Biochemical and Biophysical Research Communications 370 (2008) 566–571

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Autoproteolytic stability of a trypsin from the marine crab *Cancer pagurus*

Jan-Hendrik Hehemann [a,1], Lars Redecke [a,1], Jayaseelan Murugaiyan [b], Martin von Bergen [b], Christian Betzel [a], Reinhard Saborowski [c,*]

[a] Institute of Biochemistry and Food Chemistry, University of Hamburg, 20146 Hamburg, Germany
[b] Department of Proteomics, Helmholtz Centre for Environmental Research—UFZ, 04318 Leipzig, Germany
[c] Biologische Anstalt Helgoland, Alfred Wegener Institute for Polar and Marine Research, 27498 Helgoland, Germany

ARTICLE INFO

Article history:
Received 17 March 2008
Available online 4 April 2008

Keywords:
Crustacea
Crab
Cancer pagurus
Gastric fluid
Digestive enzymes
Trypsin
Autoproteolysis
Stability

ABSTRACT

Autoproteolytic stability is a crucial factor for the application of proteases in biotechnology. In contrast to vertebrate enzymes, trypsins from shrimp and crayfish are known to be resistant against autolysis. We show by characterisation of a novel trypsin from the gastric fluid of the marine crab *Cancer pagurus* that this property might be assigned to the entire class of crustaceans. The isolated and cloned crab trypsin (*C.p.*TryIII) exhibits all characteristic properties of crustacean trypsins. However, its overall sequence identity to other trypsins of this systematic class is comparatively low. The high resistance against autoproteolysis was determined by mass spectrometry, which revealed a low susceptibility of the N-terminal domain towards autolysis. By homology modelling of the tertiary structure, the elevated stability was attributed to the distinctly different pattern of autolytic cleavage sites, which is conserved in all known crustacean trypsin sequences.

Trypsins (E.C. 3.4.21.4) belong to the S1 family of serine endopeptidases and have been identified in organisms from various phyla [1,2]. Proteases from crustaceans have a potential for biotechnological applications, because they can be easily obtained and some Crustaceans produce high amounts with remarkable activities [3]. Crustacean trypsins were already isolated and characterised from crayfishes [4–6], shrimps [7,8], and king crab [9]. Moreover, the X-ray structures of two crustacean serine proteases [10,11] have been solved. Interestingly, crustacean trypsins turned out to be more stable than their mammalian counterparts. An important factor that determines the stability of trypsins is autolysis [12–14]. It has been suggested that the high stability of crustacean trypsins is a result of the reduced number and/or the altered localisation of autolytic cleavage sites, particularly at residue 145 [4,7,9]. However, in the meantime it was shown by site directed mutagenesis in mammalian trypsins [13] that Lys61 and Arg117 play a key role in rapid autolytic degradation. As all further N-terminal cleavage sites are located within this region, hydrolysis at these two sites exposes several other so far protected cleavage sites, resulting in complete digestion of the N-terminal domain. Since two members of the catalytic triade (His57 and Asp102) are affected by the structural destruction, a dramatic loss of activity is observed [13,15,16].

Stability of crustacean trypsins has never been mechanistically analysed. Therefore, we carried out a comparative study on the stability of a crustacean and a mammalian trypsin. The marine crab *Cancer pagurus* (Brachyura) expresses high amounts of trypsins with specific catalytic properties in terms of stability, activation by organic solvents, and high resistance against autolysis [17]. We isolated a novel trypsin from *C. pagurus* and showed by various analyses that the high autolytic resistance of this crustacean trypsin can be clearly explained by altered patterns of tryptic cleavage sites.

Materials and methods

Samples. Crabs, *C. pagurus* (Decapoda) were sampled in the North sea of Helgoland, Germany. Gastric fluid was drawn from the stomach of the crab [17] and centrifuged at 15,000*g* for 10 min. Proteins from the supernatant were precipitated with ammonium sulphate (50% w/v). After centrifugation the protein pellet was dissolved in imidazole buffer (10 mM, pH 6.8 containing 0.02% sodium azide) and stored at −80 °C.

Protein and enzyme assays. Soluble protein was determined by Coomassie dye reaction (BioRad, Hercules, USA) using bovine serum albumin (BSA) as standard. Trypsin activity was assayed with the substrate L-benzoyl-arginyl-*para*-nitroanilide (L-BAPNA, Fluka 12915) [18].

Enzyme purification and N-terminal sequencing. The dissolved protein pellet (1 ml) was loaded onto a NAP™-10 Sephadex G25 gel filtration column (GE Healthcare, NJ, USA) and eluted with 1.5 ml of 10 mM imidazole buffer (pH 6.8, buffer A). The eluate was subsequently loaded onto a pre-equilibrated (buffer A) UNO™ Q6 anion-exchange column (BioRad), which was connected to a FPLC ÄKTA Protein Puri-

* Corresponding author. Address: Biologische Anstalt Helgoland, Alfred Wegener Institute for Polar and Marine Research, Kurpromenade (Building A-107), 27498 Helgoland, Germany. Fax: +49 (0) 4725 819 3369.
E-mail address: Reinhard.Saborowski@awi.de (R. Saborowski).
[1] These authors contributed equally to this work.

0006-291X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2008.03.128

fier (GE Healthcare). Bound proteins were eluted with an increasing linear gradient from 0 to 1 M NaCl (in buffer A) into reaction tubes placed on ice. The absorbance at 280 nm was continuously recorded. Fractions with trypsin activity were pooled and dialysed overnight against buffer A at 4 °C. SDS–PAGE followed by Coomassie staining was applied to monitor the purification progress. For N-terminal sequencing, the purified enzyme was blotted (60 min, 50 mA, 4 °C) onto a PVDF membrane (Applied Biosystems, ProBlot™) in 20 mM CAPS buffer (pH 11) containing 10% methanol using a mini trans-blot cell (BioRad, Hercules, USA). Edman-degradation (10 cycles) of the Coomassie-stained and excised protein was carried out using a Protein-Sequencer Procise 494 (Applied Biosystems, Foster City, USA).

Internal sequencing by mass spectrometry. Chemical cleavage of the crab and the porcine trypsin was carried out following Fontana [19]. After lyophilization, the fragments were separated by SDS–PAGE. Protein bands of interest were cut out and digested overnight with commercial trypsin (Sigma, Seelze, Germany) [20]. The cleaved peptides were eluted, concentrated by vacuum centrifugation and separated by Agilent nano-LC (Agilent Technologies, Palo Alto, CA, USA). The peptides were identified by on-line MS/MS (LC/MSD TRAP XCT mass spectrometer, Agilent Technologies) as described elsewhere [21]. Additionally, peptides were also spotted on anchorchip target plate and identification was carried out with MALDI-TOF MS/MS (Ultraflex III TOF/TOF, Bruker Daltonics, Bremen, Germany). A database search was conducted using the MS/MS ion search (MASCOT, http://www.matrixscience.com) against all metazoan (animals) entries of NCBInr (GenBank). '*De-novo*' sequencing of these spectra was performed using the algorithm supplied by the Biotools software, and manually inspected to verify the assignment of the characteristic peptide fragment ions.

Isolation of RNA, cloning, and DNA sequencing. Total RNA was directly isolated from fresh tissue samples of *C. pagurus* midgut gland (100 mg in RNA Later buffer, Qiagen, Hilden, Germany) using the RNeasy Mini Kit (Qiagen, Hilden, Germany). Following mRNA extraction with the Oligotex mRNA mini Kit (Qiagen, Hilden, Germany), 1 pg of the isolated mRNA was transcribed by RT-PCR using an oligo-dT_{30} primer and the Thermo-X™ Reverse Transcriptase (Invitrogen, Carlsbad, USA). The *C. pagurus* trypsin gene was amplified from the cDNA by polymerase chain reaction (PCR) using a degenerated oligonucleotide primer designed on the basis of the N-terminal amino acid sequence of *C. pagurus* trypsin (forward primer, 5′-GTN GGN GGN CAR GAY ACN G-3′) and an oligo-dT_{30} primer (reverse primer). To determine the DNA sequence of *C. pagurus* trypsin, the amplification product was purified by agarose gel electrophoresis and ligated into a TOPO TA vector using the TOPO TA Cloning Kit (Invitrogen). Automated DNA sequencing was performed from both directions with M13 plasmid specific oligonucleotide primers using the Big Dye Terminator Cycle Sequencing kit (Applied Biosystems). Amplification products were analysed by electrophoresis on an ABI Prism DNA Sequencer (Applied Biosystems).

Data evaluation and model calculation. The sequences for multiplan alignments were obtained from GenBank and/or Swiss-Prot/TrEMBL by homology search (www.ncbi.nlm.nih.gov/blast [22]). The translated DNA sequence of mature *C.p.*TryIII was used as a start sequence. Alignments were performed with the program ClustalW (http://www.ebi.ac.uk/clustalw [23]) solely using the amino acid sequences of the mature trypsins starting with the N-terminal sequence IVGG. All pro-peptide sequences were ignored for these analyses. The computer generated model of *C.p.*TryIII was calculated with the program Swiss-Model (http://swissmodel.expasy.org/; [24]) using the structure of crayfish trypsin, *Pontastacus leptodactylus* (PDB code 2F91), as a template. The final alignment of both structural models using the program Turbo Frodo [25] yielded an RMSD of 0.22 Å.

Results and discussion

Isolation and characterisation of C. pagurus trypsin

We separated three distinct peaks with trypsin activity (BAPNA hydrolysis) from the gastric fluid of the marine crab *C. pagurus* by anion-exchange chromatography. They eluted at concentrations of about 0.5, 0.6, and 0.7 M NaCl, respectively. The separately combined fractions of each peak were denoted as *C.p.*TryI, *C.p.*TryII, and *C.p.*TryIII (Fig. 1). SDS–PAGE analysis revealed highest purity and homogeneity for *C.p.*TryIII showing a single protein band with an apparent molecular mass of 26 kDa (Fig. 1, inset). Therefore, we focussed all further investigations on this enzyme. *C.p.*TryIII activity was completely (>95%) inhibited by serine protease inhibitors AEBSF and SBTI and to a lesser extent by the trypsin-specific inhibitor TLCK (>50%). Aspartate (PepA) and cysteine (E64) protease inhibitors as well as the chymotrypsin inhibitor TPCK had no effect (data not shown).

We sequenced the N-terminus by Edman-degradation and obtained 10 amino acids (IVGGQDTVLG) containing the motif IVGG which is conserved for almost all mature trypsins. The transcript of *C.p.*TryIII was obtained by RT-PCR on total mRNA of *C. pagurus* midgut gland tissue and cloned into a TOPO TA vector. The open reading frame (ORF) consisted of 705 bp, coding for a mature and active protein of 235 residues with a calculated molecular mass of 25.7 kDa and a theoretical isoelectric point (p*I*) of 4.38. These parameters are close to those of mature trypsins from other crustaceans [8,10]. Almost 30 % of the deduced *C.p.*TryIII amino acid sequence was additionally confirmed by internal mass spectrometry sequencing (Fig. 2). Consequently, the cloned cDNA clearly corresponds to the purified enzyme that was classified as a trypsin.

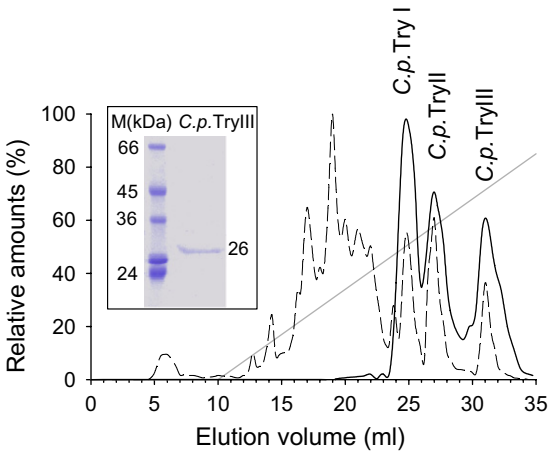

Fig. 1. Separation of gastric fluid proteins from the marine crab *Cancer pagurus* by anion-exchange chromatography. Proteins detected by UV-absorbance at 280 nm (dashed line) were eluted with increasing concentrations of NaCl (0–1 M, grey straight line). Three peaks with tryptic activity (solid black line) were detected by enzyme assays and were pooled for further analysis (*C.p.*TryI–*C.p.*TryIII). Inset: SDS–PAGE of *C.p.*TryIII shows a separate protein band with an apparent molecular mass of approximately 26 kDa.

Sequence similarity with crustacean and vertebrate trypsins

The amino acid sequence of *C.p.*TryIII was similar to other crustacean and eukaryotic trypsin sequences (Fig. 2) and displays the common features of trypsins. In this respect we found the conserved residues His57, Asp102, and Ser195 which form the catalytic triad, as well as Gly216 and Gly226, which determine the trypsin specificity. The expected evolutionary distance of *C.p.*TryIII to trypsins from mammals and other vertebrates (Table 1) was confirmed by specific features that are only characteristic for crustacean trypsins: (i) three sequence insertions; (ii) a highly acidic p*I* due to an increased amount of Asp/Glu residues, while the numbers of Lys/Arg residues were slightly reduced; (iii) a decreased and, for some species including *C. pagurus*, impaired number of Cys residues [5,7–10]. *C.p.*TryIII has eight cysteines in conserved positions, which indicate a comparable pattern of the associated disulphide bonds. The additional Cys56 is so far unique for trypsins from Brachyura (*C. pagurus* and *P. pelagicus*). However, evidence for a dimerization of *C.p.*TryIII molecules via this free cysteine in solution was not detected. Despite the structural homology, the degree of overall sequence identity with other crustacean trypsins, particularly from crayfishes (Astacura) [5,10], was comparatively low (62–65%). Merely, a trypsin sequence from the blue swimmer crab *Portunus pelagicus* matched to 81% with the sequence of *C.p.*TryIII at the time of this study. Since both species belong to the same systematic group (Brachyura), these results indicate distinct phylogenetic differences between the groups of higher crustaceans. Moreover, even the species within the same systematic group apparently retain discrete characteristics.