[Frontiers in Bioscience 13, 569-579, January 1, 2008]

The type II transmembrane serine protease Matriptase-2 – identification, structural features, enzymology, expression pattern and potential roles

Andrew J. Ramsay[1], Janet C. Reid[1], Gloria Velasco[2], James P. Quigley[3], John D. Hooper[1]

[1]Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, Queensland 4059, Australia, [2]Departamento de Bioquimica y Biologia Molecular, Instituto Universitario de Oncologia, Universidad de Oviedo, 33006 Oviedo, Spain, [3]Department of Cell Biology, The Scripps Research Institute, La Jolla, California 92037, USA

TABLE OF CONTENTS

1. ABSTRACT

Matriptase-2 (also known as TMPRSS6) is a recently identified member of the type II transmembrane serine protease (TTSP) family. Structurally this enzyme contains a short cytoplasmic amino terminal tail, a transmembrane region, a stem region containing two CUB domains and three LDL receptor class A domains, and at the carboxy terminal a trypsin-like serine protease domain. The matriptase-2 gene and encoded protein are highly conserved in mammals. Biochemically matriptase-2 has substrate specificity similar to the structurally related protein matriptase (also known as MT-SP1). Although the patho-physiological functions of matriptase-2 are not known, its high mRNA expression in liver and several cancers indicate that this enzyme, similar to other TTSPs, will likely have important cell surface associated roles in normal and disease states. Here we overview the identification of matriptase-2, summarise its structural features, biochemistry, expression pattern and disease associations and discuss its potential functions.

2. INTRODUCTION

Proteolytic events at the cell surface regulate many important cellular processes requiring transduction of signals across the cell surface, release of bioactive growth factors, cytokines and peptide hormones, as well as interactions with other cells and basement membrane and extracellular matrix proteins. As these processes are associated with normal physiological responses such as inflammation and proliferation, and are also co-opted to facilitate progression of diseases such as cancer, it is critical that we have a clear understanding of the proteases involved and the mechanisms regulating proteolytic activity.

Matriptase-2 (1, 2) is a member of the type II transmembrane serine protease (TTSP) family; a recently recognised family of cell surface proteolytic enzymes (3-6). Structurally, the TTSPs are characterised by a short amino terminal cytoplasmic tail, a membrane spanning domain, a stem region containing modular protein interacting

Table 1. Summary of matriptase-2 database entries

Organism	Chromosome Location	GenBank Accession	Amino Acids
Human (Homo sapiens)	22q12.3	NP_705837	811
Chimpanzee (Pan troglodytes)	22 (position 35944325-35945129)	XR_024662 incomplete	na
Macaque (Macaca mulatta)	10 (position 80966561-80967372)	XP_001085319	809
Cow (Bos taurus)	6	XP_871580	800
Pig (Sus scrofa)	no entry	na	na
Dog (Canis familiaris)	10	XP_531743	786
Sheep (Ovis aries)	no entry	na	na
Cat (Felis catus)	no entry	na	na
Mouse (Mus musculus)	15 E2	NP_082178	811
Rat (Rattus norvegicus)	7q34	XP_235768	811

na: not applicable

domains, a short activation domain and a carboxy terminal trypsin-like serine protease domain. It is clear that each of the functionally characterised members of the TTSP family have significant roles in normal physiology and disease. For example, the heart expressed TTSP corin (7) activates pro-atrial natriuretic peptide to the active hormone critical for maintenance of blood pressure and salt-water balance (8). In addition, expression of the gene encoding the TTSP hepsin, originally identified in liver (9), is the most highly up-regulated gene in prostate cancer (10-14). Furthermore, transgenic studies in mice indicate that hepsin has a functional role in prostate cancer progression (15). Also, enteropeptidase, expressed on the surface of enterocytes, is the physiological activator of trypsinogen and thus essential in digestion (16). Studies in knock-out mice indicate that matriptase/MT-SP1 is required for postnatal survival, epidermal barrier function, hair follicle development, and thymic homeostasis (17), while expression of this TTSP in the skin of transgenic mice causes ras-independent multistage carcinogenesis and promotes ras-mediated malignant transformation (18). Fusions of the ETS transcription factor genes *ERG* and *ETV1* with the promoter of the TTSP gene *TMPRSS2* is a common occurrence in prostate cancer with the overexpression of these ETS family members driven by androgen-responsive elements within the *TMPRSS2* promoter (19). TMPRSS2 expression is also upregulated in prostate cancer (20) and this enzyme activates the G protein coupled receptor, protease activated receptor 2, in prostate cancer cells (21). Furthermore, a role for TMPRSS2 in angiogenesis is suggested by the observation that TMPRSS2 expression is up-regulated in micro-vascular endothelial cells undergoing three dimensional tubule morphogenesis (22). The non-mammalian TTSP stubble-stubbloid is required for hormone-dependent epithelial morphogenesis of imaginal discs of Drosophila, including the formation of bristles, legs, and wings (23). Finally, insertion of beta-satellite repeats into the *TMPRSS3* gene causes both congenital and childhood onset autosomal recessive deafness (24).

Although the functions of other TTSP family members have not been determined, the restricted expression patterns and method of identification of several of these enzymes are suggestive of significant roles. For example, TMPRSS4 (formerly designated TMPRSS3) is overexpressed in pancreatic cancer (25), while the epithelial cell specific TTSP, DESC-1 (26, 27), is downregulated during squamous cell carcinoma progression and upregulated during normal epithelial differentiation (28). TMPRSS5, which is expressed in neuronal axons and at the synapses of motoneurons in the spinal cord, is thought to function at the surface of synapses to activate or inactivate other proteins (29). The human airway trypsin-like (HAT) TTSP, originally isolated from the sputum of patients with chronic airway diseases (30), has been proposed to have a role in these diseases because of its ability to mediate inflammatory responses such as increased mucus production (31) and TNFα activity (32). Other members of the TTSP family for which patho-physiological roles have not been determined include the mammalian enzymes matriptase-3 (33), polyserase-I (34), MSPL (5, 35), and the DESC/HAT-like proteases (5, 26, 36, 37) as well as the *Xenopus laevis* protease Xesp-2 (38).

3. IDENTIFICATION

The matriptase-2 cDNA was identified in human (1) and mouse (2) using *in silico* approaches. The complete human cDNA was cloned from fetal liver and named on the basis of its significant structural similarity to the TTSP matriptase/MT-SP1 (1). The mouse coding sequence was identified from an expressed sequence tag clone generated from adult liver (2). The mouse and rat encoded proteins were originally designated Tmprss6 on the basis of TTSP nomenclature (2), however, for consistency the matriptase-2 designation is now applied across all species. The complete sequence of matriptase-2 in macaque, dog and cow and a partial sequence in chimpanzee have also recently been deposited into sequence databases (Table 1).

4. STRUCTURAL FEATURES AND BIOCHEMISTRY

As shown in Figure 1 matriptase-2 has all the structural features of a TTSP including a short cytoplasmic amino terminal tail, a transmembrane region, a stem region containing two complement protein subcomponents C1r/C1s, urchin embryonic growth factor and bone morphogenetic protein 1 (CUB) domains and three low density lipoprotein receptor class A (LDLR) domains, and at the carboxy terminal a trypsin-like serine protease domain. These structural features are absolutely conserved across human, macaque, dog, cow, mouse and rat (Figure 2), with the human protein sharing 95.6% (775/811), 91.1%

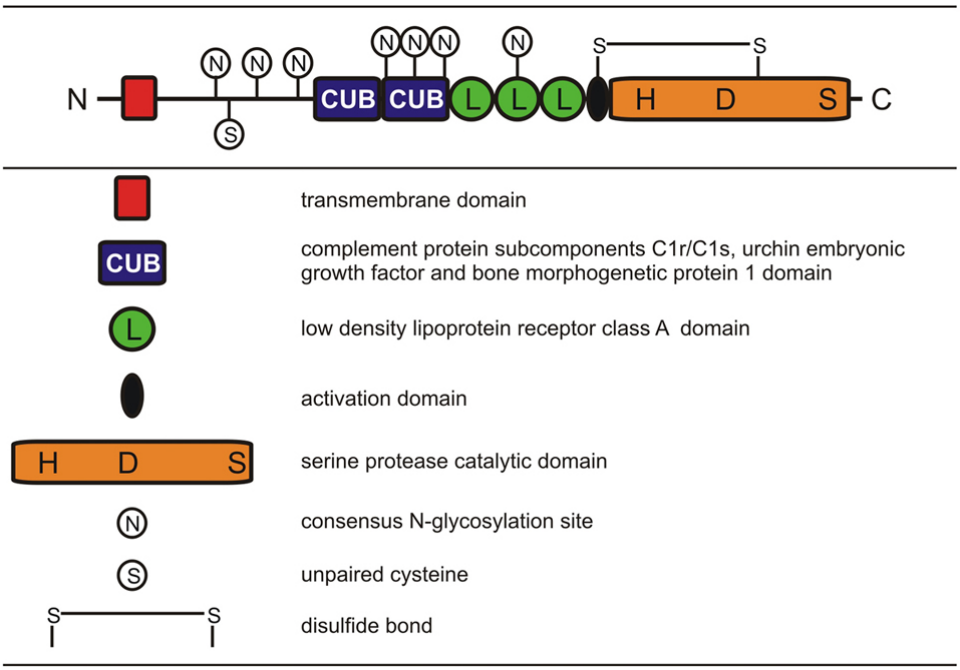


Figure 1. Matriptase-2 domain structure.

(739/811), 85.6% (694/811), 80.1% (650/811) and 80.4% (652/811) identity, respectively, to matriptase-2 from these species. Consistently, by Western blot analysis of lysates from transiently transfected cells, both human (1) and mouse (2) matriptase-2 migrate close to the predicted molecular mass of ~90kDa.

In addition to conserved structural domains, matriptase-2 contains conserved consensus N-glycosylation sites and cysteine residues. Three consensus N-glycosylation sites are present within the region between the transmembrane domain and the first CUB domain. Three other sites occur within the second CUB domain and one within LDLR2. The lack of consensus N-glycosylation sites in the serine protease domain is in contrast to the majority of other TTSPs. However, HAT (30), Hepsin (9), TMPRSS2 (20) and TMPRSS4 (25) also lack glycosylation in this region. Matriptase-2 also has a high content of cysteines - 37 extracellular cysteines are conserved in all species. On the basis of the crystal structure of the spermadhesin porcine seminal plasma (PSP)-I–PSP-II heterodimer (39), cysteines C2/C3 and C4/C5 (numbering per Figure 2) form disulfide bonds in CUB1 and cysteines C6/C7 and C8/C9 within CUB2. Similarly, on the basis of the structure of the LDLR extracellular domain (40), cysteines C10/C12, C11/C14 and C13/C15 form disulfide bonds within LDLR1, C16/C18, C17/C20 and C19/C21 within LDLR2 and C22/C24, C23/C26 and C25/C27 within LDLR3. The following disulfide bonds form within the serine protease domain C29/C30, C32/C36, C33/C34 and C35/C37 and the pro and catalytic regions are linked by the C28/C31 pair. Finally, cysteine C1 of matriptase-2 is unpaired intramolecularly and therefore has the potential to form homo- or hetero- linkages (2).

As well as conserved disulfide bond forming cysteines, the matriptase-2 proteolytic domain has all the features common to members of the serine protease S1 family. These include the serine protease triad of ^{617}H, ^{668}D, and ^{762}Ser residues (human numbering) required for catalytic activity, and an SWG motif predicted to be located at the top of the substrate S1 binding pocket positioning the scissile bond of the substrate in the correct orientation. Proteolytic activation of matriptase-2 is predicted to occur within a motif (^{576}RIVGG) at the junction of the pro- and catalytic domains which is characteristic of serine proteases and conserved across species. Consistent with the presence of an aspartate residue, 6 amino acids before the catalytic serine, which specifies preference for trypsin-like serine protease cleavage following arginine or lysine residues, the recombinant matriptase-2 protease domain cleaves following arginine but not alanine. Interestingly and in contrast to the substrate specificity of the structurally similar TTSP matriptase/MT-SP1 (1, 41), the active recombinant matriptase-2 protease domain does not cleave following lysine (1). The authors explained this difference in peptide substrate specificity using a model of matriptase-2 based on the crystal structure of matriptase/MT-SP1 (42). Their analysis showed that matriptase/MT-SP1 residue ^{190}S permits access of both arginine and lysine residues in the S1 pocket whereas an alanine at this position in matriptase-2 permits accommodation of arginine (1). Consistent with the serine protease classification of matriptase-2, its activity against tripeptide substrates was completely abolished by inhibitors of trypsin-like serine proteases (PMSF, AEBSF, leupeptin and aprotinin) but not by an inhibitor of chymotrypsin-like serine proteases (TPCK) or inhibitors of metallo (EDTA) or cysteine (E-64) proteases (1).

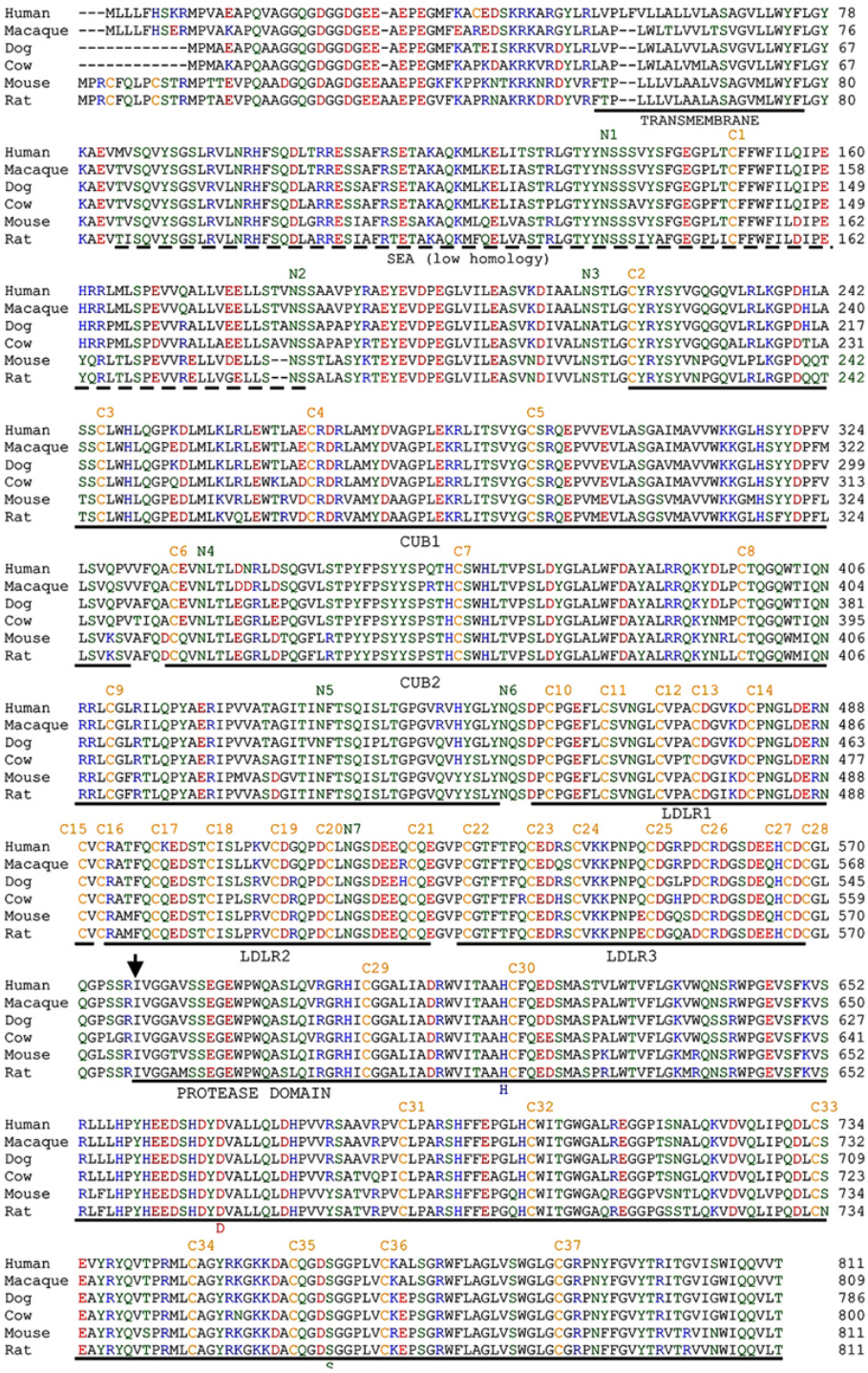


Figure 2. Multiple sequence alignment. Matriptase-2 protein sequences were aligned using the EclustalW algorithm (Australian National Genome Information Service website). The transmembrane domain was identified using the TMHMM algorithm and matriptase-2 sequences were analysed for modular protein domains using the SMART algorithm (ExPASy Proteomics Tools website). Underlined domains: CUB, complement protein subcomponents C1r/C1s, urchin embryonic growth factor and bone morphogenetic protein 1 domain; LDLR, low density lipoprotein receptor class A domain. A region with lower homology to sea urchin sperm protein, enteropeptidase, agrin (SEA) domains is underlined by a dashed line. Consensus N-glycosylation sites are indicated N1 to N7. Cysteines are numbered C1 to C37. The catalytic histidine, aspartate and serine residues necessary for serine protease activity are indicated by "H", "D" and "S" respectively. The activation site is indicated by an arrow. Sequences were obtained from the following GenBank entries: Human NP_705837; Macaque XP_001085319; Dog XP_531743; Cow XP_871580; Mouse NP_082178; Rat XP_235768.

Against macromolecular extracellular matrix (ECM) and basement membrane components, the recombinant matriptase-2 catalytic domain was able to cleave the ECM proteins fibronectin, fibrinogen and type I collagen *in vitro* but not the basement membrane protein laminin (1). In addition, the authors noted that matriptase-2 is not a *bona fide* collagenase as it did not generate the ¾ and ¼ fragments typical of fibrillar collagenases. The matriptase-2 serine protease domain was also incapable of activating the pro forms of the matrix metalloproteases MMP-2 or MMP-9 or the serine protease plasminogen, but was a low efficiency activator of pro-uPA (1). Interestingly, this domain which was generated in bacteria, was subject to autoactivation at 37°C, releasing the proteolytic region from its GST fusion (1). In contrast, in a mammalian cell system transiently over expressing mouse matriptase-2, there was no evidence by Western blot analysis under reducing conditions that the protease had been activated (2). This suggests that autoactivation of endogenous matriptase-2 may not be possible, potentially because the presence of the endogenous activation domain and stem region provide an inhibitory mechanism against auto-catalytic activity. Therefore, proteolytic activation of the matriptase-2 zymogen may represent a post-translation mechanism of controlling the catalytic activity of this TTSP.

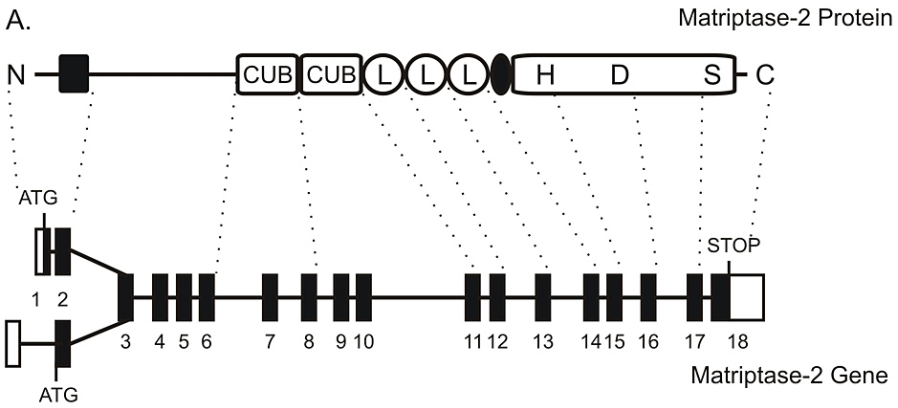
In addition to these structural domains, matriptase-2 also contains a region spanning residues 83 to 186 (human numbering) which has low identity to sea urchin sperm protein, enteropeptidase, agrin (SEA) domains. In other SEA domain containing, membrane spanning proteins, autoproteolysis, following a glycine residue within a conserved motif (eg. GSVVV (43)), releases these molecules from the cell surface. Proteolysis within this conserved motif, which is located near the middle of the SEA domain, is catalysed by conformational stress and the hydroxyl of the serine present within the motif (43). Release from the cell surface by cleavage within a similar SEA domain motif has been reported for the TTSPs porcine enteropeptidase (44) and mouse matriptase/MT-SP1 (45) at the respective cleavage sites ^{117}GSVIV and ^{149}GSVIA. Another member of the TTSP family, HAT, contains a similarly located site (^{109}GSGVR) within its SEA domain. In contrast the SEA domain-like region of matriptase-2 and the SEA domain of other TTSPs, such as mouse matriptase-3 (33) and DESC1 (26), lack a consensus cleavage site near the middle of the domain. However, the matriptase-2 SEA domain-like region contains a cleavage motif-like sequence (^{90}GSLRV) proximal to its amino terminal whereas the mouse matriptase-3 SEA domain has the sequence ^{201}GSLQG near to its carboxy terminal. Whether cleavage within these motifs mediates release of these TTSPs from the cell surface has not been examined.

The role of each of these structural features in regulating matriptase-2 function has not yet been addressed. Clearly the cell surface location and orientation of matriptase-2, which have been confirmed for both the human (1) and mouse (2) proteins by confocal microscopy analysis of transiently transfected cells, provide a level of spatial regulation of the protease domain. It is also possible that the matriptase-2 cytoplasmic tail will modulate protein function. However, none of the matriptase-2 consensus phosphorylation sites are conserved across species, suggesting that phosphorylation of this protein is either not essential to its function or is unique to each organism. Furthermore, additional regulation of matriptase-2 activity will likely be achieved via its two CUB and three LDLR domains, and possibly also involve the region spanning the low-identity SEA domain. Each of these domains is common amongst TTSP family members (3-6) and, in particular, many CUB domain containing proteins are proteases (46). Accordingly, it will be of interest to determine whether the matriptase-2 CUB domains are involved in oligomerization or recognition of substrates as is the case for other proteins (46), and similarly, whether the LDLR domains of this TTSP mediate interactions with macromolecules such as serine protease:inhibitor complexes and lipoproteins (47).

5. GENE STRUCTURE AND REGULATION

As indicated in Figure 3, the matriptase-2 gene is highly conserved across mammalian species and ranges in size from ~29 kb in mouse to ~40 kb in chimpanzee. Each gene spans 18 exons with 17 intervening introns, with the size and junctions of 14 of the exons absolutely conserved. Also, although the size of the last exon varies, the size of the coding sequence of this exon is conserved at 159 nucleotides. Each of the mammalian *matriptase-2* genes is located in a genomic region devoid of other TTSP genes. This is similar to other TTSP genes such as *corin* (7) and *hepsin* (48) and in contrast to *TMPRSS4* (25) and *TMPRSS5* (29) which both localise to human chromosome region 11q23.3, *TMPRSS2* (49) and *TMPRSS3* (24) to 21q22 and the HAT, HAT-like, DESC and DESC-like encoding genes which localise to 4q11 (5). Sequences at the 5' and 3' end of each intron conform to the GT-AG rule for splice-site recognition (50) except for the 5' end of the first human, chimpanzee and macaque intron which has GC instead of GT. All matriptase-2 protein domain boundaries correspond with intron/exon junctions of the encoding gene across all species (Figure 3A). The first CUB domain is encoded by two exons (7 and 8), and the second CUB domain by three exons (9 to11). The three LDLR domains of matriptase-2 are each encoded by separate exons (12, 13 and 14). Finally, the serine-protease domain, including the activation domain, is encoded by four exons (15 to 18).

Variation occurs between lower and higher mammals towards the 5' end of the *matriptase-2* gene. In human, chimpanzee and macaque the initiating codon is located in exon 1 whereas in rodents this codon is in exon 2. The size of the intron interrupting the first two exons in these species is also different. In human, chimpanzee and macaque the first exon is only 95 nucleotides upstream of exon 2 whereas in rodents these exons are separated by an intron greater than 4.5 kb in size. At the protein level this variation appears to have little effect. As shown in Figure 2, the rodent protein extends 3 residues beyond the amino terminal of the human and macaque protein (the complete chimpanzee protein is currently not available). In addition, both human and mouse *matriptase-2* localise efficiently to



B.

	Human	Chimpanzee	Macaque	Cow	Dog	Mouse	Rat
Exon 1	140	>100[1]	140	n/a	n/a	186	184
Intron A	95	95	95	n/a	n/a	3082	>4042
Exon 2	203	203	197	197	197	235	235
Intron B	4666	4666	4789	4834	3726	3730	4713
Exon 3	134	134	134	134	134	134	134
Intron C	1700	1417	1256	1410	1640	737	691
Exon 4	68	68	68	68	68	68	68
Intron D	557	556	562	551	519	516	520
Exon 5	185	185	185	185	185	179	179
Intron E	313	310	305	241	226	273	246
Exon 6	42	42	42	42	42	42	42
Intron F	5768	>6467	4534	4043	3128	4311	4709
Exon 7	205	205	205	205	205	205	205
Intron G	3158	>4028	2713	2721	1999	632	604
Exon 8	137	137	137	137	137	137	137
Intron H	1443	1442	1434	1356	1351	1240	1167
Exon 9	113	incomplete[2]	113	113	113	113	113
Intron I	322	incomplete	321	295	335	206	200
Exon 10	110	110	110	110	110	110	110
Intron J	9014	>9261	>9200	8427	8270	5667	5559
Exon 11	146	146	146	146	146	146	146
Intron K	426	426	445	377	360	360	320
Exon 12	99	99	99	99	99	99	99
Intron L	964	964	978	940	769	806	>959
Exon 13	114	114	114	114	114	114	114
Intron M	2499	incomplete	1527	1357	1325	1071	726
Exon 14	117	incomplete	117	117	117	117	117
Intron N	263	incomplete	272	327	277	250	249
Exon 15	169	169	169	169	169	169	169
Intron O	1139	1135	1118	1331	1308	1123	1097
Exon 16	272	272	272	272	272	272	272
Intron P	2110	2106	2022	1996	1696	1537	1532
Exon 17	137	137	137	137	137	137	137
Intron Q	587	587	581	>292	532	235	205
Exon 18[3]	803 (159)	797 (159)	803 (159)	807 (159)	694 (159)	796 (159)	692 (159)
Total	37415	>40330	>35340	>33550	>30400	28239	30692

Notes:

1. The greater than symbol indicates that the intron has not been completely sequenced.
2. "Incomplete" indicates that sequence at an exon/intron junction is not available.
3. The bracketed amount is the size of the coding sequence for exon 18.

Figure 3. Matriptase-2 gene structure. A. Protein structural features are marked in the upper panel and include transmembrane (black box), two CUB, three LDLR (L), activation (black oval) and proteolytic (boxed HDS) domains. Genomic organization is indicated in the lower panel. Coding and non-coding exon regions are represented by, respectively, black and white boxes. Upper exons 1 and 2 represent the structure of the gene in higher mammals (human, chimpanzee and macaque) and the lower exons 1 and 2 represent the gene structure in rodents. Dotted lines indicate exons encoding the protein structural domains in panel A. Not to scale. B. Comparison of sizes of exons and introns of the matriptase-2 gene in 7 mammalian species.

the cell surface in cells transiently transfected with the corresponding expression constructs (1, 2). Other contrasting features across *matriptase-2* genes are the size of exons 2 and 5. In human and chimpanzee, exon 2 is 203 nucleotides whereas in macaque, cow and dog it is 197 nucleotides and 235 nucleotides in rodents. The availability of sequence for exon 1 and intron A of cow and dog, will likely indicate the evolution of this, the major region of variation across these mammalian species. In addition, whereas exon 5 is 185 nucleotides in human, chimpanzee, macaque, cow and dog, this exon is 179 nucleotides in rodents. This variation leads to the rodent protein being 2 residues shorter in the SEA domain-like region immediately before the second matriptase-2 consensus N-glycosylation site.

There is also evidence that mammalian *matriptase-2* genes are capable of generating multiple mRNA splice variants. For example in human, several splice variants are predicted to encode matriptase-2 isoforms; one truncated after the middle of the second CUB domain (GenBank accession number EAW60141) and another with amino acid insertions before both the first CUB domain and the protease activation domain (GenBank accession number CAK54774). This last isoform is also predicted from a dog mRNA splice variant (GenBank accession number XP_850550). Another dog splice variant, generated by skipping exon 6, is predicted to translate a protein lacking 14 matriptase-2 residues immediately before the first CUB domain (GenBank accession number XP_531743). The rat gene generates mRNA splice variants encoding the full-length protein shown in Figure 2 as well as an isoform lacking the transmembrane domain predicted to be a secreted protein (2). It is clear that several of the generated protein isoforms will be functionally different to the full-length protein. In particular, these include isoforms predicted to be secreted proteins and those lacking the serine protease catalytic domain. Therefore, it will be important to determine the mechanisms regulating the generation of matriptase-2 mRNA splice variants as well as the mechanisms controlling transcription of the *matriptase-2* gene generally. Currently, there has been only one report of *matriptase-2* gene regulation and this demonstrated gonadotropin mediated down-regulation of transcription in mouse testis Leydig cells (51).

6. EXPRESSION PATTERN AND DISEASE ASSOCIATION

The expression of matriptase-2 mRNA in adult and embryonic tissues, during embryonic development and in cancers is shown in Table 2. In human, mouse and rat, liver is the common site of abundant matriptase-2 mRNA expression (1, 2). This pattern of expression in one or a few tissues, is consistent with the restricted transcription of other TTSP genes and also with the proposal that the encoded proteins have tissue-specific functions (4). For example, enteropeptidase expression is restricted to enterocytes of the proximal small intestine where it activates trypsinogen (52), while corin is predominantly produced by heart myocytes where it activates pro-ANP (7). Each of these proteolytic activation events occurs on the surface of these cells. Other TTSPs also have restricted expression patterns including TMPRSS2 in prostate and colon (49, 54), hepsin in liver and kidney (48, 55), and DESC1 in epithelial locations (27). Accordingly, it will be important to determine the role of matriptase-2 on the surface of cells in the liver.

Our analysis of the GenBank Unigene database indicated that human matriptase-2 mRNA is also expressed strongly in adrenal gland and brain with lower levels in lung, testis and uterus. Interestingly, the sites of highest expression in normal human tissues (liver, adrenal gland, brain and uterus) are also sites of high expression in cancer (Table 2). In addition, Velasco *et al.* have noted that the region containing the *matriptase-2* gene is frequently altered in other human cancers, including insulinomas, ependymomas, and colorectal and breast carcinomas (1). Significantly, two other recent reports also indicate an association of matriptase-2 with breast cancer. A proteomics approach indicated elevated levels of matriptase-2 protein in invasive ductal breast carcinoma (56), while an epidemiological approach suggested that the matriptase-2 gene is associated with breast cancer risk in an Eastern Finnish population (57). To determine whether matriptase-2 is differentially expressed in these cancers, studies are needed comparing its expression in a defined population of matched non-diseased and cancerous samples.

RT-PCR and *in situ* hybridization analyses provided significant insight into the cells expressing matriptase-2 mRNA in adult and embryonic mouse tissues (2). In adult tissues, in addition to liver, kidney was also a site of high matriptase-2 mRNA expression with lower levels in uterus and much smaller amounts detected in many other tissues. In liver matriptase-2 mRNA was expressed by hepatocytes, whereas expression was consistent throughout kidney with glandular columnar epithelial cells the predominant site in uterus. In embryos matriptase-2 mRNA expression peaked at day 13.5 *post coitus*. In addition to high expression in liver, matriptase-2 mRNA was strongly detected in olfactory epithelial cells of the nasal cavity and in pharyngo-tympanic tubes. These data suggest roles for matriptase-2 in adult tissues as well as during embryogenesis. In addition, the overlapping expression of matriptase-2 and hepsin mRNA in embryonic and adult mouse suggested some extent of functional redundancy of these proteins.

7. CONCLUSION

Although the function of matriptase-2 is not known, a number of observations point to key roles for this protein in a restricted set of tissues and, potentially, in a number of cancers. These data also indicate that matriptase-2 will function by cleaving a limited repertoire of substrates/inhibitors and by interacting with other proteins via its non-catalytic domains. For example, the ability of matriptase-2 to cleave fibrinogen suggests that this enzyme will be involved in processes requiring fibrin formation such as angiogenesis (1) which is necessary for progression of several diseases including cancer. In addition, the

Table 2. Matriptase-2 mRNA expression. Summarised from published data from Northern blot and RT-PCR analysis of human and mouse tissues, respectively, as well as from GenBank UniGene database entries for matriptase-2.

	Human		Mouse		Rat
	Northern	UniGene[1]	RT-PCR	UniGene	UniGene
Adrenal gland	x[2]	31	x	x	x
Brain/whole brain	nil	16	-/+	0	0
Colon	nil	0	x	x	x
Head and neck	x	x	x	69	0
Heart/vascular	nil	0	-/+	80	0
Intestine/GI tract[3]	nil	0	-/+	0	0
Kidney	nil	0	+++	x	0
Leukocytes/blood	nil	0	x	0	x
Liver	+++++	25	+++++	641	142
Lung	nil	5	-/+	9	0
Mammary gland	x	0	x	9	x
Muscle	nil	0	-/+	37	0
Ovary	nil	0	x	x	x
Pancreas	nil	0	-/+	0	0
Placenta	nil	0	x	x	0
Prostate	nil	0	x	0	x
Spleen	nil	0	-/+	10	x
Stomach	nil	x	-/+	x	x
Testis	nil	8	x	0	x
Thymus	nil	0	-/+	0	x
Uterus	x	8	++	x	x
Adrenal tumor	x	78	x	x	x
Glioma	x	9	x	x	x
Liver tumor	x	22	x	x	x
Uterine tumor	x	22	x	x	x
Fetal brain	nil	x	x	x	x
Fetal kidney	nil	x	x	x	x
Fetal liver	++++	x	x	x	x
Fetal lung	nil	x	x	x	x
Fetus	x	9	x	x	x
Embryo	x	0	x	0	0
Mid-gestation embryo	x	x	x	5	x
Neonate	x	0	x	36	0
Juvenile	x	0	x	20	73
Adult	x	3	x	22	0
Embryo day 7.5	x	x	-/+	x	x
Embryo day 8.5	x	x	-/+	x	x
Embryo day 9.5	x	x	-/+	x	x
Embryo day 10.5	x	x	-/+	x	x
Embryo day 12.5	x	x	++	x	x
Embryo day 13.5	x	x	+++++	x	x
Embryo day 15.5	x	x	++	x	x

Notes: 1. GenBank UniGene database - transcripts per million. 2. x = not examined by Northern blot or RT-PCR analysis; no entry in UniGene. 3. GI tract, gastrointestinal tract

abundant expression of matriptase-2 mRNA in the liver of human and rodents suggests that the encoded protein will function in this tissue and, if released from the cell surface, will also have a role as a circulating enzyme as is the case for other trypsin-like serine proteases involved in processes such as blood coagulation (58) and tissue remodelling in cancer (59,60). In fact it is possible that the cell surface location and proteolytic activity of matriptase-2 will facilitate cancer progression as is the case for other TTSPs including hepsin (15) and matriptase/MT-SP1 (18).

It is clear that the functions of matriptase-2 in normal physiology and disease will be regulated at a number of levels. At the transcriptional level the *matriptase-2* gene is capable of generating multiple splice variants, several of which, at least, will translate protein isoforms with differing functions. Also, the restricted expression pattern of matriptase-2 mRNA suggests that transcription is tightly regulated. Furthermore, it is likely that the unpaired cysteine of matriptase-2, which is located between the transmembrane domain and the first CUB domain, will be involved in homo- or hetero-meric linkages which will regulate protein function. In addition, the location of matriptase-2 at the cell surface will provide a level of regulation by restricting access of the protease to potential substrates and also by requiring a mechanism, such as proteolysis, to release the protein from the cell surface. Release from the cell surface may also be mediated via transcriptional events which generate mRNA splice variants encoding a secreted form of the protein.

As has been the case for other members of the TTSP family including corin (8), hepsin (15) and matriptase/MT-SP1 (17, 18, 61), it is likely that greatest progress on the patho-physiological roles of matriptase-2 will occur using either knock-out or transgenic mice. It has also been suggested that the overlapping expression of matriptase-2 and hepsin in mice indicates a level of functional redundancy which will require the careful analysis of matriptase-2 knock-out mice and, potentially,

the generation of matriptase-2/hepsin double knock-out mice (2).

8. ACKNOWLEDGEMENTS

AJR is supported by a PhD Scholarship from the Queensland Cancer Fund. JDH is supported by an R Douglas Wright Fellowship (number 339732) from the National Health and Medical Research Council of Australia. GV is supported by grant FISS-05-PI040225 from Ministerio de Sanidad-Spain.

9. REFERENCES

1. Velasco G, S. Cal, V. Quesada, L.M. Sanchez & C. Lopez-Otin: Matriptase-2, a membrane-bound mosaic serine proteinase predominantly expressed in human liver and showing degrading activity against extracellular matrix proteins. *J Biol Chem* 277, 37637-37646 (2002)
2. Hooper JD, L. Campagnolo, G. Goodarzi, T.N. Truong, H. Stuhlmann, & J.P. Quigley:. Mouse matriptase-2: identification, characterization and comparative mRNA expression analysis with mouse hepsin in adult and embryonic tissues. *Biochem J* 373, 689-702 (2003)
3. Wu Q: Type II transmembrane serine proteases. *Curr Top Dev Biol* 54, 167-206 (2003)
4. Hooper JD, J.A. Clements, J.P Quigley & T.M. Antalis: Type II transmembrane serine proteases. Insights into an emerging class of cell surface proteolytic enzymes. *J Biol Chem* 276, 857-860 (2001)
5. Netzel-Arnett S, J.D Hooper, R. Szabo, E.L. Madison, J.P. Quigley, T.H. Bugge & T.M. Antalis: Membrane anchored serine proteases: a rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer. *Cancer Metastasis Rev* 22, 237-258 (2003)
6. Szabo R, Q. Wu, R.B. Dickson, S. Netzel-Arnett, T.M. Antalis & Bugge TH: Type II transmembrane serine proteases. *Thromb Haemost* 90, 185-193 (2003)
7. Yan W, N. Sheng, M. Seto, J. Morser & Q. Wu: Corin, a mosaic transmembrane serine protease encoded by a novel cDNA from human heart. *J Biol Chem* 274, 14926-14935 (1999)
8. Chan JC, O. Knudson, F. Wu, J. Morser, W.P. Dole & Q. Wu: Hypertension in mice lacking the proatrial natriuretic peptide convertase corin. *Proc Natl Acad Sci U S A* 102, 785-790 (2005)
9. Leytus SP, K.R. Loeb, F.S. Hagen, K. Kurachi & E.W. Davie: A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells. *Biochemistry* 27, 1067-1074 (1988)
10. Luo J, D.J. Duggan, Y. Chen, J. Sauvageot, C.M. Ewing, M.L. Bittner, J.M. Trent & W.B. Isaacs: Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. *Cancer Res* 61, 4683-4688 (2001)
11. Welsh JB, L.M. Sapinoso, A.I. Su, S.G. Kern, J. Wang-Rodriguez, C.A. Moskaluk, H.F. Frierson Jr. & G.M. Hampton: Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res* 61,5974-5978 (2001)
12. Magee JA, T. Araki, S. Patil, T. Ehrig, L. True, P.A. Humphrey, W.J. Catalona, M.A. Watson & J. Milbrandt: Expression profiling reveals hepsin overexpression in prostate cancer. *Cancer Res* 61, 5692-5696 (2001)
13. Stamey TA, J.A. Warrington, M.C. Caldwell, Z. Chen, Z. Fan, M. Mahadevappa, J.E. McNeal, R. Nolley & Z. Zhang: Molecular genetic profiling of Gleason grade 4/5 prostate cancers compared to benign prostatic hyperplasia. *J Urol* 166, 2171-2177 (2001)
14. Dhanasekaran SM, T.R. Barrette, D. Ghosh, R. Shah, S. Varambally, K. Kurachi, K.J. Pienta, M.A. Rubin & A.M. Chinnaiyan: Delineation of prognostic biomarkers in prostate cancer. *Nature* 412, 822-826 (2001)
15. Klezovitch O, J. Chevillet, J. Mirosevich, R.L. Roberts, R.J. Matusik & V. Vasioukhin: Hepsin promotes prostate cancer progression and metastasis. *Cancer Cell* 6, 185-195 (2004)
16. Fonseca P & A. Light: The purification and characterization of bovine enterokinase from membrane fragments in the duodenal mucosal fluid. *J Biol Chem* 258, 14516-14520 (1983)
17. List K, C.C. Haudenschild, R. Szabo, W. Chen, S.M. Wahl, W. Swaim, L.H. Engelholm, N. Behrendt & T.H. Bugge: Matriptase/MT-SP1 is required for postnatal survival, epidermal barrier function, hair follicle development, and thymic homeostasis. *Oncogene* 21, 3765-3779 (2002)
18. List K, R. Szabo, A. Molinolo, V. Sriuranpong, V. Redeye, T. Murdock, B. Burke, B.S. Nielsen, J.S. Gutkin & T.H., Bugge: Deregulated matriptase causes ras-independent multistage carcinogenesis and promotes ras-mediated malignant transformation. *Genes Dev* 19, 1934-1950 (2005)
19. Tomlins SA, D.R. Rhodes, S. Perner, S.M. Dhanasekaran, R. Mehra, X.W. Sun, S. Varambally, X. Cao, J. Tchinda, R. Kuefer, C. Lee, J.E. Montie, R.B. Shah, K.J. Pienta KJ, M.A. Rubin & A.M. Chinnaiyan: Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 310, 644-648. (2005)
20. Lin B, C. Ferguson, J.T. White, S. Wang, R. Vessella, L.D. True, L. Hood L & P.S. Nelson: Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. *Cancer Res* 59, 4180-4184 (1999)
21. Wilson S, B. Greer, J. Hooper, A. Zijlstra, B. Walker, J. Quigley & S. Hawthorne: The membrane-anchored serine protease, TMPRSS2, activates PAR-2 in prostate cancer cells. *Biochem J* 388, 967-972 (2005)
22. Aimes RT, A. Zijlstra, J.D. Hooper, S.M. Ogbourne, M.L. Sit, S. Fuchs, D.C. Gotley, J.P. Quigley & T.M. Antalis: Endothelial cell serine proteases expressed during vascular morphogenesis and angiogenesis. *Thromb Haemost* 89, 561-572 (2003)
23. Appel LF, M. Prout, R.Abu-Shumays, A. Hammonds, J.C. Garbe, D. Fristrom & J. Fristrom: The Drosophila Stubble-stubbloid gene encodes an apparent transmembrane serine protease required for epithelial morphogenesis. *Proc Natl Acad Sci U S A* 90, 4937-4941 (1993)
24. Scott HS, J. Kudoh, M. Wattenhofer, K. Shibuya, A. Berry, R. Chrast, M. Guipponi, J. Wang, K. Kawasaki, S. Asakawa, S. Minoshima, F. Younus, S.Q. Mehdi, U. Radhakrishna, M.P. Papasavvas, C. Gehrig, C. Rossier, M. Korostishevsky, A. Gal, N. Shimizu, B. Bonne-Tamir &

S.E. Antonarakis: Insertion of beta-satellite repeats identifies a transmembrane protease causing both congenital and childhood onset autosomal recessive deafness. *Nat Genet* 27, 59-63 (2001)
25. Wallrapp C, S. Hahnel, F. Muller-Pillasch, B. Burghardt, T. Iwamura, M. Ruthenburger, M.M. Lerch, G. Adler & T.M. Gress: A novel transmembrane serine protease (TMPRSS3) overexpressed in pancreatic cancer. *Cancer Res* 60, 2602-2606 (2000)
26. Hobson JP, S. Netzel-Arnett, R. Szabo, S.M. Rehault, F.C. Church, D.K. Strickland, D.A. Lawrence, T.M. Antalis & T.H. Bugge: Mouse DESC1 is located within a cluster of seven DESC1-like genes and encodes a type II transmembrane serine protease that forms serpin inhibitory complexes. *J Biol Chem* 279, 46981-46994 (2004)
27. Lang JC & D.E. Schuller: Differential expression of a novel serine protease homologue in squamous cell carcinoma of the head and neck. *Br J Cancer* 84, 237-243 (2001)
28. Sedghizadeh PP, S.R. Mallery, S.J. Thompson, L. Kresty, F.M. Beck, E.K. Parkinson, J. Biancamano & J.C. Lang: Expression of the serine protease DESC1 correlates directly with normal keratinocyte differentiation and inversely with head and neck squamous cell carcinoma progression. *Head Neck* 28, 432-440 (2006)
29. Yamaguchi N, A. Okui, T. Yamada, H. Nakazato & S. Mitsui: Spinesin/TMPRSS5, a novel transmembrane serine protease, cloned from human spinal cord. *J Biol Chem* 277, 6806-6812 (2002)
30. Yamaoka K, K. Masuda, H. Ogawa, K. Takagi, N. Umemoto,& S. Yasuoka: Cloning and characterization of the cDNA for human airway trypsin-like protease. *J Biol Chem* 273, 11895-11901 (1998)
31. Chokki M, S. Yamamura, H. Eguchi, T. Masegi, H. Horiuchi, H. Tanabe, T. Kamimura & S. Yasuoka: Human airway trypsin-like protease increases mucin gene expression in airway epithelial cells. *Am J Respir Cell Mol Biol* 30, 470-478 (2004)
32. Chokki M, H Eguchi, I. Hamamura, H. Mitsuhashi & T. Kamimura: Human airway trypsin-like protease induces amphiregulin release through a mechanism involving protease-activated receptor-2-mediated ERK activation and TNF alpha-converting enzyme activity in airway epithelial cells. *Febs J* 272, 6387-6399 (2005)
33. Szabo R, S. Netzel-Arnett, J.P. Hobson, T.M. Antalis & T.H. Bugge: Matriptase-3 is a novel phylogenetically preserved membrane-anchored serine protease with broad serpin reactivity. *Biochem J* 390, 231-242 (2005)
34. Cal S, V. Quesada, C. Garabaya & C. Lopez-Otin: Polyserase-I, a human polyprotease with the ability to generate independent serine protease domains from a single translation product. *Proc Natl Acad Sci U S A* 100, 9185-9190 (20030)
35. Kim DR, S. Sharmin, M. Inoue & H. Kido: Cloning and expression of novel mosaic serine proteases with and without a transmembrane domain from human lung. *Biochim Biophys Acta* 1518, 204-209 (2001)
36. Hansen IA, M. Fassnacht, S. Hahner, F. Hammer, M. Schammann, S.R. Meyer, A.B. Bicknell & B. Allolio: The adrenal secretory serine protease AsP is a short secretory isoform of the transmembrane airway trypsin-like protease. *Endocrinology* 145, 1898-1905 (2004)
37. Puente XS, L.M. Sanchez, C.M. Overall & C. Lopez-Otin: Human and mouse proteases: a comparative genomic approach. *Nat Rev Genet* 4, 544-558 (2003)
38. Yamada K, T. Takabatake & K. Takeshima: Isolation and characterization of three novel serine protease genes from Xenopus laevis. *Gene* 252, 209-216 (2000)
39. Varela PF, A. Romero, L. Sanz, M.J. Romao, E. Topfer-Petersen & J.J. Calvete: The 2.4 A resolution crystal structure of boar seminal plasma PSP-I/PSP-II: a zona pellucida-binding glycoprotein heterodimer of the spermadhesin family built by a CUB domain architecture. J Mol Biol 274, 635-649 (1997)
40. Rudenko G & J. Deisenhofer: The low-density lipoprotein receptor: ligands, debates and lore. Curr Opin Struct Biol 13, 683-689 (2003)
41. Lee SL, R.B. Dickson & C.Y Lin: Activation of hepatocyte growth factor and urokinase/plasminogen activator by matriptase, an epithelial membrane serine protease. J Biol Chem 275, 36720-36725 (2000)
42. Friedrich R, P. Fuentes-Prior, E. Ong, G. Coombs, M. Hunter, R. Oehler, D. Pierson, R. Gonzalez, R. Huber, W. Bode & E.L. Madison: Catalytic domain structures of MT-SP1/matriptase, a matrix-degrading transmembrane serine proteinase. J Biol Chem 277, 2160-2168 (2002)
43. Macao B, D.G. Johansson, G.C. Hansson & T. Hard: Autoproteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. Nat Struct Mol Biol 13, 71-76 (2006)
44. Matsushima M, M. Ichinose, N. Yahagi, N. Kakei, S. Tsukada, K. Miki,K. Kurokawa, K. Tashiro, K. Shiokawa, K. Shinomiya, H. Umeyama, H. Inoue, T. Takahashi & K. Takahashi: Structural characterization of porcine enteropeptidase. J Biol Chem 269, 19976-19982 (1994)
45. Cho EG, M.G. Kim, C. Kim, S.R. Kim, I.S. Seong, C. Chung, R.H. Schwartz & D. Park:. N-terminal processing is essential for release of epithin, a mouse type II membrane serine protease. J Biol Chem 276, 44581-44589 (2001)
46. Blanc G, B. Font, D. Eichenberger, C. Moreau, S. Ricard-Blum, D.J. Hulmes & C. Moali: Insights into how CUB domains can exert specific functions while sharing a common fold: Conserved and specific features of the CUB1 domain contribute to the molecular basis of procollagen C-proteinase enhancer-1 activity. J Biol Chem 282, 16924-16933 (2007)
47. Strickland DK, S.L. Gonias & W.S. Argraves:. Diverse roles for the LDL receptor family. Trends Endocrinol Metab 13, 66-74 (2002)
48. Tsuji A, A. Torres-Rosado, T. Arai, M.M. Le Beau, R.S. Lemons, S.H. Chou & K. Kurachi: Hepsin, a cell membrane-associated protease. Characterization, tissue distribution, and gene localization. J Biol Chem 266, 16948-16953 (1991)
49. Paoloni-Giacobino A, H. Chen, M.C. Peitsch, C. Rossier & S.E. Antonarakis: 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 (1997)
50. Shapiro MB & P. Senapathy: RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res 15, 7155-7174 (1987)
51. Odet F, A. Verot & B. Le Magueresse-Battistoni: The mouse testis is the source of various serine proteases and serine proteinase inhibitors (SERPINs): Serine proteases

and SERPINs identified in Leydig cells are under gonadotropin regulation. *Endocrinology* 147, 4374-4383 (2006)
52. Kitamoto Y, R.A. Veile, H. Donis-Keller & J.E. Sadler: cDNA sequence and chromosomal localization of human enterokinase, the proteolytic activator of trypsinogen. Biochemistry 34, 4562-4568 (1995)
53. Takeuchi T, M.A. Shuman & C.S Craik: Reverse biochemistry: use of macromolecular protease inhibitors to dissect complex biological processes and identify a membrane-type serine protease in epithelial cancer and normal tissue. *Proc Natl Acad Sci U S A* 96, 11054-11061 (1999)
54. Jacquinet E, N.V. Rao, G.V. Rao & J.R. Hoidal: Cloning, genomic organization, chromosomal assignment and expression of a novel mosaic serine proteinase: epitheliasin. *FEBS Lett* 468, 93-100 (2000)
55. Vu TKH, R.W. Liu, C.J. Haaksma, J.J. Tomasek & E.W. Howard: Identification and cloning of the membrane-associated serine protease, hepsin, from mouse preimplantation embryos. *J Biol Chem* 272, 31315-31320 (1997)
56. Overall CM, E.M. Tam, R. Kappelhoff, A. Connor, T. Ewart, C.J. Morrison, X. Puente, C. Lopez-Otin & A. Seth: Protease degradomics: mass spectrometry discovery of protease substrates and the CLIP-CHIP, a dedicated DNA microarray of all human proteases and inhibitors. *Biol Chem* 385, 493-504 (2004)
57. Hartikainen JM, H. Tuhkanen, V. Kataja, M. Eskelinen, M. Uusitupa, V.M. Kosma & A. Mannermaa: Refinement of the 22q12-q13 breast cancer-associated region: evidence of TMPRSS6 as a candidate gene in an eastern Finnish population. *Clin Cancer Res* 12, 1454-1462 (2006)
58. Davie EW, K. Fujikawa & W. Kisiel: The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry* 30, 10363-10370 (1991)
59. Testa JE & J.P. Quigley: The role of urokinase-type plasminogen activator in aggressive tumor cell behavior. *Cancer Metastasis Rev* 9, 353-367 (1990)
60. Dano K, N. Behrendt, G. Hoyer-Hansen, M. Johnsen, L.R. Lund, M. Ploug & J. Romer: Plasminogen activation and cancer. *Thromb Haemost* 93, 676-681 (2005)
61. Netzel-Arnett S, B.M. Currie, R. Szabo, C.Y. Lin, L.M. Chen, K.X. Chai KX, T.M. Antalis, T.H. Bugge & K. List: Evidence for a matriptase-prostasin proteolytic cascade regulating terminal epidermal differentiation. *J Biol Chem* 281, 32941-32945 (2006)

Key Words: Matriptase-2, Serine Protease, Transmembrane, Cancer, Neoplasia, Tumor, Review

Send correspondence to: John Hooper, PhD; Institute of Health and Biomedical Innovation, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, Queensland 4059, Australia, Tel: 61-7-3138-6197, Fax: 61-7-3138-6030; E-mail: jd.hooper@qut.edu.au