online data

Hepatocyte growth factor is a preferred *in vitro* substrate for human hepsin, a membrane-anchored serine protease implicated in prostate and ovarian cancers

Sylvia HERTER*[1], Derek E. PIPER*, Wade AARON*, Timothy GABRIELE*, Gene CUTLER*, Ping CAO*, Ami S. BHATT†, Youngchool CHOE†, Charles S. CRAIK†, Nigel WALKER*, David MEININGER*, Timothy HOEY* and Richard J. AUSTIN*[1,2]

*Department of Biology, Amgen San Francisco, 1120 Veterans Boulevard, South San Francisco, CA 94080, U.S.A., and †University of California San Francisco, Department of Pharmaceutical Chemistry, 600 16th Street, San Francisco, CA 94143, U.S.A.

Hepsin is a membrane-anchored, trypsin-like serine protease with prominent expression in the human liver and tumours of the prostate and ovaries. To better understand the biological functions of hepsin, we identified macromolecular substrates employing a tetrapeptide PS-SCL (positional scanning-synthetic combinatorial library) screen that rapidly determines the P1–P4 substrate specificity. Hepsin exhibited strong preference at the P1 position for arginine over lysine, and favoured threonine, leucine or asparagine at the P2, glutamine or lysine at the P3, and proline or lysine at the P4 position. The relative activity of hepsin toward individual AMC (7-amino-4-methylcoumarin)-tetrapeptides was generally consistent with the overall peptide profiling results derived from the PC-SCL screen. The most active tetrapeptide substrate Ac (acetyl)-KQLR-AMC matched with the activation cleavage site of the hepatocyte growth factor precursor sc-HGF (single-chain HGF), KQLR↓VVNG (where ↓ denotes the cleavage site), as identified by a database analysis of trypsin-like precursors. X-ray crystallographic studies with KQLR chloromethylketone showed that the KQLR peptide fits well into the substrate-binding cleft of hepsin. This hepsin-processed HGF induced c-Met receptor tyrosine phosphorylation in SKOV-3 ovarian cancer cells, indicating that the hepsin-cleaved HGF is biologically active. Activation cleavage site mutants of sc-HGF with predicted non-preferred sequences, DPGR↓VVNG or KQLQ↓VVNG, were not processed, illustrating that the P4–P1 residues can be important determinants for substrate specificity. In addition to finding macromolecular hepsin substrates, the extracellular inhibitors of the HGF activator, HAI-1 and HAI-2, were potent inhibitors of hepsin activity (IC_{50} 4 ± 0.2 nM and 12 ± 0.5 nM respectively). Together, our findings suggest that the HGF precursor is a potential *in vivo* substrate for hepsin in tumours, where hepsin expression is dysregulated and may influence tumorigenesis through inappropriate activation and/or regulation of HGF receptor (c-Met) functions.

Key words: c-Met receptor activation, hepatocyte growth factor (HGF), hepsin, positional scanning-synthetic combinatorial library (PS-SCL), proteolytic cascade, substrate specificity.

INTRODUCTION

Hepsin is a serine protease that was identified as a prominently expressed protein in normal human liver [1]. Hepsin belongs to the TTSP (type II transmembrane serine protease) family [2], which includes enteropeptidase, MT-SP1 (membrane-type serine protease 1)/matriptase, HAT (human airway trypsin-like protease), corin, TMPRSS2 (transmembrane serine protease 2), TMPRSS3, TMPRSS4, TMPRSS5, MSPL (mosaic serine protease) and DESC1 (differentially expressed squamous cell carcinoma gene 1). The TTSPs all feature a short N-terminal cytoplasmic domain, a membrane-spanning region, a central region of variable length containing structural domains and a C-terminal serine protease domain with a chymotrypsin-like fold [3,4].

As with most TTSPs, the physiological roles of human hepsin are not well understood, in part because its specific *in vivo* protein substrates have not been identified. Cell culture studies with baby hamster kidney cells overexpressing human hepsin showed a calcium-dependent formation of active thrombin through cleavage of blood coagulation factor VII [5]. Although hepsin has been reported to be involved in the activation of blood coagulation, human hepatoma cell growth [6] and mouse blastocyst hatching [7], definitive *in vivo* evidence for hepsin's role in these processes is lacking. Hepsin-deficient mice developed normally, were fertile, exhibited normal blood clotting functions and liver restoration following hepatectomy, and progressed normally when compared with wild-type littermates [8–10]. Recently, a non-transmembrane isoform of hepsin (hepsin/−TM) was identified, which was found distributed within the cytoplasm and restricted in its expression to kidney, lung and brain tissue [11].

Hepsin has received attention due to its prominent over-expression in prostate and ovarian tumours [12–24] and in renal cell carcinoma *in situ* [25]. The effect of such dysregulated hepsin expression in tumours has recently been shown to promote cancer progression and metastasis in mouse models [26] but the mechanism has yet to be elucidated [27]. Interestingly, over-expression of hepsin in the prostate cancer cell line PC-3 has been reported to inhibit cell proliferation [28], whereas over-expression of hepsin/−TM was not found to cause inhibition of mammalian cell growth [11].

To model the *in vivo* functions of human hepsin, we initiated biochemical studies to examine its substrate specificity in detail,

Abbreviations used: Ac-, acetyl-; AMC, 7-amino-4-methylcoumarin; ACC, 7-amino-4-carbamoylmethylcoumarin; cmk, chloromethylketone; FBS, fetal bovine serum; HAI, hepatocyte growth factor activator inhibitor; HGF, hepatocyte growth factor; MT-SP1, membrane-type serine protease 1; PEG, poly(ethylene glycol); *p*NA, *p*-nitroanilide; PS-SCL, positional scanning-synthetic combinatorial library; sc-HGF, single-chain hepatocyte growth factor; RFU, relative fluorescence units; S, subsite, describing substrate-binding cleft of serine protease; TMPRSS, transmembrane serine protease; TTSP, type II transmembrane serine protease.

[1] Present address Research Drug Discovery Department, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080, U.S.A.
[2] To whom correspondence should be addressed (email rjaustin@amgen.com).

particularly focusing on the role of hepsin in tumours. Based on general peptide profiling and amidolytic activity measurements, we attempted to find the optimal P4–P1 cleavage motif. The data presented show that the peptide substrate spectrum of hepsin is restricted. To extend and validate these results, we focused our search for macromolecular protein substrates on precursor forms of extracellular serine-protease-fold zymogens, whose expression was coincident with hepsin in tumours and/or normal tissues. These precursors have to become activated by characteristic proteolytic cleavage events that take place in their activation domains. Hepsin cleaved precursors of blood coagulation factors and the precursor of the hepatocyte growth factor sc-HGF (single-chain HGF) *in vitro*. The hepsin-cleaved HGF induced autophosphorylation of the HGF receptor (c-Met) in SKOV-3 ovarian cancer cells, indicating that HGF is processed specifically and is functionally active. We also explored potential extracellular endogenous inhibitors of hepsin activity, and found that hepsin is potently inhibited by human HAI-1 (HGF activator inhibitor-1) and HAI-2 *in vitro*. Together, our results suggest that sc-HGF may be a preferred endogenous substrate for hepsin and provide a better understanding of how up-regulation of hepsin expression may influence tumorigenesis.

EXPERIMENTAL

Reagents and cell culture

Factors VII, IX, X, XII and Glu-plasminogen were purchased from Haematologic Technologies (Essex Junction, VT, U.S.A.). Factors VIIa and IXa were obtained from Haematologic Technologies; Factors Xa and XIIa were from Enzyme Research Laboratories (South Bend, IN, U.S.A.). Plasmin was purchased from Calbiochem (La Jolla, CA, U.S.A.). Recombinant heterodimeric human HGF and soluble human HAI-1 and HAI-2 were obtained from R&D Systems (Minneapolis, MN, U.S.A.). Neutralizing monoclonal anti-(human HGF) antibody (clone 24612.111; isotype IgG1) was purchased from Sigma–Aldrich. Epithelial SKOV-3 ovarian cancer cells were obtained from ATCC (Manassas, VA, U.S.A.) and grown on Falcon Primaria™ tissue culture plates (Becton Dickinson, Franklin Lakes, NJ, U.S.A.) in McCoy's 5A medium (modified) with L-glutamine (Gibco, Invitrogen Corporation, San Diego, CA, U.S.A.) supplemented with 10 % (v/v) FBS (fetal bovine serum) (Cellgro, Mediatech, Herndon, VA, U.S.A.).

PS-SCL (positional scanning-synthetic combinatorial library)

PS-SCLs were assayed with recombinant human hepsin. The libraries were prepared as described by Harris et al. [29] in 96-well Microfluor Black 'U'-bottom plates (Dynex Technologies, Chantilly, VA, U.S.A.). Libraries were assayed under optimal buffer conditions in 50 mM Tris/HCl, pH 8.2, 2 % (v/v) PEG [poly-(ethylene glycol)]-8000 and 0.2 % (v/v) NP40 (Nonidet P40). PS-SCLs were diluted to 0.25 mM in a final 100 μl assay volume. Thus, the individual tetrapeptides in each pool were at a concentration far below the expected K_m of hepsin, and the initial rates were proportional to the specificity constant k_{cat}/K_m (30 nM each peptide per well for P1-library). Proteolysis reactions were initiated by the addition of hepsin (1 nM). Fluorescence was monitored with a SpectraMAX Gemini fluorimetric 96-well plate reader (Molecular Devices Corporation, Sunnyvale, CA, U.S.A.) at 21 °C and the rate of substrate hydrolysis was analysed with the SOFTmax PRO software (version 3.1.1, Molecular Devices Corporation). Excitation and emission was measured at 380 nm and 460 nm respectively (21 °C for 10 min).

Determination of hepsin amidolytic activities

Individual Ac-P4-P3-P2-P1-AMC peptides (where Ac is acetyl and AMC is 7-amino-4-methylcoumarin), Ac-PVDR-AMC, Ac-KKTR-AMC, Ac-KQLR-AMC, Ac-DQLR-AMC and Ac-DPGR-AMC were synthesized by Anaspec (San Jose, CA, U.S.A.). Peptide preparations showed > 95 % purity as determined by HPLC analysis and MS. The amidolytic activities of hepsin with individual peptides were tested under the same buffer conditions as described for the PS-SCL screening, using 1 nM hepsin and 500 μM peptide. Assays contained less than 5 % (v/v) DMSO. The rate of AMC release was monitored with excitation at 380 nm and emission at 480 nm (21 °C for 30 min). The linear rates of fluorescence increase were expressed as percentage (%) activities when compared with the activity obtained with the KQLR peptide (set to 100 %). The peptides were assayed for hydrolysis by hepsin at least twice in triplicate. Rates of substrate hydrolysis were determined using the SOFTmax PRO data analysis software. To determine the steady-state kinetic specificity constant (k_{cat}/K_m) for the KQLR peptide, hepsin was diluted to 1 and 3 nM in assay buffer. The concentration of KQLR peptide was saturating at 500 μM for the determination of the k_{cat} value and seven different concentrations of peptide, ranging from 0.5 μM to 500 μM, were used for the determination of the apparent K_m value. Data were fitted to the Michaelis–Menten equation using SigmaPlot (Version 8.0, Systat, London, U.K.). The conversion factor for the correlation of RFU (relative fluorescence units) to free AMC formed was determined by titration of free AMC under given conditions to be 0.21 per nM free AMC.

Purification, crystallization and structure determination of human hepsin

Extracellular hepsin (amino acids 46–417; N112A) was expressed in *Pichia pastoris* as described previously by Somoza et al. [30] using a BioFloIV fermentation reactor (New Brunswick Scientific, Edison, NJ, U.S.A.). The fermentor supernatant was concentrated 5-fold and buffer exchanged into buffer containing 50 mM Tris/HCl, pH 7.5, and 150 mM NaCl. The buffer-exchanged supernatant was made 0.1 % (v/v) in CHAPS to facilitate hepsin activation in-*trans*. The auto-activated hepsin was loaded on to an ecotin affinity column as described by Lengyel et al. [31] and eluted with 200 mM glycine, pH 3.0. The eluted protein solution was adjusted to pH 7.5 with 1 M Tris/HCL, and NaCl was added to a final concentration of 300 mM. The protein was bound to benzamidine Sepharose 6B, equilibrated in buffer containing 50 mM Tris/HCl, pH 7.5, and 300 mM NaCl, and eluted with buffer additionally containing 100 mM benzamidine. Hepsin was then purified by size-exclusion chromatography using a Superdex 75 column equilibrated in 50 mM Tris/HCl, pH 7.5, and 50 mM NaCl. Purified hepsin was concentrated to 7 mg $\cdot$ ml^{-1} and 3-times the molar equivalent of the protease inhibitor 2-(2-hydroxyphenyl)-1*H*-benzoimidaxole-5-carboxamidine [30] was added. Initial hepsin crystals were grown using the sitting-drop method using a well solution of 200 mM ammonium fluoride, 20–25 % (v/v) PEG-3350, and 100 mM sodium cacodylate, pH 6.2 [30]. Microseeding reproducibly produced high-quality hepsin crystals with a well solution of 200 mM ammonium fluoride, 4 % (v/v) PEG-3350, and 100 mM sodium-cacodylate, pH 5.6. Diffraction data were collected at 90 K and processed using the HKL software program package (HKL Research, Charlottesville, VA, U.S.A.). The hepsin structure was solved by molecular replacement with the programme EPMR [32] using enteropeptidase [33] as the search model and refined to 1.8 Å (1 Å $\equiv$ 0.1 nm) with the program CNX [34]. The peptide inhibitor Ac-KQLR-cmk

(Ac-KQLR-chloromethylketone) was synthesized by Anaspec. The hepsin–KQLR-cmk complex was formed by soaking hepsin crystals in a solution containing the KQLR-cmk at 10 mM. The complex was refined to 1.55 Å, with the final model containing hepsin amino acid residues 49–159 and 163–417, the KQLR fragment and 461 water molecules. Data and refinement statistics are shown in Supplementary Table 1 http://www.BiochemJ.org/bj/390/bj3900125add.htm (Protein DataBank accession number 1Z8G).

Expression and purification of human recombinant wild-type and mutated sc-HGF from High Five™ insect cell culture supernatants infected with baculovirus

The baculovirus wild-type sc-HGF construct was generated using a full-length HGF clone (E30163; pDUAL vector), which was obtained from Stratagene (La Jolla, CA, U.S.A.) with C-terminal c-Myc (3×)- and His_6 tags. To subclone the sc-HGF coding sequence, including C-terminal tags, into a baculovirus transfer vector and to remove an L548R mutation present in the Stratagene clone, a two-step PCR procedure was performed. First, using the Stratagene clone as a template, two PCRs were performed using the primer pair 5′sc-HGF (5′-CGCGGATCCGCCACCATGTGGGTGACCAAACTCCTGC-3′) and HGF1622-61 (5′-GAAAGATTATGAAGCTTGGCTTGGAATTCATGATGTCCAC-3′), and the primer pairs HGF1661-22 (5′-GTGGACATCATGAATTCCAAGCCAAGCTTCATAATCTTTC-3′) and 3′pDUAL (5′-GAGTCGACTCATTAGTGATGGTGATGGTGATGTCTAG-3′). These first-round PCR products were gel purified and used as templates in a second-round PCR reaction with the 5′sc-HGF and 3′pDUAL primers. The second-round PCR product was digested with BamHI and SalI and subcloned into pFastBac1 (Invitrogen, San Diego, CA, U.S.A.), which was digested with the same enzymes: the resulting construct was termed pFB-sc-HGF.

Baculoviruses expressing sc-HGF precursor constructs with mutated activation site (P4–P1) sequences KQLQ (R494Q) or DPGR (K491D/Q492P/L493G) were generated by mutagenesis on the baculovirus transfer vector pFB-sc-HGF using the QuikChange® II Site-Directed Mutagenesis Kit (Stratagene), as described by the manufacturer. The sequences of the primers used to make the R494Q mutant (referred to as mutant 1) were 5′-GTGCCAAAACGAAACAATTGCAAGTTGTAAATGGGATTCCAAC-3′ and 5′-GTTGGAATCCCATTTACAACTTGCAATTGTTTCGTTTTGGCAC-3′. The sequences of the primers used to make the K491D/Q492P/L493G mutant (referred to as mutant 2) were 5′-CCGTAATATCTTGTGCCAAAACGGATCCAGGGCGAGTTGTAAATGGGATTCCAA-3′ and 5′-TTGGAATCCCATTTACAACTCGCCCTGGATCCGTTTTGGCACAAGATATTACGG-3′. Baculovirus expressing HGF were created in Sf9 (*Spodoptera frugiperda*) cells using the wild-type or mutated pFB-sc-HGF constructs and the techniques described in the Invitrogen Bac-to-Bac manual. For protein expression, 1 litre of High Five™ suspension-adapted cells (Invitrogen) was grown in a 2 litre Erlenmeyer flask (Corning; VWR Scientific, Batavia, IL, U.S.A.) to a density of 1×10^6 cells/ml in ESF921 medium (Expression Systems LLC, Woodland, CA, U.S.A.) supplemented with 0.1 % (v/v) Pluronic F-68 (Sigma–Aldrich), 100 units $\cdot$ ml^{-1} penicillin, 0.25 μg $\cdot$ ml^{-1} amphoteric-3 and 100 μg $\cdot$ ml^{-1} streptomycin (Cellgro, Mediatech). Cells were infected with 1 ml of wild-type or mutated sc-HGF expressing viral stock and incubated at 27 °C at 110 rev./min for 48 h in the dark before harvesting the culture medium by centrifugation at 660 ***g*** for 5 min at room temperature. For purification of wild-type sc-HGF precursor for subsequent stimulation of SKOV-3 cells (see Figures 4 and 5), the cell-free culture supernatant was adjusted to 25 mM Tris/HCl, pH 7.4, 5 mM benzamidine, 1 mM $NiCl_2$ and 5 mM $CaCl_2$ (buffer A). Precipitate was removed by centrifugation at 4660 ***g*** for 30 min at room temperature. The medium was then incubated with 5 ml of nickel-nitrilotriacetic acid Superflow resin (Qiagen, Valencia, CA, U.S.A.) at 4 °C for 1 h, which was subsequently transferred to an empty column by gravity flow. After washing with 3 column volumes of buffer containing 25 mM Tris/HCl, pH 7.4, 5 mM benzamidine, 1 mM $NiCl_2$ and 5 mM $CaCl_2$, the sc-HGF was eluted with 3 column volumes of buffer A additionally containing 250 mM imidazole, 100 mM NaCl and 10 % (v/v) glycerol. The 15 ml eluate was concentrated 12-fold to a final volume of 1.25 ml using Amicon Ultra Low Binding Regenerated Cellulose concentrators (30 kDa, Millipore, Bedford, MA, U.S.A.) at 2800 ***g***, and 4 °C. The buffer was exchanged subsequently to 25 mM Hepes/HCl, pH 7.6, 500 mM NaCl, 5 mM benzamidine and 10 % (v/v) glycerol (buffer B). The concentrate was loaded on to a HiLoad Superdex 200 preparative column 16/60 (Amersham Biosciences, GE Healthcare, Uppsala, Sweden) in buffer B at a flow rate of 1 ml $\cdot$ min^{-1} at 4 °C. Sc-HGF eluted at approx. 60 ml and was collected in fractions of 2 ml. The sc-HGF-containing elution fractions were pooled and concentrated to a final volume of 1 ml using concentrators described above. The concentrate was subjected to dialysis using Slide-A-Lyzer devices (3500 Da; Pierce, Rockford, IL, U.S.A.) into 50 mM Hepes/HCl, pH 7.6, 300 mM NaCl and 50 % (v/v) glycerol.

Precursors of wild-type and mutated sc-HGF in Figure 4(B) were purified in parallel using a modified protocol. The cell-free culture supernatants were adjusted to final concentrations of 50 mM phosphate buffer, pH 7.2, 600 mM NaCl and 1 mM benzamidine. The individual culture supernatants were then loaded on to 5 ml HiTrap Heparin HP columns (Amersham Biosciences, GE Healthcare) at a flow rate of 5 ml $\cdot$ min^{-1}. Low-heparin-binding protein contaminants were removed during a wash step over 10 column volumes with phosphate buffer, pH 7.2, containing 600 mM NaCl. The sc-HGF precursors were eluted with 10 column volumes of the same buffer and the eluates were loaded on to 1 ml Ni^{2+}-charged HiTrap Chelating HP columns (Amersham Biosciences, GE Healthcare) equilibrated with phosphate buffer, pH 7.2, additionally containing 1.2 M NaCl and 20 mM imidazole (buffer C). The columns were washed at a flow rate of 1 ml $\cdot$ min^{-1} and 50 column volumes of buffer C, followed by a second wash step with 5 column volumes of buffer C, additionally containing 50 mM imidazole. The sc-HGF precursor proteins were then eluted in 1 ml fractions with 10 column volumes of 160 mM imidazole in buffer C. Protein purity was estimated by SDS/PAGE analysis to be greater than 95 %. Immunoblot blot analysis using an anti-HGF α-chain-specific antibody were performed to ensure that there was no partly activated α/β-heterodimeric HGF present prior to stimulation of SKOV-3 cells. Sc-HGF precursors were stored at −20 °C until further use.

Digests of macromolecular target precursor proteins

Target precursor proteins (Factors VII, XII, IX and X, Glu-plasminogen and wild-type or mutated sc-HGF) were incubated at 37 °C for 2 h with recombinant human hepsin at a molar ratio of 1:30 (hepsin/precursor) in a 20 μl final assay volume additionally containing 50 mM Tris/HCl, pH 7.9, and 250 mM NaCl. Digests were analysed by SDS/PAGE under reducing conditions and products were compared with commercially available Factors VIIa, XIIa, IXa and Xa, or plasmin. In addition, N-termini of hepsin-processed factor XII (light chain N_2H-VVGG) or VII (heavy chain N_2H-IVGG) were determined by N-terminal sequencing.

Determination of N-terminal amino acid sequences

The N-terminal amino acid sequences were analysed using a Procise cLC protein sequencer (Applied Biosystems, Foster City, CA, U.S.A.) following a standard Edman degradation protocol [35]. Prior to protein analysis, samples were separated by standard SDS/PAGE under reducing conditions, and subsequently blotted on to a PVDF membrane (Millipore).

Immunoblot analysis

Protein samples were separated by SDS/PAGE under reducing conditions. Protein was subsequently transferred over 1 h at 4 °C to a PVDF membrane (80 V) using the Mini Trans-Blot Electrophoresis Transfer Cell System from Bio-Rad (Hercules, CA, U.S.A.) in Tris/glycine buffer, pH 8.4, containing 20 % (v/v) methanol. For the detection of HGF and c-Met, membranes were incubated with the following antibodies according to the suppliers' recommendations: polyclonal goat anti-(human HGF) α-chain-specific antibody (N-17) purchased from Santa Cruz Biotechnology (Santa Cruz, CA, U.S.A.), monoclonal anti-(human Met) (25H2) and polyclonal rabbit anti-(human phospho-Met) (phosphorylated on Tyr^{1234}/Tyr^{1235}) antibodies obtained from Cell Signaling Technology (Beverly, MA, U.S.A.). Horseradish peroxidase-coupled secondary antibodies from goat and donkey against rabbit and goat IgG were obtained from Bio-Rad. Goat anti-mouse IgG coupled to horseradish peroxidase was purchased from Santa Cruz Biotechnology.

Preparation of cell-free protein extracts from SKOV-3 cells

For analysis of the c-Met status in SKOV-3 cells, cell-free protein extracts were prepared at 4 °C. Adherent cells were washed twice with PBS prior to the addition of 100 μl of lysis buffer per 35 mm culture well. Lysis buffer (50 ml) contained 50 mM Hepes/HCl, pH 7.9, 250 mM NaCl, 0.1 % (v/v) Nonidet P40, 10 % (v/v) glycerol, 1 mM EDTA, one tablet of Complete™ protease inhibitor cocktail (Roche Diagnostics, Mannheim, Germany), 20 mM β-glycerolphosphate, 5 mM *p*-nitrophenyl phosphate and 1 mM sodium vanadate. Cell lysates were subjected to centrifugation at 20 800 ***g*** for 20 min. Laemmli sample buffer (Bio-Rad) including 5 % (v/v) β-mercaptoethanol was added to the protein extracts and samples were boiled for 5 min at 95 °C. Samples were stored at − 80 °C and 10 μg of protein per lane was separated by SDS/PAGE and analysed by immunoblotting.

Stimulation of SKOV-3 cells with wild-type HGF

Cells were harvested by detachment using cell dissociation solution from Sigma–Aldrich and washed twice with PBS prior to seeding 600 000 cells per 35 mm culture plate. Cells were incubated for 4 h at 37 °C in medium containing 10 % (v/v) FBS prior to stimulation. Cells were washed twice with PBS and incubated for 15 min in 1 ml of McCoy's 5A medium containing no FBS, 10 % FBS, 50 ng of heterodimeric HGF, 50 ng of sc-HGF, 50 ng of *in vitro* hepsin-digested HGF, 50 ng of heterodimeric HGF plus 500 ng of anti-HGF neutralizing antibody, or 50 ng *in vitro* hepsin-digested HGF plus 500 ng of anti-HGF neutralizing antibody.

Hepsin inhibition by recombinant HAI-1 and HAI-2

Hepsin inhibition assays were performed in 100 μl final assay volume, including 50 mM Tris/HCl, pH 7.9, 250 mM NaCl and 1 nM hepsin in 96-well round-bottom Falcon microtitre plates (Becton Dickinson). Concentrations of HAI-1 and HAI-2 were varied between 0 nM and 500 nM. Enzyme and inhibitor were pre-incubated for 2 h at room temperature prior to starting the reaction with VLR-*p*NA (H-D-Val-Leu-Arg-*p*-nitroaniline dihydrochloride) substrate (Chromogenix, Milan, Italy) at a final concentration of 500 μM (apparent K_m 300 μM). Hydrolysis of the VLR-*p*NA substrate was followed by monitoring the amount of *p*NA formed at 405 nm at 21 °C using a SpectraMax Plus 384 96-well plate reader (Molecular Devices). The linear rates of absorbance increase were expressed as percentage (%) activities. Inhibitor activities were calculated as the concentration of inhibitor giving 50 % inhibition (IC_{50}) of the uninhibited enzyme activity (100 %). At least three independent measurements were performed.

RESULTS

Substrate specificity of recombinant human hepsin as determined by PS-SCL screening

To rapidly profile the general primary (P1) and extended (P2–P4) substrate specificity of human hepsin, we screened libraries of peptide substrates previously used to investigate the specificity of other extracellular human serine proteases [36–40]. The screening was performed using PS-SCLs, which consist of peptide pools with the individual structure Ac-P4-P3-P2-P1-ACC (7-amino-4-carbamoylmethylcoumarin) [29]. A total of 157 757 peptides were tested for their ability to be cleaved by hepsin *in vitro*. Applying this peptide profiling method, co-operative binding of amino acids at P1–P4 was not addressed.

The first round of screening examined peptides with defined amino acids at the P1 position. Hepsin strongly preferred to cleave substrates with an arginine rather than a lysine at the P1 position (Figure 1). Based on this result, three additional libraries were tested, which contained an arginine at the P1 position and randomized amino acids at the P2, P3 or P4 positions in a defined manner. The results obtained indicated that the extended (P2–P4) substrate specificity of hepsin is unique when compared with profiles obtained for other extracellular serine proteases including thrombin, plasmin, tryptases βI and βII, MT-SP1/matriptase, prostasin and KLK (kallikrein) 4 [36–40]. The P2 position showed a preference for the polar amino acids threonine or asparagine and a preference for the aliphatic residues leucine, valine and isoleucine. Hepsin discriminated against smaller amino acids, such as alanine and glycine, and against aspartate and glutamate. At the P3 position, hepsin exhibited a moderate preference for the basic amino acids lysine, arginine and histidine, or polar amino acids such as glutamine, serine and threonine. Hepsin moderately discriminated against aliphatic amino acids at the P3 position, such as valine, isoleucine or leucine, and aromatic amino acids including phenylalanine, tryptophan and tyrosine. Peptides with proline or aspartate at the P3 position also made sub-optimal substrates. At the P4 position, hepsin discriminated against aspartate and glutamate and showed a preference for lysine, arginine or proline.

Amidolytic activity of hepsin towards single tetrapeptide substrates

To validate the results obtained from the PS-SCL screen and to investigate hepsin substrate specificity in more detail, a limited number of individual peptides of the general structure Ac-P4-P3-P2-P1-AMC were synthesized and tested for their ability to serve as substrates for hepsin. The amino acid sequences in these peptides were selected based on the PS-SCL screening results: lysine or proline at P4; lysine or glutamine at P3; asparagine, threonine or leucine at P2; and arginine at P1 (see Figure 1; horizontal lines mark the arbitrary threshold set to >2000 units). Thus, 12 theoretically favoured tetrapeptide combinations were extracted: KKNR, KKTR, KKLR, KQNR, KQTR, KQLR, PQNR,

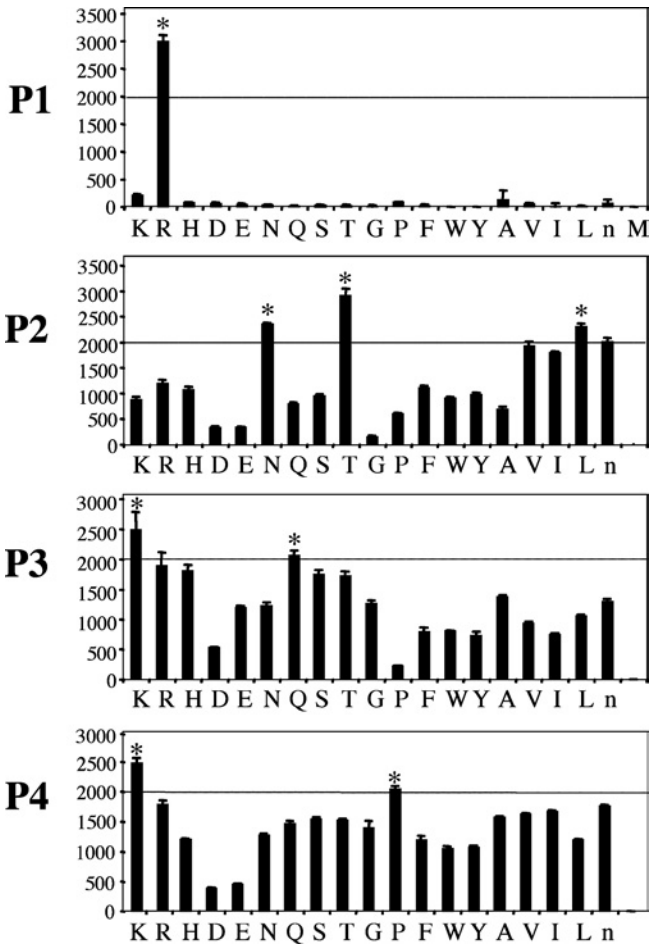

Figure 1 P1–P4 substrate specificity of recombinant human wild-type hepsin as determined using a PS-SCL screen

A PS-SCL screen was used to determine the amino acid specificity at each of the S1–S4 subsites. The *y*-axis shows pM ACC-fluorophore released/s (conversion factor RFU to pM: 0.000723). The primary (P1)-specificity was determined with a P1-complete diverse library containing 137180 different peptides. The P2–P4 specificities were determined with P1/R-fixed libraries, where P2, P3 or P4 was randomized in a defined manner. Each P1/R-P2, -P3, or -P4 library contained 6859 different peptides. A total of 157757 peptide combinations were screened. The one-letter code is used to denote the 20 natural amino acids, minus cysteine and plus norleucine (n; 2-aminohexanoic acid) on the *x*-axis. Hepsin preferred arginine over lysine in the P1 position and exhibited moderate specificity in the P2, P3 and P4 positions. Favoured amino acids are indicated by * after setting an arbitrary threshold at 2000 pM ACC-fluorophore released per s (depicted as a horizontal line).

PQTR, PQLR, PKNR, PKTR and PKLR. We assumed that hepsin is likely to activate other extracellular trypsin-like proteins by specific cleavage, a common function of serine proteases that are involved in, for example, blood coagulation. Such cleavage sites are located in the activation domains of trypsin-like precursors and are usually solvent exposed. The extracted peptide combinations obtained above were therefore run against a database of proteins that contain cleavage sites for human serine proteases (MEROPS; http://merops.sanger.ac.uk) to filter for putative biologically relevant P4–P1 recognition sites (see also Table 1). Interestingly, out of these 12 peptide combinations, only two matched with known P4–P1 sequences in serine protease-fold proteins: KQLR, the P4–P1 sequence found in the activation domain of the HGF precursor; and KKTR, the reported P4–P1 sequence for factor Va inactivation by Protein C [41].

In addition to KQLR and KKTR, a subset of tetrapeptides containing less favoured amino acids, such as aspartate at P4, valine or proline at P3 and glycine, arginine or aspartate at P2, were evaluated as amidolytic substrates: (i) DQLR, a peptide with a predicted non-favoured amino acid at P4; (ii) DPGR, a potentially non-preferred P4–P2 sequence, with P3–P1 present in the activation domain of the plasminogen precursor; and (iii) PVDR, which has the P4–P1 sequence present in the activation domain of the hepsin precursor with P3 and P2 predicted to be less favoured. The latter peptide was included to investigate if human hepsin preferentially undergoes auto-activation as has been reported to occur for mouse hepsin *in vitro* [7].

The relative amidolytic activities of hepsin for these peptides are shown in Figure 2. When compared with the KQLR peptide (relative activity set to 100 %), weaker activities were detected with the peptides DQLR (19 %) and KKTR (13 %) with little activity toward the 'hepsin-like'-peptide PVDR (6 %). Essentially no enzymatic activity was observed with the 'plasminogen-like' peptide DPGR (1 %). For the KQLR peptide, the apparent K_m value was determined to be 20 ± 6 μM. The catalytic efficiency for this peptide was high, which is illustrated by the steady-state kinetic specificity constant (k_{cat}/K_m), determined as 8.300×10^5 $M^{-1} \times s^{-1}$. Owing to the poor substrate activities, kinetic parameters for all other peptides were not determined here. We found that the 'hepsin-like' peptide PVDR was not a kinetically favoured substrate for human hepsin in this assay when compared to KQLR. Overall, the amidolytic activity of hepsin towards the individual peptides matched closely with the PS-SCL results (Figure 1), which does not address co-operative binding of amino acids at P4–P1. Although hepsin activity with the KKTR peptide was weak, this finding might be due to negative co-operativity of the two lysine residues at the P4 and P3 positions as seen previously with MT-SP1/matriptase [37].

X-ray crystal structure of human hepsin with a KQLR-cmk peptide inhibitor

To understand the specificity of hepsin for individual peptide substrates in greater detail, we determined the X-ray crystal structure of hepsin bound to the kinetically favoured KQLR peptide substrate [Figure 3 (data shown in Supplementary Table 1 http://www.BiochemJ.org/bj/390/bj3900125add.htm Protein DataBank accession number 1Z8G)].

For the purpose of this experiment, we used a KQLR-cmk derivative of the KQLR peptide that acts as an irreversible inhibitor of hepsin activity. We found that the overall structure of hepsin matched the structural features that were previously reported by Somoza et al. [30] (Figure 3A). Nucleophilic attack on the cmk moiety resulted in a covalent interaction between the peptide and catalytic residues Ser^{195} and His^{57} (chymotrypsin numbering system; Ser^{353} and His^{203} in hepsin respectively) (Figure 3B).

The first three substrate residue side chains, P1–P3, were well ordered in the crystal structure and had good electron density. The P1-arginine binds in a deep S1 pocket and interacts through a salt bridge with Asp^{189} (Asp^{347} in hepsin), which is characteristic for trypsin-like proteases. A preference for arginine over lysine at the P1 position is common for trypsin-like proteases that contain a serine-to-alanine substitution at residue 190 (Ala^{348} in hepsin) [3]. The P2 leucine residue binds into a dimple on the surface of the hepsin protein and makes van der Waals interactions with residues 57, 94, 99, 102, 214 and 215 (His^{203}, Tyr^{243}, Asn^{254}, Asp^{257}, Ser^{376} and Trp^{377} in hepsin) of the S2 pocket. This dimple is found in a large open area underneath the uncharacteristically long '90s loop' [30] of hepsin. Because of its size, this open area has the ability to accommodate and potentially interact with large amino

Table 1 Clustering of activation cleavage site sequences present in human serine-protease-fold zymogens (precursors) that were ranked based on PS-SCL results

We identified human sequences containing the trypsin domain, which represent potential macromolecular substrates for human hepsin from Pfam (http://pfam.wustl.edu) and MEROPS (http://merops.sanger.ac.uk). We selected predicted extracellular serine protease-fold zymogens that contained an arginine at the P1 position and a hydrophobic residue (I, V) at the P1′ position in their precursor activation domains. These proteins were subsequently clustered according to the total number of preferred amino acid residues at the P4 (K, P), P3 (K, Q), P2 (T, N, L), and P1 (R) position based on the PS-SCL results (as indicated by * in Figure 1) with four, three, two or one favoured amino acid residues in Clusters A, B, C and D, respectively. Note that the activation site sequence KQLR, which is present in the HGF precursor, is the only sequence found in a zymogen where all four positions are preferred†.

Precursor activation cleavage sites					Potential macromolecular substrate name (SWISS-PROT accession number)
P4	P3	P2	P1	P1′–P4′	
Cluster A†					
K	Q	L	R	VVNG	HGF precursor (P14210)
Cluster B					
P	Q	F	R	IKGG	‡Tissue-type plasminogen activator precursor (P00750)
P	Q	A	R	ITGG	Prostasin precursor (Q16651)
P	Q	G	R	IVGG	‡Coagulation factor VII precursor (P08709)
S	K	L	R	VVGG	HGF-like protein precursor (P26927)
Cluster C					
P	S	S	R	VVNG	Elastase IIIB precursor (P08861)
P	S	S	R	IVGG	Matriptase-2 precursor (Q8IU80)
P	V	Q	R	ILGG	Haptoglobin precursor (P00738)
P	V	Q	R	ILGG	Haptoglobin-related protein precursor (P00739)
P	V	D	R	IVGG	TMPRSS1 (hepsin) precursor (P05981)
P	R	G	R	ILGG	Complement factor D precursor (P00746)
K	I	K	R	IYGG	HGF activator-like protein (hyaluronan binding protein 2) precursor (Q14520)
R	Q	K	R	IIGG	Neurotrypsin precursor (P56730)
I	Q	S	R	IVGG	Tissue kallikrein 1 precursor (P06870)
I	Q	S	R	IVGG	Tissue kallikrein 2 precursor (P20151)
R	Q	S	R	IVGG	TMPRSS2 (epitheliasin) precursor (O15393)
R	Q	A	R	VVGG	MT-SP1/matriptase precursor (Q9Y5Y6)
R	K	S	R	VVDL	§Platelet-derived growth factor C (PS00249)
E	K	Q	R	IIGG	Complement C1s component precursor (P09871)
I	K	P	R	IVGG	Factor XI precursor (P03951)
S	M	T	R	VVGG	Coagulation factor XII precursor (P00748)
G	E	T	R	IIKG	Kallikrein 11 precursor (Q9UBX7)
D	F	T	R	VVGG	‡Coagulation factor IX precursor (P00740)
Y	V	T	R	VVGG	Elastase 2A precursor (P08217)
T	S	T	R	IVGG	Plasma kallikrein precursor (P03952)
N	L	T	R	IVGG	‡Coagulation factor X precursor (P00742)
Q	S	L	R	IVGG	DESC1 (differentially expressed squamous cell carcinoma gene 1) precursor (Q9UL52)
Q	L	N	R	VVGG	Brain-specific serine protease 4 precursor (Q9GZN4)
Cluster D					
Y	S	S	R	IVGG	TMPRSS3 (TAGD-12) precursor (P57727)
I	T	S	R	IVGG	Testisin precursor (Q9Y6M0)
L	A	S	R	IVGG	TMPRSS5 (spinesin) precursor (Q9H393)
I	L	S	R	IVGG	Prostate specific antigen (PSA) precursor (P07288)
R	Y	S	R	ITGG	GPI (glycoprotein 1) -SP1 precursor (Q6UWB4)
F	S	Q	R	IVNG	Chymotrypsin-like protease CTRL-1 precursor (P40313)
S	E	Q	R	ILGG	Human-airway trypsin-like protease (HAT) precursor (O60235)
L	S	A	R	VVGG	Caldecrin precursor (Q99895)
T	N	A	R	VVGG	Elastase I precursor (Q9UNI1)
L	R	P	R	IIGG[b]	HGF activator precursor (Q04756)
K	T	P	R	VVGG	TMPRSS4 (MT-SP2) precursor (Q9NRS4)
S	S	S	R	VVHG	Elastase IIIA precursor (P09093)
M	N	K	R	ILGG	Atrial natriuteric peptide-converting enzyme (corin) precursor (Q9Y5Q5)
R	R	K	R	IVGG	Complement factor I precursor (P05156)
G	G	V	R	IVGG	Acrosin precursor (P10323)
M	T	G	R	IVGG	MSPL (membrane-type mosaic serine protease) (Q9BYE2)
A	G	G	R	IVGG	Tryptase gamma precursor (Q9NRR2)
I	D	G	R	IVEG	‡Prothrombin precursor (P00734)
C	P	G	R	VVGG	‡Plasminogen precursor (P00747)

† Note that the sequence KKTR is a protease cleavage site present in factor Va (KKTR↓NLKK; amino acids 303–310) and cleavage at this site by human protein C causes inactivation of Factor Va [41].

‡ Complex activation pattern (two or more cleavage events).

§ Dimeric platelet-derived growth factor C (PDGF-CC) contains a trypsin-like cleavage site reported to be processed by tissue-type plasminogen activator (tPA) [64].

acid side chains. A glycine, on the other hand, would be the least preferred amino acid at P2 in order to build strong interactions and contribute to specificity. The P3 glutamine forms a 2.6 Å hydrogen bond with Tyr^{146} (hepsin residue Tyr^{301}).

The P4-lysine residue sits in the pocket formed above Trp^{215} (hepsin residue Trp^{377}). Even though the electron density found for the P4-lysine residue is weaker, the structure of the active-site cleft provides indications why lysine is a favoured amino acid at

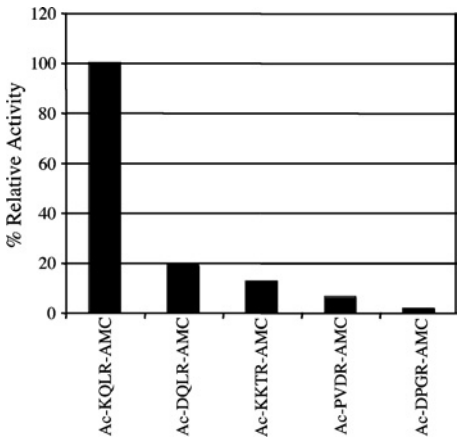

Figure 2 Relative amidolytic activity of hepsin toward selected AMC-coupled peptides

Peptides with the general structure Ac-P4-P3-P2-P1-AMC were tested in hepsin activity assays. The linear rates of fluorescence increase were expressed as percentage (%) activities when compared to the activity obtained with the Ac-KQLR-AMC peptide (set to 100 %). The peptides were assayed for hydrolysis by hepsin at least twice in triplicate.

this position compared to an aspartate or a glutamate. One side of the pocket is formed by residues Glu[97], Glu[98] and Asn[99] of the 90s loop (residues Glu[252], Glu[253], Asn[254] in hepsin), while the other side of the pocket is formed by residue 175 (residue Gln[331] in hepsin). The peptide backbone carbonyl oxygen atoms from the loops on both sides of the pocket as well as the oxygen atom of the Gln[331] side-chain point towards the S4 pocket, which gives an overall negative electrostatic potential to the S4 pocket. Although the structure does not show a specific interaction picked up by the P4 lysine residue, the electrostatic potential of the S4 pocket likely explains why a lysine is favoured at P4, while the acidic residues of aspartate and glutamate at P4 are not preferred.

Furthermore, the structure shows that the substrate peptide backbone makes hydrogen bonds to the hepsin peptide backbone in the active-site cleft. One of these hydrogen bonds is formed between the peptide bond nitrogen from the P3 position and the peptide bond carbonyl from Gly[216] (hepsin residue Gly[378]). A proline at the P3 position would be incapable of making this peptide backbone interaction and is therefore less favoured at this position.

Hepsin also features a large, hydrophobic S1′ pocket that is similar to the S1′ pockets of other trypsin-like serine proteases involved in extracellular proteolytic activation cascades, such as the blood coagulation cascade. This indicates that hepsin is likely to favour amino acids such as isoleucine or valine at the P1′ position.

In summary, a good correlation was found, with respect to the peptide substrate specificity of hepsin, between the results of the PS-SCL screen, the relative amidolytic activities for tetrapeptide-AMC substrates and the structural features of the active-site cleft. The KQLR peptide is a favoured substrate due to specific active-site interactions that are not achievable with a DQLR peptide or a DPGR peptide.

Identification and cleavage of macromolecular protein substrates

In order to identify preferred protein substrates for hepsin, we used the information obtained from the PS-SCL peptide profiling (Figure 1) and the testing of individual peptide substrates (Figure 2) to create a database of human sequences that contain a trypsin-like fold as extracted from Pfam (http://pfam.wustl.edu) and MEROPS. We then selected serine protease-fold like proteins

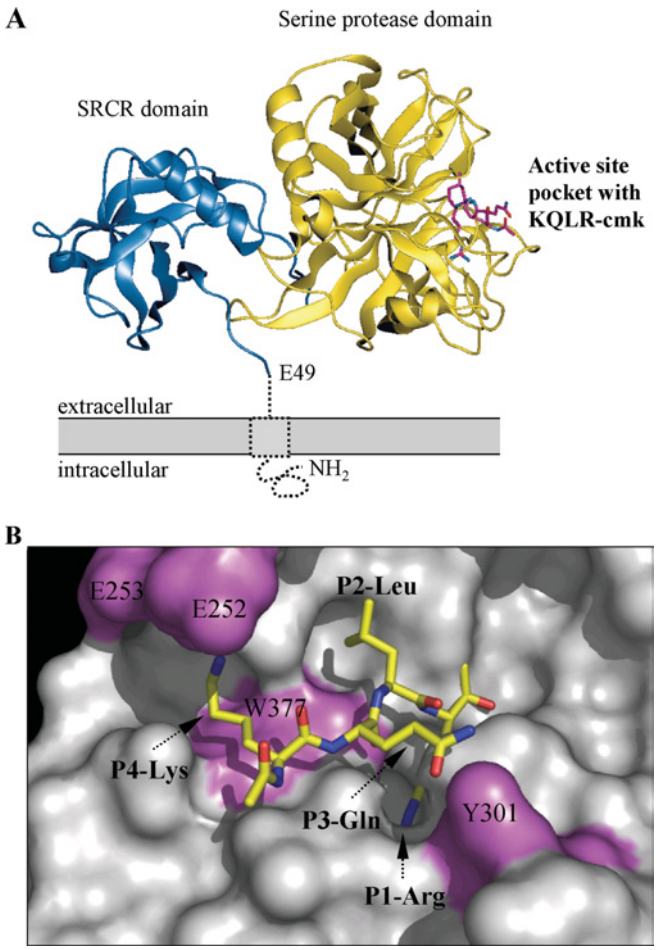

Figure 3 X-ray crystal structure of hepsin with KQLR-cmk

(**A**) A ribbon model representing the three-dimensional structure of hepsin. This structure is consistent with the report of Somoza et al. [30]. In its mature form, hepsin is a two-domain protein with a 109-amino-acid SRCR (extracellular scavenger receptor cysteine-rich domain) coupled via a single disulphide bond to a 255 amino-acid serine protease domain. The protein is predicted to be anchored to the membrane through a stretch of 27 amino acids at its N-terminus that potentially forms an α-helix (depicted as a broken line). The protease active site is formed in a cleft between the two six-stranded β-barrels. (**B**) A surface representation of the peptide-binding cleft of hepsin in complex with the KQLR-cmk peptide inhibitor. Highlighted in purple are hepsin residues Glu[252], Glu[253], Tyr[301] and Trp[377], selected amino acids of the active-site cleft. The P1-arginine is buried in the S1 pocket of hepsin, the P2-leucine lies in a dimple present in the S2 pocket, and the P3-glutamine interacts with the S3 pocket through a hydrogen bond with Tyr[301]. The P4 lysine, which shows weaker electron density, is modelled into the S4 pocket above Trp[377]. The figure was generated with PyMOL (DeLano Scientific LLC, South San Francisco, CA, U.S.A.).

that had (i) a known or predicted P4–P1 sequence in their activation domain, (ii) an arginine at the P1 position, and (iii) a hydrophobic amino acid (isoleucine or valine) at the P1′ position to bind to the large, hydrophobic S1′ pocket of hepsin.

These proteins were clustered and ranked based on matches with the preferred amino acids at the P1, P2, P3 and P4 positions as obtained by the PS-SCL screen (Table 1). Cluster A contained candidate substrates where all four cleavage site amino acids in the activation domain matched favoured ones, such as lysine or proline at P4, lysine or glutamine at P3, threonine, asparagine or leucine at P2 and arginine at P1. Clusters B and C contained precursors where only three or two positions matched a favoured amino acid respectively. Cluster D contained precursors that have an arginine at the P1 position and isoleucine or valine at the P1′ position. We then tested the ability of hepsin to cleave selected

Table 2 Hepsin cleavage of macromolecular protein substrates *in vitro*

Hepsin digests were performed on extracellular human trypsin-like precursors that were identified as candidate substrate proteins containing activation domain sequences, which were ranked according to preferred P1, P2, P3 or P4 amino acids based on PS-SCL screening results (Table 1).

Activation cleavage site of trypsin-like precursors					Potential macromolecular substrate [name (SWISS-PROT accession number)]	Cleavage*	Specificity†
P4′	P3′	P2′	P1′	P1′–P4′			
K	Q	L	R	VVNG	HGF precursor (P14210)	Yes	Yes
P	Q	G	R	IVGG	Coagulation factor VII precursor (P08709)	Yes	Yes
S	M	T	R	VVGG	Coagulation factor XII precursor (P00748)	Yes	Yes
D	F	T	R	VVGG	Coagulation factor IX precursor (P00740)	Yes	Yes
N	L	T	R	IVGG	Coagulation factor X precursor (P00742)	No	–
P	V	D	R	IVGG	Hepsin (TMPRSS1) precursor (P05981)	Yes‡	Yes
C	P	G	R	VVGG	Plasminogen precursor (P00747)	No	–

* Cleavage was determined by *in vitro* digest and assessed by SDS/PAGE analysis under reducing conditions.

† Specificity was determined by N-terminal sequencing of cleavage products and comparison with commercially available activated forms (factor VIIa, XIIa, IXa, Xa, and plasmin). The N-terminus of hepsin-processed factor XII light chain was N_2H-VVGG and that of the hepsin-processed VII heavy chain was N_2H-IVGG.

‡ Predominantly in the presence of detergents and/or high hepsin concentrations during purification (R. J. Austin, unpublished work).

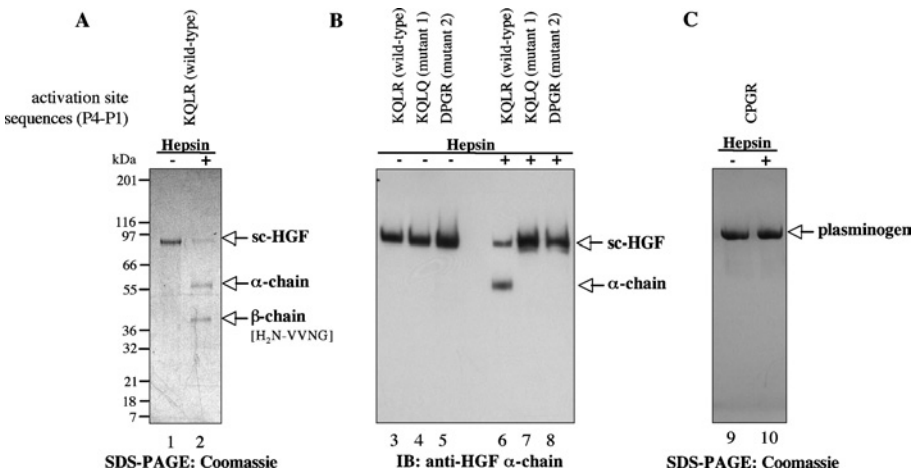

Figure 4 Hepsin digests of wild-type sc-HGF precursor (A, B), activation site mutant HGF precursors (B) and plasminogen (C)

Precursors were incubated for 2 h at 37 °C in the absence (−) (lanes 1, 3, 4, 5, 9) or in the presence (+) (lanes 2, 6, 7, 8) of hepsin. The cleavage products were analysed by SDS/PAGE under reducing conditions. (**A**) Wild-type HGF precursor (lane 1; 94–85 kDa; activation domain sequence KQLR↓VVNG) is converted into α/β-heterodimeric HGF (54 kDa α-chain; 26 kDa β-chain) by hepsin (lane 2). The N-terminal amino acid sequence of the β-chain fragment was determined as H_2N-VVNG, which indicates that the protein was specifically processed through cleavage between the P1 and P1′ residues in the activation domain. (**B**) Immunoblot (IB) analysis using an anti-HGF α-chain-specific antibody: two activation domain mutants of the HGF precursor (mutant 1, KQLQ↓VVNG; lanes 4 and 7, mutant 2, DPGR↓VVNG, lanes 5 and 8) are not processed by hepsin. In a parallel digest, wild-type HGF precursor was converted into α/β-heterodimeric HGF (lanes 3 and 6). (**C**) Plasminogen (lane 9; activation domain sequence CPGR-VVGG) is not converted to plasmin when incubated with hepsin (lane 10). The non-cleavage of plasminogen or the sc-HGF mutant precursors shows that the P4–P1 residues are specificity determinants for HGF processing by hepsin.

candidate substrates from each cluster that were reported to be localized in the extracellular environment of the adult human liver, such as the HGF precursor, factors VII, XII, X, IX and plasminogen. These candidates were subjected to standard *in vitro* digests at a molar ratio of 1:30 (hepsin/precursor protein). The digests were analysed by SDS/PAGE under reducing conditions and the cleavage site on the macromolecular protein substrates was determined by N-terminal amino acid sequencing of the detected protein bands. The results are summarized in Table 2.

As expected from the ranking of potential macromolecular substrates, the sc-HGF (cluster A) and factor VII (cluster B), which contain either the KQLR↓VVNG or the PQGR↓IVGG sequence in their activation domain cleavage sites, made excellent protein substrates for human hepsin, as seen by the appearance of two major protein bands in the gel (Figure 4A, factor VII digest not shown). Additional data confirmed that the cleavage of hepsin is specific and that sc-HGF is not processed elsewhere. The HGF β-chain (apparent molecular mass 26 kDa) was confirmed by sequencing to have the correct N-terminal amino acids (H_2N-VVNG). Due to an N-terminal pyroglutamate [42], the N-terminus of the HGF α-chain cannot be sequenced under the standard conditions. However, digested HGF precursor was analysed by immunoblotting using an anti-HGF α-chain-specific antibody, which detected the presence of an intact HGF α-chain fragment (apparent molecular mass 54 kDa). The wild-type HGF precursor was the only suggested natural candidate substrate with all four positions matching favoured amino acids as predicted from the PS-SCL screen (cluster A) and that contained valine at the P1′ position.

To examine in detail the interaction between hepsin and a protein substrate and to investigate the importance of P1–P4 residues for specificity, two activation domain mutants of

sc-HGF were created. 'Mutant 1' contained a mutated P1 residue, KQLQ (R494Q), and 'mutant 2' the P4–P1 ('plasminogen-like')-sequence DPGR (K491D/Q492P/L493G). Neither of these precursors was processed by hepsin *in vitro* (Figure 4B). These findings are consistent with the peptide profiling results and the kinetics for the individual peptide substrates, and indicate that the macromolecular substrate specificity of hepsin for sc-HGF is determined to a great extent by the P4–P1 amino acid residues.

The ability of hepsin to cleave precursors in clusters C and D (Table 1) was also examined. Factors XII, X and IX were three representative candidate substrates from cluster C, where only two positions had to match PS-SCL results. Specific cleavage in the activation domain of factor XII was observed, with an N-terminal sequence of the light chain of H_2N-VVGG, whereas factor X was not cleaved at all (Table 2). Factor IX was processed by hepsin, and the products formed corresponded to the band pattern obtained with commercially available factor IXa. Plasminogen, a precursor from cluster D, contains the non-preferred sequence CPGR↓VVGG (where ↓ denotes the cleavage site) in its activation domain, a P4–P1 sequence similar to the sc-HGF 'mutant 2'. Like this DPGR mutant, plasminogen was not cleaved under the conditions described (Figure 4C).

The results indicate that the macromolecular substrate profile of hepsin is restricted and that the presence of specific amino acids at the P4–P1 positions of the activation domain cleavage site is critical for processing by hepsin.

Hepsin-cleaved wild-type HGF activates the HGF receptor (c-Met) in SKOV-3 ovarian cancer cells

To model the *in vivo* functions of hepsin, we speculated that the potential role of dysregulated hepsin expression might be due mainly to its potential to activate the trypsin-like growth factor HGF, which contains the kinetically and structurally preferred peptide sequence KQLR in its activation domain. Furthermore, the HGF precursor might be an endogenous substrate for hepsin, because it shows the potential to co-localize with hepsin to the extracellular matrix of ovarian and prostate cancers. HGF is a serine-protease-like growth factor and is the ligand of the c-Met receptor tyrosine kinase [43]. Specific cleavage of the sc-HGF precursor by hepsin at the activation site should result in α/β-heterodimeric HGF that is capable of causing autophosphorylation of the c-Met receptor. To demonstrate that the *in vitro* hepsin-processed HGF (Figure 4A) is functionally active, the hepsin-cleaved and mock-cleaved HGF were assessed for their ability to stimulate tyrosine autophosphorylation of the c-Met receptor in SKOV-3 ovarian cancer cells. As shown in Figure 5(A), hepsin-cleaved HGF, but not mock-digested HGF, caused c-Met-specific tyrosine phosphorylation at residues 1234 and 1235. This phosphorylation was specifically due to HGF activity, because the addition of an anti-HGF neutralizing antibody prevented c-Met phosphorylation (Figure 5B). These findings suggest that the HGF precursor might be an endogenous protein substrate for hepsin, especially with respect to hepsin over-expression in ovarian tumours [12,21].

Hepsin is inhibited by human recombinant HAI-1 and HAI-2

To explore further factors that might affect the activity of hepsin in the extracellular environment of cells, we investigated the ability of recombinant HAIs to inhibit hepsin activity *in vitro*. HAI-1 and HAI-2 are known Kunitz-type inhibitors of extracellular serine proteases that are generally considered to regulate HGF activation [44–46]. These HAIs exhibit potent inhibition of hepsin activity with IC_{50} values of 12 nM (HAI-1) and 4 nM (HAI-2)

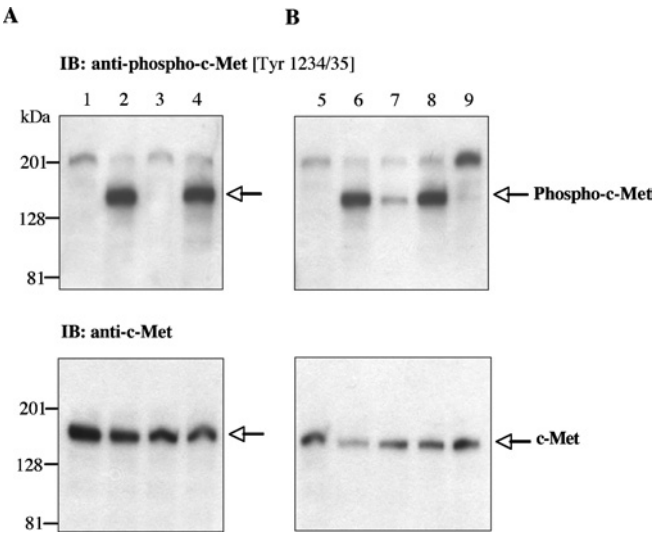

Figure 5 Hepsin-processed wild-type HGF stimulates c-Met receptor tyrosine auto-phosphorylation in SKOV-3 ovarian cancer cells

(**A**) SKOV-3 cells were treated for 15 min at 37 °C with (lane 1) 10 % FBS-containing medium, or in serum-free medium with heterodimeric recombinant HGF (lane 2), sc-HGF (lane 3), and hepsin-processed HGF (lane 4). (**B**) SKOV-3 cells in serum-free medium (lane 5) were treated with heterodimeric recombinant HGF (lane 6), heterodimeric HGF plus anti-HGF neutralizing antibody (lane 7), hepsin-processed HGF (lane 8), and (lane 9) hepsin-processed HGF plus anti-HGF neutralizing antibody (lane 9). Immunoblot (IB) analysis was performed using anti-phospho-tyrosine c-Met antibody or anti-c-Met antibody. The arrow points to the 145 kDa β-chain of c-Met receptor (active form) under reducing conditions.

Table 3 Inhibition of human hepsin by recombinant human HAIs

The serine protease inhibitors HAI-1 and HAI-2 exhibited complete inhibition (99 %) of hepsin serine protease activity towards the peptide substrate VLR-*p*NA *in vitro*. The values shown represent data fitted to a four-parameter equation, from which the IC_{50} and maximal percentage inhibition were calculated.

Kunitz-type inhibitor	IC_{50} (nM)
HAI-2	4 ± 0.2
HAI-1	12 ± 0.5

(Table 3). Therefore HAI-1 and HAI-2 are potential endogenous hepsin inhibitors that can modulate extracellular hepsin activity.

DISCUSSION

To better understand the biological functions of the human serine protease hepsin *in vivo*, we identified potential natural protein substrates using a synthetic PS-SCL screen. Hepsin almost exclusively cleaved substrates with an arginine at the P1 position (Figure 1), which had been described previously [2,47,48]. The screen further elucidated an amino acid preference profile for tetrapeptide substrates, where threonine, asparagine or leucine were most favoured in the P2 position, lysine or glutamine in the P3 position and lysine or proline in the P4 position. Hepsin discriminated against smaller amino acids (alanine or glycine) in the P2 position and against the acidic amino acids aspartate or glutamate in the P4 position. The extended substrate specificity of hepsin was unique when compared to previously obtained PS-SCL profiles of other extracellular serine proteases [36–40].

The amino acid preference profile for the P4–P1 positions was used to identify preferred macromolecular protein substrates for hepsin. We assumed that hepsin is likely to cleave other human

trypsin-like precursor proteins that contain a hydrophobic amino acid, such as isoleucine or valine, at the P1′ position and co-localize with hepsin in the extracellular environment. We searched the MEROPS and Pfam databases for proteins that matched these criteria and clustered them according to the number of matches with the favoured P1, P2, P3 or P4 amino acids (Table 1). The subsequent ranking of the clusters allowed us to identify preferred macromolecular substrate candidates, such as sc-HGF (cluster A, where all four positions matched), factor VII (cluster B, where three position matched) and non-substrate candidates, such as plasminogen (cluster D, where only P1 matched).

Based on this ranking of clustered candidates, we determined the relative amidolytic activity of hepsin towards individual AMC-coupled tetrapeptides such as KQLR (in sc-HGF) and KKTR (in factor Va), as well as the peptide substrates DPGR (P1–P3 in plasminogen), PVDR (in pro-hepsin) and DQLR, which contained the non-preferred aspartate in the P4 position. Although the factor Va cleavage site does not continue with isoleucine or valine at P1′ [41], the KKTR peptide was also tested. The amidolytic activities of hepsin towards these peptide substrates correlated well with the PS-SCL screen results. Hepsin showed the highest activity with the peptide KQLR (100 % activity). The DPGR peptide, where P1–P3 matched with plasminogen, showed the weakest substrate activity of all (1.0 % activity relative to KQLR), although it contained an arginine at P1. As an exception, KKTR (13 % activity relative to KQLR), a combination predicted to be a relatively preferred amidolytic substrate, showed weaker activity when compared to DQLR (19 % activity relative to KQLR). The effects of non-constructive co-operativity between neighbouring residues was not considered in the PS-SCL screen, but could be addressed by substrate phage display [49].

In the past, several chromogenic and fluorogenic peptide substrates have been described for human and rat hepsin [2,47,48]. For comparison, the tripeptide QRR has been reported by Zhukov et al. [48] as a relatively favoured amidolytic peptide substrate for rat hepsin. In order to be consistent with our analysis, we did not compare the activity of tetrapeptides with tripeptide substrates. Furthermore, there were no human trypsin-like precursors found that contained an arginine at the P2 position in their activation domains (see Table 1).

Consistent with the amidolytic substrate activity of the KQLR peptide, hepsin processed the macromolecular substrate sc-HGF specifically in the activation domain (Table 2, Figure 4). The processing of such trypsin-like growth factors was first suggested by Wu [10]. Furthermore, hepsin specifically cleaved the coagulation Factors VII, IX and XII, but not factor X, even though it also possesses two favoured amino acids (Table 1: cluster C). Our results show that P1–P4 amino acids are important for hepsin–substrate interactions (Figure 4). When compared to factor XII or IX cleavage results, the factor X result suggests that additional structural elements of proteins can cause steric hindrance and therefore affect their ability to be hepsin substrates. Based on our ranking of potential natural substrates (Table 1), it would be interesting to test whether hepsin cleaves other candidate substrates derived from cluster B, such as the precursor of prostasin, a protease found to be up-regulated in prostate and ovarian cancer [18,50].

We also investigated hepsin auto-activation *in vitro* as this was reported previously to occur for mouse hepsin [7]. The kinetic parameters obtained for the individual 'hepsin-like' peptide PVDR showed that this peptide sequence, although a poor substrate for human hepsin, can be turned over (Figure 2). Accordingly, active hepsin is capable of activating other hepsin zymogens, but only *in vitro* in the presence of detergent or at high protein concentrations (Table 2). However, transactivation might be more favorable *in vivo* because hepsin is membrane-bound and therefore may co-localize with other hepsin molecules.

In accordance with the finding that the individual 'plasminogen-like' peptide DPGR is not a hepsin substrate, plasminogen was also not processed by hepsin (Figure 4C). Moreover, two sc-HGF mutants containing either a mutated P1 residue, KQLQ (R494Q) or the 'plasminogen-like' sequence DPGR (K491D/Q492P/L493G) were not processed by hepsin (Figure 4B). This demonstrates that macromolecular substrates of hepsin are greatly selected by their P4–P1 residues, and maybe to a lesser extent by macromolecular exosite interactions with hepsin. Thus the PS-SCL screen was useful to identify kinetically favoured candidate substrates, in this case sc-HGF. Nonetheless, other extracellular subtrates, that have not been listed or tested, cannot be excluded.

The X-ray crystal structure of hepsin with the KQLR-cmk inhibitor (Figure 3) showed that the amino acid residues bind well in the corresponding S4-S1 pockets of the substrate-binding cleft of hepsin (Figure 3B). The P2-leucine and the P3-glutamine make especially good interactions with the amino acids of the S2 and S3 pockets. The overall negative electrostatic potential of the S4 pocket explains why a basic residue at P4 (lysine) was favoured, while acidic residues, such as aspartate or glutamate, were not preferred. This suggests an explanation for our finding that hepsin shows a 5-fold decreased relative amidolytic activity towards the DQLR peptide when compared with the KQLR peptide. Theoretically, the structure of the active-site cleft explains the discrimination of the DPGR peptide as a substrate. A glycine at the P2 position is the least likely to interact with the residues in the S2 pocket and therefore cannot contribute to specificity. Similarly, a proline at the P3 position would be impaired in making peptide backbone interactions with the hepsin peptide backbone and is therefore less preferred in this position.

To model the putative *in vivo* functions of hepsin, particularly in tumorigenesis, we focused on the secreted precursor of the serine-protease-like growth factor HGF, which potentially co-localizes with the membrane-anchored hepsin to the extracellular environment. Inactive sc-HGF is processed by various serine proteases [42,51–55] to its active α/β-heterodimeric form, which acts as a pleiotropic cytokine by signalling exclusively through the proto-oncogenic c-Met receptor. Numerous studies have shown that the HGF (c-Met) system influences tumorigenesis, metastasis, tissue regeneration and morphogenesis (reviewed in [43]). Namely, ovarian and prostate cancers have been reported to frequently over-express the c-Met receptor (reviewed in [56]) and elevated levels of HGF in ascitic fluids of ovarian cancer patients have been found previously [57,58].

In the present study, we demonstrated that the wild-type HGF precursor is specifically cleaved by hepsin *in vitro* to form a disulphide-linked α/β-heterodimer (Figure 4A). The hepsin-processed wild-type HGF is functionally active in that it induces c-Met receptor tyrosine auto-phosphorylation in SKOV-3 epithelial ovarian cancer cells (Figure 5).

We showed that hepsin activity can be potently inhibited by the epithelial expressed Kunitz-type HAI-1 and HAI-2 (Table 3). HAIs inhibit a number of serine proteases involved in various extracellular protease cascades, and HAI-1 splice variants have been suggested to be potential endogenous regulators of cell-surface proteolysis that modulate the activity of the HGF (c-Met) receptor system [59,60] or could influence ovarian cancer patient outcome [44,45]. Recent findings confirm inhibition of hepsin by HAI-1 and HAI-2 and also document cleavage of sc-HGF by hepsin [61,62].

In summary, the observation that hepsin cleaves precursors of trypsin-like blood coagulation factors and sc-HGF specifically,

but not other trypsin-like precursors containing an arginine at the P1 position of their activation domains, shows that the macromolecular substrate profile of hepsin is restricted. Furthermore, we provide evidence for a potential mechanism by which hepsin proteolytic activity can promote ovarian and prostate tumorigenesis. These tumour types frequently over-express hepsin, HGF and c-Met, thus providing an environment for potential co-localization and dysregulated action. Furthermore, Klezovitch et al. [26] have shown recently that over-expression of mouse hepsin in prostate epithelium *in vivo* can cause the disorganization of the basement membrane. They also showed that hepsin over-expression causes metastasis to the liver, lung and bone in a mouse model of non-metastasizing prostate cancer. Our findings suggest that over-expressed hepsin may regulate processes mediating growth factor function, thereby influencing events in the extracellular matrix, and possibly mediating the interactions between epithelial and mesenchymal cells. Hepsin can thereby contribute to tumour maintenance, progression of cancer and metastasis. In addition, the activation of blood coagulation factors, due to dysregulated hepsin expression in tumours, might contribute to venous thrombo-embolism, a complication commonly associated with malignant diseases [63]. It has been shown that hepsin over-expressing cells can cause thrombin formation and blood coagulation [5]. However, whether hepsin over-expression has an essential role in cancer-associated proteolytic cascades still needs to be determined. In addition to hepsin's putative functions in tumours, the prominent expression of human hepsin in the normal adult liver suggests roles for human hepsin that are not yet revealed in hepsin-deficient mice [8–10]. The lack of striking phenotypes of hepsin-deficient mice might be explained by redundant roles for hepsin in haemostasis, which can be compensated theoretically by other factors. Dissection of hepsin's biological activities in tumour and normal cells will require the development of inhibitors specific to hepsin as well as other cell-surface proteases.

This work was supported by a Cancer Grant (NIHCA72006) to C. S. C., A. S. B. and Y. C. from the National Institutes of Health.

REFERENCES

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

2 Kurachi, K., Torres-Rosado, A. and Tsuji, A. (1994) Hepsin. Methods Enzymol. **244**, 100–114

3 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

4 Wu, Q. (2003) Type II transmembrane serine proteases. Curr. Opin. Dev. Biol. **54**, 167–206

5 Kazama, Y., Hamamoto, T., Foster, D. C. and Kisiel, W. (1995) Hepsin, a putative membrane-associated serine protease, activates human factor VII and initiates a pathway of blood coagulation on the cell surface leading to thrombin formation. J. Biol. Chem. **270**, 66–72

6 Torres-Rosado, A., O'Shea, K. S., Tsuji, A., Chou, S. H. and Kurachi, K. (1993) Hepsin, a putative cell-surface serine protease, is required for mammalian cell growth. Proc. Natl. Acad. Sci. U.S.A. **90**, 7181–7185

7 Vu, T. K., Liu, R. W., Haaksma, C. J., Tomasek, J. J. and Howard, E. W. (1997) Identification and cloning of the membrane-associated serine protease, hepsin, from mouse preimplantation embryos. J. Biol. Chem. **272**, 31315–31320

8 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

9 Yu, I. S., Chen, H. J., Lee, Y. S., Huang, P. H., Lin, S. R., Tsai, T. W. and Lin, S. W. (2000) Mice deficient in hepsin, a serine protease, exhibit normal embryogenesis and unchanged hepatocyte regeneration ability. Thromb. Haemostasis **84**, 865–870

10 Wu, Q. (2001) Gene targeting in hemostasis: Hepsin. Front. Biosci. **6**, D192–D200

11 Li, Y., Yu, Z., Zhao, X. and Shen, S. H. (2005) Identification and characterization of hepsin/-TM, a non-transmembrane hepsin isoform. Biochim. Biophys. Acta **1681**, 157–165

12 Tanimoto, H., Yan, Y., Clarke, J., Korourian, S., Shigemasa, K., Parmley, T. H., Parham, G. P. and O'Brien, T. J. (1997) Hepsin, a cell surface serine protease identified in hepatoma cells, is overexpressed in ovarian cancer. Cancer Res. **57**, 2884–2887

13 Dhanasekaran, S. M., Barrette, T. R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K. J., Rubin, M. A. and Chinnaiyan, A. M. (2001) Delineation of prognostic biomarkers in prostate cancer. Nature (London) **412**, 822–826

14 Magee, J. A., Araki, T., Patil, S., Ehrig, T., True, L., Humphrey, P. A., Catalona, W. J., Watson, M. A. and Milbrandt, J. (2001) Expression profiling reveals hepsin overexpression in prostate cancer. Cancer Res. **61**, 5692–5696

15 Luo, J., Duggan, D. J., Chen, Y., Sauvageot, J., Ewing, C. M., Bittner, M. L., Trent, J. M. and Isaacs, W. B. (2001) Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. Cancer Res. **61**, 4683–4688

16 Stamey, T. A., Warrington, J. A., Caldwell, M. C., Chen, Z., Fan, Z., Mahadevappa, M., McNeal, J. E., Nolley, R. and Zhang, Z. (2001) Molecular genetic profiling of Gleason grade 4/5 prostate cancers compared to benign prostatic hyperplasia. J. Urol. **166**, 2171–2177

17 Welsh, J. B., Sapinoso, L. M., Su, A. I., Kern, S. G., Wang-Rodriguez, J., Moskaluk, C. A., Frierson, Jr, H. F. and Hampton, G. M. (2001) Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. Cancer Res. **61**, 5974–5978

18 Rhodes, D. R., Barrette, T. R., Rubin, M. A., Ghosh, D. and Chinnaiyan, A. M. (2002) Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer Res. **62**, 4427–4433

19 Chen, Z., Fan, Z., McNeal, J. E., Nolley, R., Caldwell, M. C., Mahadevappa, M., Zhang, Z., Warrington, J. A. and Stamey, T. A. (2003) Hepsin and maspin are inversely expressed in laser capture microdissectioned prostate cancer. J. Urol. **169**, 1316–1319

20 Riddick, A. C., Barker, C., Sheriffs, I., Bass, R., Ellis, V., Sethia, K. K., Edwards, D. R. and Ball, R. Y. (2003) Banking of fresh-frozen prostate tissue: methods, validation and use. BJU. Int. **91**, 315–323

21 Adib, T. R., Henderson, S., Perrett, C., Hewitt, D., Bourmpoulia, D., Ledermann, J. and Boshoff, C. (2004) Predicting biomarkers for ovarian cancer using gene-expression microarrays. Br. J. Cancer **90**, 686–692

22 Huppi, K. and Chandramouli, G. V. (2004) Molecular profiling of prostate cancer. Curr. Urol. Rep. **5**, 45–51

23 Lai, Y., Wu, B., Chen, L. and Zhao, H. (2004) Statistical method for identifying differential gene-gene coexpression patterns. Bioinformatics **20**, 3146–3155

24 Stephan, C., Yousef, G. M., Scorilas, A., Jung, K., Jung, M., Kristiansen, G., Hauptmann, S., Kishi, T., Nakamura, T., Loening, S. A. and Diamandis, E. P. (2004) Hepsin is highly over expressed in and a new candidate for a prognostic indicator in prostate cancer. J. Urol. **171**, 187–191

25 Zacharski, L. R., Ornstein, D. L., Memoli, V. A., Rousseau, S. M. and Kisiel, W. (1998) Expression of the factor VII activating protease, hepsin, in situ in renal cell carcinoma. Thromb. Haemostasis **79**, 876–877

26 Klezovitch, O., Chevillet, J., Mirosevich, J., Roberts, R. L., Matusik, R. J. and Vasioukhin, V. (2004) Hepsin promotes prostate cancer progression and metastasis. Cancer Cell **6**, 185–195

27 Vasioukhin, V. (2004) Hepsin paradox reveals unexpected complexity of metastatic process. Cell Cycle **3**, 1394–1397

28 Srikantan, V., Valladares, M., Rhim, J. S., Moul, J. W. and Srivastava, S. (2002) HEPSIN inhibits cell growth/invasion in prostate cancer cells. Cancer Res. **62**, 6812–6816

29 Harris, J. L., Backes, B. J., Leonetti, F., Mahrus, S., Ellman, J. A. and Craik, C. S. (2000) Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. Proc. Natl. Acad. Sci. U.S.A. **97**, 7754–7759

30 Somoza, J. R., Ho, J. D., Luong, C., Ghate, M., Sprengeler, P. A., Mortara, K., Shrader, W. D., Sperandio, D., Chan, H., McGrath, M. E. and Katz, B. A. (2003) The structure of the extracellular region of human hepsin reveals a serine protease domain and a novel scavenger receptor cysteine-rich (SRCR) domain. Structure **11**, 1123–1131

31 Lengyel, Z., Pal, G. and Sahin-Toth, M. (1998) Affinity purification of recombinant trypsinogen using immobilized ecotin. Protein Expression Purif. **12**, 291–294

32 Kissinger, C. R., Gehlhaar, D. K. and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search. Acta Crystallogr. Sect. D: Biol. Crystallogr. **55**, 484–491

33 Lu, D., Futterer, K., Korolev, S., Zheng, X., Tan, K., Waksman, G. and Sadler, J. E. (1999) Crystal structure of enteropeptidase light chain complexed with an analog of the trypsinogen activation peptide. J. Mol. Biol. **292**, 361–373

34 Brunger, A. T., Clore, G. M., Gronenborn, A. M. and Karplus, M. (1986) Three-dimensional structure of proteins determined by molecular dynamics with interproton distance restraints: application to crambin. Proc. Natl. Acad. Sci. U.S.A. **83**, 3801–3805

35 Matsudaira, P. (1990) Limited N-terminal sequence analysis. Methods Enzymol. **182**, 602–613
36 Backes, B. J., Harris, J. L., Leonetti, F., Craik, C. S. and Ellman, J. A. (2000) Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin. Nat. Biotechnol. **18**, 187–193
37 Takeuchi, T., Harris, J. L., Huang, W., Yan, K. W., Coughlin, S. R. and Craik, C. S. (2000) Cellular localization of membrane-type serine protease 1 and identification of protease-activated receptor-2 and single-chain urokinase-type plasminogen activator as substrates. J. Biol. Chem. **275**, 26333–26342
38 Harris, J. L., Niles, A., Burdick, K., Maffitt, M., Backes, B. J., Ellman, J. A., Kuntz, I., Haak-Frendscho, M. and Craik, C. S. (2001) Definition of the extended substrate specificity determinants for beta-tryptases I and II. J. Biol. Chem. **276**, 34941–34947
39 Shipway, A., Danahay, H., Williams, J. A., Tully, D. C., Backes, B. J. and Harris, J. L. (2004) Biochemical characterization of prostasin, a channel activating protease. Biochem. Biophys. Res. Commun. **324**, 953–963
40 Matsumura, M., Bhatt, A. S., Andress, D., Clegg, N., Takayama, T. K., Craik, C. S. and Nelson, P. S. (2005) Substrates of the prostate-specific serine protease prostase/KLK4 defined by positional-scanning peptide libraries. Prostate **62**, 1–13
41 Kalafatis, M., Rand, M. D. and Mann, K. G. (1994) The mechanism of inactivation of human factor V and human factor Va by activated protein C. J. Biol. Chem. **269**, 31869–31880
42 Peek, M., Moran, P., Mendoza, N., Wickramasinghe, D. and Kirchhofer, D. (2002) Unusual proteolytic activation of pro-hepatocyte growth factor by plasma kallikrein and coagulation factor XIa. J. Biol. Chem. **277**, 47804–47809
43 Birchmeier, C., Birchmeier, W., Gherardi, E. and Vande Woude, G. F. (2003) Met, metastasis, motility and more. Nat. Rev. Mol. Cell Biol. **4**, 915–925
44 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. and Sellar, G. C. (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
45 Kirchhofer, D., Peek, M., Li, W., Stamos, J., Eigenbrot, C., Kadkhodayan, S., Elliott, J. M., Corpuz, R. T., Lazarus, R. A. and Moran, P. (2003) Tissue expression, protease specificity, and Kunitz domain functions of hepatocyte growth factor activator inhibitor-1B (HAI-1B), a new splice variant of HAI-1. J. Biol. Chem. **278**, 36341–36349
46 Parr, C., Watkins, G., Mansel, R. E. and Jiang, W. G. (2004) The hepatocyte growth factor regulatory factors in human breast cancer. Clin. Cancer Res. **10**, 202–211
47 Tsuji, A., Torres-Rosado, A., Arai, T., Le Beau, M. M., Lemons, R. S., Chou, S. H. and Kurachi, K. (1991) Hepsin, a cell membrane-associated protease. Characterization, tissue distribution, and gene localization. J. Biol. Chem. **266**, 16948–16953
48 Zhukov, A., Hellman, U. and Ingelman-Sundberg, M. (1997) Purification and characterization of hepsin from rat liver microsomes. Biochim. Biophys. Acta **4**, 85–95
49 Matthews, D. J. and Wells, J. A. (1993) Substrate phage: selection of protease substrates by monovalent phage display. Science (Washington D.C.) **260**, 1113–1117
50 Mok, S. C., Chao, J., Skates, S., Wong, K., Yiu, G. K., Muto, M. G., Berkowitz, R. S. and Cramer, D. W. (2001) Prostasin, a potential serum marker for ovarian cancer: identification through microarray technology. J. Natl. Cancer Inst. **93**, 1458–1464
51 Miyazawa, K., Shimomura, T., Kitamura, A., Kondo, J., Morimoto, Y. and Kitamura, N. (1993) Molecular cloning and sequence analysis of the cDNA for a human serine protease reponsible for activation of hepatocyte growth factor. Structural similarity of the protease precursor to blood coagulation factor XII. J. Biol. Chem. **268**, 10024–10028
52 Miyazawa, K., Shimomura, T. and Kitamura, N. (1996) Activation of hepatocyte growth factor in the injured tissues is mediated by hepatocyte growth factor activator. J. Biol. Chem. **271**, 3615–3618
53 Naldini, L., Vigna, E., Bardelli, A., Follenzi, A., Galimi, F. and Comoglio, P. M. (1995) Biological activation of pro-HGF (hepatocyte growth factor) by urokinase is controlled by a stoichiometric reaction. J. Biol. Chem. **270**, 603–611
54 Shimomura, T., Miyazawa, K., Komiyama, Y., Hiraoka, H., Naka, D., Morimoto, Y. and Kitamura, N. (1995) Activation of hepatocyte growth factor by two homologous proteases, blood-coagulation factor XIIa and hepatocyte growth factor activator. Eur. J. Biochem. **229**, 257–261
55 Lee, S. L., Dickson, R. B. and Lin, C. Y. (2000) Activation of hepatocyte growth factor and urokinase/plasminogen activator by matriptase, an epithelial membrane serine protease. J. Biol. Chem. **275**, 36720–36725
56 Trusolino, L. and Comoglio, P. M. (2002) Scatter-factor and semaphorin receptors: cell signalling for invasive growth. Nat. Rev. Cancer **2**, 289–300
57 Sowter, H. M., Corps, A. N. and Smith, S. K. (1999) Hepatocyte growth factor (HGF) in ovarian epithelial tumour fluids stimulates the migration of ovarian carcinoma cells. Int. J. Cancer **83**, 476–480
58 Baykal, C., Demirtas, E., Al, A., Ayhan, A., Yuce, K., Tulunay, G. and Kose, M. F. (2004) Comparison of hepatocyte growth factor levels of epithelial ovarian cancer cyst fluids with benign ovarian cysts. Int. J. Gynecol. Cancer **14**, 152–156
59 Shimomura, T., Denda, K., Kitamura, A., Kawaguchi, T., Kito, M., Kondo, J., Kagaya, S., Qin, L., Takata, H., Miyazawa, K. and Kitamura, N. (1997) Hepatocyte growth factor activator inhibitor, a novel Kunitz-type serine protease inhibitor. J. Biol. Chem. **272**, 6370–6376
60 Kataoka, H., Miyata, S., Uchinokura, S. and Itoh, H. (2003) Roles of hepatocyte growth factor (HGF) activator and HGF activator inhibitor in the pericellular activation of HGF/scatter factor. Cancer Metastasis Rev. **22**, 223–236
61 Shia, S., Stamos, J., Kirchhofer, D., Fan, B., Wu, J., Corpuz, R. T., Santell, L., Lazarus, R. A. and Eigenbrot, C. (2005) Conformational lability in serine protease active sites: structures of hepatocyte growth factor activator (HGFA) alone and with the inhibitory domain from HGFA inhibitor-1B. J. Mol. Biol. **346**, 1335–1349
62 Kirchhofer, D., Peek, M., Lipari, M. T., Billeci, K., Fan, B. and Moran, P. (2005) Hepsin activates pro-hepatocyte growth factor and is inhibited by hepatocyte growth factor activator inhibitor-1B (HAI-1B) and HAI-2. FEBS Lett. **579**, 1945–1950
63 Lee, A. Y. (2003) Epidemiology and management of venous thromboembolism in patients with cancer. Thromb. Res. **110**, 167–172
64 Fredriksson, L., Li, H., Fieber, C., Li, X. and Eriksson, U. (2004) Tissue plasminogen activator is a potent activator of PDGF-CC. EMBO J. **23**, 3793–3802

Received 23 November 2004/4 April 2005; accepted 20 April 2005
Published as BJ Immediate Publication 20 April 2005, doi:10.1042/BJ20041955