Matriptase-3 is a novel phylogenetically preserved membrane-anchored serine protease with broad serpin reactivity

Roman SZABO*, Sarah NETZEL-ARNETT†, John P. HOBSON*, Toni M. ANTALIS† and Thomas H. BUGGE*[1]

*Proteases and Tissue Remodeling Unit, Oral and Pharyngeal Cancer Branch, National Institute of Dental and Craniofacial Research, National Institutes of Health, 30 Convent Drive, Bethesda, MD 20892, U.S.A., and †Departments of Physiology and Surgery, University of Maryland School of Medicine, Rockville, MD 20855, U.S.A.

We report in the present study the bioinformatic identification, molecular cloning and biological characterization of matriptase-3, a novel membrane-anchored serine protease that is phylogenetically preserved in fish, birds, rodents, canines and primates. The gene encoding matriptase-3 is located on syntenic regions of human chromosome 3q13.2, mouse chromosome 16B5, rat chromosome 11q21 and chicken chromosome 1. Bioinformatic analysis combined with cDNA cloning predicts a functional TTSP (type II transmembrane serine protease) with 31 % amino acid identity with both matriptase/MT-SP1 and matriptase-2. This novel protease is composed of a short N-terminal cytoplasmic region followed by a transmembrane domain, a stem region with one SEA, two CUB and three LDLRa (low-density lipoprotein receptor domain class A) domains and a C-terminal catalytic serine protease domain. Transcript analysis revealed restricted, species-conserved expression of *matriptase-3*, with the highest mRNA levels in brain, skin, reproductive and oropharyngeal tissues. The full-length *matriptase-3* cDNA directed the expression of a 90 kDa N-glycosylated protein that localized to the cell surface, as assessed by cell-surface biotin labelling. The purified activated matriptase-3 serine protease domain expressed in insect cells hydrolysed synthetic peptide substrates, with a strong preference for Arg at position P_1, and showed proteolytic activity towards several macromolecular substrates, including gelatin, casein and albumin. Interestingly, activated matriptase-3 formed stable inhibitor complexes with an array of serpins, including plasminogen activator inhibitor-1, protein C inhibitor, α1-proteinase inhibitor, α2-antiplasmin and antithrombin III. Our study identifies matriptase-3 as a novel biologically active TTSP of the matriptase subfamily having a unique expression pattern and post-translational regulation.

Key words: extracellular protease, inhibitor complex, matriptase-3, pericellular proteolysis, serpin, type II transmembrane serine protease (TTSP).

INTRODUCTION

Extracellular proteolysis is an integral part of tissue development and homoeostasis, and is causally associated with a number of pathological processes, including cancer. Likely to encompass more than 200 different enzymes [1], extracellular proteases cleave structural components of the extracellular matrix and basement membranes, liberate growth factors and receptor ligands from the extracellular matrix and latency-associated conformations, and directly activate receptors on the cell surface [2–4]. In so doing, extracellular proteases trigger multiple signalling pathways that regulate cell adhesion and detachment, migration, invasion, proliferation and differentiation [5–7].

The study of the role of extracellular proteases in development and cancer has traditionally involved the matrix metalloproteases, the plasminogen activation system and the cathepsins [8–10]. However, the recent sequencing of multiple vertebrate genomes revealed a surprisingly vast array of novel candidate genes that are predicted to encode proteolytic enzymes. Of these, increasing attention is being paid to the previously unappreciated family of TTSPs (type II transmembrane serine proteases). All members of this novel gene family are predicted to encode transmembrane proteases that contain an N-proximal transmembrane domain and a C-terminal serine protease catalytic domain of the chymotrypsin (S1) fold, separated by an extracellular stem region that contains a diverse range of potential protein–protein interacting modules [11–14]. The ultimate size of the TTSP family in vertebrates is unknown, but more than 20 candidate TTSPs have been identified to date in humans and mice [1,12,13,15]. The capacity of most of the predicted TTSP genes to encode biologically relevant gene products is untested. However, initial studies of a subset of TTSP genes strongly suggest that the TTSP family directs the expression of functional transmembrane proteases having critical roles in development, tissue homoeostasis and human pathogenesis [12–14,16,17].

The matriptases constitute a subfamily of the TTSPs with a unique stem composition and phylogenetically related serine protease domains [12,13]. Well-characterized members of the matriptase subfamily include matriptase/MT-SP1, matriptase-2 and polyserase-1 [18–21]. Matriptase/MT-SP1, the prototypic member of the matriptase subfamily, was originally identified as a novel gelatinolytic activity in human breast cancer cells [22], and was subsequently shown to be overexpressed in a remarkably vast array of human tumours of epithelial origin. These include prostate, ovarian, uterine, colon, epithelial-type mesothelioma, cervical and head and neck squamous cell carcinoma [23–31]. Moreover, high expression of matriptase/MT-SP1 was reported to be an indicator of poor prognosis in ovarian and breast cancers [27,28]. We recently showed that mice deficient in matriptase/MT-SP1 die perinatally due to severe defects in epidermal development, directly implicating a member of the matriptase subfamily in epithelial genesis [32,33]. Mouse or human null

Abbreviations used: α1-PI, α1-proteinase inhibitor; AT III, antithrombin III; α2-AP, α2-antiplasmin; DESC, differentially expressed in squamous cell carcinoma; E-64, *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane; endo H, endoglycosidase H; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HA, haemagglutinin; HAI, hepatocyte growth factor activator inhibitor; HC II, heparin cofactor II; LDS, lithium dodecyl sulphate; α_2M, α_2-macroglobulin; PAI, plasminogen activator inhibitor; PCI, protein C inhibitor; Suc, 3-carboxy-propionyl; RT, reverse transcriptase; TTSP, type II transmembrane serine protease.

[1] To whom correspondence should be addressed (email thomas.bugge@nih.gov).

mutations in matriptase-2 and polyserase-1 have not yet been reported. However, the divergent patterns of expression of matriptase/MT-SP1 and matriptase-2 would indicate separate biological functions [20,34].

In the present study, we describe the identification, cloning and biological characterization of a new member of the matriptase subfamily that is present in fish, birds, rodents, canines and primates, which we term *matriptase-3*. We show that *matriptase-3* encodes a TTSP that has high amino acid identity but a widely divergent expression profile to matriptase/MT-SP1 and other members of the matriptase subfamily. Furthermore, we show that *matriptase-3* encodes a biologically active TTSP that is present on the cell surface and forms stable inhibitor complexes with a wide range of serpins.

EXPERIMENTAL

Bioinformatic analysis

A mouse matriptase/MT-SP1 homology search was performed by tBLASTn analysis of the NCBI nucleotide sequence database (http://www.ncbi.nlm.nih.gov/BLAST). Chromosome localization and a complete cDNA sequence of the mouse *matriptase-3* gene were identified by BLASTn search of a partial *matriptase-3* cDNA sequence against mouse genomic DNA using the Celera Discovery System mouse genome database (http://www.celeradiscoverysystem.com). Exon–intron boundaries of the mouse *matriptase-3* gene were predicted with Fuzzy Finder (http://www.genome.nci.nih.gov/tools). The sequences of human, chimpanzee, dog, rat, chicken, Danio zebrafish, spotted green pufferfish and tiger pufferfish matriptase-3 were identified by screening the NCBI, Celera Discovery System, Ensembl and MEROPS (http://merops.sanger.ac.uk) databases for proteins homologous with the predicted mouse matriptase-3 sequence. Domain structure and putative signal peptide sequences were analysed with SMART (http://www.smart.ox.ac.uk) and SignalP 2.0 (http://www.cbs.dtu.dk) software, and protein topology was predicted using the transmembrane hidden Markov model (TMHMM) 2.0 program (http://www.cbs.dtu.dk/services/TMHMM-2.0/). Potential N-glycosylation sites were analysed with NetNGlyc1.0 (http://www.cbs.dtu.dk). Multiple alignments of TTSP protein sequences were generated with Clustal W (version 1.7) and a phylogenetic tree was constructed using MEGA2.1 software (http://www.megasoftware.net).

Molecular cloning of the full-length mouse matriptase-3 cDNA

Total RNA was extracted from the skin and testis of an adult male 129/BlackSwiss mouse using TRIzol® reagent (Invitrogen, Carlsbad, CA, U.S.A.). cDNA was prepared from 1 μg of total RNA using BD SMART cDNA Synthesis kit (BD Biosciences–Clontech, Palo Alto, CA, U.S.A.). Gene-specific primers Mat3A (5′-ATGGATAAAGAAAAAAGTGATCC-3′) and Mat3B (5′-TACAGTTACAAAAGAGAAGGGAC-3′) were used to amplify the mouse matriptase-3 cDNA using the Expand Long DNA Template kit (Roche, Indianapolis, IN, U.S.A.) and a GeneAmp 9700 thermocycler (Applied Biosystems, Foster City, CA, U.S.A.) with 30 cycles of 10 s denaturation at 94 °C, 30 s annealing at 51 °C and 4 min elongation at 68 °C. The PCR product was cloned into the TOPO-TA vector (Invitrogen), and verified by complete sequencing of both DNA strands of the insert.

Analysis of matriptase-3 expression in human and mouse tissues

Primary mouse keratinocytes and fibroblasts were isolated from the epidermis of 1-day-old 129/BlackSwiss mice and cultured as described previously [35]. The cells were grown for 3 days at 37 °C in 5 % CO_2, and the total RNA from the cells was isolated using the TRIzol® reagent. Total RNA from adult mouse liver, brain, thymus, heart, lung, spleen, testis, ovary and kidney was purchased from Ambion (Austin, TX, U.S.A.). Total RNA from different sections of an adult mouse brain was obtained from Cemines (Evergreen, CO, U.S.A.). Total RNA from mouse skin and eye was extracted from tissues of 1-month-old male 129/BlackSwiss mice using the TRIzol® reagent. First-strand cDNA synthesis using an oligo(dT) primer was carried out with 1 μg of total RNA using the RETROscript kit (Ambion) according to the manufacturer's instructions. PCR was performed with Taq PCR Master Mix kit (Qiagen, Valencia, CA, U.S.A.) using gene-specific primers mat51 (5′-CTCATGTTGGTGACA-CTGAAGTCTCC-3′) and mat31 (5′-GGAAGATGCTGCTGT-TGCAGGCAGG-3′). Ribosomal protein S15 gene-specific primers S15fwd (5′-TTCCGCAAGTTCACCTACC-3′) and S15rev (5′-CGGGCCGGCCATGCTTTACG-3′) were used for positive control reactions. All PCRs were run for 35 cycles (94 °C for 1 min, 59 °C for 1 min and 72 °C for 1 min).

A mouse RNA Master Blot membrane containing an adult mouse cDNA array was purchased from BD Biosciences–Clontech, and FirstChoice™ Northern Human Blot 2 was from Ambion. The membranes were prehybridized for 1 h at 65 °C with ExpressHyb (BD Biosciences–Clontech) containing 100 μg/ml sheared salmon sperm DNA (Phoenix Biotechnologies, Huntsville, AL, U.S.A.). The PCR-amplified DNA (50–100 ng) corresponding to nt 967–1549 of the mouse *matriptase-3* open reading frames or nt 911–1359 of the human *matriptase-3* open reading frames was labelled with ^{32}P-dCTP using Ready-to-Go-dCTP labelling beads (Amersham Biosciences, Piscataway, NJ, U.S.A.). For hybridization, 1–2 $\times$ 10^6 c.p.m./ml of the probe was added directly to the prehybridization solution and incubated at 65 °C overnight. The blot was washed twice for 5 min with 2 $\times$ SSC/0.1 % SDS, twice for 15 min with 0.1 $\times$ SSC/0.1 % SDS, and exposed to a PhosphorImager screen (Molecular Dynamics, Sunnyvale, CA, U.S.A.). Quantification of the hybridization signal was performed with ImageQuant software (Amersham Biosciences).

Expression of mouse matriptase-3 in mammalian cells

COS7 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 8 % (v/v) fetal calf serum at 37 °C in 5 % CO_2. Cells (8×10^5) were seeded on 10 cm cell-culture dishes 1 day before transfection. To transfect the cells, 5 μg of plasmid DNA containing the full-length mouse *matriptase-3* cDNA cloned into the pMH expression vector (Roche) was mixed with Polyfectamine reagent (Qiagen) and was added to the cells following the manufacturer's instructions. The cells were washed twice with ice-cold PBS 24–32 h after transfection, and lysed using 0.3 ml of RIPA buffer [150 mM NaCl, 10 mM Tris/HCl, pH 7.4, 0.1 % SDS, 1 % Triton X-100, 0.5 % sodium deoxycholate, 5 mM EDTA and 1 $\times$ protease inhibitor cocktail (Sigma, St. Louis, MO, U.S.A.)]. The cell lysate was then clarified by centrifugation at 12 000 ***g*** at 4 °C, and the protein concentration was determined using RC DC Protein Assay reagents (Bio-Rad, Hercules, CA, U.S.A.).

Total protein (60–80 μg) was boiled for 5 min in 1 $\times$ NuPAGE LDS loading buffer [564 mM Tris base, 8 % LDS (lithium dodecyl sulphate), 40 % glycerol, 2.04 mM EDTA, 0.88 mM SERVA® Blue G250 and 0.7 mM Phenol Red, pH 8.5] (Invitrogen) containing 0.25 M 2-mercaptoethanol, separated on 4–12 % NuPAGE LDS gels (Invitrogen), and analysed by Western blotting using primary anti-HA (where HA stands for haemagglutinin) antibody (Covance Laboratories, Richmond, CA, U.S.A.) and

a secondary alkaline phosphatase-conjugated antibody (Dako-Cytomation, Carpinteria, CA, U.S.A.), both at a dilution of 1:1000.

Deglycosylation of the mouse matriptase-3 protein

Protein lysate (50 μg) was denatured for 10 min at 99 °C in 1× glycoprotein denaturing buffer (0.5 % SDS and 1 % 2-mercaptoethanol) (New England Biolabs, Beverly, MA, U.S.A.). Deglycosylation was performed in G5 endo H (endoglycosidase H) reaction buffer (New England Biolabs) containing 50 mM sodium citrate (pH 5.5), by adding 2000 units of endo H and incubating at 37 °C for 2 h. Samples were then denatured for 5 min at 99 °C in 1× NuPAGE LDS loading buffer containing 0.25 M 2-mercaptoethanol and analysed by Western blotting as indicated above.

Biotinylation of cell-surface proteins

COS7 cells transfected with either the full-length *matriptase-3* cDNA, a full-length mouse DESC1 (differentially expressed in squamous cell carcinoma 1) cDNA [15] or an empty pMH vector were cultured for 36 h in Dulbecco's modified Eagle's medium supplemented with 8 % fetal calf serum at 37 °C in 5 % CO_2. Cells were washed three times with ice-cold PBS and incubated with 2 mM membrane non-permeable EZ-Link Sulfo-NHS-Biotin (where NHS stands for *N*-hydroxysuccinimido) reagent (Pierce, Rockford, IL, U.S.A.) for 1 h at room temperature (15–20 °C). The biotin was aspirated, the cells were washed twice with PBS with 100 mM glycine to remove residual biotin, lysed in 0.3 ml of RIPA buffer, and clarified by centrifugation for 10 min at 10000 ***g***. The volume of supernatant corresponding to 500 μg of total protein was diluted with RIPA buffer to a final concentration of 1 μg/μl, and incubated with 100 μl of agarose-conjugated avidin (Sigma) for 4 h at 4 °C. The beads were then washed twice with 1 ml of RIPA buffer, twice with 1 ml of 150 mM NaCl, 10 mM Tris/HCl (pH 7.4) and 5 mM EDTA. Finally, the beads were resuspended in 100 μl of 4× NuPAGE LDS loading buffer with or without 1 M 2-mercaptoethanol and boiled for 5 min. After spinning, 40 μl of the supernatant was analysed by Western blotting using anti-HA antibody. Rabbit anti-mouse anti-GAPDH (glyceraldehyde-3-phosphate dehydrogenase) antibody (Santa Cruz Biotechnology, Santa Cruz, CA, U.S.A.) was used to detect GAPDH in cell lysates.

Expression and purification of the matriptase-3 serine protease domain

The region of the mouse *matriptase-3* cDNA encoding the serine protease domain (amino acids 581–829) was amplified by PCR using Mat3SPD5 (5′-AGCTGCATGCTAGGCTGTAGCAGG-AGTTCTTCC-3′) and Mat3SPD3 (5′-TAGACTCGAGCAAAA-GAGAAGGGACATACTTATG-3′) primers containing SphI and XhoI restriction sites respectively, and was cloned into the pMIB/V5HisA expression vector (Invitrogen) in frame with an N-terminal honeybee mellitin secretion signal peptide and a C-terminal V5 epitope. The construct was verified by DNA sequencing. *Spodoptera frugiperda* Sf9 cells were grown in Excell 420 SFM medium (JRH Biosciences, Lenexa, KA, U.S.A.) at 27 °C in a 5 % CO_2-buffered humidified incubator. The cells were transfected with pMIB/V5HisA/Mat3 or a control plasmid DNA (5 μg of plasmid DNA/35 mm plastic dish), using Celfectin reagent (Invitrogen). To select for stably transfected insect cell lines, Blastcidin (50 μg/ml; Invitrogen) was added 48 h after transfection. Conditioned medium was collected from Blastcidin-resistant Sf9 cells, and sodium phosphate buffer (pH 7.5) and NaCl were added to a final concentration of 20 mM and 0.5 M respectively. The solution was filtered through a 0.22 μm filter (Corning, Corning, NY, U.S.A.) and applied to a Co^{2+}-charged HiTrap chelating column (Amersham Biosciences) using the Pharmacia AKTA FPLC system. Chelated proteins were eluted with a linear gradient of 20–200 mM imidazole, and the positive fractions were identified by Western blotting using a 1:5000 dilution of the anti-V5 antibody (Invitrogen). The positive fractions were dialysed overnight at 4 °C against 200 vol. of 50 mM Tris/HCl (pH 8.0), 1 mM $CaCl_2$ and 0.1 % Tween 20 using Slide-A-Lyzer® dialysis cassettes (Pierce). To investigate the proteolytic activity of mouse matriptase-3, 80 μg of gelatin obtained from bovine skin (Sigma), heat-denatured casein (Safeway, Tempe, AZ, U.S.A.) or heat-denatured BSA (Sigma) was incubated for 6 h at 37 °C with 50 nM purified, activated matriptase-3 or with 100 nM trypsin (Sigma) in 50 mM Tris/HCl (pH 8.0), 1 mM $CaCl_2$ and 0.1 % Tween 20 reaction buffer, then separated on a 4–12 % NuPAGE LDS gel and stained with 0.1 % Coomassie Brilliant Blue R250 (Sigma) in 30 % (v/v) methanol and 10 % (v/v) acetic acid.

Mouse matriptase-3–inhibitor complex formation

Purified, active mouse matriptase-3 serine protease domain (30 nM) was incubated for 30 min in 50 mM Tris/HCl (pH 7.5), 150 mM NaCl and 2.5 mM $CaCl_2$ buffer at room temperature with 30 nM human α_2M (α_2-macroglobulin; Sigma) or 100 nM human serpins. Where indicated, reactions also included 20 μg/ml heparin (Sigma). Experiments with α_2M included control reaction mixtures preincubated for 5 min with 1 mM of the serine protease inhibitor PMSF (Sigma) before the complex formation assay. Purified serpins were as follows: PAI-1 (plasminogen activator inhibitor-1; Molecular Innovations, Southfield, MI, U.S.A.); PCI (protein C inhibitor; R&D Systems, Minneapolis, MN, U.S.A.); HAI-1 (hepatocyte growth factor activator inhibitor; R&D Systems); α2-AP (α2-antiplasmin; Calbiochem, San Diego, CA, U.S.A.), AT III (antithrombin III; ICN Biomedicals, Irvine, CA, U.S.A.); α1-PI (α1-proteinase inhibitor; a gift from Annemarie Ralston, American Red Cross, Rockiville, MD, U.S.A.); and HC II (heparin cofactor II) [36]. PAI-1 mutant serpins $PAI\text{-}1_A$ [R346A ($Arg^{346} \rightarrow Ala$)] and $PAI\text{-}1_R$ (T333R/A335R) were gifts from Dr D. A. Lawrence (University of Michigan, Ann Arbor, MI, U.S.A.). Complex formation was analysed by SDS/PAGE and Western blotting using anti-V5 antibody.

To evaluate the effectiveness of different serpin inhibitors towards matriptase-3, 10 nM purified recombinant enzyme was incubated in triplicate with 200 μM Suc-Ala-Ala-Pro-Arg-pNA (where Suc stands for 3-carboxy-propionyl) chromogenic peptide in the presence of 100, 50, 20, 10, 5, 2 and 0 nM purified recombinant PAI-1, PCI, α2-AP, AT III or α1-PI.

Chromogenic peptide hydrolysis assay

The peptides Suc-Ala-Ala-Pro-Val-pNA, Suc-Ala-Ala-Pro-Phe-pNA, Suc-Ala-Ala-Pro-Lys-pNA, Suc-Ala-Ala-Pro-Arg-pNA, H-Glu-Gly-Arg-pNA, Z-Gly-Gly-Arg-pNA and Z-Phe-Val-Arg-pNA (where Z stands for benzyloxycarbonyl) were obtained from Bachem Bioscience (King of Prussia, PA, U.S.A.). PMSF, aprotinin, leupeptin, E-64 [*trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane] and egg white trypsin inhibitor (ovomucoid) were obtained from Sigma, human recombinant PAI-1 was from Molecular Innovations and mouse recombinant HAI-1 was from R&D Systems. Recombinant active mouse matriptase-3 (10 nM) or thrombin (10 nM) was incubated with 100 μM chromogenic peptide substrate in 50 mM Tris/HCl (pH 8.0), 150 mM NaCl, 0.01 % Tween 20 and 0.01 % BSA buffer at 35 °C over a 20 min period. Changes in absorbance A at 405 nm were monitored on a PerkinElmer HTS 7000 Bio Assay reader. To

test the inhibitory profile of mouse matriptase-3, hydrolysis of the Suc-Ala-Ala-Pro-Arg-pNA substrate was tested also in the presence of 1 mM PMSF, 3 μM aprotinin, 20 μM leupeptin, 30 μM E-64, 100 nM egg white trypsin inhibitor, 2.5 mM EDTA, 50 nM PAI-1 or 50 nM HAI-1.

For the determination of the K_m of chromogenic substrate cleavage by the matriptase-3 catalytic domain, the purified enzyme (10 nM) was incubated with 500, 200, 100, 50, 20 and 10 μM Suc-Ala-Ala-Pro-Lys-pNA and Suc-Ala-Ala-Pro-Arg-pNA chromogenic peptides. The K_m values for both peptides were determined from the double-reciprocal plot from triplicate measurements.

RESULTS

Identification of a putative new membrane-anchored serine protease

To identify putative homologues of *matriptase*/*MT-SP1* in vertebrate genomes, the NCBI nucleotide sequence database was screened against the mouse matriptase/MT-SP1 protein sequence using tBLASTn. This identified two non-overlapping partial cDNA sequences (now updated in NCBI under accession no. NM_172455), putatively encoding parts of a novel matriptase/MT-SP1-like protein. Both cDNA fragments mapped to the 16B5 mouse chromosomal region, and a detailed *in silico* analysis of this region revealed the presence of an as yet unidentified candidate gene predicted to consist of 18 exons and spanning almost 36 kb of genomic DNA (Figure 1A). Splicing of the primary transcript would give rise to a 2490 bp open reading frame with the capacity to encode a protein with a molecular mass of 92.6 kDa (Figure 1B). The sequence of the predicted new protein is most similar to the previously identified mouse matriptase/MT-SP1 and matriptase-2 proteins, displaying 31 % amino acid identity to both proteins. Therefore we named the new gene *matriptase-3*. The predicted mouse matriptase-3 protein has a single transmembrane domain (amino acids 63–85) and a type II membrane topology, as determined by the transmembrane hidden Markov model. Domain structure analysis of the predicted amino acid sequence indicated that matriptase-3 has a structure that is typical of a TTSP, with a short N-terminal cytosolic domain, an N-proximal transmembrane domain – an extracellular SEA domain, two CUB domains, three LDLRa (low-density lipoprotein receptor domain class A) repeats – and a C-terminal serine protease domain of the chymotrypsin (S1) fold (Figure 1A). A search for matriptase-3-like proteases in other species revealed the presence of highly homologous predicted protein sequences in the human (GenBank® accession no. XM_293599), chimpanzee (GenBank® accession no. XM_516646), dog (GenBank® accession no. XM_545095), rat (GenBank® accession no. XM_221464), chicken (GenBank® accession no. XM_416635), zebrafish (Ensembl accession no. ENSDARG00000012004), spotted green pufferfish (Ensembl accession no. GSTENG-00023863001) and tiger pufferfish (Ensembl accession no. SINFRUG00000149004) genomes. A comparison of the deduced serine protease domain amino acid sequences of matriptase-3 from these species to matriptase/MT-SP1 and matriptase-2 is shown in Figure 2(A). The human, mouse, rat and chicken genes all localized to syntenic regions of human chromosome 3q13.2, mouse chromosome 16B5, rat chromosome 11q21 and chicken chromosome 1, further supporting that all four genes are true orthologues.

Analysis of the amino acid sequence of the putative mouse matriptase-3 protein revealed five potential N-glycosylation sites at residues 196, 401, 465, 512 and 538 (Figures 1A and 1B). The catalytic domain of mouse matriptase-3 has all the known essential features of a functional serine protease (Figure 2A), including (i) the catalytic triad residues His^{632}-Asp^{680}-Ser^{776}, all within their consensus motifs, (ii) the residues Asp^{770}, Gly^{794} and Gly^{809} forming the substrate pocket, (iii) eight cysteine residues predicted to form four disulphide bridges essential for proper folding, as well as (iv) a conserved activation cleavage site R^{591}IVGG that is required for the conversion of an inactive zymogen into an active enzyme. These findings indicate that the *matriptase-3* gene has the potential to encode a functional serine protease. Constructing a neighbour-joining phylogenetic tree of the serine protease domains of all known mouse members of the TTSP family clearly placed matriptase-3 within the matriptase subfamily of TTSPs (Figure 2B) and it suggests that matriptase-3 may be the closest homologue of matriptase/MT-SP1 identified to date.

Human and mouse matriptase-3 display a similar tissue expression pattern

To determine whether *matriptase-3* is an actively transcribed gene, an RT (reverse transcriptase)–PCR analysis was performed on total RNA extracted from a series of adult mouse tissues (Figure 3A). This analysis revealed that *matriptase-3* mRNA was expressed in the mouse brain, testis, skin and eye (Figure 3A, lanes 2, 7, 10 and 11 respectively). A complementary dot-blot hybridization analysis of *matriptase-3* expression confirmed the expression of *matriptase-3* in the brain, eye and testis. In addition, high *matriptase-3* expression was detected in the epididymis and salivary gland, and lower levels of expression were detected in the heart, skeletal muscle, thymus, ovary, prostate and uterus (Figure 3B). To explore further the expression of *matriptase-3* in the central nervous system, the RT–PCR analysis was extended using RNA isolated from different brain structures. The results suggest a uniform expression of *matriptase-3* in all the analysed parts of the brain (Figure 3C). The critical role of *matriptase*/*MT-SP1* in keratinocyte differentiation [32,33] prompted further analysis of *matriptase-3* expression in the skin. Interestingly, although amply expressed in whole skin (Figure 3A, lane 10, and Figure 3D, lane 1), no *matriptase-3* transcript could be detected in either cultured primary fibroblasts or cultured primary keratinocytes (Figure 3D, lanes 2 and 3 respectively). This suggests that *matriptase-3* is expressed in a different cell population within the mouse skin, or that the cultivation of the two cell types *in vitro* does not induce expression of the novel protease.

To compare the expression of *matriptase-3* in humans and mice, an extensive survey of *matriptase-3* mRNA expression in human tissues was carried out by an RT–PCR analysis. As in the mouse, *matriptase-3* transcripts were present in the brain, ovary, testis and salivary gland (Figure 4A, lanes, 3, 11, 17 and 21 respectively). *Matriptase-3* mRNA was also expressed in the human trachea (not analysed in the mouse), and small amounts of mRNA were present in human lung tissue (Figure 4A, lanes 20 and 10 respectively). A complementary Northern-blot analysis of mRNA from human tissues revealed a specific hybridization signal only in the testis (Figure 4B, lane 9). The approx. 2.5 kb size of the major hybridizing transcript in the testis is in excellent agreement with the predicted human *matriptase-3* mRNA transcript size of 2553 bp.

Taken together, the analysis shows that *matriptase-3* displays an evolutionarily conserved, tissue-restricted pattern of expression that differs from the rather unrestricted expression of *matriptase*/*MT-SP1* in epithelium-containing tissues [37], the liver-specific expression of *matriptase-2* [20], and the expression of *polyserase-1* restricted to skeletal muscle, liver, placenta and heart [21].

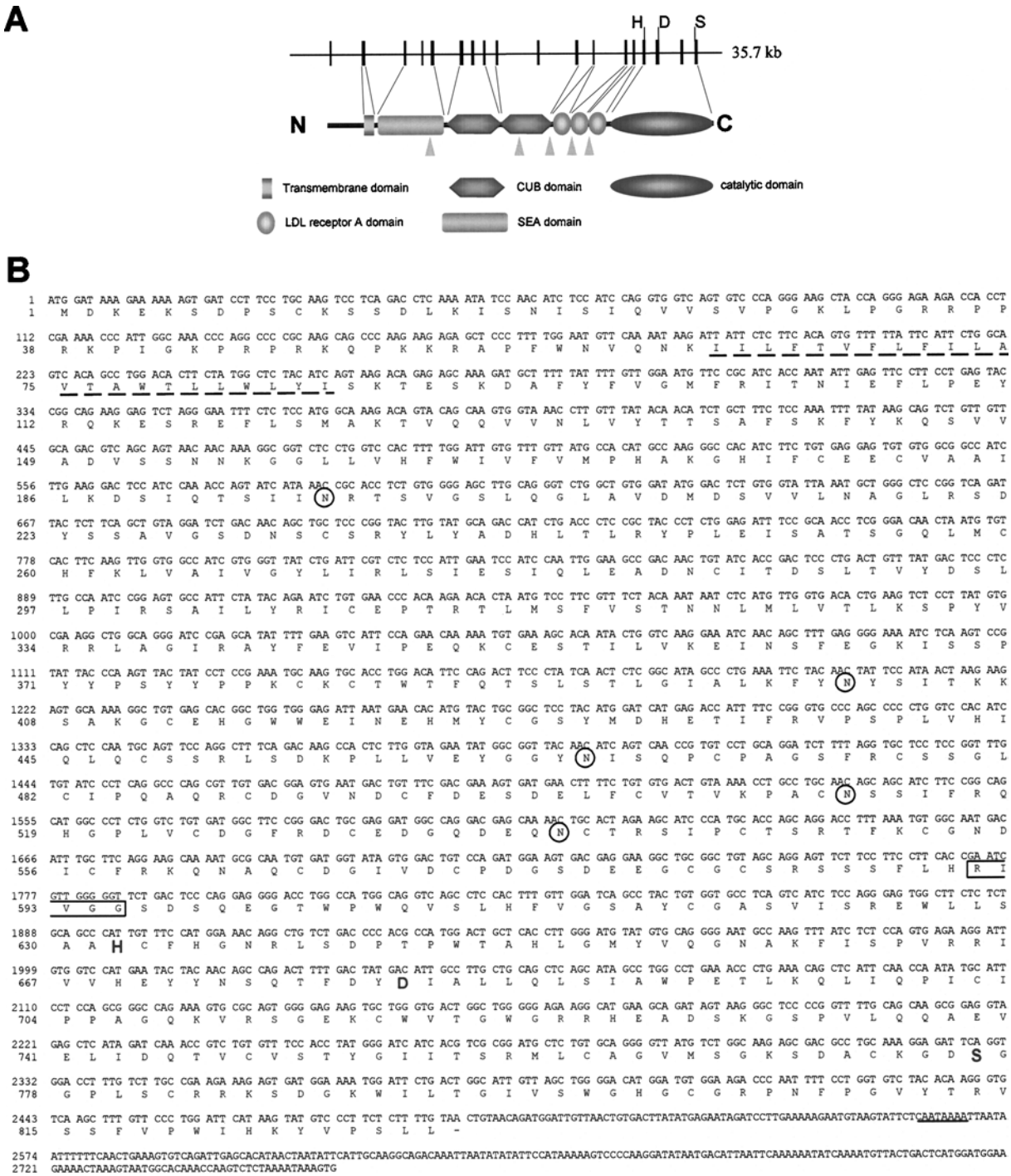

Figure 1 Gene organization, protein domain structure, nucleotides and amino acid sequence of mouse *matriptase-3*

(**A**) Exon–intron structure of the mouse *matriptase-3* gene and domain architecture of the deduced protein. The locations of the five potential N-glycosylation sites are indicated by arrowheads. The exons encoding the catalytic residues His^{632} (H), Asp^{680} (D) and Ser^{776} (S) are shown. (**B**) Nucleotide sequence of the mouse *matriptase-3* cDNA, as predicted by bioinformatic analysis and verified by direct molecular cloning, and the deduced amino acid sequence of the encoded protein. Amino acid residues encoding the predicted transmembrane domain are underlined, the activation cleavage site is boxed, the residues constituting the catalytic triad are highlighted in boldface and enlarged font, the potential N-glycosylation sites are encircled, and the predicted polyadenylation signal sequence in the 3′-untranslated region is underlined.

Matriptase-3 directs the expression of an N-glycosylated cell-surface protein

To characterize the protein product produced by the *matriptase-3* gene, a full-length mouse *matriptase-3* cDNA was generated by high-fidelity RT–PCR from a mouse newborn skin cDNA library. DNA sequencing of the cloned RT–PCR product confirmed the cDNA identity and exon–intron structure of the *matriptase-3* gene that was predicted by the *in silico* analysis (results not shown). The full-length cDNA was then inserted into the expression vector

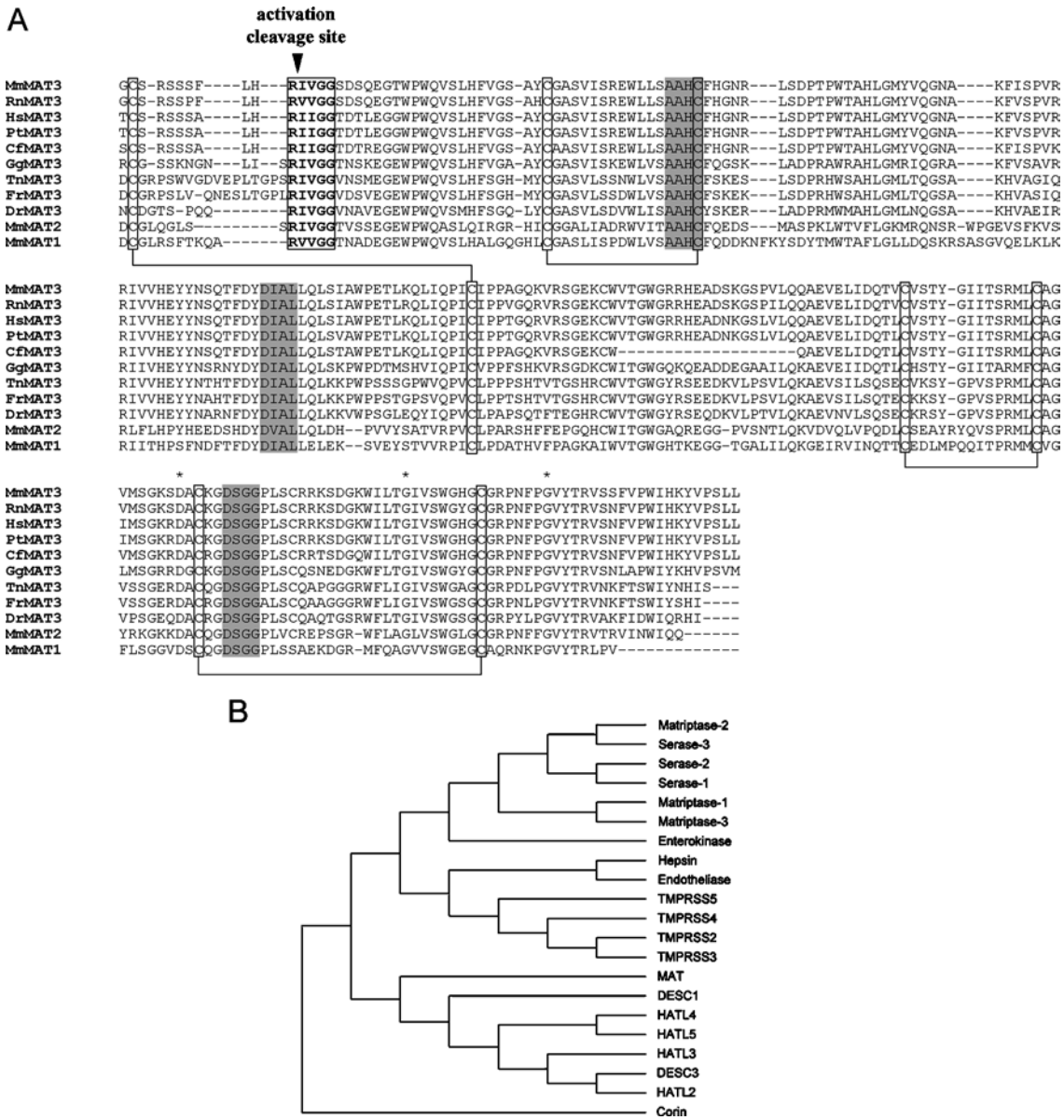

Figure 2 Phylogenetic relationship of matriptase-3 proteins to other type II transmembrane proteases

(**A**) The predicted amino acid sequences of the serine protease domains of mouse (MmMAT3), rat (RnMAT3), human (HsMAT3), chimpanzee (PtMAT3), dog (CfMAT3), chicken (GgMAT3), tiger pufferfish (TnMAT3), spotted green pufferfish (FrMAt3), Danio zebrafish (DrMAT3) matriptase-3, mouse matriptase/MT-SP1 (MmMAT1) and mouse matriptase-2 (MmMAT2) were aligned with the Clustal W program. Amino acids essential for serine protease activity are highlighted. The activation cleavage site is boxed and highlighted in boldface, disulphide bridges between the eight conserved Cys residues (boxed) are indicated by solid lines, catalytic residues within their canonical motifs are highlighted in grey, and residues forming the substrate-binding site are indicated by asterisks. (**B**) Phylogenetic tree of the 19 known murine members of the TTSP family. Serase-1, -2 and -3 represent individual serine protease subunits of polyserase-1 [21]. The amino acid sequences of the catalytic domains were aligned with the Clustal W program using default settings. The phylogenetic tree was constructed with MEGA2.1 software using the neighbour-joining method and was tested for validity by 1000 bootstrap repetitions.

pMH to allow expression of C-terminally HA-tagged matriptase-3 in mammalian cells. COS7 cells transfected with the *matriptase-3* expression vector (Figure 5A, lane 3), but not those transfected with the empty vector (Figure 5A, lane 1), showed the expression of an approx. 90 kDa HA-immunoreactive protein. Treatment of the matriptase-3–HA cell lysate with endo H decreased the molecular mass of the HA-immunoreactive protein to approx. 80 kDa (Figure 5A, lane 4), demonstrating that matriptase-3 is an N-glycosylated protein. Deglycosylation of matriptase-3 yielded a protein with an apparent molecular mass that was substantially smaller than the predicted molecular mass of 92.6 kDa. This would suggest that the novel protease either exhibits an abnormal migration on SDS/polyacrylamide gels or is being proteolytically processed after synthesis similar to matriptase/MT-SP1 [38]. To distinguish between the two possibilities, we compared the migration of matriptase-3 proteins furnished with either N- or C-terminal HA epitope tags after expression in COS-7 cells and analysis of cell extracts by Western blotting. This analysis showed that N- and C-terminal HA-tagged matriptase-3 migrates identically in SDS/polyacrylamide gels, suggesting that matriptase-3 does not undergo any proteolytic processing, but rather displays an abnormal migration in SDS/polyacrylamide gels (Figure 5B, lanes 2 and 3). *In vitro* transcription/translation of the *matriptase-3* cDNA using reticulocyte lysates

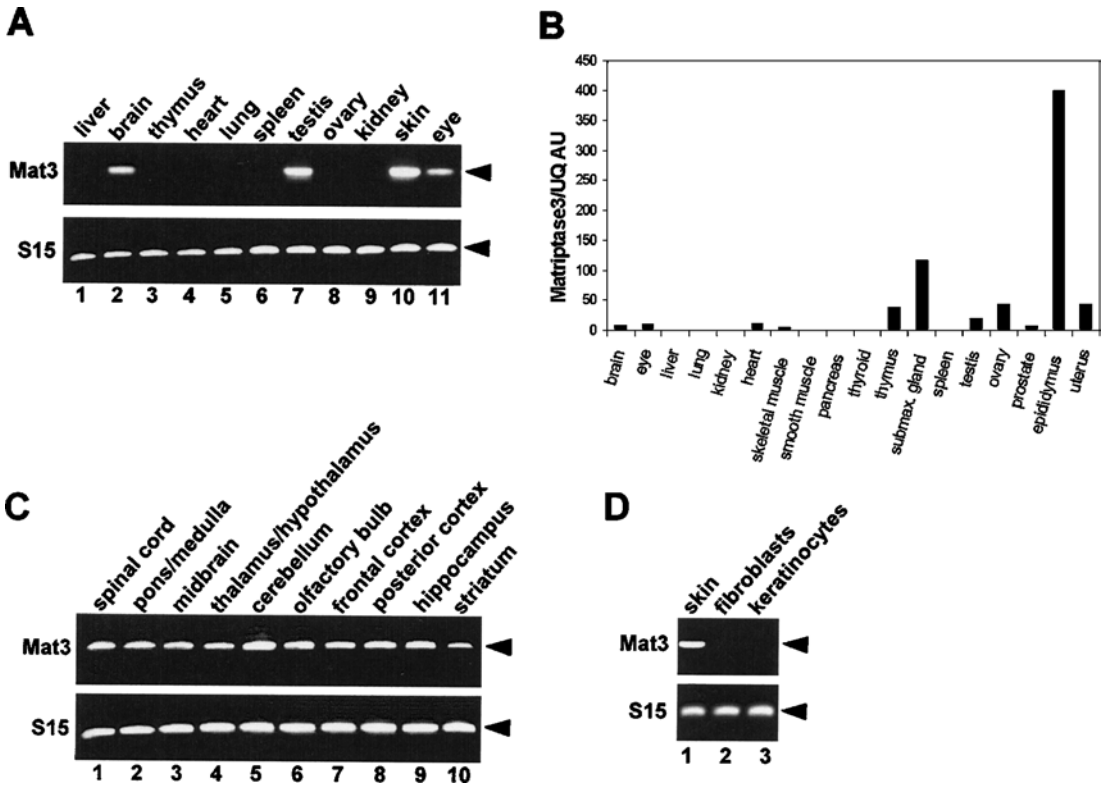

Figure 3 Matriptase-3 mRNA expression in the mouse

(**A**) RT–PCR analysis demonstrating *matriptase-3* mRNA in brain, testis, skin and eye of adult mice. RNA was isolated from mouse liver (lane 1), brain (lane 2), thymus (lane 3), heart (lane 4), lung (lane 5), spleen (lane 6), testis (lane 7), ovary (lane 8), kidney (lane 9), skin (lane 10) and eye (lane 11), reverse transcribed, amplified by PCR with a mouse *matriptase-3*-specific primer pair (Mat3, upper panel) or a ribosomal protein S15 specific primer pair (S15, lower panel), and the amplification products were analysed by agarose gel electrophoresis. (**B**) Analysis of mouse *matriptase-3* expression by dot-blot hybridization confirms the expression of *matriptase-3* mRNA in brain, testis and eye, reveal high expression of *matriptase-3* mRNA in the epididymis and salivary gland, and lower expression in the heart, skeletal muscle, thymus, ovary, prostate and uterus. The relative ratio of matriptase-3 expression to ubiquitin (UQ) expression is shown. Au, arbitrary units. (**C**) RT–PCR analysis of RNA isolated from spinal cord (lane 1), pons/medulla (lane 2), midbrain (lane 3), thalamus/hypothalamus (lane 4), cerebellum (lane 5), olfactory bulb (lane 6), frontal cortex (lane 7), posterior cortex (lane 8), hippocampus (lane 9), and striatum (lane 10) shows expression of *matriptase-3* mRNA in all parts of the mouse brain. (**D**) RT–PCR analysis of RNA isolated from whole mouse skin (lane 1), primary mouse fibroblasts (lane 2) and primary mouse keratinocytes (lane 3) reveals the expression of *matriptase-3* mRNA in whole mouse skin, but not in cultured keratinocytes or fibroblasts.

yielded a protein with a molecular mass of approx. 80 kDa, even in the presence of a protease inhibitor cocktail, further supporting this conclusion (results not shown).

Analysis of the mouse and human matriptase-3 protein sequences revealed the presence of a single putative transmembrane domain, indicating that the matriptase-3 protein, similar to previously characterized members of the TTSP family, may be located on the cell surface. To investigate the subcellular localization of mouse matriptase-3, COS7 cells expressing the HA-tagged full-length matriptase-3 were labelled with a cell-membrane non-permeable biotinylation agent. Proteins in the cell lysate were then separated into biotinylated and non-biotinylated fractions by avidin–agarose precipitation and analysed by Western blotting. This analysis showed that a fraction of matriptase-3 was biotinylated by the treatment (Figure 5C, top panel). The fraction of biotinylated matriptase-3 was proportionally similar to the fraction of biotinylated DESC1, which was analysed in parallel (Figure 5C, middle panel) and served as a structurally related control protein with validated transmembrane topology [15]. In contrast, the cytosolic protein GAPDH was not labelled by the treatment, showing that the biotinylation of matriptase-3 was not caused by the biotinylation agent permeating the plasma membrane (Figure 5C, bottom panel).

Western-blot analysis of the biotinylated fraction of matriptase-3 performed under reducing and non-reducing conditions revealed only one immunoreactive protein corresponding to the predicted size of the full-length matriptase-3 (results not shown). This indicates that, at least on the surface of COS7 cells, matriptase-3 remains in its unprocessed, zymogen form.

In conclusion, these experiments demonstrate that the *matriptase-3* gene directs the synthesis of a 90 kDa transmembrane glycoprotein that is expressed on the cell surface.

Matriptase-3 is a catalytically active serine protease that can cleave macromolecular substrates

To investigate the enzymatic properties of the new enzyme, the serine protease domain, including the activation cleavage site, was furnished with a C-terminal V5 epitope tag and a polyhistidine tract to facilitate immunological detection and purification by metal ion-chelation chromatography respectively. The chimaeric protein was expressed in insect cells and purified to homogeneity from conditioned medium. The recombinant matriptase-3 serine protease domain produced this way would be a single-chain proenzyme that typically requires proteolytic cleavage within a highly conserved activation motif (R^{591}IVGG) to become active. Similar to matriptase/MT-SP1 [18], matriptase-2 [20] and TMPRSS2 [39], the matriptase-3 zymogen appeared to undergo activation site cleavage during the purification process, as evidenced by a slight increase in the electrophoretic mobility (Figure 6A). The

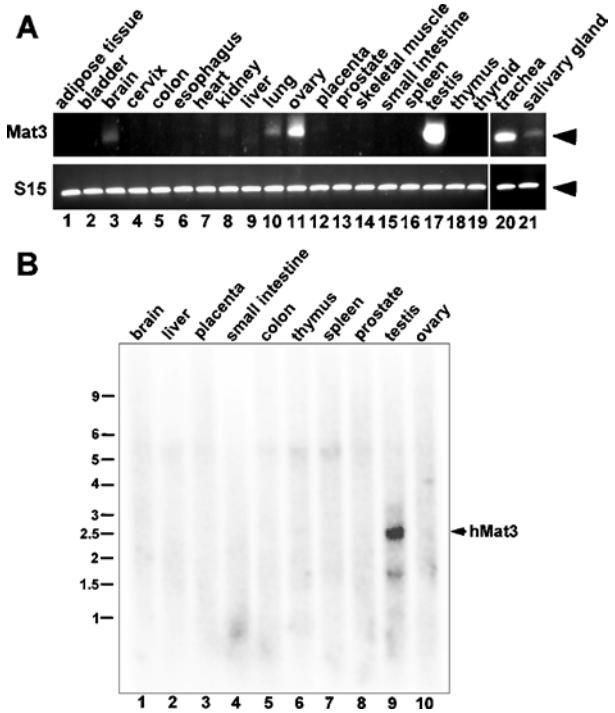

Figure 4 Analysis of *matriptase-3* mRNA expression in human tissues shows evolutionarily conserved expression in mice and humans

(**A**) RT–PCR analysis of human tissue samples reveals *matriptase-3* mRNA expression in brain, lung, ovary, testis, trachea and salivary gland. RNA from human adipose tissue (lane 1), bladder (lane 2), brain (lane 3), cervix (lane 4), colon (lane 5), oesophagus (lane 6), heart (lane 7), kidney (lane 8), liver (lane 9), lung (lane 10), ovary (lane 11), placenta (lane 12), prostate (lane 13), skeletal muscle (lane 14), small intestine (lane 15), spleen (lane 16), testis (lane 17), thymus (lane 18), thyroid (lane 19), trachea (lane 20) and salivary gland (lane 21) was reverse-transcribed, amplified by PCR with a human *matriptase-3*-specific primer pair (Mat3, upper panel) or a human ribosomal protein S15 specific primer pair (S15, lower panel), and the amplification products were analysed by agarose gel electrophoresis. (**B**) Northern-blot analysis of *matriptase-3* expression in human tissues. Polyadenylated RNA (2 μg/lane) from tissue samples of human brain (lane 1), liver (lane 2), placenta (lane 3), small intestine (lane 4), colon (lane 5), thymus (lane 6), spleen (lane 7), prostate (lane 8), testis (lane 9) and ovary (lane 10) was separated by agarose gel electrophoresis, immobilized to a membrane and probed with a ^{32}P-labelled human *matriptase-3* cDNA probe that contained nt 464–980 of the predicted open reading frame. The positions of RNA markers (kb) are indicated on the left.

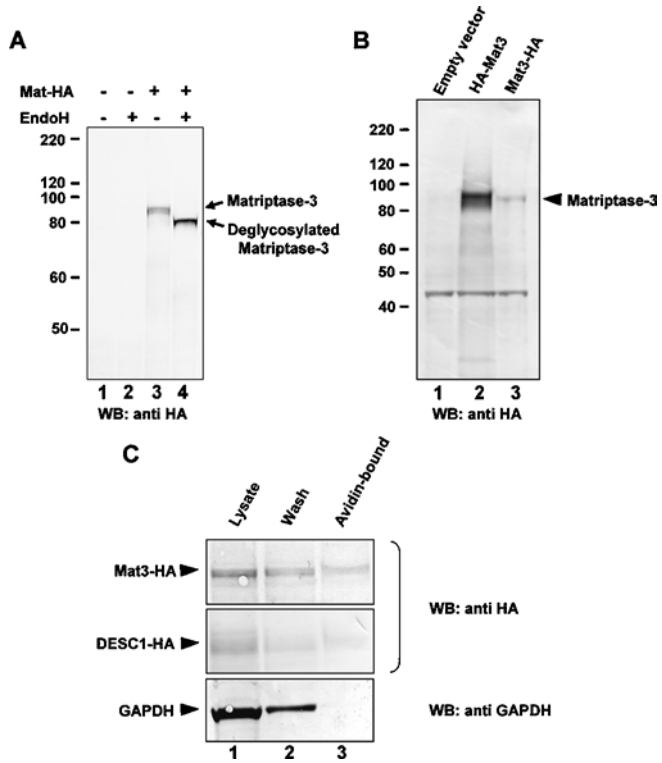

Figure 5 Mouse matriptase-3 is a 90 kDa transmembrane glycoprotein that is expressed on the cell surface

(**A**) N-glycosylation of mouse matriptase-3. Western-blot analysis with HA antibodies of whole cell lysates of COS7 cells transfected with empty expression vector (lanes 1 and 2) or a full-length, C-terminally HA-tagged mouse *matriptase-3* cDNA expression plasmid (Mat3–HA; lanes 3 and 4). The protein extracts in lanes 2 and 4 were treated with endo H before SDS/PAGE, and the protein extracts in lanes 1 and 3 were treated with deglycosylation buffer (50 mM sodium citrate, pH 5.5) without endo H. The upper arrow indicates the position of the N-glycosylated form of mouse matriptase-3 (90 kDa), and the lower arrow indicates the position of the N-deglycosylated mouse matriptase-3 protein (80 kDa). The positions of molecular-mass standards (kDa) are indicated on the left. (**B**) The abnormal migration of mouse matriptase-3 in SDS/polyacrylamide gels is not caused by proteolytic processing. Whole cell lysates of COS7 cells transfected with empty expression vector (lane 1), a full-length, N-terminally HA-tagged mouse *matriptase-3* cDNA expression plasmid (HA–Mat3; lane 2), or a full-length, C-terminally HA-tagged mouse *matriptase-3* cDNA expression plasmid (Mat3–HA) were analysed by Western blotting using HA antibodies. The arrow indicates the position of N- and C-terminally tagged mouse matriptase-3 (90 kDa). The positions of molecular-mass standards (kDa) are indicated on the left. (**C**) Mouse matriptase-3 is expressed on the plasma membrane. COS7 cells transfected with a full-length, C-terminally HA-tagged mouse *matriptase-3* cDNA (Mat3–HA; top panel) or with a full-length, C-terminally HA-tagged mouse *DESC1* cDNA (DESC1–HA; middle panel) were treated with Sulfo-NHS-Biotin to biotinylate cell-surface proteins. Whole cell lysates were prepared and either analysed directly by Western blotting using HA antibodies (lysate, lane 1) or separated into non-biotinylated proteins (wash, lane 2) and biotinylated proteins (avidin-bound, lane 3) by avidin–agarose batch affinity chromatography. Bottom panel: lysates from mouse *matriptase-3* cDNA expression plasmid-transfected cells were analysed by Western blotting against the cytoplasmic protein GAPDH. The absence of biotinylated GAPDH shows that the biotinylation agent did not penetrate the plasma membrane.

mobility shift typically occurred after dialysis of the purified matriptase-3 was performed to remove the imidazole used for protein elution after metal ion-chelation chromatography, and was time- and temperature-dependent (results not shown). To determine whether the activation site cleavage resulted in the formation of a catalytically active serine protease, an α_2M capture assay was employed [40,41]. Incubation of the matriptase-3 serine protease domain with α_2M after, but not before, the activation site cleavage did indeed result in the formation of multiple matriptase-3-containing high molecular mass complexes, consisting of matriptase-3 covalently linked to 90 and 180 kDa α_2M subunits (Figure 6B, lane 5). Several additional high molecular mass complexes were apparent, probably caused by partial cleavage by uncomplexed matriptase-3 serine protease domains. Preincubation of the activated matriptase-3 with the active-site serine protease inhibitor PMSF prevented the formation of these higher molecular mass complexes (Figure 6B, lane 6), showing that matriptase-3–α_2M complex formation was dependent on matriptase-3 serine protease activity. To document further the activity of the recombinant protein, the purified activated matriptase-3 catalytic domain was incubated with denatured collagen (gelatin), casein or BSA, and the resulting reactions were then analysed by SDS/PAGE. As shown in Figure 6(C, lanes 2, 5 and 8), the matriptase-3 serine protease domain clearly exhibited proteolytic activity against all three proteins, providing additional evidence that mouse matriptase-3 is a functional serine protease.

Mouse matriptase-3 preferentially hydrolyses peptides with Arg at position P_1

To investigate the substrate preference of matriptase-3, the recombinant protein was incubated with different synthetic chromogenic peptides. As shown in Figure 7(A), matriptase-3 showed a

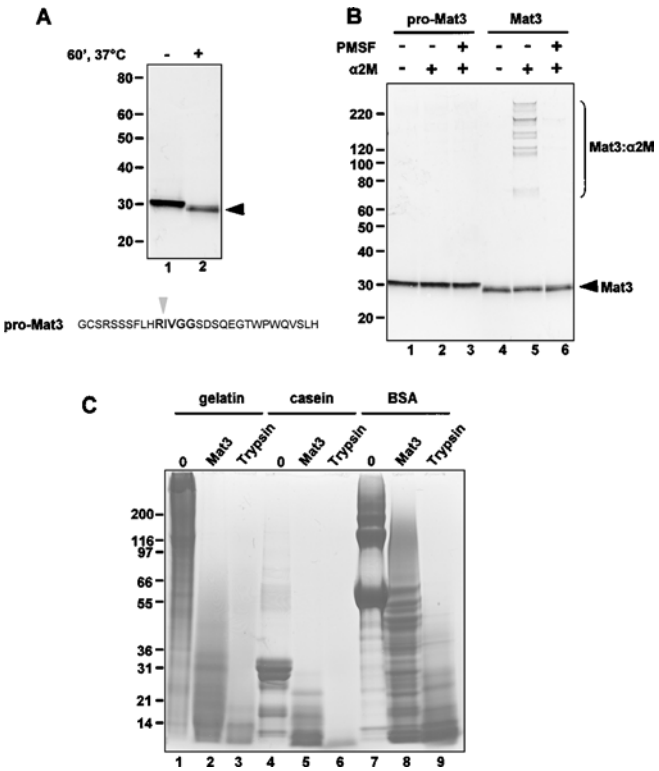

Figure 6 Mouse matriptase-3 is an active serine protease with gelatinolytic and caseinolytic activity

(**A**) Top: C-terminally V5-tagged mouse matriptase-3 serine protease domain was purified to homogeneity from insect cell conditioned medium and analysed by Western blotting using anti-V5 antibodies before (lane 1) and after (lane 2) incubation for 60 min at 37 °C in 50 mM Tris/HCl (pH 8.0), 1 mM $CaCl_2$ and 0.1 % Tween 20. The position of the activation site-cleaved mouse matriptase-3 serine protease domain is indicated. The positions of molecular-mass standards (kDa) are indicated on the left. Bottom: amino acid sequence of the N-terminus of the recombinant matriptase-3 zymogen (first 30 amino acids are shown). Amino acids forming the predicted activation cleavage site are highlighted in boldface, and the P_1–P_1' bond is indicated with an arrowhead. (**B**) Activated mouse matriptase-3 serine protease domain cleaves and forms SDS-stable complexes with α_2M. Purified inactive (lanes 1–3) or activated (lanes 4–6) mouse matriptase-3 serine protease domain was pretreated with PMSF (lanes 3 and 6) or solvent (lanes 1, 2, 4 and 5) and incubated with α_2M (lanes 2, 3, 5 and 6) or buffer (lanes 1 and 4). Complex formation was analysed by Western blotting using anti-V5 antibodies under reducing conditions. The positions of mouse matriptase-3–α_2M complexes (Mat3:α_2M) and the uncomplexed activated mouse matriptase-3 serine protease domain (Mat3) are indicated. The position of molecular-mass standards (kDa) are shown on the left. (**C**) Matriptase-3 shows proteolytic activity against denatured collagens, casein and BSA. Gelatin (80 μg; lanes 1–3), denatured casein (lanes 4–6) or denatured BSA (lanes 7–9) were incubated with the buffer (lanes 1, 4 and 7), the purified activated mouse matriptase-3 serine protease domain (lanes 2, 5 and 8) or trypsin (lanes 3, 6 and 9). All reactions were analysed on SDS/PAGE followed by a Coomassie Brilliant Blue staining to visualize substrate cleavage.

strong preference for peptides containing Arg at position P_1, with the peptides Suc-Ala-Ala-Pro-Arg-pNA and Z-Phe-Val-Arg-pNA showing the highest rate of conversion. Lys at position P_1 was significantly less effective, and no activity has been detected when this position was occupied by either Val (preferred by leucocyte elastase) or Phe (preferred by chymotrypsin). The preference for Arg at position P_1 is consistent with the primary structure of the catalytic domain of matriptase-3. The P_1 specificity of serine proteases has been correlated with the amino acid residue positioned six amino acids N-terminal of the active-site serine residue. A negatively charged Asp^{770} at this position indicates preference for substrates with positively charged Arg or Lys at position P_1, whereas the following Ala^{771} would favour P_1-Arg over Lys. A more detailed analysis of the kinetic properties of matriptase-3 confirmed that the recombinant enzyme exhibited significantly

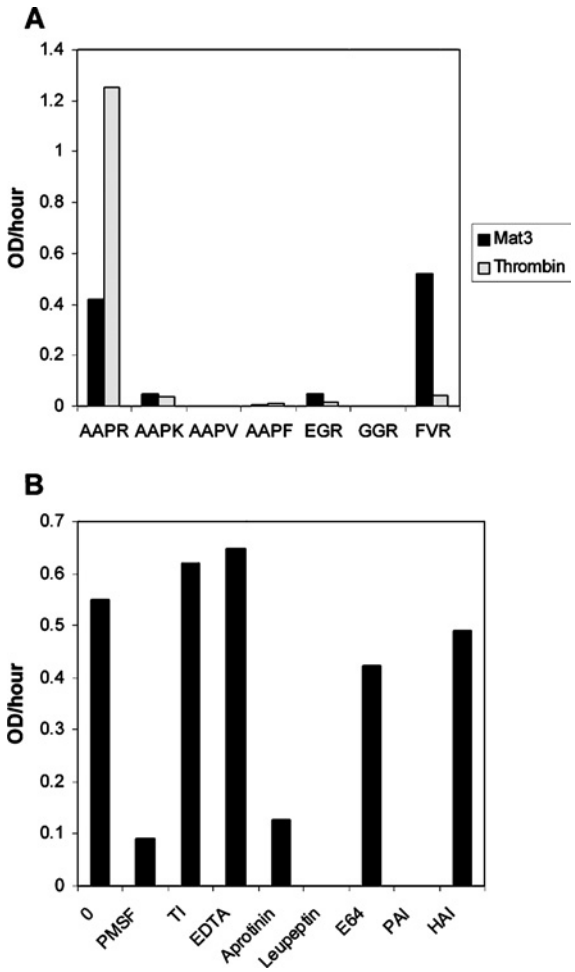

Figure 7 Analysis of substrate preference and inhibitory profile of the mouse matriptase-3

(**A**) Purified recombinant mouse matriptase-3 (10 nM) or thrombin (10 nM) were incubated with 100 μM of Suc-Ala-Ala-Pro-Arg-pNA (AAPR), Suc-Ala-Ala-Pro-Lys-pNA (AAPK), Suc-Ala-Ala-Pro-Val-pNA (AAPV), Suc-Ala-Ala-Pro-Phe-pNA (AAPF), H-Glu-Gly-Arg-pNA (EGR), Z-Gly-Gly-Arg-pNA (GGR), and Z-Phe-Val-Arg-pNA (FVR) at 35 °C, and the changes in A at 405 nm with time were monitored. (**B**) Purified recombinant mouse matriptase-3 (10 nM) was incubated with the synthetic peptide Suc-Ala-Ala-Pro-Arg-pNA in the absence or presence of PMSF (1 mM), egg white trypsin inhibitor (TI; 100 nM), EDTA (2.5 mM), aprotinin (3 μM), leupeptin (20 μM), E-64 (30 μM), human recombinant PAI-1 (50 nM) or soluble mouse recombinant HAI-1 (50 nM). The conversion rate in each reaction is presented as the change in A_{405} with time (OD/hour, absorbance units/h).

lower K_m for peptides bearing P_1-Arg compared with P_1-Lys (240 ± 83 nM for Suc-Ala-Ala-Pro-Arg-pNA, compared with 534 ± 135 nM for Suc-Ala-Ala-Pro-Lys-pNA; results not shown).

The hydrolytic activity of the matriptase-3 serine protease domain towards Suc-Ala-Ala-Pro-Arg-pNA was reduced by generic small molecule inhibitors of serine proteases (aprotinin, leupeptin and PMSF) and, interestingly, by the serpin PAI-1, but not by inhibitors of either cysteine proteases or metalloproteases, E-64 and EDTA respectively (Figure 7B). On the other hand, HAI-1, an established inhibitor of matriptase/MT-SP1, did not show significant inhibitory activity towards matriptase-3 in this assay.

Mouse matriptase-3 forms inhibitory complexes with the serpins PAI-1, PCI, α1-PI, α2-AP and AT III

Matriptase/MT-SP1 forms stable inhibitory complexes with HAI-1, and HAI-1 is a principal matriptase/MT-SP1 inhibitor

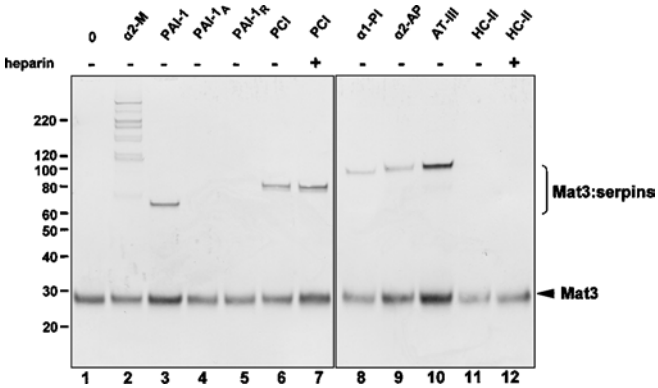

Figure 8 Mouse matriptase-3 forms stable inhibitory complexes with serpins

The active mouse matriptase-3 catalytic domain was incubated for 30 min at room temperature with buffer (lane 1), 30 nM α_2M (lane 2), 100 nM human PAI-1 (lane 3), PAI-1_A P_1 mutant (lane 4), PAI-1_R insertion loop mutant (lane 5), PCI (lane 6), PCI with heparin (lane 7), α1-PI (lane 8), α2-AP (lane 8), AT III (lane 10), HC II (lane 11) or HC II with heparin (lane 12). The positions of the activated mouse matriptase-3 serine protease domain (Mat3) and the complexes of the protease with serpins (Mat3:serpins) are indicated. The positions of molecular-mass standards (kDa) are indicated on the left.

Table 1 Serpin specificity of matriptase-3

Clade	Serpin	Reactive centre loop (P_4–P_4')	Matriptase-3 complex	IC_{50} (nM)
SERPINA1	α1-PI	AIPM-SIPP	Yes	No inhibition
SERPINA5	PCI	FTFR-SARL	Yes	18.3 ± 1.2
SERPINC1	HC II	FMPL-STQV	No	Not tested
SERPIND1	AT III	IAGR-SLNP	Yes	44.0 ± 8.5
SERPINE1	PAI-1	VSAR-MAPE	Yes	4.5 ± 0.5
SERPINF2	α2-AP	AMSR-MSLS	Yes	56.7 ± 15.1

in vivo [42,43]. Furthermore, we have previously shown that DESC1 forms stable inhibitory complexes with PAI-1 and PCI, suggesting that the TTSPs may be novel targets for serpin inhibition [15]. To test the ability of matriptase-3 to form complexes with serine protease inhibitors, the active, purified matriptase-3 catalytic domain was incubated with a number of serpins. Inhibitory complex formation was analysed by SDS/PAGE and Western blotting using an antibody raised against the V5 epitope tag of matriptase-3. Interestingly, the matriptase-3 serine protease domain formed stable inhibitory complexes with PAI-1 (Figure 8, lane 3), PCI (lane 6 and 7), α1-PI (lane 8), α2-AP (lane 9) and AT III (lane 10), but not with HC II (lanes 11 and 12). Inhibitory complex formation of matriptase-3 with PAI-1 was completely blocked by the mutation of the P_1 amino acid of the PAI-1 reactive site loop that creates a mutant resistant to the protease cleavage required for the complex formation (PAI-1_A; Figure 8, lane 4). Similarly, no complex formation was detected with a PAI-1 mutant defective in loop insertion and the associated conformational changes required for the formation of the complex (PAI-1_R; Figure 8, lane 5).

The efficiency with which the five serpins, found to form SDS-stable complexes with recombinant matriptase-3, inhibited the enzyme activity was further tested in a chromogenic peptide assay. As shown in Table 1, PAI-1 was the most efficient inhibitor of matriptase-3 activity with IC_{50} values in the low nanomolar range, followed by PCI, AT III and α2-AP. On the other hand, no appreciable inhibition of matriptase-3 activity was detected with up to 100 nM concentrations of α1-PI, the only serpin that forms a complex with matriptase-3, which does not have Arg at position P_1 of its reactive centre loop (Table 1).

DISCUSSION

A systematic bioinformatic search for new mammalian genes and gene families has recently become possible by advances in vertebrate genome sequencing projects. This has led not only to a great expansion in the number of members of previously characterized protease families, but even to the definition of entire families of new proteolytic enzymes [12–14,44,45]. Particularly striking in this respect was the unveiling of the TTSP family, of which a large majority of the > 20 predicted family members have now been identified by genome and cDNA database homology searches.

In this paper, we described the identification and characterization of the fourth member of the matriptase subfamily of the TTSPs, which we have termed *matriptase-3*. Bioinformatic analysis supported by direct cDNA cloning showed that *matriptase-3* would encode a protein with a high overall similarity to matriptase/MT-SP1 and matriptase-2, including the presence of an N-proximal transmembrane domain, followed by an SEA domain, two CUB domains, three LDLRa domains and a C-terminal catalytic domain of S1-chymotrypsin fold. The bioinformatic prediction of matriptase-3 as a novel functional TTSP was supported by exhaustive analysis of the *matriptase-3* mRNA distribution and biological characterization of the protein product. Thus the *matriptase-3* mRNA transcripts were present in several murine and human tissues. In addition, the cloned murine *matriptase-3* cDNA directed the synthesis of a full-length protein that was properly targeted to the secretory pathway, N-glycosylated, and expressed on the cell surface. Finally, the purified matriptase-3 serine protease domain formed stable complexes with the trap inhibitor α_2M, displayed proteolytic activity towards many macromolecular substrates, hydrolysed small peptide substrates with preference for Arg at position P_1, and was susceptible to inhibition by generic serine protease inhibitors, but not by inhibitors of other classes of proteases.

The highest relative expression of the mouse *matriptase-3* mRNA was found in the epididymis, salivary gland, thymus, testis, ovary, uterus and central nervous system. From the panel of human tissues tested by Northern blotting and/or RT–PCR, the highest levels of *matriptase-3* transcript were detected in testis and ovary, and additional sites of expression included brain, salivary gland, lung and trachea. Although more detailed delineation of *matriptase-3* expression awaits further studies when suitable reagents are developed, these results suggest that the mouse and human *matriptase-3* orthologues are expressed in highly overlapping subsets of tissues, and therefore are likely to perform analogous functions in both species.

Previous analysis of the gene and/or protein expression of other members of matriptase subfamily showed that matriptase/MT-SP1 is broadly expressed in the epithelial lining of almost all tissues examined [37], whereas matriptase-2 is highly liver-specific, with lower levels of gene expression detected also in kidney and uterus [20,34], and polyserase-1 expression is restricted to skeletal muscle, liver, placenta and heart [21]. It therefore appears that the pattern of expression of *matriptase-3* is different from that of *matriptase*/*MT-SP1* and *matriptase-2*, suggesting that despite the substantial sequence and structural similarities between the three proteins, matriptase-3 is likely to perform a unique, non-redundant physiological function(s).

Studies of the post-translational regulation of TTSPs, including their interaction with serine protease inhibitors, are in their infancy. A pioneering series of studies by Lin and Dickson has

established that the Kunitz-type inhibitor, HAI-1, plays a critical role in regulating the activity of matriptase/MT-SP1 at multiple levels [42,43,48]. In this respect, it is noteworthy that we were unable to document either an inhibitory activity of HAI-1 towards matriptase-3 or any effect of its co-expression on the intracellular trafficking of matriptase-3 full-length protein (results not shown). Although more detailed studies are necessary, this would tentatively suggest that HAI-1 is not a major regulator of matriptase-3 and, thus, that the biological regulation of the two closely related proteases could be fundamentally different. In notable contrast, we found that the activated matriptase-3 serine protease domain formed bona fide inhibitor complexes with an array of serpin-type inhibitors, including PAI-1, PCI, α1-PI, α2-AP and AT III. Of the five, PAI-1 was the most effective in inhibiting matriptase-3 enzymatic activity *in vitro*, displaying an IC_{50} in the low nanomolar range. This finding is analogous to our recent report that PAI-1 and PCI formed stable inhibitor complexes with both the isolated serine protease domain and a soluble full-length form of the TTSP DESC1 [15]. Although it is likely that some of these interactions do not take place *in vivo*, due to lack of co-expression of the serpin with matriptase-3, steric interactions or unfavourable kinetics, the combined data, nevertheless, open the intriguing possibility that inhibition of TTSPs may be a new and biologically important function of the large clade of serpin-type serine protease inhibitors.

Elucidation of the physiological functions of TTSPs is the focus of attention of an increasing number of researchers. Null mutations have been identified in two human TTSP genes (TMPRSS3 and enteropeptidase), and mice with a targeted inactivation of three additional TTPS genes (hepsin, corin and matriptase/MT-SP1) have been reported to date [16,17,32,46,47]. It is noteworthy that all five of these TTSP genes were found to have functions in either basic developmental or post-natal homoeostatic processes that are essential to all vertebrate classes, including epidermal differentiation, inner ear development, blood pressure regulation, food digestion and liver homoeostasis. Moreover, bioinformatic analysis presented in this paper identified orthologues of *matriptase-3* in all the nine vertebrate genomes that were analysed, including teleost fish, birds, canines, rodents and primates, and predicted that all nine orthologues would encode functional serine proteases. Taken together, this would predict a species-conserved function of *matriptase-3* and suggests that genetic studies aimed at understanding the physiological functions of *matriptase-3* could be advantageously carried out in non-mammalian vertebrates having a short life cycle and external embryonic development, such as pufferfish or zebrafish.

In summary, in the present study, we have identified, cloned and biologically characterized *matriptase-3*, a novel member of the matriptase subfamily of TTSPs that is present in all higher vertebrates analysed. Our results show that *matriptase-3* encodes a functional, cell-surface-associated serine protease with an expression pattern and an inhibitor susceptibility that is different from that of matriptase/MT-SP1.

We thank Drs R. C. Angerer, M. J. Danton and J. S. Gutkind (National Institute of Dental and Craniofacial Research, National Institutes of Health) for critically reading this paper. We also thank Dr D. A. Lawrence (University of Michigan) for kindly providing PAI-1 mutants. This work was supported by the National Institutes of Health (grant no. CA098369; to T. M. A.), the Lance Armstrong Foundation (to T. M. A.) and by a grant from the U.S. Department of Defense (DAMD-17-02-1-0693; to T. H. B.).

REFERENCES

1 Puente, X. S., Sanchez, L. M., Overall, C. M. and Lopez-Otin, C. (2003) Human and mouse proteases: a comparative genomic approach. Nat. Rev. Genet. **4**, 544–558
2 Egeblad, M. and Werb, Z. (2002) New functions for the matrix metalloproteinases in cancer progression. Nat. Rev. Cancer **2**, 161–174
3 Turk, V., Kos, J. and Turk, B. (2004) Cysteine cathepsins (proteases) – on the main stage of cancer? Cancer Cell **5**, 409–410
4 White, J. M. (2003) ADAMs: modulators of cell-cell and cell-matrix interactions. Curr. Opin. Cell Biol. **15**, 598–606
5 McCawley, L. J. and Matrisian, L. M. (2001) Matrix metalloproteinases: they're not just for matrix anymore! Curr. Opin. Cell Biol. **13**, 534–540
6 Sternlicht, M. D. and Werb, Z. (2001) How matrix metalloproteinases regulate cell behavior. Annu. Rev. Cell Dev. Biol. **17**, 463–516
7 Mackie, E. J., Pagel, C. N., Smith, R., de Niese, M. R., Song, S. J. and Pike, R. N. (2002) Protease-activated receptors: a means of converting extracellular proteolysis into intracellular signals. IUBMB Life **53**, 277–281
8 McCawley, L. J. and Matrisian, L. M. (2000) Matrix metalloproteinases: multifunctional contributors to tumor progression. Mol. Med. Today **6**, 149–156
9 Werb, Z. (1997) ECM and cell surface proteolysis: regulating cellular ecology. Cell (Cambridge, Mass.) **91**, 439–442
10 Andreasen, P. A., Egelund, R. and Petersen, H. H. (2000) The plasminogen activation system in tumor growth, invasion, and metastasis. Cell. Mol. Life Sci. **57**, 25–40
11 Wu, Q. (2003) Type II transmembrane serine proteases. Curr. Top. Dev. Biol. **54**, 167–206
12 Szabo, R., Wu, Q., Dickson, R. B., Netzel-Arnett, S., Antalis, T. M. and Bugge, T. H. (2003) Type II transmembrane serine proteases. Thromb. Haemostasis **90**, 185–193
13 Netzel-Arnett, S., Hooper, J. D., Szabo, R., Madison, E. L., Quigley, J. P., Bugge, T. H. and Antalis, T. M. (2003) Membrane anchored serine proteases: a rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer. Cancer Metastasis Rev. **22**, 237–258
14 Hooper, J. D., Clements, J. A., Quigley, J. P. and Antalis, T. M. (2001) Type II transmembrane serine proteases. Insights into an emerging class of cell surface proteolytic enzymes. J. Biol. Chem. **276**, 857–860
15 Hobson, J. P., Netzel-Arnett, S., Szabo, R., Rehault, S. M., Church, F. C., Strickland, D. K., Lawrence, D. A., Antalis, T. M. and Bugge, T. H. (2004) 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
16 Wu, Q., Yu, D., Post, J., Halks-Miller, M., Sadler, J. E. and Morser, J. (1998) Generation and characterization of mice deficient in hepsin, a hepatic transmembrane serine protease. J. Clin. Invest. **101**, 321–326
17 Chan, J. C., Knudson, O., Wu, F., Morser, J., Dole, W. P. and Wu, Q. (2005) Hypertension in mice lacking the proatrial natriuretic peptide convertase corin. Proc. Natl. Acad. Sci. U.S.A. **102**, 785–790
18 Takeuchi, T., Shuman, M. A. and Craik, C. S. (1999) 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
19 Lin, C. Y., Anders, J., Johnson, M., Sang, Q. A. and Dickson, R. B. (1999) Molecular cloning of cDNA for matriptase, a matrix-degrading serine protease with trypsin-like activity. J. Biol. Chem. **274**, 18231–18236
20 Velasco, G., Cal, S., Quesada, V., Sanchez, L. M. and Lopez-Otin, C. (2002) 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
21 Cal, S., Quesada, V., Garabaya, C. and Lopez-Otin, C. (2003) 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
22 Shi, Y. E., Torri, J., Yieh, L., Wellstein, A., Lippman, M. E. and Dickson, R. B. (1993) Identification and characterization of a novel matrix-degrading protease from hormone-dependent human breast cancer cells. Cancer Res. **53**, 1409–1415
23 Benaud, C. M., Oberst, M., Dickson, R. B. and Lin, C. Y. (2002) Deregulated activation of matriptase in breast cancer cells. Clin. Exp. Metastasis **19**, 639–649
24 Bhatt, A. S., Takeuchi, T., Ylstra, B., Ginzinger, D., Albertson, D., Shuman, M. A. and Craik, C. S. (2003) Quantitation of membrane type serine protease 1 (MT-SP1) in transformed and normal cells. Biol. Chem. **384**, 257–266
25 Hoang, C. D., D'Cunha, J., Kratzke, M. G., Casmey, C. E., Frizelle, S. P., Maddaus, M. A. and Kratzke, R. A. (2004) Gene expression profiling identifies matriptase overexpression in malignant mesothelioma. Chest **125**, 1843–1852
26 Johnson, M. D., Oberst, M. D., Lin, C. Y. and Dickson, R. B. (2003) Possible role of matriptase in the diagnosis of ovarian cancer. Expert Rev. Mol. Diagn. **3**, 331–338
27 Kang, J. Y., Dolled-Filhart, M., Ocal, I. T., Singh, B., Lin, C. Y., Dickson, R. B., Rimm, D. L. and Camp, R. L. (2003) Tissue microarray analysis of hepatocyte growth factor/Met pathway components reveals a role for Met, matriptase, and hepatocyte growth factor activator inhibitor 1 in the progression of node-negative breast cancer. Cancer Res. **63**, 1101–1105

28 Oberst, M. D., Johnson, M. D., Dickson, R. B., Lin, C. Y., Singh, B., Stewart, M., Williams, A., al-Nafussi, A., Smyth, J. F., Gabra, H. et al. (2002) Expression of the serine protease matriptase and its inhibitor HAI-1 in epithelial ovarian cancer: correlation with clinical outcome and tumor clinicopathological parameters. Clin. Cancer Res. **8**, 1101–1107
29 Santin, A. D., Cane, S., Bellone, S., Bignotti, E., Palmieri, M., De Las Casas, L. E., Anfossi, S., Roman, J. J., O'Brien, T. and Pecorelli, S. (2003) The novel serine protease tumor-associated differentially expressed gene-15 (matriptase/MT-SP1) is highly overexpressed in cervical carcinoma. Cancer **98**, 1898–1904
30 Santin, A. D., Zhan, F., Bellone, S., Palmieri, M., Cane, S., Bignotti, E., Anfossi, S., Gokden, M., Dunn, D., Roman, J. J. et al. (2004) Gene expression profiles in primary ovarian serous papillary tumors and normal ovarian epithelium: identification of candidate molecular markers for ovarian cancer diagnosis and therapy. Int. J. Cancer **112**, 14–25
31 Tanimoto, H., Underwood, L. J., Wang, Y., Shigemasa, K., Parmley, T. H. and O'Brien, T. J. (2001) Ovarian tumor cells express a transmembrane serine protease: a potential candidate for early diagnosis and therapeutic intervention. Tumour Biol. **22**, 104–114
32 List, K., Haudenschild, C. C., Szabo, R., Chen, W., Wahl, S. M., Swaim, W., Engelholm, L. H., Behrendt, N. and Bugge, T. H. (2002) Matriptase/MT-SP1 is required for postnatal survival, epidermal barrier function, hair follicle development, and thymic homeostasis. Oncogene **21**, 3765–3779
33 List, K., Szabo, R., Wertz, P. W., Segre, J., Haudenschild, C. C., Kim, S. Y. and Bugge, T. H. (2003) Loss of proteolytically processed filaggrin caused by epidermal deletion of Matriptase/MT-SP1. J. Cell Biol. **163**, 901–910
34 Hooper, J. D., Campagnolo, L., Goodarzi, G., Truong, T. N., Stuhlmann, H. and Quigley, J. P. (2003) Mouse matriptase-2: identification, characterization and comparative mRNA expression analysis with mouse hepsin in adult and embryonic tissues. Biochem. J. **373**, 689–702
35 Netzel-Arnett, S., Mitola, D. J., Yamada, S. S., Chrysovergis, K., Holmbeck, K., Birkedal-Hansen, H. and Bugge, T. H. (2002) Collagen dissolution by keratinocytes requires cell surface plasminogen activation and matrix metalloproteinase activity. J. Biol. Chem. **277**, 45154–45161
36 Griffith, M. J., Noyes, C. M. and Church, F. C. (1985) Reactive site peptide structural similarity between heparin cofactor II and antithrombin III. J. Biol. Chem. **260**, 2218–2225
37 Oberst, M. D., Singh, B., Ozdemirli, M., Dickson, R. B., Johnson, M. D. and Lin, C. Y. (2003) Characterization of matriptase expression in normal human tissues. J. Histochem. Cytochem. **51**, 1017–1025
38 Cho, E. G., Kim, M. G., Kim, C., Kim, S. R., Seong, I. S., Chung, C., Schwartz, R. H. and Park, D. (2001) N-terminal processing is essential for release of epithin, a mouse type II membrane serine protease. J. Biol. Chem. **276**, 44581–44589
39 Afar, D. E., Vivanco, I., Hubert, R. S., Kuo, J., Chen, E., Saffran, D. C., Raitano, A. B. and Jakobovits, A. (2001) Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. Cancer Res. **61**, 1686–1692
40 Feldman, S. R., Gonias, S. L. and Pizzo, S. V. (1985) Model of alpha 2-macroglobulin structure and function. Proc. Natl. Acad. Sci. U.S.A. **82**, 5700–5704
41 Barrett, A. J. and Starkey, P. M. (1973) The interaction of alpha 2-macroglobulin with proteinases. Characteristics and specificity of the reaction, and a hypothesis concerning its molecular mechanism. Biochem. J. **133**, 709–724
42 Lin, C. Y., Anders, J., Johnson, M. and Dickson, R. B. (1999) Purification and characterization of a complex containing matriptase and a Kunitz-type serine protease inhibitor from human milk. J. Biol. Chem. **274**, 18237–18242
43 Oberst, M. D., Williams, C. A., Dickson, R. B., Johnson, M. D. and Lin, C. Y. (2003) The activation of matriptase requires its noncatalytic domains, serine protease domain, and its cognate inhibitor. J. Biol. Chem. **278**, 26773–26779
44 Tang, B. L. (2001) ADAMTS: a novel family of extracellular matrix proteases. Int. J. Biochem. Cell Biol. **33**, 33–44
45 Bugge, T. H. (2003) Proteolysis in carcinogenesis. In Head and Neck Cancer: Emerging Perspectives (Ensley, J. F., Gutkind, J. S., Jacobs, J. R. and Lippman, S. M., eds.), pp. 137–149, Academic Press, San Diego, CA
46 Wattenhofer, M., Di Iorio, M. V., Rabionet, R., Dougherty, L., Pampanos, A., Schwede, T., Montserrat-Sentis, B., Arbones, M. L., Iliades, T., Pasquadibisceglie, A. et al. (2002) Mutations in the TMPRSS3 gene are a rare cause of childhood nonsyndromic deafness in Caucasian patients. J. Mol. Med. **80**, 124–131
47 Rutgeerts, L. and Eggermont, E. (1976) Human enterokinase. Tijdschr. Gastroenterol. **19**, 231–246
48 Oberst, M. D., Chen, L. Y., Kiyomiya, K. I., Williams, C. A., Lee, M. S., Johnson, M. D., Dickson, R. B. and Lin, C. Y. (2005) HAI-1 regulates activation and expression of matriptase, a membrane-bound serine protease. Am. J. Physiol. Cell Physiol. **289**, C462–470

Received 16 February 2005/22 April 2005; accepted 27 April 2005
Published as BJ Immediate Publication 27 April 2005, doi:10.1042/BJ20050299