AD_____

Award Number: DAMD17-96-1-6238

TITLE: Regulation of Growth Factor Release by Protease Inhibitors in Breast Cancer

PRINCIPAL INVESTIGATOR: Thomas H. Finlay, Ph.D.

CONTRACTING ORGANIZATION: New York University Medical Center
New York, New York 10016

REPORT DATE: September 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1999	Annual (15 Aug 98 -14 Aug 99)

4. TITLE AND SUBTITLE
Regulation of Growth Factor Release by Protease Inhibitors in Breast Cancer

5. FUNDING NUMBERS
DAMD17-96-1-6238

6. AUTHOR(S)
Thomas H. Finlay, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
New York University Medical Center
New York, New York 10016

E-MAIL:
finlat1@mrcr0.med.nyu.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Authorized for Release;
Distribution Unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

Epithin, a potential TGFα cleaving enzyme has been cloned from MCF-7 breast cancer cells. The cDNA has a deduced 855 amino acid sequence consistent with that of a multi-domain, type II membrane protease. Epithin shows both membrane spanning and serine protease domains. It also has putative CUB and LDRA domains, which may be important for regulation and/or substrate binding. Northern blotting showed highest expression in epithelial cells particularly breast and kidney. Epithin antibodies blocked release of TGFα from MCF-7 cells suggesting that epithin may be involved proTGFα processing. Incubation of MCF-7 cells with the metalloprotease inhibitor BB-3103 blocked epithin processing suggesting a metalloprotease may needed for epithin activation. Membrane fractions from MCF-7 cells incubated with a labeled serine protease inhibitor showed a 33 kDa labeled band, which was recognized by the epithin antibody. Epithin also shows gelatinase activity suggesting that it may be directly involved in invasion. Immunohistochemistry of paraffin sections from normal human breast tissue showed that the epithin protein was restricted to the surface of ductal epithelial cells where it co-localized with the epidermal growth factor receptor. Epithin expression appeared to be elevated in DCIS and in invasive breast carcinoma.

14. SUBJECT TERMS
Breast Cancer, TGF-alpha, Serine Protease, Protease inhibitor, invasion

15. NUMBER OF PAGES
16

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

X In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

PI - Signature

9/1/99

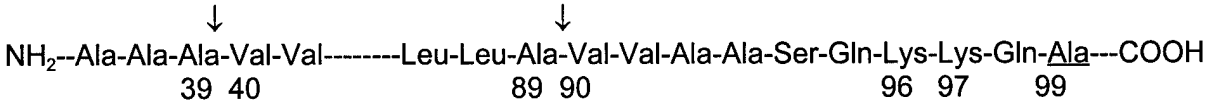
Date

Table of Contents

INTRODUCTION

The broad aim of this proposal is to test the hypothesis that the coordinate action of proteases and protease inhibitors are responsible for growth factor activation and release in the human breast cancer cell, and for the progression of breast cancer *in vivo*. One-third of all cases of advanced breast cancer are estrogen responsive, and recent epidemiological studies and studies using human breast cancer cells in culture strongly suggest a correlation between estrogens and the pathogenesis of breast cancer (Harris et al 1992). The mechanisms by which this occurs are not entirely clear. Certainly, estrogens have a direct effect on cell growth. In addition, they can stimulate the expression and release of a variety of polypeptide growth factors. It is highly likely that the tummorigenic effects of estrogens are due, at least in part, to the autocrine/paracrine action of these factors. Several of these polypeptides, including epidermal growth factor (EGF), and its analogs heregulin and transforming growth factor-α (TGFα) and the insulin-like growth factors (IGF-I and IGF-II), have been shown to require pericellular proteolysis for activation or release (Massagué and Pandiella 1993, Hooper et al 1997). To achieve homeostasis in a normal breast epithelial cell, levels of these pericellular growth factor activating proteases also must be regulated. We hypothesized that this was accomplished by the action of locally synthesized protease inhibitors. Thus an imbalance in the ratio between local levels of particular proteases and protease inhibitors could be responsible for increases in tumorigenic potential.

TGFα, a peptide structurally and functionally related to EGF, interacts with the EGF receptor (EGFR) and elicits a mitogenic response in a variety of cells (Lee et al 1995). TGFα expression occurs in normal breast tissue, breast tumors and breast cancer cells in culture and TGFα has been proposed to act as a major autocrine mediator of estrogen-stimulated growth in estrogen-dependent breast cancer cells (Harris et al 1992). Expression of the TGFα/EGFR pair have been shown to be associated with proliferation and angiogenesis in invasive breast cancer (De Jong et al 1998). TGFα is synthesized as part of a 20-22 kDa glycosylated, type I membrane protein precursor (proTGFα) which can be processed intracellularly and extracellularly by glycosylation and proteolysis to yield a family of polypeptides of from 6 to 17 Kda in size. (Massagué and Pandiella 1993, Baselga et al, 1996). Proteolytic processing occurs in two steps, the first results in cleavage between Ala^{39} and Val^{40} (Scheme 1). The second, occurring closer to the cell membrane, results in release of the 6 KDa mature TGFα peptide. While it is generally accepted that the second cleavage occurs between Ala^{89} and Val^{90}, this does not rule out cleavage at another site (e.g between Lys^{96} and Lys^{97}). Complete processing does not occur to the same extent in all tissues and it is certainly possible that shedding of TGFα from the cell surface may be accomplished by different proteases in different tissues (Hooper et al 1997, Arribas et al 1996). While all TGFα forms appear to possess some degree of biological activity, there is good evidence that particular biological actions may depend on the degree of proteolytic processing.



Scheme 1. Cleavage sites in ProTGFα yielding mature 50 aa TGFα peptide. Arrows indicate the normal cleavage sites. Ala^{99} begins the hydrophobic membrane-spanning region.

Results from this laboratory have clearly demonstrated a relationship between tumorigenicity of MCF-7 human breast cancer cells (as measured by growth in soft agar), endogenous synthesis of the protease inhibitor α_1-antitrypsin (α_1-AT) and release of TGFα (Tamir et al 1990, Finlay et al 1993a,b). Growth in soft agar was blocked by α_1-AT whether added to the tissue culture media or synthesized by the tumor cell itself. A useful tool in these studies was a new MCF-7 cell subline, producing 10-fold higher levels of α_1-AT than its parental cell line, constructed by stable transfection with an α_1-AT cDNA (Yavelow et al 1997). Growth in soft agar and release of TGFα was decreased in cells transfected with

the α_1-AT cDNA when compared to cells transfected with vector alone. Consistent with the above we had identified a serine protease with elastase-like activity, capable of forming a stable complex with α_1-AT, on the MCF-7 cell surface.

Our specific aims as stated in our grant application were to:

1. To identify and clone the growth-modulating pericellular proteases from MCF-7 cells, particularly, the elastase-like enzyme(s) that are able to effect the release of TGFα from the tumor cell surface.
2. To show that the ability of MCF-7 sublines to form colonies in soft agar and tumors in nude mice is a function of their expression of α_1-AT and specific pericellular proteases.
3. To extend our observations relating to TGFα release/activation made in MCF-7 human breast cancer cells in culture to other breast cancer cell lines, and to normal and malignant human breast tissue.
4. To test the hypothesis that the pericellular protease:protease inhibitor ratio responsible for TGFα release can be regulated by ST-3 and other potential effectors such as antiestrogens, phorbol esters and SEC receptor agonists. This information may provide insight into mechanisms by which protease and protease inhibitor levels may be independently controlled.

To accomplish the above specific aims, we proposed to carry out the sequence of studies described in our Statement of Work. While the project has not proceeded in exactly the order as originally anticipated, I believe that we are fairly well on track. Our progress is outlined below.

YEAR	STATUS OF PROJECTED STUDY
1-2	Isolate, characterize and clone proteases from MCF-7 cells. This phase of the work has been completed. A revised manuscript describing the cloning and properties of the TGFα cleavage protease (which we now call human epithin) has been submitted for publication.
1-3	Compare production of protease and protease inhibitors by MCF-7 cell sublines with their ability to form colonies in soft agar and cause tumor formation in nude mice. Because of difficulties in generating the MCF-7 cell clones hyper-producing pro-TGFα, PCR-7 protease, and α_1-antitrypsin-(Pittsburgh) and because of problems with the identity of the TGFα cleavage enzyme, this project is a little behind schedule. Studies of colony formation in soft agar have been completed. We have just initiated the nude mouse experiments. This phase of the work should be completed within 60 days.
1-3	Extend our observations relating to TGFα release/activation made in MCF-7 human breast cancer cells in culture to other breast cancer cell lines, and to normal and malignant human breast tissue. We have demonstrated production of PCR-7 protease in normal and malignant breast epithelial tissue and anticipate looking at TGFα shedding in other breast cancer cell lines and in breast epithelial tissue over the next several months. Breast cancer cell studies using the non-transformed cell line MCF-10A, the ER-negative cell lines BT-20 and T47D and a second ER-positive cell line, ZR-65-1 are being completed in my laboratory.
2-4	Examine production and localization of proteases, protease inhibitors, growth factor receptors and sites of growth factor activation in normal and malignant breast tissue. In collaboration with Dr. Helen Feiner (Department of Anatomic Pathology) we have begun to look at histochemical localization of PCR-7 protease, EGF receptor and ProTGFα in normal and maligmnant breast tissue. These studies are continuing.

1-4 Test the hypothesis that the pericellular protease:protease inhibitor ratio responsible for TGFα release can be regulated by ST-3 and other potential effectors Identify potential modulators of protease inhibitor: protease ratio and growth factor activation/release in breast cancer cells and in ductal epithelium from human breast. These studies have been initiated

RESULTS AND DISCUSSION

1. Cloning of Human Epithin: a Potential TGFα Releasing Serine Protease.

Degenerate oligonucleotides based on the conserved sequences about the his^{57}, asp^{102} and ser^{195} residues in mammalian serine proteases (chymotrypsin) were used to clone a potential TGFα-releasing serine protease from an MCF-7 breast cancer cell cDNA library. In the initial step, a series of approximately 500 bp fragments between the his^{57} and ser^{195} sites were amplified by PCR. The amplified sequences were then cloned into a PCR cloning vector, which was used to construct a mini-cDNA library. Clones from the mini library were selected by Southern blotting using a ^{32}P-labeled degenerate oligonucleotide probe based on the sequence about asp^{102}. Several positive clones were sequenced. The 460 bp sequence of one strongly hybridizing clone, designated PCR-7, showed a high degree of homology to known serine proteases (64% identity in a 220 bp region at the 3' terminus to human trypsinogen-B and a 60% identity in a 120 bp region around the 5' terminus to human pancreatic protease). Northern blot analysis, using a PCR-7 cDNA probe, showed the expression of an approximately 4.2 kb mRNA species in total RNA from MCF-7 cells (**Fig. 1**). Similar mRNA species were expressed in a several different human breast cancer cell lines and in normal human breast tissue, in trophoblast from term human placenta and in proliferative phase uterine glandular epithelium.

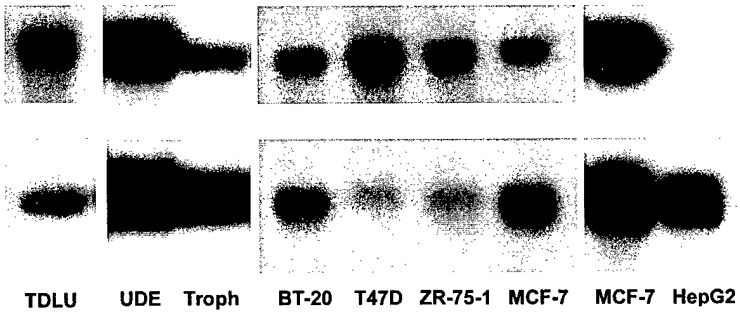


Figure 1. Northern blot analysis using the 500 b PCR-7 cDNA probe. Total RNA, prepared from various human tissues and cell lines as indicated below, were electrophoresed on a 1.5% agarose/ formaldehyde gel, transferred to GeneScreen nylon membrane. Membranes were probed with a ^{32}P-labelled PCR-7 cDNA insert and then reprobed with a ^{32}P-labelled actin cDNA. Lane 1, Terminal duct lobular Unit from human breast ductal epithelium; lane 2; human ductal glandular epithelial cells; lane 3, trophoblast cells from term human placenta; lanes 4-7, various human breast cancer cell line; lanes 8-9, MCF-7 breast cancer cells and HepG2 human liver cancer cell lines

Two cDNAs of approximately 3.8 and 3.2 kb containing the PCR-7 sequence were cloned from a second MCF-7 cell cDNA library using the PCR-7 sequence as a probe. Both clones hybridized to the same sized mRNA as did the PCR-7 sequence used to screen the library. Both cDNAs have been sequenced and except for an additional 600 nucleotides in the 5' terminus of the larger clone and some minor differences in the 3' sequence, the two sequences appear to be identical where they overlap. Because of inconsistencies between the MW of the TGFα releasing protease calculated from the deduced amino acid sequence and the apparent MW observed on western blotting, the larger, 3.8 kb, cDNA has been resequenced for a third time and we believe all ambiguities have finally been resolved (**Fig. 2**). The sequence has a 2565 b open reading frame coding for an 855 amino acid protein of approximately 95 kDa and a 542 b 5' sequence following the stop codon at position 2721. A Kozack consensus sequence for initiation of eukaryotic translation is present at the putative methionine initiation codon at position 156 (Kozack et al 1990).

Figure 2. Complete nucleotide sequence of human epithin/matriptase cDNA and its deduced amino acid sequence. The potential membrane-spanning region is doubly underlined. The two putative CUB domains have black letters on a gray background while the four putative LDRa domains have white letters on a black background. The poteintial protease activation site after Arg^{614} is marked by an arrow. The residues at the putative catalytic site (His^{656}, Asp^{711} and Ser^{805}) are underlined and in boldface.

```
   1  gagcgcgccagggcgagggcaccgccgccggtcgggcgcgctgggcctgcccggaatcccgccgcctgcgccccgcgccccgcgccctgcgggccatgggagccggccgccggcagggac  120

 121  gacgcctgtgagacccgcgagcggcctcggggaccatgggtagcgatcgggcccgcaagggcggaggggccgaaggacttcggcgcgggactcaagtacaactcccggcacgagaaag  240
   1                                  M  G  S  D  R  A  R  K  G  G  G  G  P  K  D  F  G  A  G  L  K  Y  N  S  R  H  E  K  V   29

 241  tgaatggcttggaggaaggcgtggagttcctgccagtcaacaacgtcaagaaggtggaaaagcatggcccggggcgctgggtggtgctggcagccgtgctgatcggcctcctcttggtct  360
  30   N  G  L  E  E  G  V  E  F  L  P  V  N  N  V  K  K  V  E  K  H  G  P  G  R  W  V  V  L  A  A  V  L  I  G  L  L  L  V  L   69

 361  tgctggggatcggcttcctggtgtggcatttgcagtaccgggacgtgcgtgtccagaaggtcttcaatggctacatgaggatcacaaatgagaattttgtggatgcctacgagaactcca  480
  70   L  G  I  G  F  L  V  W  H  L  Q  Y  R  D  V  R  V  Q  K  V  F  N  G  Y  M  R  I  T  N  E  N  F  V  D  A  Y  E  N  S  N  109

 481  actccactgagtttgtaagcctggccagcaaggtgaaggacgcgctgaagctgctgtacagcggagtcccattcctgggcccctaccacaaggagtcggctgtgacggccttcagcgagg  600
 110   S  T  E  F  V  S  L  A  S  K  V  K  D  A  L  K  L  L  Y  S  G  V  P  F  L  G  P  Y  H  K  E  S  A  V  T  A  F  S  E  G  149

 601  gcagcgtcatcgcctactactggtctgagttcagcatcccgcagcacctggtggaggaggccgagcgcgtcatggccgaggagcgcgtagtcatgctgcccccgcgggcgcgctccctga  720
 150   S  V  I  A  Y  Y  W  S  E  F  S  I  P  Q  H  L  V  E  E  A  E  R  V  M  A  E  E  R  V  V  M  L  P  P  R  A  R  S  L  K  189

 721  agtcctttgtggtcacctcagtggtggctttccccacggactccaaaacagtacagaagacccaggacaacagctgcagctttggcctgcacgcccgcggtgtggagctgatgcgcttca  840
 190   S  F  V  V  T  S  V  V  A  F  P  T  D  S  K  T  V  Q  K  T  Q  D  N  S  C  S  F  G  L  H  A  R  G  V  E  L  M  R  F  T  229

 841  ccacgcccggcttccctgacagcccctaccccgctcatgcccgctgccagtgggccctgcgggggacgccgactcagtgctgagcctgaccttccgcagctttgaccttgcgtcctgcg  960
 230   T  P  G  F  P  D  S  P  Y  P  A  H  A  R  C  Q  W  A  L  R  G  D  A  D  S  V  L  S  L  T  F  R  S  F  D  L  A  S  C  D  269

 961  acgagcgcggcagcgacctggtgacggtgtacaacaccttgagccccatggagccccacgccctggtgcagttgtgtggcacctaccctccctcctacaacctgaccttccactcctccc  1080
 270   E  R  G  S  D  L  V  T  V  Y  N  T  L  S  P  M  E  P  H  A  L  V  Q  L  C  G  T  Y  P  P  S  Y  N  L  T  F  H  S  S  Q  309

1081  agaacgtcctgctcatcacactgataaccaacactgagcggcggcatcccggctttgaggccaccttcttccagctgcctaggatgagcagctgtggaggccgcttacgtaaagcccagg  1200
 310   N  V  L  L  I  T  L  I  T  N  T  E  R  R  H  P  G  F  E  A  T  F  F  Q  L  P  R  M  S  S  C  G  G  R  L  R  K  A  Q  G  349

1201  ggacattcaacagcccctactacccaggccattacccacccaacattgactgcacatggaacattgaggtgcccaacaaccagcatgtgaaggtgcgcttcaaattcttttacctgctgg  1320
 350   T  F  N  S  P  Y  Y  P  G  H  Y  P  P  N  I  D  C  T  W  N  I  E  V  P  N  N  Q  H  V  K  V  R  F  K  F  F  Y  L  L  E  389

1321  agcccggcgtgcctgcgggcacctgccccaaggactacgtggagatcaacggggagaaatactgcggagagaggtcccagttcgtcgtcaccagcaacagcaacaagatcacagttcgct  1440
 390   P  G  V  P  A  G  T  C  P  K  D  Y  V  E  I  N  G  E  K  Y  C  G  E  R  S  Q  F  V  V  T  S  N  S  N  K  I  T  V  R  F  429

1441  tccactcagatcagtcctacaccgacaccggcttcttagctgaatacctntcctacgactccagtgacccatgcccggggcagttcacgtgccgcacggggcggtgtatccggaaggagc  1560
 430   H  S  D  Q  S  Y  T  D  T  G  F  L  A  E  Y  L  S  Y  D  S  S  D  P  ' P  G  Q  F  T  C  R  T  G  R  C  I  R  K  E  I  469

1561  tgcgctgtgatggctgggccgactgcaccgaccacagcgatgagctcaactgcagttgcgacgccggccaccagttcacgtgcaagaacaagttctgcaagccccttcttctgggtctgcg  1680
 470   N  C  D  G  W  A  D  C  T  D  H  S  D  E  L  N  C  S  ' D  A  G  H  Q  F  T  C  K  N  K  F  C  K  P  L  F  W  V  C  D  509

1681  acagtgtgaacgactgcggagacaacagcgacgagcaggggtgcagttgtCCGGCCCAGACCTTCAGGTGTTCCAATGGGAAGTGCcTcTCGAAAAGCCAGCAGTGCAATGGGAAGGACG  1800
 510   S  V  N  D  C  G  D  N  S  D  E  Q  G  C  S  ' P  A  Q  T  F  R  C  S  N  G  K  C  L  S  K  S  Q  Q  C  N  G  K  D  D  549

1801  ACTGTGGGGACGGTTCCGACGAGGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCTACCGCTGCCTCAATGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACG  1920
 550   '  G  D  G  S  D  E  A  S  C  I  K  V  N  V  V  T  ' T  K  H  T  Y  R  C  L  N  G  L  C  L  S  K  G  N  P  E  C  D  G  589

1921  GGAAGGAGGACTGTAGCGACGGCTCAGATGAGAAGGACTGCGACTGTTGGCTGCGGTCATTCACGAGACAGGCTCGTGTTGTTGGGGGCACGGATGCGGATGAGGGCGAGTGGCCCTGGC  2040
 590   K  E  D  C  S  D  G  S  D  E  K  D  C  I  C  W  L  R  S  F  T  R  Q  A  R  V  V  G  G  T  D  A  D  E  G  E  W  P  W  Q  629
                                                                                  ↑

2041  AGGTAAGCCTGCATGCTCTGGGCCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGCTGGTCTCTGCCGCACACTGCTACATCGATGACAGAGGATTCAGGTACTCAGACC  2160
 630   V  S  L  H  A  L  G  Q  G  H  I  C  G  A  S  L  I  S  P  N  W  L  V  S  A  A  H  C  Y  I  D  D  R  G  F  R  Y  S  D  P  669

2161  CCACGCAGTGgacggccttccTGGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGGAGCGCAGGCTCAAGCGCATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCG  2280
 670   T  Q  W  T  A  F  L  G  L  H  D  Q  S  Q  R  S  A  P  G  V  Q  E  R  R  L  K  R  I  I  S  H  P  F  F  N  D  F  T  F  D  709

2281  ACTATGACATCGCGCTGCTGGAGCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGACGCCTCCCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGG  2400
 710   Y  D  I  A  L  L  E  L  E  K  P  A  E  Y  S  S  M  V  R  P  I  C  L  P  D  A  S  H  V  F  P  A  G  K  A  I  W  V  T  G  749

2401  GCTGGGGACACACCCAGTATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACCACCTGCGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGA  2520
 750   W  G  H  T  Q  Y  G  G  T  G  A  L  I  L  Q  K  G  E  I  R  V  I  N  Q  T  T  C  E  N  L  L  P  Q  Q  I  T  P  R  M  M  789

2521  TGTGCGTGGGCTTCCTCAGCGGCGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGTGGAGGCGGATGGGCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACG  2640
 790   C  V  G  F  L  S  G  G  V  D  S  C  Q  G  D  S  G  G  G  P  L  S  S  V  E  A  D  G  R  I  F  Q  A  G  V  V  S  W  G  D  G  829

2641  GCTGCGCTCAGAGGAACAAGCCAGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAGAGAACACTGGGGTATAGGGGCCGGGGCCACCCAAATGTGTACACCTGCGGGGCC  2760
 830   C  A  Q  R  N  K  P  G  V  Y  T  R  L  P  L  F  R  D  W  I  K  E  N  T  G  V                                          855

2761  ACCCATCGTCCACCCCAGTGTGCACGCCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCCCCAGAACATACACTGTGAACTCAATCTCCAGGGCTCCAAATCTGCCTAGAAAA  2880

2881  CCTCTCGCTTCCTCAGCCTCCAAAGTGGAGCTGGGAGGTAGAAGGGTTCTACTGACCCAACTGGGGGCAAAGGTTTGAAGACACAGCTTCCCCCGCCAGCCCCAAGCTGGGCCGAGGCGC  3000

3001  GTTTGTGTATATCTGCCTCCCCTGTCTGTAAGGAGCAGCGGGAACGGAGCTTCGGAGCCTCCTCAGTGAAGGTGGTGGGGCTGCCGGATCTGGGCTGTGGGGCCCTTGGGCCACGCTCTT  3120

3121  GAGGAAGCCCAGGCTCGGAGGACCCTGGAAAACAGACGGGTCTGAGACTGAAATTGTTTTACCAGCTCCCAGGGTGGACTTCAGTGTGTGTATTTGTGTAAATGAGTAAAACATTTTATT  3240

3241  TNCTTTTTAAAAAAAAAAAAAAA  3263
```

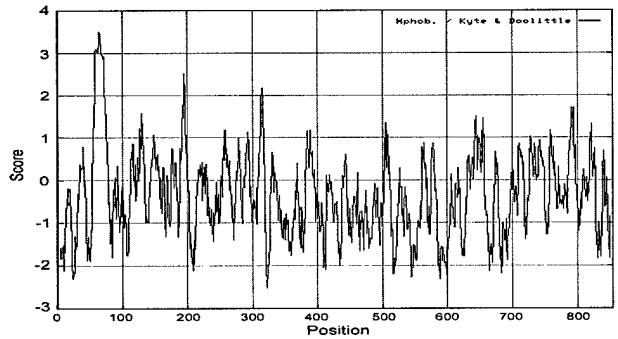
The deduced amino acid sequence from the open reading frame contains the His^{57}, Asp^{102} and Ser^{195} motifs characteristic of serine proteases (Val^{652}-Cys^{657}, Asp^{711}-Leu^{714} and Asp^{799}-Ser^{810} respectively) and is compatible with the S1 family of the SA clan of serine-type petidases (Rawlings and Barrett, 1994). A potential serine protease activation sequence (RVVGG) is present at residues

R^{614}-G^{618} suggesting that human epithin is synthesized as a single chain zymogen. While there is no apparent signal peptide sequence, Kyle-Doolittle hydropathy plots using the GCG Peptide Structure program (**Fig. 3A**) indicated the presence of a highly hydrophobic region between residues V^{57} and W^{77} consistent with a transmembrane spanning region. Similar results were obtained using the TMpred (Hofman & Stoffel,1993) and TopPred 2 programs (Claros & von Heijne, 1994). A ProfileScan of the ORF against the Prosite library showed two potential CUB domains (residues 214-334 and 340-447) and four potential LDL-receptor class A domains (residues 452-487, 487-524, 524-560 and 566-603). Epithin has 40 putative external cysteine residues. Each LDR repeat contains six cysteine residues presumably in 3 internal disulfide linkages (Brown et al 1997). Each of the two CUB domains contains four cysteine residues in disulfide linkage which, also most likely are in internal disulfide linkages (Bork & Beckman 1993). The catalytic domain has 8 cysteine residues in four disulfide linkages analogous to chymotrypsin (i.e. Cys^{604}-Cys^{732}, Cys^{641}-Cys^{657}, Cys^{776}-Cys^{790}, Cys^{801}-Cys^{830}). Cys^{604}-Cys^{732} serves to link the catalytic and regulatory domains. A schematic representation of the human epithin domain structure is shown in **Fig. 3B.**

Using BLAST, an 80.3% identity in 843 amino acid overlap was found between the deduced amino acid sequence and the deduced sequence of epithin a putative protease cloned from a mouse thymocyte cDNA library (Kim et al 1999). An almost perfect identity at the cDNA level was found between the MCF-7 cell protease and matriptase, a 683 amino acid serine protease recently cloned from T47-D breast cancer cells (Lin et al 1999). A 45-55% identity at the amino acid level was found between the human epithin and the human serine proteases (or their zymogens) enterokinase precursor (Kitamoto et al 1995), hepsin (Leytus et al 1988), prekalikrein (Chung et al 1986), TMPRSS2 protease (Paolino-Giacobino, et al 1997), prostasin (Yu et al 1995) and drosophila protease stubble (Appel et al 1993) (**Fig. 3C**). Significantly, enterokinase precursor, hepsin, prekalikrein, and prostasin are all cell membrane-bound proteases. Epithin most likely shows specificity for cleavage of peptide bonds after Lys or Arg residues as it, like trypsin, hepsin and enterokinase contains an Asp (Asp^{266}) at the base of the specificity pocket (S1 subsite). Elastase and chymotrypsin-like enzymes have cysteine and serine residues, respectively, at this site.

Figure 3. A. Hydropathy plot of the deduced human epithin amino acid sequence by the Kyle-Doolittle method. **B.** Schematic representation of the multi-domain structure of human epithin. Numbers correspond to the deuced amino acid sequence derived from the full-length cDNA shown in Fig. 2. **C.** Comparison of the amino acid sequences of human epithin and related proteins.

A.



B.

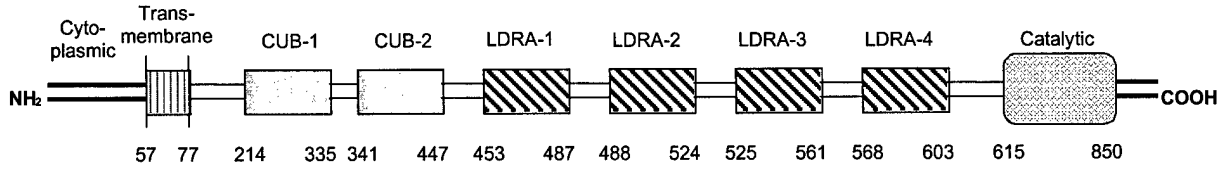
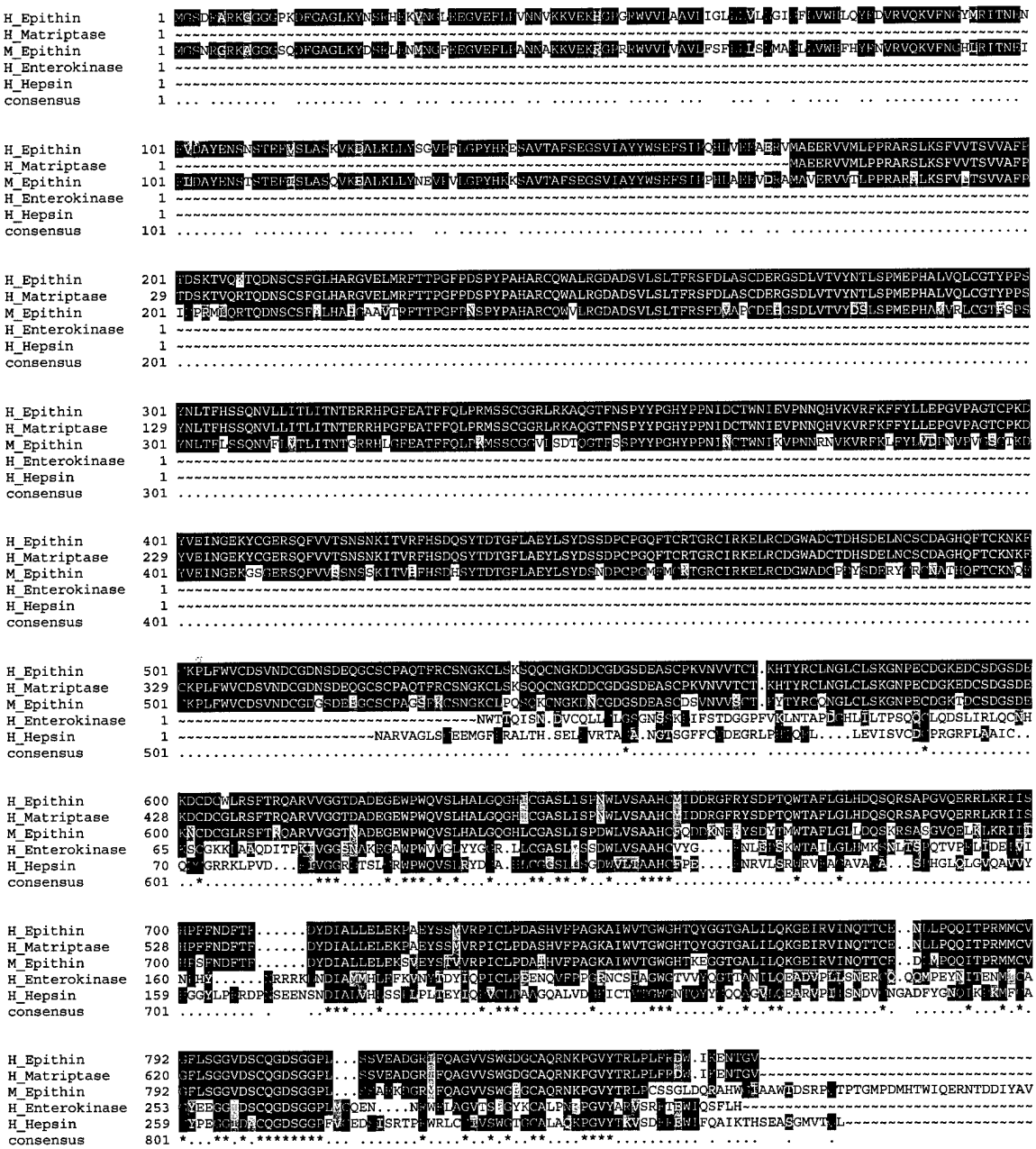


Figure 3C.



2. Activation and Release of Human Epithin from MCF-7 Cells

Western blotting shows that both the parental MCF-7 subline and a subline stably transfected with a full-length from epithin cDNA release the same amount of an approximately 84 kDa fragment into the media (**Fig. 4A**). The nature of the fragment is unclear although it must contain the catalytic domain as this region was used to prepare the antibody. Whether the fragment is generated by proteolytic cleavage or results from altered splicing is presently under investigation, although the later possibility appears to be remote as only a single 4.2 kb transcript is seen on Northern blots (**Fig. 1).** There is a prominent 33 kDa band in reduced samples in the spent media from transfected cells which is barely visible in spent media from the parental cells suggesting that it may be a consequence of

autoprocessing. Undoubtedly, this is the catalytic domain as membrane fractions from MCF-7 cells labeled with the serine protease inhibitor [^{3}H]-diisopropyl flurophosphate, showed a 33 kDa [^{3}H]-labeled band, which also was recognized by the epithin antibody after western blotting (**Fig. 5**). The fact that treatment with phorbol 12-myristate 13-acetate (PMA) was without effect suggests that epithin activation and release of TGFα may not be directly connected. The observation that addition of the serine protease inhibitor phenylmethylsulfonyl fluoride (DIFP) to the spent media after collection also had no effect suggests that activation occurs rapidly.

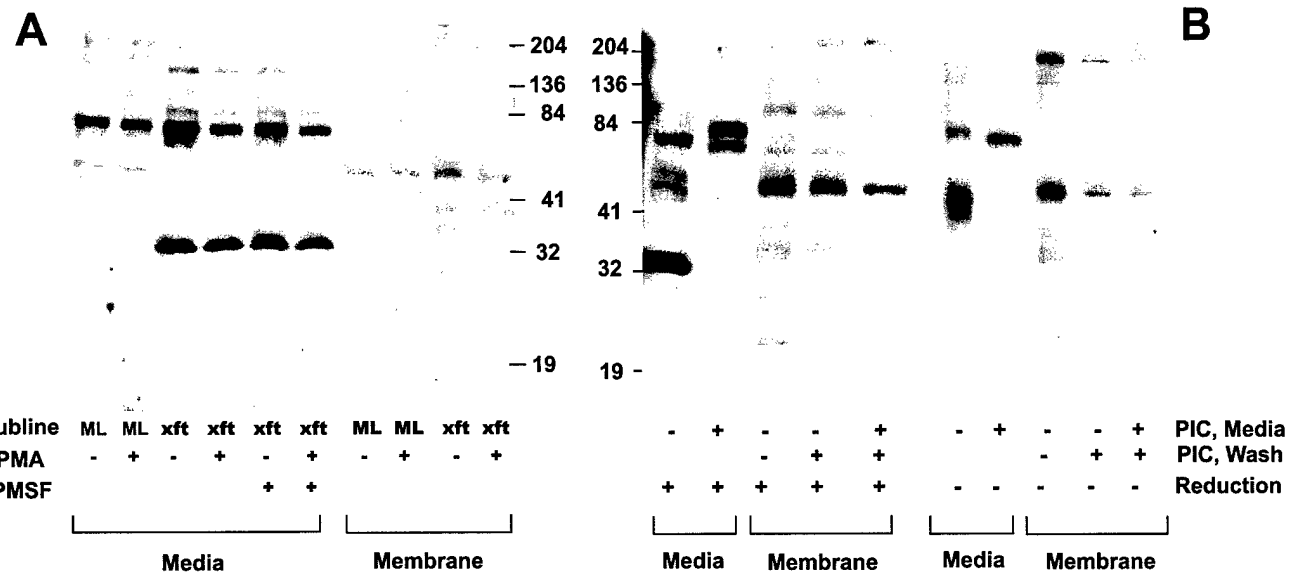


Figure 4. Identification of Epithin on MCF-7 Cell Membranes or Released into the Media. A. Confluent MCF-7 ML (-xfct) cells and MCF-7 ML cells transfected with a full length epithin cDNA (+xfct) were washed 2 times with serum-free media before incubation with serum-free media containing 10^{-8} M estradiol. After 21 hours PMA (50 ng/ml) was added to half the flasks and the incubation was allowed to continue for an additional 4 hours. Conditioned media were collected, clarified by centrifugation and ammonium sulfate was added to 65% saturation with continuous mixing. After overnight incubation at 4° the samples were centrifuged (14,000 x g, 20 minutes), the pellets resuspended in 10 mM phosphate buffer (pH 6.8) and dialyzed overnight against the same buffer. Membrane fractions were prepared by washing cells 3 times with PBS, scraping in 20mM Tris-HCl, pH 7.4 and homogenizing with a Dounce homogenizer. After centrifugation at 600 x g for 10 minutes the supernatants were centrifuged at 20,000 x g for 20 minutes. The resulting pellets were suspended in 20 mM Tris-HCl, pH 7.4. 1% Triton-X-100. Aliquots of precipitated media and membrane fractions were electrophoresed on a 12 % SDS-polyacrylamide gel under reducing conditions, transferred to a PVDF membrane and probed with the antibody to PCR-7. Some aliquots of media from transfected cells were subjected to treatment with 0.5 mM PMSF before electrophoresis as indicated. **B.** Transfected MCF-7 ML cells were washed 2 times with serum-free media then incubated for 24 hours with serum-free media ± PIC (protease inhibitor cocktail: Complete™,Mini EDTA-free tablets, Boehringer Mannheim, plus BB-3103, British Biotech). Media were collected and treated as above except that ammonium sulfate precipitates from media containing protease inhibitor were suspended in and dialyzed against 10 mM phosphate buffer containing 10 mM EDTA. Membrane fractions from cells not treated with protease inhibitor were prepared ± PIC in wash and suspension buffer, while cells incubated with protease inhibitors in the media were treated with protease inhibitors throughout the purification procedure. Aliquots of media and membrane fractions were subjected to western blot analysis under both reducing and non-reducing conditions.

Western blotting of spent media from MCF-7 cells recognized an approximately 45 kDa polypeptide without, and 33 kDa with reduction (**Fig. 5B**). The 45 kDa fragment most likely contains the activated catalytic chain and an approximately 12 kDa fragment from the regulatory chain. The nature of the prominent 45 kDa band seen in all of the membrane fractions, and whether it is identical to the slightly larger poly peptide seen in the spent media is presently under investigation. Significantly, in spent media from cells incubated with the metalloprotease inhibitor BB-3103, only a 84 kDa band is evident suggesting a metalloprotease may needed for initial epithin activation.

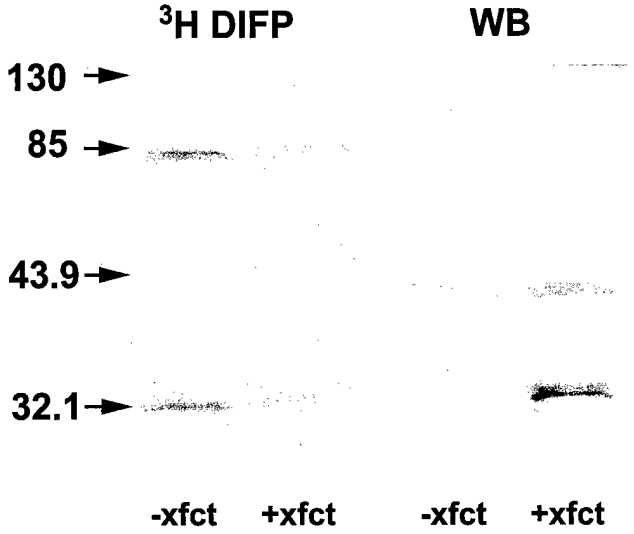


Figure 5. Labeling of MCF-7 cell membrane proteases with ^{3}H DIFP. Confluent MCF-7 ML (-xfct) cells and MCF-7 ML cells transfected with a full length human epithin cDNA (+xfct) in T-25 flasks, were washed 3 times with serum-free media and then incubated for 30 min with media containing 10^{-8} M estradiol and 50 ng/ml PMA. After 30 minutes, ^{3}H-DIFP (12.5 μCi/ml) was added to each flask and the incubation was continued for an additional 60 minutes. Media was then removed and the cells washed 5 times with PBS and lysed with buffer A (0.5% Triton-X-100, 150 mM NaCl, 5 mM EDTA, 50 mM Tris-HCl, pH 7.4 and 0.02% NaN_3). **A.** Aliquots of the lysed cells were electrophoresed on a 12% polyacrylamide gel under reducing conditions and transferred to a PVDF membrane. The membrane was dried and exposed to Kodak Biomax MS film under a LE intensifying screen for 6 days at -70°. **B.** The membrane was then rehydrated with methanol and subjected to Western blot analysis using the antibody to PCR-7. Proteins were detected by ECL.blot analysis using the antibody to PCR-7. Proteins were detected by ECL.

Gelatin zymography was used to confirm that human epithin, like matriptase has activity against matrix proteins (**Fig. 6**). In the presence of Ca^{++}, required for the action of metalloproteases, two bands are apparent (**Fig. 6 left panel**) in both the parental and epithin-transfected cell lines. The higher, 92 kDa band is present at the level in both cell lines The lower, 84 kDa band, is considerably more pronounced in the transfected cells. When the zymograms were incubated in the presence of EDTA, an inhibitor of matrix metalloproteases, the upper band, most likely the MMP9 gelatinase, disappeared. The would suggest that the lower band is epithin or some other non-metalloprotease. That this is the case is shown in the **right panel** where the 84 kDa band was made to disappear by incubation with DIFP whether in the presence or absence of EDTA. Consistent with the western blotting experiments, PMA had no effect. The observation that the 84 kDa activity is inhibited by DIPF but not by EDTA and that it is elevated in transfected cells strongly suggests that the 84 kDa activity is indeed epithin and that epithin may be directly involved in invasion.

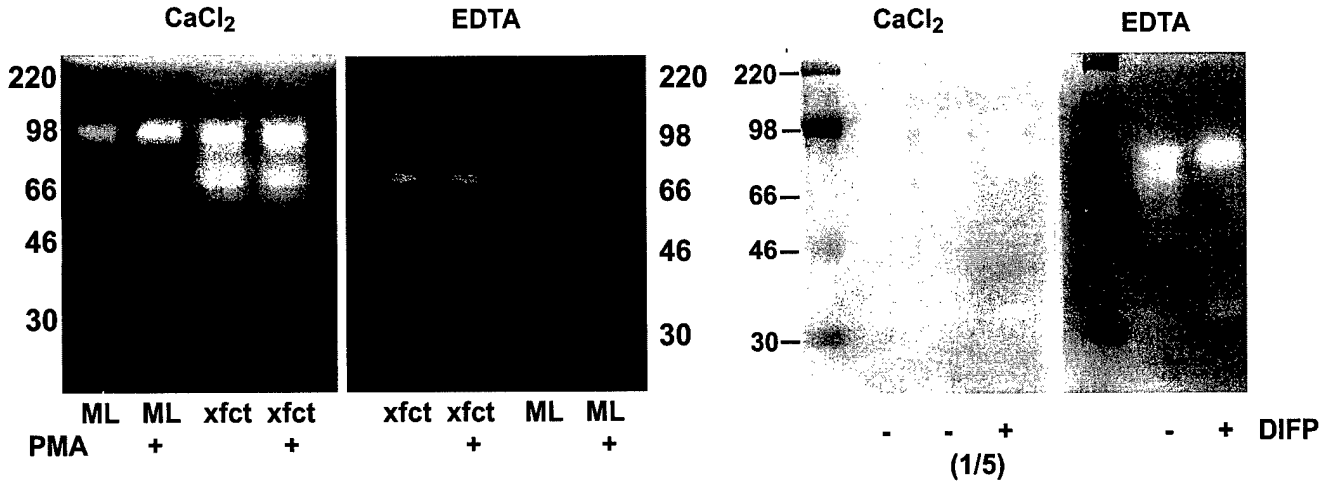


Figure 6. Effect of protease inhibitors on epithin gelatinase activity. Left panel: Confluent MCF-7 ML (-xfct) cells and MCF-7 ML cells transfected with a full length epithin cDNA (+xfct) were washed 2 times with serum-free media before incubation with serum-free media containing 10^{-8} M estradiol. After 21 hours PMA (50 ng/ml) was added to half the flasks and the incubation was allowed to continue for an additional 4 hours. Conditioned media were collected, clarified by centrifugation and ammonium sulfate was added to 65% saturation with continuous mixing. After overnight incubation at 4° the samples were centrifuged (14,000 x g, 20 minutes), the pellets resuspended in 10 mM phosphate buffer (pH 6.8) and dialyzed overnight against the same buffer. Aliquots were electrophoresed on 10% zymogram gels containing gelatin. Gels were incubated overnight, washed with several changes of 2.5% Triton-X-100 and incubated at 37° overnight in 50mM Tris-HCl, pH 7.5, 200 mM Nacl, 0.02% Brij-35 and either 5mM $CaCl_2$ or 10 mM EDTA. Gels were then stained with Coomassie Blue R-250. **Right panel:** Ammonium sulfate precipitated conditioned media from transfected cells were incubated 5 min minus or plus 10 mM DIFP before electrophoresis on 10% zymogram gels containing gelatin. Gels were treated as in the left panel.

3. Comparative Expression of Epithin, ProTGFα and α_1-Antitrypsin by Various Breast Cancer Cell Lines. Three breast cancer cell lines in addition to MCF-7 cells were examined for the expression of epithin, a potential proTGFα cleaving enzyme, proTGFα and α_1-antitrypsin (α_1-AT), a potential regulator of TGFα cleaving enzyme activity. Preliminary data from this laboratory suggested that TGFα from MCF-7 cells could be blocked by α_1-AT and α_1-AT could complex with a cell surface protease on MCF-7 cells. If our hypothesis is correct, then all four cell lines should express the three proteins and TGFα release and tumorigenicity should be related to the ratio of the expression of epithin: α_1-AT. To test this hypothesis we have first compared steady-state mRNA levels of epithin (shown as PCR-7), α_1-AT and proTGFα in the four breast cancer cell lines.

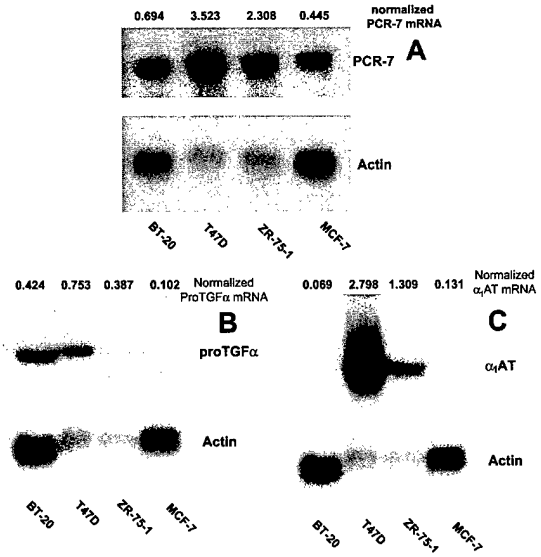


Figure 7 Steady state mRNA levels of α_1-AT, epithin (PCR-7), and proTGFα in breast cancer cell lines. Total RNA was isolated and subjected to Northern blot analysis with the indicated ^{32}P-labeled cDNA probes. Blots were then cleared and reprobed with a β-actin cDNA. Autoradiograms were quantified by densitometry.

CONCLUSSIONS

- Degenerate oligonucleotides based on the conserved sequences about the his^{57} and ser^{195} residues in mammalian serine proteases were used to clone a potential TGFα cleaving enzyme from MCF-7 breast cancer cells. The 3270 bp cDNA has a deduced 855 amino acid sequence consistent with that of a multi-domain, type II membrane protease and shows considerable homology to epithin, a recently described protein from mouse thymocytes. Mouse and human epithin have similar membrane spanning and trypsin-like serine protease domains. Each also has two putative CUB and four low-density lipoprotein receptor domains, which may be important for regulation and/or substrate binding. Matriptase, a recently described cDNA from T47-D breast cancer cells may be a partial epithin clone.
- Northern blotting showed highest expression of epithin in human epithelial cells particularly breast, kidney, trophoblast, and uterine glandular epithelium. Southern analysis indicated the presence of sequences homologous to human epithin in baboon, mouse, and rabbit but not in chicken or drosophila DNA.
- A polyclonal antibody to the epithin active site domain inhibited phorbol ester-induced TGFα release from MCF-7 cells by >50% suggesting that epithin is involved in proTGFα processing.
- Western blotting of spent media from MCF-7 cells recognized an approximately 42 kDa polypeptide without, and 33 kDa with reduction. However, in media from cells incubated with the metalloprotease inhibitor BB3103, only a 84 kDa band was evident suggesting a metalloprotease may needed for epithin activation. Membrane fractions from MCF-7 cells labeled with the serine protease inhibitor [^{3}H]-diisopropyl flurophosphate, showed a 33 kDa [^{3}H]-labeled band, which also was recognized by the epithin antibody after western blotting. On gelatin zymography, two prominent bands were apparent in the presence of Ca^{++}, the higher MW form disappeared in the presence of EDTA, an inhibitor of matrix metalloproteases, the lower in the presence of inhibitors of serine proteases. These results suggest that epithin may be directly involved in invasion.
- Immunohistochemistry of paraffin sections from normal human breast tissue showed that the epithin protein was restricted to the surface of ductal epithelial cells where it co-localized with the epidermal growth factor receptor. Epithin expression appeared to be elevated in DCIS and in invasive breast carcinoma.

REFERENCES

Appel, L.F., Prout, M., Abu-Shumays, R., Hammonds, A., Garbe, J.C., Fristrom, D.K. Fristrom, J.W. 1993 Proc Natl Acad Sci USA 90:4937-4941.

Arribas, J., Coodly, L., Vollmer, P. Kishimoto, T. K., Rose-John, S. Massagué, J. 1996 Diverse cell surface protein ectodomains are shed by a system sensitive to metalloprotease inhibitors. J Biol Chem 271: 11376-11382.

Arribas, J., Massagué, J. 1995 Transforming growth factor-alpha and beta-amyloid precursor protein share a secretory mechanism. J Cell Biol 128:433-441

Baselga, J., Mendelsohn, J., Kim, Y-M. and Pandiella, A. 1996 Autocrine regulation of membrane transforming growth factor-α cleavage. J Biol Chem 271:3279-3284.

Bork, P. and Beckman G. 1993 The CUB domain. A widespread module in developmentally regulated proteins. J Mol Biol 231: 539-545.

Brown, M.S., Herz, J., and Goldstein, J.L. 1997 Nature 388:629-630

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter W.J. 1979 Biochemistry 18; 5294-5299.

Chung, D.W., Fujikawa,K., McMullen,B.A. and Davie,E.W. (1986) Human plasma prekallikrein, a zymogen to a serine protease that contains four tandem repeats. Biochemistry 25: 2410-2417.

Claros, M.G. and von Heijne G 1994 TopPred II: an improved software for for membrane protein structure predictions. Comput. Appl Biosci 10:685-686.

De Jong, J.S., Van Diest, P.J., Van der Valk, P. and Baak, J.P. 1998 Expression of growth factors, growth inhibiting factors and their receptors in invasive breast cancer. II: Correlations with proliferation and angiogenesis. J Pathol 184:53-57.

Elvin, C.M., Whan, V. and Riddles, P.W. 1993 A Family of Serine Protease Genes Expressed in Adult Buffalo Fly. (*Haematobia irritans exigua*) Mol. Gen. Genet. 240:132-139.

Finlay, T.H, Tamir, S., Kadner, S.S., Yavelow, J. and Levitz, M. 1993a α_1-Antitrypsin and Anchorage-Independent Growth of MCF-7 Breast Cancer Cells. Endocrinology 133:996-1002.

Finlay, T.H. Kadner, S.S. and Tamir, S. 1993b Protease Inhibitor Synthesis by MCF-7 Breast Cancer Cells. *in* Protease Inhibitors as Cancer Chemopreventative Agents. (W. Troll and A. Kennedy, Eds.). Plenum Press, New York. pp 141-159.

Frangioni, J.V. and Neel, B.G. (1993) Solubilization and Purification of Enzymatically Active Glutathione S-Transferase (pGEX) Fusion Proteins. Analyt Biochem 210:179-187.

Fujikawa, K., Chung, D.W., Hendrickson, L.E., Davie, E.W. (1986) Amino acid sequence of human factor XI, a blood coagulation factor with four tandem repeats that are highly homologous with plasma prekallikrein. Biochemistry. 25:2417-2424.

Harris, J.R., Lippman, M.E., Veronesi, U.V., Willett, W. 1992 Breast Cancer N. Engl J Med 327:319-328, 390-398, 473-480.

Hofmann, K. & Stoffel 1993 Tmbase - A database of mambrane spanning proteins. Biol. Chem. Hoppe-Seyler 347:166-168.

Hooper, N.M, Karran, E.H. and Turner A.J. 1997 Membrane protein secretases. Biochem J 321:265-279.

Kitamoto,Y., Veile,R.A., Donis-Keller,H. and Sadler,J.E. 1995 cDNA sequence and chromosomal localization of human enterokinase the proteolytic activator of trypsinogen. Biochemistry 34: 4562-4568.

Kozak, C.A., Peyser, M., Krall, M., Mariano, T.M., Kumar, C.S., Pestka, D., and Mock, B.A. 1990 Molecular genetic markers spanning mouse chromosome 10. Genomics 8:519-524.

Lee, D.C., Fentón, S.E., Berkowitz, E.A. and Hissong, M.A. 1995 Transforming Growth Factor α: Expression, Regulation, and Biological Activities. Pharm. Rev. 47:51-85.

Leytus, S.P., Loeb, K.R., Hagen, F.S., Kurachi, K., Davie, E.W. (1988) A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells. Biochemistry. 27:1067-1074.

Massagué, J. and Pandiella, A. 1993 Membrane-Anchored Growth Factors. Ann Rev Biochem 62:515-541.

Miyazawa, K., Shimomura,T ., Kitamura, A., Kondo, J., Morimoto, Y. and Kitamura, N. (1993) Molecular cloning and sequence analysis of the cDNA for a human serine protease reponsible for activation of hepatocyte growth factor. Structural similarity of the protease precursor to blood coagulation factor XII. J. Biol. Chem. 268:10024-10028.

Paolino-Giacobino, A., Chen, H., Peitsch, M.C., Rossier, C. and Antronarakis, S.C. 1997 Cloning of the TMPRSS2 gene, which encodes a novel serine protease with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3 Genomics 44:309-320.

Rawlings ND, Barrett AJ 1994 Families of serine peptidases. Meth Enzymol 244:19-61

Sakanari, J.A., Staunton, C.E., Eakin, A.E, Craik, C.S. and McKerrow, J.H. (1989) Serine Proteases from Nematode and Protozoan Parasites: Isolation of Sequence Homologs Using Generic Molecular Probes. Proc. Natl. Acad. Sci. USA. 86:4863-4867.

Sambrook, J., Fritsch, E.F., Maniatis, T. Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, New York, 1989.

Sokolova, E.A., Starkova, N.N., Vorotynseva, T.I. and Zamolodchikova, T.S. 1998 A seine protease from the bovine duodenal mucosa, chymotrypsin-like duodenase. Eur J Biochem 255:501-507.

Tamir, S., Kadner, S.S., Katz, J. and Finlay, T.H. 1990 Regulation of Antitrypsin and Antichymotrypsin Synthesis by MCF-7 Breast Cancer Cell Lines. Endocrinology 127:1319-1328.

Yavelow J, Tuccillo A, Kadner SS, Katz J, Finlay TH. 1997 α_1-Antitrypsin Blocks Release of TGFα from MCF-7 Human Breast Cancer Cells. J. Clin. Endo. Metabol 82:745-752.

Yu, J.X., Chao, L. and Chao, J. 1995 Molecular Cloning, tissue-specific expression and cellular localization of human prostasin mRNA. J. Biol. Chem 270:13483-13489.