Minireview

Type II Transmembrane Serine Proteases

INSIGHTS INTO AN EMERGING CLASS OF CELL SURFACE PROTEOLYTIC ENZYMES

Published, JBC Papers in Press, November 1, 2000, DOI 10.1074/jbc.R000020200

John D. Hooper†, Judith A. Clements‡, James P. Quigley§, and Toni M. Antalis¶

From the Centre for Molecular Biotechnology, Queensland University of Technology, Gardens Point, Brisbane 4000, Australia, §The Scripps Research Institute, La Jolla, California 92037, and the Cellular Oncology Laboratory, University of Queensland and the Queensland Institute of Medical Research, Brisbane 4029, Australia

Cell surface proteolysis has emerged as an important mechanism for the generation of biologically active proteins that mediate a diverse range of cellular functions. The proteolytic activities of membrane-anchored proteins, such as ADAMs (1) and MT-MMPs (2), are thought to play central roles in cell surface-activating events. In contrast, most of the members of the serine protease family, one of the oldest characterized and largest multigene proteolytic families, are either secreted enzymes or sequestered in cytoplasmic storage organelles awaiting signal-regulated release. These serine proteases have well characterized roles in diverse cellular activities, including blood coagulation, wound healing, digestion, and immune responses, as well as tumor invasion and metastasis. However, during the last few years there has been an explosion in the identification of transmembrane proteins containing C-terminal extracellular serine protease domains. These enzymes are ideally positioned to interact with other proteins on the cell surface as well as soluble proteins, matrix components, and proteins on adjacent cells. In addition, these membrane-spanning proteases have cytoplasmic N-terminal domains, suggesting possible functions in intracellular signal transduction. This review delineates for the first time this emerging class of cell surface proteolytic enzymes, the type II transmembrane serine proteases (TTSPs), to highlight their structural features, expression profiles, and possible roles in mediating cell surface proteolytic events.

Structural Features of TTSPs

In mammals the TTSPs currently consist of 17 members (Table I), of which seven are found in man. Enteropeptidase (also known as enterokinase) (3), because of its essential role in the processing of digestive proteases, was the first member of this group to be discovered nearly a century ago. The other more recently identified members include hepsin (4), human airway trypsin-like protease (HAT) (5), corin (6), MT-SPI (7) (also known as matriptase (8)), TMPRSS2 (9), and most recently TMPRSS4 (10). The only non-mammalian TTSP identified to date is the Drosophila protease stubble-stubboid (st-sb) (11). Mammalian orthologues have been reported for enteropeptidase (mouse (12), rat (13), cow (14), and pig (15)), hepsin (mouse (16) and rat (17)), corin (mouse, also known as LRPI (18), MT-SPI (mouse, also known as epithelin (19)), and TMPRSS2 (mouse, also known as epithelin (20)) (Table I). The TTSPs share a number of common structural features including (i) a proteolytic domain, (ii) a transmembrane domain, (iii) a short cytoplasmic domain, and (iv) a variable length stem region containing modular structural domains, which links the transmembrane and catalytic domains (Fig. 1). It is this unique combination of domains that suggests novel roles for the TTSPs at the cell surface.

Proteolytic Domains—As is the case for the wider family of enzymes of the chymotrypsin (S1) fold, the proteolytic domains of the TTSPs share a high degree of amino acid sequence identity. In particular, the histidine, aspartate, and serine residues necessary for catalytic activity are present in highly conserved motifs. TTSPs are synthesized as single chainzymogens and are likely activated by cleavage following an arginine or lysine present in a highly conserved activation motif. Based on the predicted presence of a conserved disulfide bond linking the pro- and catalytic domains (Fig. 1), the TTSPs are likely to remain membrane-bound following activation. However, the isolation of soluble forms of enteropeptidase (21, 22), HAT (23), and MT-SPI (24) suggests that the extracellular domains of at least some of the TTSPs may also be shed from the cell surface. Other cysteine residues conserved among the TTSPs include six cysteines predicted to form three intraprotease domain disulfide bonds. Enteropeptidase and hepsin each have one and two additional predicted disulfide linkages within the catalytic domain. The presence of an asparagine six residues before the catalytic serine, which in the activated TTSP would be positioned at the bottom of the S1 substrate binding pocket, is indicative that all of the TTSPs have preference for substrates containing an arginine or lysine in the P1 amino acid position (S1 and P1 designations are described (25)). The cleavage specificities and candidate physiological substrates for some of the TTSPs have been elucidated. The predicted cleavage specificity following basic amino acids indicates that the TTSPs are likely to have a degree of autocalytic activity. Indeed truncated mouse hepsin lacking cytoplasmic and transmembrane domains (16) and the human MT-SPI proteolytic domain (7) are capable of autoactivation. In contrast, bovine enteropeptidase has extremely low autocalytic activity (26). Interestingly, the proteolytic domain of bovine enteropeptidase has an additional role in the targeting of enteropeptidase to the apical membrane of enterocytes (27).

Transmembrane Domains—Each of the TTSPs contains a hydrophobic domain near the N terminus. This domain is predicted to span the plasma membrane in such a way that the proteolytic domain lies extracellularly, presumably to localize TTSP proteolytic activity in close proximity to target substrates and/or to permit regulated release of the protein from the cell surface. Cell surface localization has been experimentally demonstrated for enteropeptidase, hepsin (28, 29), MT-SPI (30, 31), TMPRSS2 (20), and TMPRSS3 (10).

Cytosplasmic Domains—The cytoplasmic domains of the TTSPs (Fig. 1) range in length from 12 amino acids for HAT to 112 amino acids for murine corin. Whether these domains have the potential to support interactions with cytoskeletal components and signaling molecules is not yet known. However, a number of the TTSPs including corin, MT-SPI, st-sb, and TMPRSS2 contain consensus...
Minireview: Type II Transmembrane Serine Proteases

The abbreviations used are: b, brain; bl, bladder; bp, Drosophila 36-h pupae; c, colon; de, Drosophila 12–18-h embryo; dp, Drosophila early prepupa; e, esophagus; h, heart; int, intestine; k, kidney; l, lung; le, leukocytes; li, liver; p, pancreas; pl, placenta; pr, prostate; psi, proximal small intestine (si); s, spleen; st, stomach; t, testes; th, thymus; tr, trachea.

<table>
<thead>
<tr>
<th>Name</th>
<th>Organism</th>
<th>Accession Number</th>
<th>Gene Location</th>
<th>Expression Pattern</th>
<th>mRNA Size (kb)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corin</td>
<td>Human</td>
<td>AF133845</td>
<td>1q12-13</td>
<td>h</td>
<td>-</td>
<td>5 (6, 56)</td>
</tr>
<tr>
<td>Corin</td>
<td>Mouse</td>
<td>AB519574</td>
<td>-</td>
<td>i,k</td>
<td>5 (16)</td>
<td></td>
</tr>
<tr>
<td>Enteropeptidase</td>
<td>Human</td>
<td>U098560</td>
<td>2q11</td>
<td>psi</td>
<td>-</td>
<td>4,4 (3)</td>
</tr>
<tr>
<td>Enteropeptidase</td>
<td>Mouse</td>
<td>U09856a</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Enteropeptidase</td>
<td>Rat</td>
<td>U73389</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Enteropeptidase</td>
<td>Porcine</td>
<td>D30799</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AST</td>
<td>Human</td>
<td>AF133886</td>
<td>1q25</td>
<td>c,si,si,pr,l,i,p,th</td>
<td>3,3 (7, 8, 31)</td>
<td></td>
</tr>
<tr>
<td>Mouse</td>
<td>81</td>
<td>94.4</td>
<td>9s</td>
<td>int,k</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mouse</td>
<td>AF042862</td>
<td>1q25</td>
<td>c,si,si,pr,l,i,p,th</td>
<td>3,3 (7, 8, 31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enteropeptidase</td>
<td>Human</td>
<td>AB002134</td>
<td>1q12</td>
<td>tr</td>
<td>-</td>
<td>0.5, 1.9, 3.0 (5)</td>
</tr>
<tr>
<td>Hepsin</td>
<td>Human</td>
<td>M19390</td>
<td>11q13</td>
<td>i,i</td>
<td>-</td>
<td>1.8 (4)</td>
</tr>
<tr>
<td>Mouse</td>
<td>AF030065a</td>
<td>11q13</td>
<td>i,i</td>
<td>-</td>
<td>1.8, 1.9 (16)</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>88</td>
<td>44.7</td>
<td>k,l</td>
<td>1.8 (16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shubbe-Stubrod</td>
<td>Drosophila</td>
<td>L11451</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.8 (11)</td>
</tr>
<tr>
<td>TMPRSS2</td>
<td>Human</td>
<td>UT58299</td>
<td>19q13</td>
<td>c,k,h</td>
<td>3.8 (9)</td>
<td></td>
</tr>
<tr>
<td>Mouse</td>
<td>77</td>
<td>53.8</td>
<td>21q23.3</td>
<td>pr</td>
<td>-</td>
<td>3.8 (9)</td>
</tr>
<tr>
<td>Mouse</td>
<td>EPH8</td>
<td>16G2</td>
<td>1q12</td>
<td>k,l</td>
<td>1.5, 2.8 (20)</td>
<td></td>
</tr>
<tr>
<td>TMPRSS4</td>
<td>Human</td>
<td>AF172224</td>
<td>1q23.3</td>
<td>c,k,h,si</td>
<td>2.3 (10)</td>
<td></td>
</tr>
<tr>
<td>TMPRSS4</td>
<td>Mouse</td>
<td>AF172224</td>
<td>1q23.3</td>
<td>c,k,h,si</td>
<td>2.3 (10)</td>
<td></td>
</tr>
</tbody>
</table>

* Splice variants have been identified. † Experimentally derived molecular weight. ‡ V5/His6-tagged protein. ¶ Putative assignment based on our unpublished observation that LRIP4 sequences have greater than 96% identity with mouse chromosome 5 BAC RP23-294A15 sequences deposited in the GenBank™ hgs database (GenBank™ accession no. AC006146). § Closest linkage to the Fli1 gene.

Tissue Expression of TTSPs

Although a few of the TTSPs are expressed across several tissue and cell types, in general these enzymes demonstrate relatively restricted expression patterns, indicating that they may have tissue-specific functions (Table I). Enteropeptidase shows a very narrow expression pattern, being restricted in normal tissues to enteroctyes of the proximal small intestine (12). Corin expression is also quite specific, with corin mRNA highly expressed in human heart (6) and corin protein expression localized to cardiac myocytes (43). HAT is predominantly expressed in trachea (5, 23). Human TMPRSS2 expression is predominantly associated with prostate (9, 44). Hepsin, originally identified from liver, is highly expressed in fetal liver and kidney (45). Hepsin mRNA has been reported to be overexpressed by ovarian tumors (46), and protein expression has been localized to tumor cell membranes in renal cell carcinoma (29). TMPRSS4 has only recently been characterized and was identified as a consequence of its strong up-regulation in pancreatic tumors (10). While TMPRSS4 was not detected in normal pancreas, very low level TMPRSS4 mRNA expression was detected in tissues of the gastrointestinal tract and in some tissues of the urogenital tract (10). MT-SP1 was originally identified from a human breast cancer line (30) but shows the broadest pattern of expression of the TTSPs being detected in a wide range of both human (7) and murine tissues (19).

Biochemical Data and Pathophysiological Roles

The majority of the TTSPs have been identified relatively recently and consequently have not been extensively characterized. Enteropeptidase is somewhat of an exception. Although the enzymatic activity ascribed to enteropeptidase was first identified almost a century ago (47) it has been only recently that the complete amino acid sequence was described (3). Enteropeptidase functions near the apex of the digestive enzymatic cascade activating the digestive protease trypsinogen to trypsin, which subsequently activates other enzymes including chymotrypsinogen, proelastase, prolipases, and procarboxypeptidases. Enteropeptidase possesses extremely low autocalytic activity, and it has been proposed that the serine protease duodenase, secreted by duodenal epithelocytes, may be its physiological activator (48). Active enteropeptidase con...

* The Northern blot data reported (9) are incorrectly labeled due to inversion of the membranes (Stylianos Antonarakis, personal communication).

phosphorylation sites for either or both of protein kinase C and casein kinase II. In addition, based on the cellular sorting of other integral membrane proteins (32) it is likely that the cytoplasmic and transmembrane domains also contribute to the targeting of the TTSPs to a particular cell surface in polarized cells.

**Stem Regions**—The stem regions of the TTSPs contain as many as 11 structural domains that may serve as regulatory and/or binding domains (Fig. 1). These include low density lipoprotein (LDL) receptor class A domains, Group A scavenger receptor (SR) domains, frizzled domains, CUB/Cls/Cls, enarch embryonic growth factor and bone morphogenetic protein 1 (CUB) domains, sea urchin sperm protein, enterokinase, agrin (SEA) domains, a meprin, A5 antigen, and receptor protein phosphatase (MAM) domain, and a disulfide knotted domain. Hepsin is the only TTSP that does not possess an identified structural domain within its stem region. Although functional roles for individual stem region domains have not been demonstrated, the stem region of bovine enteropeptidase has been shown to be required for efficient cleavage of its physiological substrate trypsino-gen (26). In addition, the N terminus of the stem region of this protein is required for delivery of enteropeptidase to the apical surface of polarized Madin-Darby canine kidney cells (27).

The most common stem region structural domain is the LDL receptor class A domain: corin contains eight, MT-SP1 four, enteropeptidase two, and TMPRSS2 and TMPRSS4 one each (Fig. 1). Although the function of these domains in the TTSPs has not been demonstrated, in other proteins they bind Ca^2+ ions and mediate the internalization of macromolecules including serine protease inhibitor complexes and lipoproteins (33–35). In addition, although LDL receptor domains also function in the uptake of LDLs, increased LDL uptake could not be demonstrated following expression of murine corin in COS cells (18).

Six other structural domains that are thought to be involved in protein-protein interactions or protein-ligand interactions are found in various TTSPs. SR domains (36) are present in corin, enteropeptidase, TMPRSS2, and TMPRSS3; frizzled domains (37) are present in corin; CUB domains (38) are present in enteropeptidase and MT-SP1; SEA domains (39) are present in HAT and enteropeptidase; a MAM domain (40) is present in enteropeptidase; a disulfide knotted domain (41) is present in st-sb (Fig. 1). In addition to these structural domains, human and mouse MT-SP1s possess a conserved RGD motif (42) present in the first CUB domain. Interestingly, truncated human MT-SP1 lacking cytoplasmic and transmembrane domains remains bound to the cell surface of COS cells (31). Binding may be mediated via an interaction between the MT-SP1 RGD motif and an integrin protein or another cell surface protein. Alternatively, the mode of attachment could be via a direct link such as a hydrocarbon chain.
Minireview: Type II Transmembrane Serine Proteases

589

The human airway TTSP, HAT, was originally purified as a surface where it is processed to a soluble form, and then released from tracheal serous glands as part of the host immune defense system (5).

Significantly, the human heart TTSP, corin, is an in vitro activator of pro-atrial natriuretic peptide (ANP), a cardiac hormone essential for the regulation of blood pressure (56), suggesting that corin is the long sought pro-ANP convertase. This proteolytic cleavage is critical for the regulation of ANP activity (57); thus, corin may well prove to be an important factor in the regulation of major cardiovascular diseases. Dysfunctional corin was proposed to be a candidate for the rare congenital heart disease, total anomalous pulmonary venous return (TAPVR), as the corin gene colocalizes to the TAPVR locus on human chromosome 4p12–13 (6). In addition to heart, murine corin is expressed by chondrocytes in a differentiation-stage-specific manner during mouse development, suggesting that this protease may play a role during chondrocyte differentiation/bone formation (6). However, while human and murine corin share high homology, common structural features, expression profiles, and syntenic chromosomal locations, these proteases are variant in the lengths of their cytoplasmic domains (45 residues in human and 112 in mouse) and show no conservation in amino acid sequence in this domain. This may indicate that murine and human corin have different but perhaps overlapping species-specific roles, or alternatively the cytoplasmic domain is not essential for corin functions.

In other significant recent experiments it has been shown that MT-SP1 may be involved in initiating signaling and proteolytic cascades via the activation of the cell surface-associated proteases PAR2 and pro-uPA (31). Interestingly, MT-SP1 from breast cancer cells is detected largely as an uncomplexed protein, whereas in milk it is present mainly as a complex with the Kunitz-type serine protease inhibitor hepatocyte growth factor inhibitor-1 (24). It will be important to identify the inhibitor binding domains of MT-SP1 and the function of the protease-inhibitor complex.

TMPRSS2 and TMPRSS4 have been identified through association with cancer. TMPRSS2 is thought to play a role in epithelial cell lysis and its association with prostate carcinogenesis has led to the proposal that it may be a diagnostic or therapeutic target for prostate cancer (44). TMPRSS2 has been proposed to be part of an enzymatic cascade involving the serine proteases prostate-specific antigen and human kallikrein K2 in a manner analogous to the fibrinolytic and blood coagulation cascades (44). TMPRSS4 is over-expressed in pancreatic cancers; however, its functional significance remains unclear (10).

The Drosophila serine protease st-sb is one of a number of proteases involved in fly morphogenesis (11) and has a proteolytic function in detaching imaginal disks from extracellular matrices. In addition, the phenotype of st-sb mutants has led to speculation that the encoded protein is involved in outside to inside signal transduction via its cytoplasmic domain, thus resulting in cytoskeletal reorganization and changes in cell shape during morphogenesis (11).

**Analogous Membrane-associated Proteolytic Systems**

In contrast to the traditional protein catabolic functions of many of the secreted members of the serine protease family and based on the presence of multiple structural domains in the TTSPs, it is tantalizing to speculate that the TTSPs function as key regulators of signaling events at the plasma membrane. Precedents for such functions come from other well characterized membrane-associated proteolytic systems such as the ADAMs (1), the MT-MMPs (2), and the uPA/uPA receptor system (58). The ADAMs have recognized and proposed roles in the proteolysis of extracellular matrix (ECM) components and cell surface proteins, in mediating cell adhesion via integrin binding, in cell fusion and signaling via cell surface-associated proteins, and in RGD-mediated interactions with integrins (59–61). The TTSPs are similarly positioned at the plasma membrane to release ECM components and to proteolytically activate cell surface proteins such as PRRs, growth factors, and cytokines, and to interact with cell surface and soluble ligands. In addition, the presence of the cytoplasmic domains indicates that the TTSPs may be capable of interacting with the cytoskeleton and/or with cellular signaling molecules.

The MT-MMPs function in pericellular cascades to activate other MMPs involved in the cleavage of ECM components. The TTSPs may well perform similar functions in activating proteolytic cascades on...
the plasma membrane. Indeed, this function has been demonstrated for enteropeptidase in the activation of digestive proteases. Moreover, there is increasing evidence for cross-talk between proteolytic systems. The uPA/uPAR receptor system of cell surface-localized proteolytic activity has a recognized role in the initial stage of MMP activation (62), and other serine proteases are also capable of in vitro MMP activation (63, 64). The TTSPs could play a direct role in MMP activation or an indirect role in localizing and activating other serine proteases more directly associated with MMP activation. The activation of uPA by MT-SPI (31) and subsequent downstream MMP activation could be an example of such cross-talk.

Several other parallels may also be drawn from the uPA/uPAR receptor system. That the TTSPs are directly anchored to the plasma membrane implies that they have potential to mimic localization of the uPA/uPAR system to the leading edge of migrating tumor cells (65). Furthermore, the interaction of the uPA/uPAR system, via a urokinase proteolytic mechanism, in mediating cell-cell contacts through association with integrins may also parallel TTSP properties. Indeed the multimodal structure of the TTSPs indicates their capacity to interact with multiple partners and suggests the possibility that these membrane proteins may form part of a signalosome-like complex, thereby mediating at the cell surface multiple signaling pathways as is the case for the uPA/uPAR system (58).

**Concluding Remarks**

What is known about the TTSPs is that they are function or have the structural motifs necessary to function as serine proteases. What can be speculated upon is that their numerous and varied nonproteinoglycan domains are likely to interact with proteolytic substrates and inhibitors as well as other proteins and ligands. Such interactions will potentially regulate the proteolytic activity of the catalytic domain but perhaps may also have functions quite independent of this domain. Furthermore, given the integral plasma membrane nature of the TTSPs, it is tempting to speculate that at least some of the TTSPs will function directly in transducing signals across the plasma membrane, as has been suggested for the Drosophilal TTSP st-sb (11). There is clearly a need for a greater understanding of the biological and physiological functions of this group of unique proteases to obtain a better picture of the dynamics occurring on the cell surface.

Because of the mosaic structure of the TTSPs it will be important to understand the role of their individual domains as well as the role of each protein in toto.

---

**Note Added in Proof**—Two cDNAs encoding the putative TTSPs Xesp-2 and XMT-SPI have recently been identified from Xenopus laevis (67).
Type II Transmembrane Serine Proteases: INSIGHTS INTO AN EMERGING CLASS OF CELL SURFACE PROTEOLYTIC ENZYMES*
John D. Hooper, Judith A. Clements, James P. Quigley and Toni M. Antalis

doi: 10.1074/jbc.R000020200 originally published online November 1, 2000

Access the most updated version of this article at doi: 10.1074/jbc.R000020200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 65 references, 33 of which can be accessed free at http://www.jbc.org/content/276/2/857.full.html#ref-list-1