Biol. Chem., Vol. 388, pp. 1243–1253, November 2007 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/BC.2007.143

Activity of ulilysin, an archaeal PAPP-A-related gelatinase and IGFBP protease

Cynthia Tallant[1,a], Raquel García-Castellanos[1,a], Aniebrys Marrero[1], Francesc Canals[2], Yongzheng Yang[3], Jean-Louis Reymond[3], Maria Solà[1], Ulrich Baumann[1,3,*] and F. Xavier Gomis-Rüth[1,*]

[1] Departament de Biologia Estructural, Institut de Biologia Molecular de Barcelona, C.S.I.C., c/ Jordi Girona, 18-26, E-08034 Barcelona, Spain, and Parc Científic de Barcelona, c/ Josep Samitier, 1-5, E-08028 Barcelona, Spain

[2] Proteomics Laboratory, Medical Oncology Research Program, Vall d'Hebron University Hospital Research Institute, Psg. Vall d'Hebron, 119-129, E-08035 Barcelona, Spain

[3] Departement für Chemie und Biochemie, Universität Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

* Corresponding authors
e-mail: ulrich.baumann@ibc.unibe.ch; xgrcri@ibmb.csic.es

Abstract

Human growth and development are conditioned by insulin-like growth factors (IGFs), which have also implications in pathology. Most IGF molecules are sequestered by IGF-binding proteins (IGFBPs) so that exertion of IGF activity requires disturbance of these complexes. This is achieved by proteolysis mediated by IGFBP proteases, among which the best characterised is human PAPP-A, the first member of the pappalysin family of metzincins. We have previously identified and studied the only archaeal homologue found to date, *Methanosarcina acetivorans* ulilysin. This is a proteolytically functional enzyme encompassing a pappalysin catalytic domain and a pro-domain involved in maintenance of latency of the zymogen, proulilysin. Once activated, the protein hydrolyses IGFBP-2 to -6 and insulin chain β *in vitro*. We report here that ulilysin is also active against several other substrates, *viz* (azo)casein, azoalbumin, and extracellular matrix components. Ulilysin has gelatinolytic but not collagenolytic activity. Moreover, the proteolysis-resistant skeletal proteins actin and elastin are also cleaved, as is fibrinogen, but not plasmin and α1-antitrypsin from the blood coagulation cascade. Ulilysin develops optimal activity at pH 7.5 and strictly requires peptide bonds preceding an arginine residue, as determined by means of a novel fluorescence resonance energy transfer assay, thus pointing to biotechnological applications as an enzyme complementary to trypsin.

[a] These two authors contributed equally to this work.

Keywords: adamalysin; ADAMs; astacin; IGF; IGFBP protease; metalloprotease; metzincin; pappalysin.

Introduction

Studies of the effect of pituitary gland activity on body growth led to the discovery of the insulin-like growth factors (IGF)-I and -II in the late 1950s (Salmon and Daughaday, 1956; Daughaday, 1998). These are small peptide hormones that are highly similar to pro-insulin and likewise act on glucose levels in the blood. They are present at very high concentrations throughout the body as circulating mitogens and function through interaction with the insulin receptor and IGF-receptor (IGFR)-I, both of which exert tyrosine-kinase activity (Yu and Rohan, 2000). A third receptor, IGFR-II, lacks kinase activity and is thought to function as a clearance receptor for IGF-II (LeRoith and Roberts, 2003). IGFs regulate cell processes such as survival and apoptosis, migration, differentiation, and proliferation, as well as the cell cycle, and thus act on somatic growth and development, which are dependent on nutritional conditions. In addition, IGFs intervene in energy metabolism, reproduction and longevity (Baxter, 2000; Pollak et al., 2004). As shown for many other growth factors, IGF activity also correlates with several diseases, including atherosclerosis, cardiovascular disease, diabetes and cancer (Pollak et al., 2004; Bach et al., 2005). In addition, IGFs have a potent mitogenic and antiapoptotic effect on various cancer cell types. In combination with other growth factors and steroids, they neutralise the effect of antiproliferative molecules on cancer growth, thus contributing to metastatic events (Bahr and Groner, 2005). Moreover, epidemiological studies have unambiguously correlated high levels of circulating IGF-I with prostate, breast, colorectal and lung cancers (Holly, 2004; LeRoith and Helman, 2004; Durai et al., 2005).

IGFs are present at ca. 100 nM in circulation, a concentration that is approximately 1000-fold higher than that of insulin and of most peptide hormones (Holly, 2004). This happens because 99% of the molecules are sequestered in binary complexes with IGF-binding proteins (IGFBP-1 to -6) (Zapf et al., 1975; Baxter, 1994). These recruiters are soluble ubiquitous 23–53-kDa factors that consist of conserved N-terminal and C-terminal domains, the latter bearing similarity with thyroglobulin-type I domains. IGFBPs bind IGFs (K_d ca. 0.1 nM) with a 2- to 50-fold higher affinity than their cognate receptors. Binding is exerted through both domains, which are linked by a non-conserved connecting mid-region or L-domain (Hwa et al., 1999). IGFBPs modulate IGF transport and IGFR binding, dramatically slowing down IGF clearance, but also hindering their mitogenic activity (Burger et al., 2005). Among these proteins, IGFBP-2 and

source: https://doi.org/10.7892/boris.23636 | downloaded: 12.10.2024

-6 have a higher affinity for IGF-II than for IGF-I, while IGFBP-1, -3, -4 and -5 prefer IGF-I over IGF-II (Durai et al., 2005) and none of them binds insulin. IGFBP-3 is the most abundant of the IGFBPs and accounts for more than 90% of IGF binding. It is the only IGFBP that recruits an additional element, an 80-kDa liver-derived growth-hormone-regulated glycoprotein termed the acid-labile subunit, as a third binding partner to produce a ternary complex that cannot traverse the capillary. In this way, IGFBP-3-bound IGF is kept within the vascular system (Durai et al., 2005). In the aggregate to IGF-binding, a number of IGF-independent functions have been described for IGFBPs, many of which occur through interaction with extracellular matrix (ECM) components (Miyakoshi et al., 2001; Bach et al., 2005). Alternatively, IGFBPs can translocate to the nucleus with the help of importin β and act as transcriptional regulators (Mohan and Baylink, 2002; Burger et al., 2005). Such IGF-independent functions have implications not only in cell growth inhibition, the promotion of apoptosis and the modulation of cell adhesion and migration, but also in cancer.

To perform their physiological tasks, IGFs must be liberated from their IGFBP complexes and released to the circulation. Mechanisms for this include binding of IGFBPs to ECM components, IGFBP phosphorylation or glycosylation, and proteolytic inactivation. Among these options, only the last has been demonstrated to date for all six IGFBPs, each one being the target of distinct IGFBP proteases (IGFBPPs). Sensitivity towards proteolysis of IGFBPs may depend on the presence or absence of bound ligand (Bunn and Fowlkes, 2003) and cleavage can be non-specific or following limited proteolysis, mostly in the IGFBP mid-regions. IGFBPPs are present both in circulation and in interstitial fluids, and their activity is kept in check by endogenous protease activators and inhibitors (Conover et al., 2001; Bunn and Fowlkes, 2003; Holly, 2004). Among the IGFBPPs reported are serine proteinases (complement protein 1s; kallikreins such as prostate-specific antigen, human kallikrein 2 and 7S nerve growth factor; plasmin; a HtrA-related protease; seminal plasma trypsin; cathepsin G and neutrophil elastase; and thrombin), cysteine proteinases (cathepsins and calpain), and metalloproteases (MPs) from the adamalysin/ADAM (ADAM-9, -12 and -28) and matrix metalloprotease (MMP) families (MMP-1, -2, -3, -7, -9, -11 and -19) (Fowlkes et al., 1994; Maile and Holly, 1999; Bunn and Fowlkes, 2003; Fang et al., 2004; Burger et al., 2005; Nakamura et al., 2005). The latter two families are involved in the regulation of other growth factor families through the shedding of ectodomains of membrane-anchored growth factors, cytokines and receptors to increase their circulating forms (Gomis-Rüth et al., 1998; Overall and López-Otín, 2002; Arribas and Merlos-Suárez, 2003; Blobel, 2005).

Another MP, distinct from MMPs and adamalysins/ADAMs, specifically inactivates IGFBPs, namely human PAPP-A, also termed pappalysin-1 or IGFBP-4 protease (Lawrence et al., 1999; Boldt et al., 2001; Oxvig et al., 2004). This MP was originally identified in 1972 as an antigen present in human pregnancy plasma (Gall and Halbert, 1972). It is ubiquitously expressed and plays central roles in ovarian follicular development (Conover et al., 2001), myogenesis (Kumar et al., 2005), human embryo implantation (Giudice et al., 2002) and wound healing (Chen et al., 2003), as well as in atherosclerosis (Lawrence et al., 1999; Bayes-Genis et al., 2001a,b; Thorn and Khan, 2007), early foetal stages of Down's syndrome development and cancer. Mature PAPP-A is a glycosylated multidomain protein of 1547 residues, among them 82 cysteines [SwissProt (SP) sequence database access code Q13219]. In pregnancy, the protein is primarily present in circulation as a 500-kDa disulfide-linked 2:2 complex, with two subunits of the 206-residue glycosylated proform of the eosinophil major basic protein that acts as a covalent proteinase inhibitor. Outside pregnancy, PAPP-A circulates as a 400-kDa homodimer (Oxvig et al., 1993, 2004; Lawrence et al., 1999). PAPP-A is the biochemically most thoroughly studied specific IGFBPP and has been shown to specifically hydrolyse human IGFBP-4 in an IGF-dependent manner. In addition, PAPP-A cuts human IGFBP-5 and in this case IGF-II has been reported to have a slightly inhibitory function (Laursen et al., 2001). In contrast, cleavage of IGFBP-2 is enhanced by IGF binding (Monget et al., 2003; Kumar et al., 2005). PAPP-A proteolytic activity is inhibited by basic peptides derived from heparin-binding domains (Monget et al., 2003). As a complement to PAPP-A's proteolytic role, other domains were identified by sequence similarity searches. These are a LamG-like jellyroll-fold domain, as found in proteins of unknown function from metazoa and bacteria; three lin12-notch repeats (LNR-1, -2 and -3); and five ca. 65-residue sushi domains, also known as complement control protein modules (CCP1–5), encountered in a wide variety of complement and adhesion proteins such as P-selectin. Accordingly, PAPP-A probably has several still undiscovered functions, thus providing an explanation for its participation in a number of physiological processes and its ubiquity (Laursen et al., 2002b; Boldt et al., 2004).

PAPP-A has been designated the first member of the pappalysin family of MPs (Boldt et al., 2001), which further comprises the paralogue PAPP-A2, also known as PAPP-E or pappalysin-2 (SP Q9BXP8). For this protein, which is also found in mammals, specific IGF-independent cleavage of IGFBP-5 identical to PAPP-A and partial activity against IGFBP-3 have been reported (Farr et al., 2000; Bayes-Genis et al., 2001b; Overgaard et al., 2001). These pappalysins contain the sequence motif His-Glu-Xxx-Xxx-His-Xxx-Xxx-Gly-Xxx-Xxx-His/Asp (Xxx for any residue), a zinc-binding consensus sequence (ZBCS) that is the hallmark of the metzincin clan of MPs. In these metalloproteins, the three histidine/aspartate residues are engaged in coordination of the catalytic zinc ion and the glutamate is the general base/acid required for enzyme activity. Sequence analysis of PAPP-A and PAPP-A2 further revealed the presence of a conserved methionine residue, as found in what is known as the Met-turn in metzincins. These findings led pappalysins to be assigned to the latter clan (Bode et al., 1993; Stöcker et al., 1995; Gomis-Rüth et al., 1998; Boldt et al., 2001; Gomis-Rüth, 2003; Oxvig et al., 2004).

Probably the most interesting function of pappalysins is their proteolytic action on IGFBPs because of the

implications for mechanistic studies and drug design. An attempt to delimit the protease domain of PAPP-A within the 170-kDa protein was made on the basis of the three-dimensional (3D) structure of the catalytic domains of other metzincins and secondary structure predictions. These studies suggested an approximately 300-residue protease domain from approximately Lys352 to Tyr664 (in all cases, residue numbering is according to the respective SP entry), but no experimental evidence is available to date, as this segment could not be produced independently in a functional state. To overcome these difficulties, we used bioinformatic searches in a novel approach to identify a whole series of novel putative homologues of PAPP-A including at least the catalytic domain (Tallant et al., 2006; García-Castellanos et al., 2007). These can be classified into two major groups: one features bacterial, archaeal and fungal sequences that potentially comprise only a catalytic domain and a pro-domain (short pappalysins), while the other includes PAPP-A and PAPP-A2 and consists of potential multi-domain proteins from mammals, fish, birds, amphibians and a collagenase precursor from *Cytophaga* sp. (large pappalysins). This is reminiscent of other metzincin families such as astacins, MMPs and adamalysins/ADAMs, for which single- and multi-domain sequences have been reported (Bode et al., 1993; Stöcker et al., 1995; Gomis-Rüth, 2003).

We have produced and studied a short family member, *Methanosarcina acetivorans* ulilysin, the only archaeal pappalysin found. As previously reported, its 3D crystal structure uncovered a novel fold under maintenance of certain structural elements found in other metzincin prototypes such as MMPs and adamalysins/ADAMs (Tallant et al., 2006). We report here the full cloning, overexpression and purification procedures for two mutagenic variants of the ulilysin precursor, proulilysin, as well as additional comprehensive functional studies *in vitro* of the mature form, to contribute to the characterisation of this paradigm of the protease domain of pappalysins. These findings reveal unexpected features that may be relevant to biotechnology and have consequences for other short and large pappalysins, and pave the way for new strategies for the study of the catalytic domain of PAPP-A.

Results and discussion

Heterologous recombinant protein overexpression and purification

As pointed out previously (Tallant et al., 2006) and discussed hereafter in more detail, several constructs and plasmids were assayed to produce the native full-length 38-kDa protein. These produced either no expression, protein aggregation or protein that would not crystallise reproducibly. We suspected that some of the five cysteine residues present in ulilysin could be responsible for these problems. Therefore, a systematic cysteine-to-alanine scan was attempted. Eventually, a clone with a mutation at position 269 (C269A) was identified that overcame these obstacles. In addition, a potential active-site mutant was generated, in which the putative general base/acid residue Glu229 imbedded in the ZBCS was replaced by alanine (E229A). The alanine mutation of the equivalent glutamate residue in human PAPP-A (Glu483) had been reported to completely abolish activity and similar findings for other metzincins have been published (Boldt et al., 2001; Gomis-Rüth, 2003). The E229 variant did not lead to the above-mentioned biophysical inconveniences of the wild-type construct. For both mutant clones, *Escherichia coli*-based expression systems were developed that yielded soluble and non-aggregating N-terminally hexahistidine-tagged protein, which was purified by affinity chromatography. The tags were efficiently removed with thrombin and the proteins obtained were subjected to an anionic exchange chromatography step to remove thrombin and other impurities from the non-tagged proulilysin. Fractions containing the protein were pooled and subjected to a final gel filtration chromatography step for polishing. Typical protein production rendered yields of ca. 0.6 mg of pure protein starting from 500 ml of cell culture.

Calcium-mediated autolytic activation and pH optimum for activity

As reported earlier, the 38-kDa form encompassing the full-length protein sequence is actually the pro-enzyme, which is stable but undergoes autolytic activation at two specific sites to produce a fragment of 29 kDa, mature ulilysin (Figure 1A, lanes 2–4; see also Tallant et al., 2006), in the presence of calcium. In contrast, the 38-kDa E229A mutant was inactive and non-activatable by calcium and did not render a short form (Figure 1A, lanes 5–7), providing additional proof that the proteolytic event is autolytic.

To further study the activation and activity on a quantitative basis, a fluorescence resonance energy transfer (FRET) reaction experiment was carried out using a fluorigenic custom-synthesised peptide, DABCYL-Leu-Ala-Arg-Val-Glu-EDANS, that mimics the C-terminal cleavage sequence within proulilysin (Tallant et al., 2006). Mass spectrometry analysis of the untreated peptide revealed the presence of a peak corresponding to the expected M+H value (1086.57 Da), as well as some minor lower-molecular-weight impurities (Figure 2A). The sequence of the peptide was confirmed by MALDI-TOF/TOF fragmentation analysis (Figure 2B,C). The sample treated with ulilysin C269A in the presence of calcium exhibited a peak of mass 454.20 Da corresponding to the fragment DABCYL-Leu-Ala (Figure 2D). Fragmentation of the latter ion produced a further peak of mass 252 Da, attributable to the DABCYL moiety. These results confirm that the fluorigenic peptide had been cleaved before the arginine residue. Active 29-kDa ulilysin C269A showed straightforward activity on this peptide from the beginning of the reaction, as visualised by a hyperbolic fluorescence curve (Tallant et al., 2006). This proved that this pentapeptide sequence was a good substrate for the active enzyme. Determination of the kinetic parameters governing cleavage of this substrate by mature ulilysin C269A revealed an apparent second-order rate constant k_{cat}/K_m of 2.867×10^6 s^{-1} M^{-1} (Table 1), indicative of good affinity for and rapid turnover of this fluorigenic peptide containing an endogenous sequence.

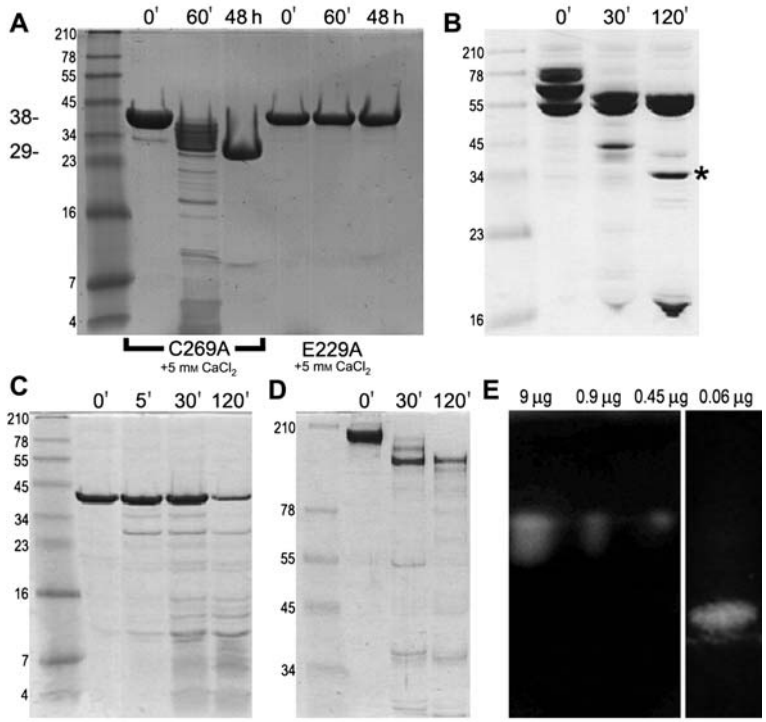

Figure 1 Proteolytic activity of ulilysin.
(A) Stability of purified proulilysin (0.9 mg/ml) mutant C269A (lanes 2–4) and mutant E229A (lanes 5–7) in 20 mM Tris-HCl, pH 7.5, 0.2 M NaCl, 5 mM $CaCl_2$ over time. (B) Limited proteolysis by ulilysin C269A of human platelet fibrinogen over time. The band marked with an asterisk was subjected to N-terminal Edman degradation (see Figure 4). (C) Proteolytic activity of ulilysin C269A on bovine muscle actin. (D) Proteolytic activity of ulilysin C269A on bovine plasma fibronectin. (E) Casein (left) and gelatin (right) zymograms in the presence of 5 mM $CaCl_2$ with the amount of ulilysin C269A indicated on top.

The optimal pH for catalysis was determined by a gelatinase assay kit that contains DQ™ gelatin as a fluorescein conjugate. This assay is based on the fact that gelatin is so heavily labelled with fluorescein that the fluorescence is quenched. In the presence of a gelatinolytic active moiety, highly fluorescent peptides are generated in proportion to the proteolytic activity. These studies revealed that ulilysin is most active at pH 7.5 under the assay conditions used (Figure 3). This is consistent with human PAPP-A, for which optimal cleavage of IGFBP-4 and -5 was reported at pH 8.0 (Laursen et al., 2002a) and for which cleavage experiments *in vitro* are routinely carried out at pH 7.5 (Overgaard et al., 2000; Gyrup and Oxvig, 2007). These values, in turn, corroborate the observation that zinc, when bound by three histidine imidazole groups at the active site of an enzyme, is activated to provide the Zn^{2+}-OH^{-} species required for nucleophilic attack of a scissile peptide bond at a pH near 7 (Fraústo da Silva and Williams, 2001).

Proteolytic assays of protein substrates and fluorigenic peptide substrates

Mature ulilysin mainly cleaved IGFBP-2 to -6 in a specific manner, but not IGFBP-1 and IGF-II. These studies suggested lower efficiency *in vitro* of PAPP-A and PAPP-A2 against IGFBP-4 and -5 than ulilysin when using similar detection techniques (Tallant et al., 2006). For the human enzymes, standard assay conditions entailed a protease/substrate ratio of approximately 1:10 or 1:20 and incubation times of up to 20 h (Boldt et al., 2001; Overgaard et al., 2001). In contrast, ulilysin produced comparable cleavage levels (as inferred from SDS-PAGE) in our hands at a weight ratio of 1:300 and reaction times of 2–4 h. However, more recent detailed kinetic studies following a different analytical approach revealed that PAPP-A cleaved IGFBP-4 and -5 in a much more efficient manner than ulilysin and that the hydrolysis of IGFBP-4 was dependent on IGF binding for both proteases (Gyrup and Oxvig, 2007). In addition to IGFBPs, ulilysin could cut insulin chain β, but not chain α (Tallant et al., 2006). Besides IGF-axis proteins, ulilysin was shown to process a whole series of substrates, mainly in a selective manner, including the general substrates (azo)casein (as mentioned above) and azoalbumin, as well as ECM (derived) components such as azocollagen, gelatine (including DQ™ gelatin; see above) and fibronectin (Figure 1). In this context, it should be noted that ECM components have been shown to interact with the ulilysin substrate IGFBP-5 and to modulate its activity *in vivo* (Bach et al., 2005). Activity on casein and gelatin was further corroborated by zymography studies, which revealed that gelatin was a better substrate than casein (Figure 1E). Natural type-I collagens from human placenta and type-IV collagen from calf skin were also assayed, but no cleavage was detected (Tallant et al., 2006). Type-I collagen from kangaroo tail was cleaved at 37°C, but not at room temperature. However, this substrate was also cleaved by trypsin under the assay conditions at 37°C, although not at room temperature, as was azocollagen. This pointed to a certain content of gelatinous material in the purchased samples. Accordingly, we suggest that

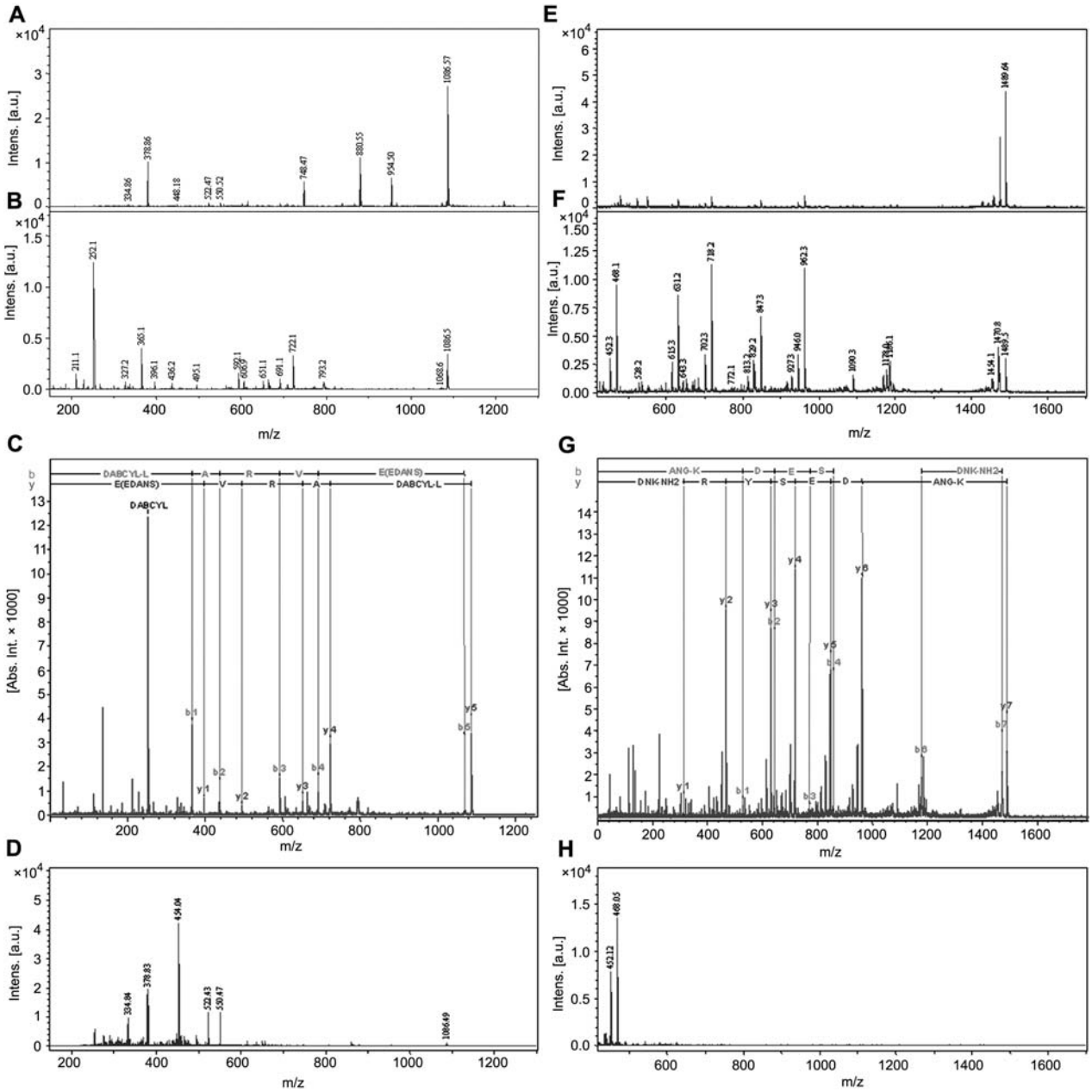

Figure 2 Mass spectrometry analyses of cleaved fluorigenic peptides.
(A) Spectrum of the intact DABCYL-Leu-Ala-Arg-Val-Glu-EDANS peptide and (B) after MALDI-TOF/TOF fragmentation. (C) Fragmentation allows the chemical sequence of the peptide to be assessed. (D) Spectrum of the cleaved peptide, with the 454-Da peak attributable to the DABCYL-Leu-Ala moiety. (E) Spectrum of the intact ANG-Lys-Asp-Glu-Ser-Tyr-Arg-DNK-NH_2 peptide and (F) after MALDI-TOF/TOF fragmentation. (G) Fragmentation confirms the chemical sequence of the peptide. (H) Spectrum of the cleaved peptide, with the 468-Da peak assignable to the Arg-DNK-NH_2 cleavage product.

Table 1 Ulilysin activity on fluorigenic peptides.

Substrate	Cleavage	Specific cleavage	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
Synthetic fluorigenic peptides					
FALGPA	–				
DABCYL-Leu-Ala-Arg-Val-Glu-EDANS	+	+	173.7	497.9	2.867×10^6
ANG-Lys-Asp-Glu-Ser-Tyr-Arg-DNK-NH_2 (peptide 1)	+	+	220.1	555.5	2.524×10^6
ANG-Ala-Val-Pro-Glu-Arg-DNK-NH_2 (peptide 2)	+	+	184.2	670.7	3.642×10^6
ANG-Glu-Phe-Val-Gly-Ser-Asp-DNK-NH_2 (peptide 3)	–				
ANG-Tyr-Ala-Arg-Lys-Leu-Phe-DNK-NH_2 (peptide 4)	+	+	88.0	192.2	2.185×10^6
ANG-Leu-Lys-Tyr-Phe-Asp-Ile-DNK-NH_2 (peptide 5)	–				

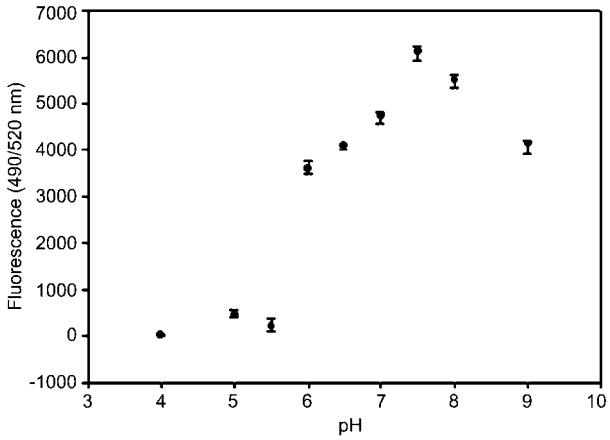

Figure 3 pH optimum for activity.
pH dependence of ulilysin activity monitored via the fluorigenic EnzCheck DQ gelatinase assay.

ulilysin has gelatinolytic but not collagenolytic activity. Moreover, the rather proteolysis-resistant skeletal proteins actin and elastin were also cleaved, as was fibrinogen (chain α) (Figure 1), but not plasmin nor α1-antitrypsin, from the blood-coagulation cascade (Tallant et al., 2006).

As mentioned above, a fluorigenic peptide mimicking the putative C-terminal autolytic cleavage site was successfully employed to study activity. In addition, the specific *Clostridium histolyticum* collagenase IV fluorigenic substrate FALGPA, which mimics the sequence of the α chains in triple-helical collagen, was not cut (Table 1). This correlates well with the above-mentioned results and points to a lack of collagenolytic activity. Further to these analyses, a FRET assay was employed to study proteolysis (Table 1). This assay is a further development of one already reported and consists of five fluorigenic peptides with a variable amino-acid sequence containing 25 different types of dipeptides in a domino arrangement (Yang and Reymond, 2005). The peptides possess a 7-glutaramidonaphthalene-1,3-disulfonate fluorophore (ANG) that acts as a donor attached to the peptide N-terminus and the last residue is always a dinitrophenyl lysine with a C-terminal carboxamide (DNK-NH_2), which acts as a fluorophore quencher. Cleavage of a peptide bond produces an increase in fluorescence as a result of donor-acceptor separation. Three of these peptides were cut and two were not (Table 1). MALDI-TOF mass spectrometry analysis of the cleavage reaction of ANG peptide 1 (Figure 2E–H) revealed that the untreated sample showed a very clear peak corresponding to M+H at 1489.64 Da (Figure 2E). In addition, some lower-molecular-weight peaks corresponding to spontaneous fragmentation were observed. In a MALDI-TOF/TOF spectrum, these fragments were also observed and the sequence could be confirmed, including the corresponding ANG and DNK-NH_2 modifications of the first and last amino acids (Figure 2G,H). In the treated sample, an ion corresponding to 468.05 Da was observed, together with a fragmentation ion of type M-16. This correlates with the mass of fragment Arg-DNK-NH_2, suggesting that cleavage had occurred between the amino acids tyrosine and arginine. The 468.05-Da peptide was subjected to further fragmentation. This reaction gave rise to peaks present in the TOF/TOF fragmentation spectrum of the entire peptide (Figure 2F), thus ruling out possible artefacts. Determination of the kinetic parameters governing cleavage of ANG peptide 2 by ulilysin rendered an apparent second-order rate constant k_{cat}/K_m of 3.642×10^6 s^{-1} M^{-1} (Table 1), indicative of very close affinity for and rapid turnover of this substrate. Similarly, ANG peptides 1 and 4, which also contain an arginine residue, were also very efficiently processed, with k_{cat}/K_m values of 2.524×10^6 and 2.185×10^6 s^{-1} M^{-1}, respectively, while ANG peptides 3 and 5 remained unaffected.

Search for a specificity pattern for ulilysin

Initial cleavage assays with ulilysin based on the recognised sequence motifs around the N- and C-terminal autolytic activation sites, as well as using human fibrinogen chain α and bovine insulin chain β as substrates, revealed an apparent substrate preference following the pattern Bbb-Xxx-↓-Arg-Bbb-Glu/Gln (Bbb, bulky hydrophobic or aromatic) (Tallant et al., 2006). However, current information based on its behaviour towards the ANG peptides enabled us to establish that flanking sequences distinct from those indicated were also very efficiently cleaved, at least in peptide substrates, so that no additional positions seem to be specifically recognised besides $P_{1'}$ (Figure 4). The very high specificity of ulilysin for an arginine residue at $P_{1'}$ is further underlined by the fact that ANG peptide 5 was not cleaved at all, despite including a lysine. Accordingly, ulilysin can be envisaged as a very efficient protease with a strict specificity for substrates that inject an arginine residue side chain into the specificity pocket, thus pointing to interesting biotechnological applications of this enzyme in protein processing as a complement to trypsin and other proteases that cleave after a basic residue, as utilised during proteomic studies.

Inferences for the pappalysin family

The results reported here for ulilysin may have consequences for functional studies of other pappalysins such as human PAPP-A. This enzyme specifically hydrolyses human IGFBP-4 at the Met155-Lys156 bond (see SP P22692) (Laursen et al., 2001), human IGFBP-5 (SP

Autoactivation site I	I-S-✂-R-M-E
Autoactivation site II	L-A-✂-R-V-E
Human fibrinogen chain α	V-E-✂-R-H-Q
Bovine insulin chain β	G-E-✂-R-G-F
Synthetic fluorigenic peptide	DABCYL-L-A-✂-R-V-E(EDANS)-OH
Reymond-peptide 1	ANG-K-D-E-S-Y-✂-R-DNK
Reymond-peptide 2	ANG-A-V-P-E-✂-R-DNK
Reymond-peptide 4	ANG-Y-A-✂-R-K-L-F-DNK

Figure 4 Cleavage sequences.
Sequences recognised by ulilysin C269A for cleavage as determined by N-terminal Edman-degradation or MALDI-TOF/TOF fragmentation. The first five cleavage sites were reported previously (Tallant et al., 2006). The only position absolutely conserved is an arginine residue at $P_{1'}$.

24593) at Ser163-Lys164 (Laursen et al., 2001) and human IGFBP-2 (SP P18065) at Gln165-Met166 (Monget et al., 2003; Kumar et al., 2005). PAPP-A also undergoes autolytic cleavage at Phe386-Asn387, which causes a loss of activity, and no other substrates for PAPP-A have yet been found. These cleavage sites do not evince a common pattern for substrate recognition. In contrast, ulilysin is a very efficient, broad-spectrum enzyme with a stringent preference for peptide bonds preceding an arginine residue.

An accurate structure-based sequence alignment of ulilysin with the region thought to encompass the catalytic domain of PAPP-A (for data on the ulilysin structure and a sequence alignment see Tallant et al., 2006; García-Castellanos et al., 2007) reveals that residues that are key for catalysis and substrate binding and residues required for structural stability in the archaeal enzyme are fully conserved in the human protein. The sequence conservation is especially noteworthy within the C-terminal sub-domain in charge of substrate binding and thus of specificity modulation. Among these residues are Asp295 (ulilysin numbering hereafter) at the bottom of the specificity pocket; Asn288, placed at the beginning of the Met-turn and hydrogen-bonded with Gln305, which is required for structural integrity; Arg308, engaged in hydrogen bonding of the Asn235 main chain; and two residue duos required for structural integrity in ulilysin, Thr302-Asp258 and Trp165-Asn172. Another similarity between the two enzymes is that activity is dependent on calcium (Boldt et al., 2004; Tallant et al., 2006). Inspection of the residues and protein atoms participating in ulilysin in the coordination of the two structurally indispensable calcium cations reveals that most ligands are provided by solvent molecules or main-chain carbonyl oxygen atoms (see Figure 1 in Tallant et al., 2006; García-Castellanos et al., 2007). Only three side chains are actively engaged in ulilysin, Asp254, Thr259 and Glu243. The first two are conserved in PAPP-A, while the third one could easily be replaced by a vicinal glutamate, Glu580 (PAPP-A sequence). These common features strongly suggest that the catalytic domain of PAPP-A is structurally similar to ulilysin and likewise evinces an active-site cleft with a specificity pocket prepared to snugly accommodate arginine residues. This hypothesis is further supported by the very recent finding that PAPP-A cleaves a fluorigenic pentapeptide mimicking the C-terminal autoactivation site of proulilysin, DABCYL-Leu-Ala-Arg-Val-Glu-EDANS (Gyrup and Oxvig, personal communication).

Accordingly, the apparent paradox of probable structural similarity of the PAPP-A catalytic domain with the efficient and $P_{1'}$-specific ulilysin, on the one hand, and PAPP-A's lack of specificity and narrow substrate palette, on the other, supports a hypothesis that steric regulation might be a mechanism to limit substrate recognition in PAPP-A and other large pappalysins, possibly mediated by other domains preceding and/or following the catalytic protease domain (Boldt et al., 2001, 2004; Laursen et al., 2002a). In particular, the three LNR motifs are attractive candidates for such a restricting function, as they have been reported to be indispensable for proteolytic specificity in PAPP-A (Boldt et al., 2001, 2004; Laursen et al., 2002a). These repeats are modules of ca. 35–40 residues that were first described as regulating ligand-induced proteolytic cleavage of the Notch receptor (Aster et al., 1999). They are absent from short pappalysins such as ulilysin (Tallant et al., 2006; García-Castellanos et al., 2007), although two of them can be envisaged as a ca. 66-residue insertion replacing the superficial 10-residue LNR-like loop observed between two β-strands in ulilysin. A sequence alignment of pappalysins reveals that homologues from *Gloeobacter violaceus* and *Magnaporthe grisea* probably lack this LNR-like loop, so that this structural feature may be optional for the distinct family members. It may occur with distinct lengths, with maintenance of the overall protease structure, and may be surface-located, which positions it for carrying out its binding functions. In large pappalysins, these LNR motifs could restrict the access of substrates to the active site, thus limiting the substrate palette and specificity. The presence of such additional modules inserted in catalytic domains has been described for MMP-2 and -9, for which three consecutive fibronectin type-II domains were found (Morgunova et al., 1999). These are involved in substrate binding and are absent in the other MMP catalytic domains (Maskos and Bode, 2003). Another example is provided by bone morphogenetic protein-1, a multidomain procollagen C-proteinase of the astacin family of metzincins, which also shows a much broader substrate profile and higher efficiency in cleaving a variety of ECM proteins for its isolated catalytic domain than the full-length protein. This difference was also attributed to additional C-terminal domains present in the latter, which modulate and restrict the substrate specificity through their potential protein-binding or steric-hindering competence (Hartigan et al., 2003; Garrigue-Antar et al., 2004; Hintze et al., 2006). A similar scenario is conceivable for PAPP-A and other large pappalysins with a multidomain structure, where the proteolytic potential would be kept in check by additional domains *via* exosite binding and/or steric hindrance of substrate binding.

Materials and methods

Cloning and recombinant overexpression of *M. acetivorans* proulilysin C269A and E229A mutants

The procedures used were already introduced briefly in a preceding publication and are presented here in full detail for better reproducibility (Tallant et al., 2006). A DNA fragment of 1029 bp encoding the chemical sequence of the full-length 342-aa 38 370-Da protein MA3214 (SP Q8TL28; hereafter referred to as proulilysin) was amplified from total genomic DNA of *M. acetivorans* strain C2A (ATCC no. 35395) using oligonucleotides 5′-AGG GAT GTA CAT ATG GCA GAA AAA TTT GAA AG-3′ as the forward primer and 5′-CAC ACT CGA GTT ACC TTG GCA TAG GCA TTT C-3′ as the reverse primer (purchased from Microsynth; Balgach, Switzerland) and $Vent_R$ DNA polymerase (New England Biolabs, Ipswich, MA, USA). The reaction yielded a fragment containing the correct coding sequence within recognition sites for the restriction endonucleases NdeI (5′) and XhoI (3′). After digestion with these enzymes (Fermentas, Vilnius, Lithuania), the fragment was cloned into different expression vectors employing T4 DNA ligase (Roche, Indianapolis, IN, USA). The

point mutation C269A was introduced into the wild-type clone as inserted into plasmid pET-28a (Novagen; San Diego, CA, USA) with 5′-TCA GGC AGA TCC TAA TTT TGG AGC TCC AAG TTA TCC ACA TGT CAG C-3′ as the forward primer and 5′-GCT GAC ATG TGG ATA ACT TGG AGC TCC AAA ATT AGG ATC TGC CTG A-3′ as the reverse primer using a QuickChange PCR-based mutagenesis reaction kit and Pfu polymerase (Stratagene, La Jolla, CA, USA). The clone was sequenced to confirm the presence of the desired mutation and to rule out PCR artefacts. Plasmid pET-28a provides kanamycin resistance to the host and adds an N-terminal hexahistidine tag and a thrombin cleavage recognition site to the protein sequence. This means that the tripeptide Gly-Ser-His precedes the first endogenous methionine after digestion with thrombin. The recombinant plasmid was introduced into *Escherichia coli* BL21 Star™ cells by heat-shock treatment for heterologous overexpression. A typical protein production protocol stipulated growing 500-ml bacterial cultures in Luria-Bertani medium for 22 h at 20°C after induction with 0.25 mM isopropyl-β-D-thiogalactopyranoside at an optical density of 1.2. Bacterial cultures were collected and centrifuged at 3000 *g* for 30 min and the pellets were stored at -20°C for further processing. Proulilysin E229A was obtained by site-directed mutagenesis using the wild-type expression vector as a template and 5′-CCA TGC AAT CGG ACA CTG GTT A-3′ and 5′-CCG ATT GCA TGG GTT GCC GT-3′ as mutagenic forward and reverse primers, respectively. Protein overexpression proceeded as described for proulilysin C269A.

Purification of proulilysin C269A and E229A

Pellets were gently resuspended in 50 ml of buffer A (10 mM Tris-HCl, pH 8, 20 mM imidazole, 0.5 M NaCl, 10 mM β-mercaptoethanol), supplemented with one EDTA-free protease inhibitor cocktail tablet (Complete; Roche, Basel, Switzerland). Cells were disrupted using a cell disruptor (Constant Systems, Northants, UK) and centrifuged at 35 000 *g* for 30 min. The supernatant was filtered (0.22-μm pore size; Millipore, Billerica, MA, USA) and incubated at 4°C with 2 ml of Nickel-NTA resin (Qiagen, Venlo, The Netherlands) previously equilibrated with buffer A. After 2 h, the sample was applied to a batch purification column (Bio-Rad, Hercules, CA, USA) and washed with 15×4 ml of buffer A. Subsequently, the protein was first washed with solutions containing 10% and 20% of buffer B (buffer A containing 400 mM imidazole) and then eluted with a 50% (proulilysin C269A) or 25% (proulilysin E229A) solution of buffer B. The purity of the samples was checked by 16% tricine SDS-PAGE. The purified hexahistidine-tagged protein samples were pooled and thrombin was added for tag removal (Amersham, Little Chalfont, UK; 1 U/mg protein). This reaction mixture was dialysed overnight at 4°C against buffer C [10 mM Tris-HCl, pH 7.5, 2 mM 1,4-dithio-D,L-threitol (DTT), 10 mM NaCl] to enhance cleavage. The reaction was stopped by the addition of phenylmethylsulfonyl fluoride (Sigma, Madrid, Spain) to the protein/protease mixture, which was then dialysed against buffer D (buffer C adjusted to pH 6.3). Further purification was by anionic exchange chromatography on a Pharmacia UPC-900 FPLC apparatus employing a MonoQ HR 5/5 column (Amersham) equilibrated with buffer E (20 mM Tris-HCl, pH 6.3, 2 mM DTT). Elution was via a linear gradient within 30 column volumes from 0% to 100% buffer F (buffer E plus 1 M NaCl). Fractions containing tag-free 38-kDa proulilysin C269A and E229A were identified by SDS-PAGE and pooled and subjected to a final size-exclusion chromatography step on a Superdex 75 HR 10/30 (Amersham) column equilibrated with buffer G (30 mM Tris-HCl, pH 7.5, 2 mM DTT, 100 mM NaCl). The central fractions of the peak were pooled and concentrated using Centricon-30 centrifugal filter devices (Amicon, Billerica, MA, USA). Concentrated protein samples (5–20 mg/ml) were stored at 4°C.

Activity and pH-optimum assessment *in vitro*

The proteolytic activity and its pH dependence were assessed using a fluorimetric EnzCheck gelatinase assay kit containing DQ™ gelatin from pig skin as a fluorescein conjugate (Invitrogen, Carlsbad, CA, USA) at λ_{EX}=490 nm and λ_{EM}=520 nm in a microplate fluorescence reader (Victor, Perkin Elmer, Waltham, MA, USA). The reaction mixture contained 5 μl of DQ™ gelatin (1 mg/ml) and 3 μl of ulilysin C269A (0.9 mg/ml) in 100 mM NaH_2PO_4/100 mM Na_2HPO_4 buffer adjusted to pH values from 4 to 9. Each measurement was repeated three times.

Proteolytic assays on protein substrates and peptide-derived fluorigenic substrates

Unless otherwise indicated, proteolytic assays were performed at a C269A concentration of 0.65 mg/ml in 50 mM Tris-HCl, pH 7.5, 5 mM $CaCl_2$ at 37°C for 2–4 h and at a protease/substrate weight ratio of 1:100 or 1:200. Proteins assayed for proteolytic susceptibility comprised bovine plasma fibronectin and human plasma fibrinogen, bovine muscle actin (all three assayed at a weight ratio of 1:125 by mixing 450 μl of substrate in 50 mM Tris-HCl, pH 8 with 6 μl of enzyme at 0.9 mg/ml in 50 mM Tris-HCl, pH 7.5, 5 mM $CaCl_2$). The fluorigenic peptide substrate *N*-[3-(2-furyl)acryloyl]-Leu-Gly-Pro-Ala (FALGPA, Sigma) was assessed for cleavage by ulilysin C269A at a molar ratio of 1:7 in 50 mM Tris-HCl, pH 7.5, 0.4 M NaCl by monitoring the fluorescence change at 324 nm at room temperature over 4 h on an UV/Vis Ultraspec 2000 spectrophotometer. DABCYL-Leu-Ala-Arg-Val-Glu-EDANS was assayed at a molar ratio of 1:7 in 20 mM Tris-HCl, pH 7.5. In addition, a fluorescent assay of five peptides of sequence ANG-Lys-Asp-Glu-Ser-Tyr-Arg-DNK-NH_2 (ANG peptide 1), ANG-Ala-Val-Pro-Glu-Arg-DNK-NH_2 (ANG peptide 2), ANG-Glu-Phe-Val-Gly-Ser-Asp-DNK-NH_2 (ANG peptide 3), ANG-Tyr-Ala-Arg-Lys-Leu-Phe-DNK-NH_2 (ANG peptide 4) and ANG-Leu-Lys-Tyr-Phe-Asp-Ile-DNK-NH_2 (ANG peptide 5), where ANG is glutirylaminonaphthalene-1,3-disulfonate-fluorophore and DNK-NH_2 is dinitrophenyl lysine with a C-terminal carboxamide, was further used for cleavage studies. These peptides were conditioned as 1 mM stock solutions in 1:1 dimethylformamide/water (v/v) by dissolving 1–2 mg of solid peptide in an appropriate volume. For the assays, the stock solutions were diluted with 20 mM Tris-HCl, pH 8. The assay was carried out by mixing 5 μl of ulilysin C269A solution (0.9 mg/ml) with the respective peptide solution and monitoring the fluorescence change (λ_{EX}=355 nm; λ_{EM}=460 nm) using a microplate fluorescence reader at 30°C for 2 h.

Kinetic studies on synthetic substrates

Kinetic parameters governing ulilysin C269A cleavage were determined for DABCYL-Leu-Ala-Arg-Val-Glu-EDANS and for ANG-peptides 1, 2 and 4. In all cases, the reactions were carried out at the pH of maximal protease activity in 20 mM Tris-HCl, pH 7.5 at 30°C in a microplate. Initial reaction velocities were determined at two different substrate concentrations to ascertain the region attributable to a single displacement reaction following simple Michaelis-Menten kinetics (Polgár, 1999). The initial rate of reaction (ν), the substrate concentration ([S]), the rate of reaction at substrate saturation (V_{max}) and the total enzyme concentration ($[E_0]$) were assumed to be related via the first-order rate constant k_{cat} and the pseudo-equilibrium Michaelis constant K_m (the substrate concentration when $\nu=V_{max}/2$), according to the following equation:

$$\nu=V_{max}\cdot[S]/(K_m+[S])=k_{cat}\cdot[E_0][S]/(K_m+[S]). \quad (1)$$

In a first instance, the fluorescence variation at the corresponding λ_{EX} and λ_{EM} values (see above) was monitored against time

for a fixed enzyme concentration (1 μM) and different initial substrate concentrations (3, 5, 10, 20, 50, 70 and 100 μM for DAB-CYL-Leu-Ala-Arg-Val-Glu-EDANS; 7, 17, 33, 67 and 133 μM for the ANG peptides). The slope of the linear part of each hyperbolic curve allowed determination of the associated value of ν. Subsequently, the values of ν (ordinate axis) were plotted versus [S] (abscissa) and a hyperbolic equation of type y=ax/(b+x) was adjusted using the program SigmaPlot. This allowed calculation of V_{max}, K_m and k_{cat} according to Eq. (1).

Zymographic analyses

Precast Tris/glycine Ready Gel systems (Bio-Rad) containing either gelatin (10%) or casein (12%) were employed for zymography studies. Aliquots of 10 μl of purified protease were incubated with 10 μl of 2× sample buffer [0.125 M Tris-HCl, pH 6.8, 20% (v/v) glycerol, 4% (w/v) sodium dodecyl sulfate, 0.005% bromophenol blue] at room temperature for 5 min and then loaded onto the gels, which were run at 125 V with 1× running buffer [20 mM Tris-HCl, pH 8.3, 0.2 M glycine, 0.1% (w/v) sodium dodecyl sulfate], then washed twice with renaturing buffer (2.5% Triton X-100) for 2 h, and finally equilibrated with developing buffer [40 mM Tris-HCl, pH 7.5, 0.2 M NaCl, 0.02% (w/v) Brij 35] with/without 5 mM $CaCl_2$ at 37°C for 20 h under gentle agitation.

Mass spectrometry studies

Matrix-assisted laser desorption/ionisation time of flight (MALDI-TOF) and MALDI-TOF/TOF analyses were performed on a Bruker Ultraflex (Bremen, Germany) mass spectrometer equipped with a pulsed nitrogen laser (337 nm). Samples were prepared by mixing 0.5 μl of the peptide solution with 1 μl of the matrix solution (0.3 mg/ml α-cyano-4-hydroxycinnamic acid in ethanol/acetone, 2:1 v/v). The mixture was then spotted on an anchorchip target (Bruker) and allowed to evaporate to dryness at room temperature. Spectra were acquired in positive reflectron mode using an acceleration voltage of 25 kV. A mixture of standard peptides of known molecular mass in the range 750–3500 Da (Bruker) was used for external calibration of the mass spectra. The BioTools 3.0 software (Bruker) was used for MALDI-TOF/TOF spectra fragment analysis.

Miscellaneous

Chemical sequencing by N-terminal Edman degradation was performed at the Proteomics Unit of the Technical and Scientific Service of the Barcelona Science Park, Barcelona, Spain.

Acknowledgements

This study was supported by the following grants and fellowships: SAB2002-0102 from the former Spanish Ministry for Education, Culture and Sports; BIO2003-00132, GEN2003-20642 and BIO2004-20369-E from the former Spanish Ministry for Science and Technology; BIO2006-02668, BFU2006-09593 and CONSOLIDER-INGENIO 2010 Project 'La Factoría de Cristalización' (CSD2006-00015) from the Spanish Ministry for Education and Science; EU FP6 Integrated Project LSHC-CT-2003-503297 'CANCERDEGRADOME' and EU FP6 Strep Project 18830 'CAMP'; 'AVON-Project' 2005X0648 from the Scientific Foundation of the Spanish Association Against Cancer; Grant 2005SGR00280 from the Generalitat of Catalunya; and grants 31-67253.01 and 3100AO-108262/1 from the Swiss National Science Foundation. Y.Y. and J.-L.R. acknowledge Protéus S.A., Nîmes (France) and the Swiss National Science Foundation. M.S. is a beneficiary of the 'Ramón y Cajal' Program of the Spanish Ministry for Science and Education. The Proteomics Laboratory at VHUH-RI belongs to the Spanish National Institute of Proteomics, PROTEORED. We are further indebted to C. Gyrup and C. Oxvig, University of Aarhus, Denmark, for providing IGFBP-5 and for assaying the cleavage of a fluorigenic peptide with PAPP-A.

References

Arribas, J. and Merlos-Suárez, A. (2003). Shedding of plasma membrane proteins. Curr. Top. Dev. Biol. *54*, 125–144.

Aster, J.C., Simms, W.B., Zavala-Ruiz, Z., Patriub, V., North, C.L., and Blacklow, S.C. (1999). The folding and structural integrity of the first LIN-12 module of human Notch1 are calcium-dependent. Biochemistry *38*, 4736–4742.

Bach, L.A., Headey, S.J., and Norton, R.S. (2005). IGF-binding proteins – the pieces are falling into place. Trends Endocrinol. Metab. *16*, 228–234.

Bahr, C. and Groner, B. (2005). The IGF-1 receptor and its contributions to metastatic tumor growth – novel approaches to the inhibition of IGF-1R function. Growth Factors *23*, 1–14.

Baxter, R.C. (1994). Insulin-like growth factor binding proteins in the human circulation: a review. Horm. Res. *42*, 140–144.

Baxter, R.C. (2000). Insulin-like growth factor (IGF)-binding proteins: interactions with IGFs and intrinsic bioactivities. Am. J. Physiol. Endocrinol. Metab. *278*, E967–E976.

Bayes-Genis, A., Conover, C.A., Overgaard, M.T., Bailey, K.R., Christiansen, M., Holmes, D.R. Jr., Virmani, R., Oxvig, C., and Schwartz, R.S. (2001a). Pregnancy-associated plasma protein A as a marker of acute coronary syndromes. N. Engl. J. Med. *345*, 1022–1029.

Bayes-Genis, A., Schwartz, R.S., Lewis, D.A., Overgaard, M.T., Christiansen, M., Oxvig, C., Ashai, K., Holmes, D.R. Jr., and Conover, C.A. (2001b). Insulin-like growth factor binding protein-4 protease produced by smooth muscle cells increases in the coronary artery after angioplasty. Arterioscler. Thromb. Vasc. Biol. *21*, 335–341.

Blobel, C.P. (2005). ADAMs: key components in EGFR signalling and development. Nat. Rev. Mol. Cell. Biol. *6*, 32–43.

Bode, W., Gomis-Rüth, F.X., and Stöcker, W. (1993). Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. FEBS Lett. *331*, 134–140.

Boldt, H.B., Overgaard, M.T., Laursen, L.S., Weyer, K., Sottrup-Jensen, L., and Oxvig, C. (2001). Mutational analysis of the proteolytic domain of pregnancy-associated plasma protein-A (PAPP-A): classification as a metzincin. Biochem. J. *358*, 359–367.

Boldt, H.B., Kjaer-Sorensen, K., Overgaard, M.T., Weyer, K., Poulsen, C.B., Sottrup-Jensen, L., Conover, C.A., Giudice, L.C., and Oxvig, C. (2004). The Lin12-notch repeats of pregnancy-associated plasma protein-A bind calcium and determine its proteolytic specificity. J. Biol. Chem. *279*, 38525–38531.

Bunn, R.C. and Fowlkes, J.L. (2003). Insulin-like growth factor binding protein proteolysis. Trends Endocrinol. Metab. *14*, 176–181.

Burger, A.M., Leyland-Jones, B., Banerjee, K., Spyropoulos, D.D., and Seth, A.K. (2005). Essential roles of IGFBP-3 and IGFBP-rP1 in breast cancer. Eur. J. Cancer *41*, 1515–1527.

Chen, B.K., Leiferman, K.M., Pittelkow, M.R., Overgaard, M.T., Oxvig, C., and Conover, C.A. (2003). Localization and regulation of pregnancy-associated plasma protein A expression in healing human skin. J. Clin. Endocrinol. Metab. *88*, 4465–4471.

Conover, C.A., Faessen, G.F., Ilg, K.E., Chandrasekher, Y.A., Christiansen, M., Overgaard, M.T., Oxvig, C., and Giudice, L.C. (2001). Pregnancy-associated plasma protein-A is the insulin-like growth factor binding protein-4 protease secreted

by human ovarian granulosa cells and is a marker of dominant follicle selection and the corpus luteum. Endocrinology *142*, 2155–2155.

Daughaday, W.H. (ed.) (1998). From Sulfation Factor to IGF-1, 40 Years of Research on the Regulation of Cartilage Growth (Amsterdam, The Netherlands: Elsevier).

Durai, R., Yang, W., Gupta, S., Seifalian, A.M., and Winslet, M.C. (2005). The role of the insulin-like growth factor system in colorectal cancer: review of current knowledge. Int. J. Colorectal Dis. *20*, 203–220.

Fang, P., Hwa, V., and Rosenfeld, R. (2004). IGFBPs and cancer. Novartis Found. Symp. *262*, 215–234.

Farr, M., Strube, J., Geppert, H.G., Kocourek, A., Mahne, M., and Tschesche, H. (2000). Pregnancy-associated plasma protein-E (PAPP-E). Biochim. Biophys. Acta *1493*, 356–362.

Fowlkes, J.L., Enghild, J.J., Suzuki, K., and Nagase, H. (1994). Matrix metalloproteinases degrade insulin-like growth factor-binding protein-3 in dermal fibroblast cultures. J. Biol. Chem. *269*, 25742–25746.

Fraústo da Silva, J.J.R. and Williams, R.J.P. (2001). The biological chemistry of the elements: the inorganic chemistry of life (New York, USA: Oxford University Press Inc.).

Gall, S.A. and Halbert, S.P. (1972). Antigenic constituents in pregnancy plasma which are undetectable in normal non-pregnant female or male plasma. Int. Arch. Allergy Appl. Immunol. *42*, 503–515.

García-Castellanos, R., Tallant, C., Marrero, A., Solà, M., Baumann, U., and Gomis-Rüth, F.X. (2007). Substrate specificity of a metalloprotease of the pappalysin family revealed by an inhibitor and a product complex. Arch. Biochem. Biophys. *457*, 57–72.

Garrigue-Antar, L., Francois, V., and Kadler, K.E. (2004). Deletion of epidermal growth factor-like domains converts mammalian tolloid into a chordinase and effective procollagen C-proteinase. J. Biol. Chem. *279*, 49835–49841.

Giudice, L.C., Conover, C.A., Bale, L., Faessen, G.H., Ilg, K., Sun, I., Imani, B., Suen, L.F., Irwin, J.C., Christiansen, M., et al. (2002). Identification and regulation of the IGFBP-4 protease and its physiological inhibitor in human trophoblasts and endometrial stroma: evidence for paracrine regulation of IGF-II bioavailability in the placental bed during human implantation. J. Clin. Endocrinol. Metab. *87*, 2359–2366.

Gomis-Rüth, F.X. (2003). Structural aspects of the metzincin clan of metalloendopeptidases. Mol. Biotech. *24*, 157–202.

Gomis-Rüth, F.X., Meyer, E.F., Kress, L.F., and Politi, V. (1998). Structures of adamalysin II with peptidic inhibitors. Implications for the design of tumor necrosis factor α convertase inhibitors. Protein Sci. *7*, 283–292.

Gyrup, C. and Oxvig, C. (2007). Quantitative analysis of insulin-like growth factor (IGF) modulated proteolysis of IGF binding protein-4 and -5 by pregnancy-associated plasma protein-A. Biochemistry *46*, 1972–1980.

Hartigan, N., Garrigue-Antar, L., and Kadler, K.E. (2003). Bone morphogenetic protein-1 (BMP-1). Identification of the minimal domain structure for procollagen C-proteinase activity. J. Biol. Chem. *278*, 18045–18049.

Hintze, V., Höwel, M., Wermter, C., Grosse Berkhoff, E., Becker-Pauly, C., Beermann, B., Yiallouros, I., and Stöcker, W. (2006). The interaction of recombinant subdomains of the procollagen C-proteinase with procollagen I provides a quantitative explanation for functional differences between the two splice variants, mammalian tolloid and bone morphogenetic protein 1. Biochemistry *45*, 6741–6748.

Holly, J. (2004). Physiology of the IGF system. Novartis Found. Symp. *262*, 19–35.

Hwa, V., Oh, Y., and Rosenfeld, R.G. (1999). The insulin-like growth factor-binding protein (IGFBP) superfamily. Endocr. Rev. *20*, 761–787.

Kumar, A., Mohan, S., Newton, J., Rehage, M., Tran, K., Baylink, D.J., and Qin, X. (2005). Pregnancy-associated plasma protein-A regulates myoblast proliferation and differentiation through an insulin-like growth factor-dependent mechanism. J. Biol. Chem. *280*, 37782–37789.

Laursen, L.S., Overgaard, M.T., Soe, R., Boldt, H.B., Sottrup-Jensen, L., Giudice, L.C., Conover, C.A., and Oxvig, C. (2001). Pregnancy-associated plasma protein-A (PAPP-A) cleaves insulin-like growth factor binding protein (IGFBP)-5 independent of IGF: implications for the mechanism of IGFBP-4 proteolysis by PAPP-A. FEBS Lett. *504*, 36–40.

Laursen, L.S., Overgaard, M.T., Nielsen, C.G., Boldt, H.B., Hopmann, K.H., Conover, C.A., Sottrup-Jensen, L., Giudice, L.C., and Oxvig, C. (2002a). Substrate specificity of the metalloproteinase pregnancy-associated plasma protein-A (PAPP-A) assessed by mutagenesis and analysis of synthetic peptides: substrate residues distant from the scissile bond are critical for proteolysis. Biochem. J. *367*, 31–40.

Laursen, L.S., Overgaard, M.T., Weyer, K., Boldt, H.B., Ebbesen, P., Christiansen, M., Sottrup-Jensen, L., Giudice, L.C., and Oxvig, C. (2002b). Cell surface targeting of pregnancy-associated plasma protein A proteolytic activity. Reversible adhesion is mediated by two neighboring short consensus repeats. J. Biol. Chem. *277*, 47225–47234.

Lawrence, J.B., Oxvig, C., Overgaard, M.T., Sottrup-Jensen, L., Gleich, G.J., Hays, L.G., Yates, J.R. III, and Conover, C.A. (1999). The insulin-like growth factor (IGF)-dependent IGF binding protein-4 protease secreted by human fibroblasts is pregnancy-associated plasma protein-A. Proc. Natl. Acad. Sci. USA *96*, 3149–3153.

LeRoith, D. and Helman, L. (2004). The new kid on the block(ade) of the IGF-1 receptor. Cancer Cell *5*, 201–202.

LeRoith, D. and Roberts, C.T. Jr. (2003). The insulin-like growth factor system and cancer. Cancer Lett. *195*, 127–137.

Maile, L.A. and Holly, J.M.P. (1999). Insulin-like growth factor binding protein (IGFBP) proteolysis: occurrence, identification, role and regulation. Growth Horm. IGF Res. *9*, 85–95.

Maskos, K. and Bode, W. (2003). Structural basis of matrix metalloproteinases and tissue inhibitors of metalloproteinases. Mol. Biotechnol. *25*, 241–266.

Miyakoshi, N., Richman, C., Kasukawa, Y., Linkhart, T.A., Baylink, D.J., and Mohan, S. (2001). Evidence that IGF-binding protein-5 functions as a growth factor. J. Clin. Invest. *107*, 73–81.

Mohan, S. and Baylink, D.J. (2002). IGF-binding proteins are multifunctional and act via IGF-dependent and -independent mechanisms. J. Endocrinol. *175*, 19–31.

Monget, P., Mazerbourg, S., Delpuech, T., Maurel, M.C., Maniere, S., Zapf, J., Lalmanach, G., Oxvig, C., and Overgaard, M.T. (2003). Pregnancy-associated plasma protein-A is involved in insulin-like growth factor binding protein-2 (IGFBP-2) proteolytic degradation in bovine and porcine preovulatory follicles: identification of cleavage site and characterization of IGFBP-2 degradation. Biol. Reprod. *68*, 77–86.

Morgunova, E., Tuuttila, A., Bergmann, U., Isupov, M., Lindqvist, Y., Schneider, G., and Tryggvason, K. (1999). Structure of human pro-matrix metalloproteinase-2: activation mechanism revealed. Science *284*, 1667–1670.

Nakamura, M., Miyamoto, S., Maeda, H., Ishii, G., Hasebe, T., Chiba, T., Asaka, M., and Ochiai, A. (2005). Matrix metalloproteinase-7 degrades all insulin-like growth factor binding proteins and facilitates insulin-like growth factor bioavailability. Biochem. Biophys. Res. Commun. *333*, 1011–1016.

Overall, C.M. and López-Otín, C. (2002). Strategies for MMP inhibition in cancer: innovations for the post-trial era. Nat. Rev. Cancer *2*, 657–672.

Overgaard, M.T., Haaning, J., Boldt, H.B., Olsen, I.M., Laursen, L.S., Christiansen, M., Gleich, G.J., Sottrup-Jensen, L., Conover, C.A., and Oxvig, C. (2000). Expression of recombinant human pregnancy-associated plasma protein-A and identification of the proform of eosinophil major basic protein as its physiological inhibitor. J. Biol. Chem. *275*, 31128–31133.

Overgaard, M.T., Boldt, H.B., Laursen, L.S., Sottrup-Jensen, L., Conover, C.A., and Oxvig, C. (2001). Pregnancy-associated plasma protein-A2 (PAPP-A2), a novel insulin-like growth factor-binding protein-5 proteinase. J. Biol. Chem. *276*, 21849–21853.

Oxvig, C., Sand, O., Kristensen, T., Gleich, G.J., and Sottrup-Jensen, L. (1993). Circulating human pregnancy-associated plasma protein-A is disulfide-bridged to the proform of eosinophil major basic protein. J. Biol. Chem. *268*, 12243–12246.

Oxvig, C., Overgaard, M.T., and Sottrup-Jensen, L. (2004). 220. Pappalysin-1 (pregnancy-associated plasma protein-A). In: Handbook of Proteolytic Enzymes Vol. 1, A.J. Barrett, N.D. Rawlings and J.F. Woessner Jr., eds. (London, UK: Elsevier), pp. 754–757.

Polgár, L. (1999). Basic kinetic mechanisms of proteolytic enzymes. In: Proteolytic Enzymes – Tools and Targets, E.E. Sterchi and W. Stöcker, eds. (Berlin/Heidelberg, Germany: Springer Verlag), pp. 148–166.

Pollak, M.N., Schernhammer, E.S., and Hankinson, S.E. (2004). Insulin-like growth factors and neoplasia. Nat. Rev. Cancer *4*, 505–518.

Salmon, W.D. and Daughaday, W.H. (1956). Sulfation factor, a serum component mediating the action of growth hormone in stimulating incorporation of sulfate into cartilage. J. Clin. Invest. *35*, 733–733.

Stöcker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Rüth, F.X., McKay, D.B., and Bode, W. (1995). The metzincins – topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. Protein Sci. *4*, 823–840.

Tallant, C., García-Castellanos, R., Seco, J., Baumann, U., and Gomis-Rüth, F.X. (2006). Molecular analysis of ulilysin, the structural prototype of a new family of metzincin metalloproteases. J. Biol. Chem. *281*, 17920–17928.

Thorn, E.M. and Khan, I.A. (2007). Pregnancy-associated plasma protein-A: an emerging cardiac biomarker. Int. J. Cardiol. *117*, 370–372.

Yang, Y. and Reymond, J.L. (2005). Protease profiling using a fluorescent domino peptide cocktail. Mol. Biosyst. *1*, 57–63.

Yu, H. and Rohan, T. (2000). Role of the insulin-like growth factor family in cancer development and progression. J. Natl. Cancer Inst. *92*, 1472–1489.

Zapf, J., Waldvogel, M., and Froesch, E.R. (1975). Binding of nonsuppressible insulin-like activity to human serum. Evidence for a carrier protein. Arch. Biochem. Biophys. *168*, 638–645.

Received May 8, 2007; accepted July 12, 2007

Addendum

Below are some unintentionally omitted key references regarding the structural characterization of the IGF/IGFBP axis:

Kalus, W., Zweckstetter, M., Renner, C., Sanchez, Y., Georgescu, J., Grol, M., Demuth, D., Schumacher, R., Dony, C., Lang, K., and Holak, T.A. (1998). Structure of the IGF-binding domain of the insulin-like growth factor-binding protein-5 (IGFBP-5): implications for IGF and IGF-I receptor interactions. EMBO J. *17*, 6558–6572.

Kamionka, M., Rehm, T., Beisel, H.G., Lang, K., Engh, R.A., and Holak, T.A. (2002). *In silico* and NMR identification of inhibitors of the IGF-I and IGF-binding protein-5 interaction. J. Med. Chem. *45*, 5655–5660.

Renner, C. and Holak, T. (2001). NMR ^{15}N relaxation of the insulin-like growth factor (IGF)-binding domain of IGF binding protein-5 (IGFBP-5) determined free in solution and in complex with IGF-II. Eur. J. Biochem. *268*, 1058–1065.

Sitar, T., Popowicz, G.M., Siwanowicz, I., Huber, R., and Holak, T.A. (2006). Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins. Proc. Natl. Acad. Sci. USA *103*, 13028–13033.

Siwanowicz, I., Popowicz, G.M., Wisniewska, M., Huber, R., Kuenkele, K.P., Lang, K., Engh, R.A., and Holak, T.A. (2005). Structural basis for the regulation of insulin-like growth factors by IGF binding proteins. Structure *13*, 155–167.

Zeslawski, W., Beisel, H.G., Kamionka, M., Kalus, W., Engh, R.A., Huber, R, Lang, K., and Holak, T.A. (2001). The interaction of insulin-like growth factor-I with the N-terminal domain of IGFBP-5. EMBO J. *20*, 3638–3644.