lines and Strains

A polyclonal antibody was produced in rabbits against a hydrophilic peptide which was not highly conserved among other serine proteases (^{73}GKHNLRQRESSQEQS^{87}). The Western blot (Fig. 8) shows that the antibody detects a protein of 37 kD in total cell lysates of the normal mammary epithelial cell strain 81N and in the primary tumor cell line 21NT. Protease M protein is not detected in the metastatic breast cell line MDA-MB-435. In other Western blots (not shown), the antibody detected a 37-kD protein in the normal strains 70N and 76N, as well as the primary tumor cell line 21PT, but not in the metastatic cell lines T47D and MCF-7. Up to 1 ml of conditioned media from 70N and 21NT was examined in Western blots probed with this antibody, and no reacting proteins were detected (not shown). This result suggests that the protein is primarily localized intracellularly and not secreted. The protein detected by the antibody is 37

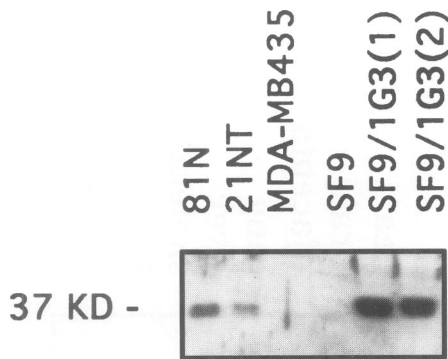

FIG. 8. Expression of protease M protein in mammary cell lines and insect cells infected with recombinant protease M baculovirus

Fifty micrograms of total cell lysates from mammary cell lines, uninfected insect cells (SF9), and insect cells infected with either 4.5 μl of recombinant protease M baculovirus (SF9/1G3 [1]) or 22.5 μl of recombinant baculovirus (SF9/IG3 [2]) were run on a 12% polyacryamide/SDS gel, transferred to PDVF membrane, and reacted with protease M polyclonal anti-peptide antibody as the primary antibody and horseradish peroxidase-conjugated anti-rabbit IgG secondary antibody. Bands were detected with ECL.

TABLE 1. Expression of protease M mRNA and protein in mammary cells

Cell Line	RNA[a]	Protein[b]
70N	5	100
81N	4	60
16-1-1 (76N/HPV16)	4	64
21NT	85	47
21PT	100	76
MDA.MB435	0	0
T47D	0	0
MCF-7	0	0

[a]RNA values were obtained by running 10 μg of total RNA on a Northern blot, by hybridizing it to ^{32}P-labeled protease M probes, and by quantitating the resulting autoradiograms. The most intense band was set equal to 100 and the other values normalized accordingly.

[b]Protein values were obtained by running 50 μg of total cell lysates on a Western blot and probing it with the protease M antibody, as described in Materials and Methods. The 37-kD bands on the autoradiograms were quantitated; the most intense band was set equal to 100 and the other values normalized accordingly.

kD, while the amino acid sequence predicts a protein of about 27 kD. The potential glycosylation site at (^{134}Asn-Thr-Thr^{136}) might explain this size discrepancy.

Table 1 shows that the RNA levels for the serine protease are not always correlated with the protein levels. While the primary tumor cell lines (21NT and 21PT) have 20 to 100 times more protease M mRNA than normal cell strains (70N, 76N, and 81N), the protein detected on Western blots is equal to or somewhat lower in the primary tumor cell lines than in the normal cell strains.

The antipeptide polyclonal protease M antibody has been used successfully in Western blots but does not seem to work in cellular immunofluorescence studies in which the antibody has given a high background with MDA-MB-435 cells, which do not express the protease M message.

Production of Recombinant Protein

Extensive efforts were made to produce recombinant protein for further study of the protease. As discussed below, production neither in *E. coli* as a GST-fusion protein nor in baculovirus as a pure protein was successful in providing more than minimal amounts of the protease. As a result, the products that were recovered were used primarily to verify the specificity of the antibody preparations.

In a further effort to obtain recombinant protein, transfectants were produced expressing protease M in the mammary tumor cell line MDA-MB-435. Transfectants were screened initially for protein production and, as shown below, the results demonstrated that only 5 of the 76 transfectants produced any protein and this at low levels.

Production of GST Fusion Protein

The expected 52-kD GST/protease M fusion protein was purified and yielded approximately 600 μg of fusion protein per 500 ml culture. When the fusion protein was cleaved by incubation with thrombin, the protease M fragment was degraded, even at limiting dilutions, while only the GST portion remained intact. At least 1 μg of fusion protein was required to obtained a detectable signal on Western blots.

Production of Baculovirus Recombinant Protein

A Western blot containing 50 μg of lysates prepared from SF9 cells infected with an amplified

stock of protease M recombinant baculovirus was probed with anti-protease M antibody (Fig. 8). Whereas reacting proteins were not detected in the lysate from uninfected SF9 cells, a protein of 39 kD was detected in lysates of SF9 infected with recombinant baculovirus. Sf9/1G3 (1) had approximately 50% infected, enlarged cells, and SF9/1G3 (2), which was infected with five times as much virus, had nearly 100% infected cells. The amount of recombinant protein was, however, quite low, and we were not able to detect a band of 39 kD on Commassie blue-stained gels (not shown).

We attempted to purify recombinant protease M from lysates. By using *p*-aminobenzamidine agarose affinity chromatography followed by concanavalin A agarose, recombinant protease M was purified approximately 80-fold. The protein was, however, still only 10% pure, as determined from silver-stained gels, and the yield was less than 1 μg/10^8 cells. Using this data, we calculated that 50 μg of lysate contains 15 ng of protease M or 0.03% of the total protein. Furthermore, by comparing the amount of the 39-kD band on silver-stained gels of the 80-fold purified protease M with Western blots of the purified protein, we determined that the antibody can detect 5 ng of protease M protein as a lower limit.

MDA-MB435 Transfectants

A pCMV/neo/protease M construct and a neo-vector control were transfected into MDA-MB-435 cells (5×10^6 cells for each construct) by electroporation. Eighty colonies of protease transfected clones and 20 colonies of vector transfected clones were transferred to 24-well dishes. The protease-transfected cells grew more slowly and had more enlarged, dying cells than the vector controls. Total cell lysates were prepared from the 76 protease transfectants when the cells were approximately 70% confluent. Western blots, prepared from 50 μg of the lysate from the 76 transfectants as well as 50 μg of lysate from 70N (positive control), were probed with the protease M antibody. Only 2 of 42 fast-growing clones and 3 of 34 slow-growing clones expressed any detectable protein (data not shown). Furthermore, the level of protein expressed by these positive clones was, in all cases, considerably less than in 70N cells.

Table 2 shows that protease M RNA was found in clones expressing protein as well as the majority of those not expressing protein. Thus, in MDA-MB-435 cells either the message is not translated efficiently or the translated protein is extremely unstable.

TABLE 2. Analysis of protease M RNA and protein expression in MDA-MB-435 tranfectants

Cell Line	RNA[a]	Protein[b]
70N	12	100
MDA-MB-435	0	0
Protease M transfectant		
#13	4	0
#19	10	0
#42	96	25
#44	61	12
#53	0	0
#58	100	0
#59	6	0
#64	22	0
#65	44	25
#66	55	63
#75	22	0
#86	0	0

[a,b]These values were determined as in the footnote to Table 1.

DISCUSSION

In a search for novel genes involved in metastasis, we have isolated by differential display mRNAs whose expression differs in a primary breast tumor cell line and a metastatic cell line derived from the same patient. Here, we describe the isolation of an mRNA that encodes a novel member of the serine protease family closely related in sequence to trypsin and to PSA. This novel gene, protease M, is strongly expressed in primary breast cancer cell lines and in primary ovarian cancers but is down-regulated in cell lines derived from breast tumor metastases. This expression pattern suggests that protease M may be important in establishing breast and ovarian primary tumors, and may function later in progression as a potential metastasis inhibitor.

The projected sequence of the encoded protein had about 40% amino acid identity to trypsin and members of the kallikrein family (glan-

dular kallikrein, pancreatic kallikrein, and prostate-specific antigen). Structural features important for serine protease activity, such as the catalytic triad, the residues lining the binding cleft, and the cysteine bridges, were almost perfectly conserved. Unlike the members of the kallikrein family, protease M and trypsin lack the kallikrein loop at amino acid residues 109–119, which is important for kallikrein specificity. The size of the detected protein is 36 kD rather than the predicted size of 27 kD. This size discrepancy could be accounted for by glycosylation at Asn^{134} as in PSA (9).

The protease M gene was mapped by somatic cell hybrid and FISH analyses to chromosome 19q13.3. Several other members of the serine protease family, including pancreatic/renal kallikrein (KLK1), glandular kallikrein (KLK2), and PSA, map to 19q13.3 (39). Physical mapping of this region has revealed that all three kallikrein genes are clustered within a 60-kb region and that the order is KLK1-PSA-KLK2 (40). Our mapping data suggest that protease M may also be part of this gene cluster and that all four genes may have originated from a single ancestral precursor gene. Trypsin 1 (TRY1) and other more distantly related serine proteases, such as granzyme B (14q11.2), cathepsin G (14q11.2), coagulation factor VII (13q34), and protein C (2q13–q21), however, are all on other chromosomes, indicating that during evolution this gene cluster has been split up to several human chromosomes. Interestingly, chromosome band 19q13 is nonrandomly rearranged in a variety of human solid tumors, including pancreatic carcinomas, astrocytomas (grades III and IV), hemangiopericytomas, ovarian cancers, and thyroid tumors (41). Whether any of these rearrangements affect protease M remains to be determined.

Protease M mRNA was expressed in normal breast and prostate cell strains but not in cell lines derived from metastatic tumors. The basis for this down-regulation is not known; however, since Southern blots show that the gene is neither lost nor grossly rearranged in breast tumor cell lines (data not shown), it may be transcriptional. While protease M mRNA was expressed in normal human brain, kidney, and pancreatic tissue, it was absent in heart, placenta, lung, liver, and skeletal muscle. This expression pattern resembles that of the kallikrein HKLK1 (3). In contrast, PSA is expressed primarily in the prostrate, though low levels have recently been detected in breast tissue (42).

We have produced a polyclonal anti-peptide antibody to protease M and have detected the protein in whole cell lysates, but not in conditioned growth media. The question of whether protease M is secreted, however, remains open. Based on its sequence, the protein appears capable of secretion. The amino-terminal sequences encode putative *pre* and *pro* regions, and the related pancreatic kallikrein and trypsin proteins are secreted. Since our anti-peptide antibody is rather weak (a lower limit for detectability of 5 ng on Western blots), detection of low amounts of secreted protein in conditioned media and biological fluids will have to await the production of a high-affinity antibody to use, for example, in radioimmunoassays.

The protease M protein was detected in lysates of normal breast epithelial cell lines, but not in breast metastatic cell lines, and correlated with the mRNA expression levels. In the primary breast tumor cell line 21PT, however, very high mRNA levels were observed, though the protein level was low. This lack of correlation of mRNA and protein levels was also observed in MDA-MB-435 cells that were transfected with a protease M expression construct. Although many of the tranfectants expressed high levels of the mRNA, the protein was absent or barely detectable. These observations suggest that the expression of protease M is regulated both at the transcriptional and translational level. Whether the low protein levels are due to inefficient translation or rapid degradation of the translated protein is not known. Pulse-chase labeling experiments, followed by immunoprecipitation and Western blot analysis of protease M might elucidate the mechanism of translational down-regulation.

In breast, while the primary tumor cell lines (21 NT and 21 PT) expressed high levels of protease M message, another primary tumor cell line (BT474) did not express this message. When breast tissue samples were examined, two of the four primary tumor biopsies produced high levels of protease M message, whereas samples from normal adjacent tissue or reduction mammoplasties produced lower levels of the message (data not shown). Therefore, high protease M mRNA might serve as a marker for a subset of primary tumors. This subset of primary tumors could be distinguished from normal tissue by the higher levels of protease M mRNA. Protease M mRNA was expressed in the majority of primary ovarian tumor cell lines and tissues, but not in borderline ovarian tumor tissue, normal cell lines, or normal tissues.

These observations suggest that protease M might be useful as a diagnostic marker for primary epithelial carcinomas. Furthermore, the close sequence relatedness to trypsin and the presence of a signal peptide in the N terminus raise the possibility that the protein may have valuable medical applications.

ACKNOWLEDGMENTS

We thank Dr. Shijie Sheng for her advice on the protein studies and Tom Graf for his assistance with computer analysis. Binh Truong is thanked for her expert technical assistance. This work was supported in part by NIH grants CA57517 and CA61253.

REFERENCES

1. Schachter M. (1980) Kallikreins (kininogenases)—A group of serine proteases with bioregulatory actions. *Pharmacol. Rev.* **31:** 1–17.
2. Berg T, Bradshaw RA, Carretero OA, et al. (1992) A common nomenclature for members of the tissue (glandular) kallikrein gene families. Recent progress on kinins. *Agents Actions* **38**(Suppl 1): 19–25.
3. Fukushima D, Kitamura N, Nakanishi S. (1985) Nucleotide sequence of cloned cDNA for human pancreatic kallikrein. *Biochemistry* **24:** 8037–8043.
4. Schedlich LJ, Bennetts BH, Morris BJ. (1987) Primary structure of a human glandular kallikrein gene. *DNA* **6:** 429–437.
5. Riegman PH, Vlietstra RJ, Van der Korput HA, et al. (1991) Identification and androgen-regulated expression of two major human glandular kallikrein–1 (hGK-1) mRNA species. *Mol. Cell Endocrinol.* **76:** 181–190.
6. Lilja H. (1985) A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. *J. Clin. Invest.* **76:** 1899–1903.
7. Lilja H, Oldbring J, Rannevik G, et al. (1987) Seminal vesicle-secreted proteins and their reactions during gelation and lique-faction of human semen. *J. Clin. Invest.* **80:** 281–285.
8. McCormack RT, Rittenhouse HG, Finlay JA, et al. (1995) Molecular forms of prostate-specific antigen and the human kallikrein gene family: A new era. *Urology* **45:** 729.
9. Liang L, Pardee AB. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257:** 967–970.
10. Liang L, Averboukh L, Pardee AB. (1993) Distribution of eukaryotic mRNAs as 2-D arrays by means of differential display. *Nucleic Acids Res.* **21:** 3267–3275.
11. Sager R, Anisowicz A, Neveu M, et al. (1993) Identification by differential display of alpha 6 integrin as a candidate tumor suppressor gene. *FASEB J.* **7:** 964–970.
12. Band V, Sager R. (1989) Distinctive traits of normal and tumor-derived human mammary epithelial cells expressed in a medium that supports long-term growth of both cell types. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 1249–1253.
13. Band V, Zajchowski D, Swisshelm K, et al. (1990) Tumor progression in four mammary epithelial cell lines derived from the same patient. *Cancer Res.* **50:** 7351–7357.
14. Tsao SW, Mok SC, Fey E, et al. (1995) Characterization of human ovarian surface epithelial cells immortalized by human papilloma viral oncogenes (HPV-E6E7 ORFs). *Exp. Cell Res.* **218:** 499–507.
15. Mok SC, Tsao SW, Knapp RC, et al. (1992) Unifocal origin of advanced human epithelial ovarian cancers. *Cancer Res.* **52:** 5119–5122.
16. Feinberg AP, Vogelstein B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments in high specific activity. *Anal. Biochem.* **132:** 6–13.
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.
18. Sambrook J, Fritsch EF, Maniatis T. (1989) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19. Masiakowski P, Breathnach R, Bloch J, et al. (1982) Cloning of cDNA sequences of hormone-regulated genes from the MCF-7 human breast cancer cell line. *Nucleic Acids Res.* **10:** 7895–7903.
20. Roijer E, Kas K, Klavitz I, Bullerdiek J, Van de Ven W, Stenman G. (in press) Identification of a YAC spanning the chromosome 8q12 translocation breakpoint in pleomorphic adenomas with t(3;8)(p21;q12). *Genes Chromosom Cancer.*
21. Tam JP. (1988). Synthetic peptide vaccine design: Synthesis and properties of a high-

density multiple antigenic peptide system. *Proc. Natl. Acad. Sci. U.S.A.* **85:** 5409–5413.
22. Smith DB, Johnson KS. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67:** 31–40.
23. Gruenwald S, Heitz J. (1993) *Baculovirus Expression Vector System: Procedures and Methods Manual.* 1st ed. Pharmingen, San Diego, CA, pp. 34–55.
24. Sheng S, Pemberton PA, Sager, R. (1994) Production, purification, and characterization of recombinant maspin proteins. *J. Biol. Chem.* **269:** 30988–30993.
25. Tomasetto C, Neveu MJ, Daley J, et al. (1993) Specificity of gap junction communication among human mammary cells and connexin transfectants in culture. *J. Cell Biol.* **122:** 57–167.
26. Kozak M. (1984) Compilation and analysis of sequences upstream from the translational start site eukaryotic mRNAs. *Nucleic Acids Res.* **12:** 857–872.
27. Emi M, Nakamura Y, Ogawa M, et al. (1986) Cloning, characterization and nucleotide sequences of two cDNAs encoding human pancreatic trypsinogens. *Gene* **41:** 305–310.
28. Baker A, Shine J. (1985) Human kidney kallikrein: cDNA cloning and sequence analysis. *DNA* **4:** 445–450.
29. Takahashi S, Irie A, Miyake Y. (1988) Primary structure of human urinary prokallikrein. *J. Biochem.* **104:** 22–29.
30. Lu HS, Lin F, Chao L, et al. (1989) Human urinary kallikrein: Complete amino acid sequence and site of glycosylation. *Int. J. Pept. Protein Res.* **33:** 237–249.
31. Angermann A, Bergmann C, AppelHans H. (1989) Cloning and expression of human salivary-gland kallikrein in *Escherichia coli. Biochem. J.* **262:** 787–793.
32. Watt KWK, Lee PJ, M'Timkulu T, et al. (1986) Human prostate-specific antigen: Structural and functional similarities with serine proteases. *Proc. Natl. Acad. Sci. U.S.A.* **83:** 3166–3170.
33. Lundwall A, Lilja H. (1987) Molecular cloning of human prostate specific antigen cDNA. *FEBS Lett.* **214:** 317–322.
34. Schaller J, Akiyama K, Tsuda R, et al. (1987) Isolation, characterization and amino-acid sequence of *g*-seminoprotein, a glycoprotein from human seminal plasma. *Eur. J. Biochem.* **170:** 111–120.
35. Riegman PHJ, Klaassen P, Van der Korput, et al. (1988) Molecular cloning and characterization of novel prostate antigen cDNA's. *Biochem. Biophys. Res. Commun.* **155:** 181–188.
36. Henttu P, Vihko P. (1989) cDNA coding for the entire human prostate specific antigen shows high homologies to the human tissue kallikrein genes. *Biochem. Biophys. Res. Commun.* **60:** 903–910.
37. Dayhoff MO. (1978) Atlas of protein sequence and structure. *Natl. Biomed. Res. Found.* (Wash DC) **5**(Suppl 3): 79–81.
38. Ashley PL, MacDonald RJ. (1985) Kallikrein-related mRNAs of the rat submaxillary gland: Nucleotide sequences of four distinct types including tonin. *Biochemistry* **24:** 4512–4520.
39. Ropers HH, Mohrenweiser H. (1993) Report of the committee on the genetic constitution of chromosome 19. Chromosome Coordinating Meeting (1992). In: Cuticchia AJ, Person PL, Klinger HP (eds). *Genome Priority Reports*. Basel, Karger. Vol. 1.
40. Riegman PHJ, Vliestra RJ, Suurmeijer L, Cleutjens CBJ, Trapman J. (1992) Characterizations of the human kallekrein locus. *Genomics* **14:** 6–11.
41. Mitelman F. (1994) *Catalog of Chromosome Aberrations in Cancer*. 5th ed. Wiley-Liss, New York, pp. 3067–3198.
42. Yu H, Diamandis EP, Sutherland DJ. (1994) Immunoreactive prostate-specific antigen levels in female and male breast tumors and its association with steroid hormone receptors and patient age. *Clin. Biochem.* **27:** 75–79.

Contributed by R. Sager on June 21, 1996.