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.

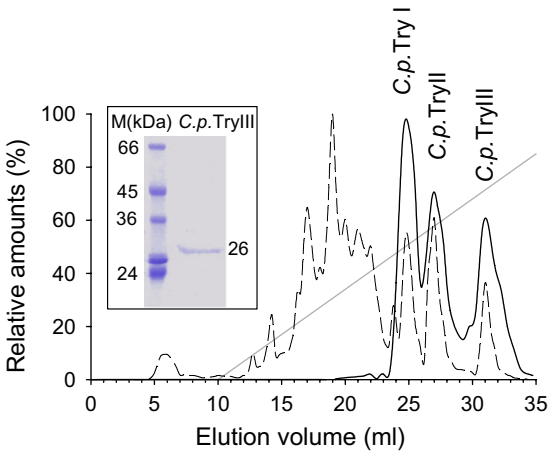

Fig. 1. Separation of gastric fluid proteins from the marine crab *Cancer pagurus* by anion-exchange chromatography. Proteins detected by UV-absorbance at 280 nm (dashed line) were eluted with increasing concentrations of NaCl (0–1 M, grey straight line). Three peaks with tryptic activity (solid black 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.

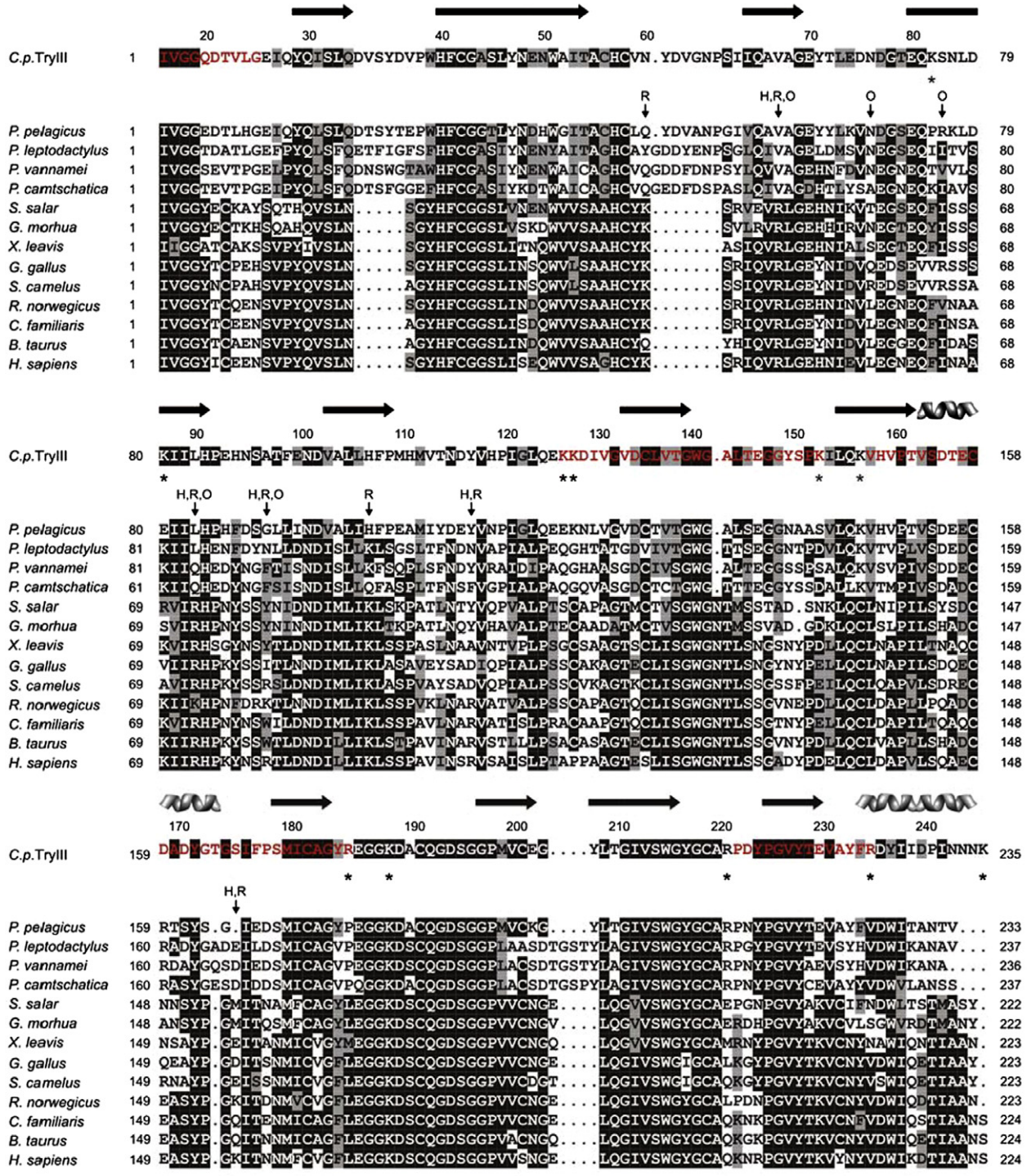

Fig. 2. Comparison of the deduced amino acid sequence of *C.p.*TryIII with known crustacean, fish, frog, and vertebrate trypsin sequences by multiple alignment. Swiss-Prot/TrEMBL accession numbers of the proteins are listed in Table 1. Identical amino acids are highlighted in black, homologous residues in grey. Regions of the deduced *C.p.*TryIII sequence that were confirmed by mass spectrometry sequencing are marked in red. Autolysis sites already reported for human (H), rat (R), and ostrich (O) trypsins are labelled with arrows, while potential autolysis sites within the sequence of the *C. pagurus* trypsin (Arg/Lys residues) are marked by asterisks. Numbers in the first row show the conventional chymotrypsin numbering, while numbers beside the sequences indicate the exact amino acid position. The location of the secondary structure elements derived from the 3D models of human (PDB code 1TRN), bovine (PDB code 1D80), salmon (PDB code 1HJ8), and crayfish (PDB code 2F91) trypsins is schematically indicated above the sequences. (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

Enzyme stability and autoproteolysis

In agreement with previous analyses of the crude gastric fluid of *C. pagurus* [17], an increased stability compared to widely used commercial mammalian trypsins was also confirmed for purified *C.p.*TryIII. After incubation at RT for 21 days, the activity of *C.p.*TryIII decreased to about 60% of the initial activity, while porcine trypsin lost about 85% of its activity (Fig. 3A). However, the primary structure of *C.p.*TryIII remained almost intact. Only a slight degradation of about 1 kDa was observed by SDS–PAGE, which was completed after four weeks of incubation (Fig. 3B). The N-termini of both proteins were identical as verified by Edman sequencing.

Table 1
Similarity of selected anionic trypsins from crustacean and vertebrates with *C.p.*TryIII as obtained from BLAST database search

Species	Common name	UniProtKB/ TrEMBL entry	Similarity (%)	Identity (%)
Portunus pelagicus	Blue swimmer crab	A2I7J3	81	71
Pennaeus vannamei	Shrimp	O62561	66	56
Pontastacus leptodactylus	Crayfish	Q52V24	65	54
Paralithodes camtchatica	King crab	Q8WR10	65	52
Salmo salar	Salmon	P35031	56	40
Gadus morhua	Cod	P16049	54	38
Xenopus laevis	Frog	P19799	56	39
Gallus gallus	Chicken	Q90629	56	41
Struthio camelus	Ostrich	Q6GYJ5	57	41
Rattus norvegicus	Rat	P00763	55	42
Canis familiaris	Dog	P06872	57	43
Bos taurus	Cattle	Q29463	58	42
Homo sapiens	Man	P07478	56	43

Therefore, we conclude that the loss of mass is caused by hydrolysis at the exposed site Arg235 of the C-terminus, which liberates a peptide of 11 amino acids obviously without a strong effect on the activity of *C.p.*TryIII. Additional autolysis fragments were only detected by highly sensitive mass spectrometry after 3 h of incubation at RT (Fig. 3C). The most abundant peptide with a mass of 3191.2 *m*/*z* represents the residues V158–R186 within the C-terminal domain of *C.p.*TryIII. Further peptides in the spectrum were not unambiguously identified. In contrast, porcine trypsin exhibited a more complex peptide pattern in mass spectrometry analysis (Fig. 3D). The N-terminal peptides V118–R125 (*m*/*z* 842.50) and L108–R117 (*m*/*z* 1045.56) were clearly identified by MS^2-analysis. Moreover, fragmentation analysis identified the precursor masses of 1427.86 *m*/*z* (LSS-sequence tag, residues L108–S110) and 1666.01 *m*/*z* (VAT tag, residues V118–T120), which are also located within the N-terminal domain. Only one peptide with a mass of 1541.87 *m*/*z* (GVYT tag, residues G226–T229) could be assigned to the C-terminal part of porcine trypsin.

For interpretation of this data, we used computer generated homology modelling to propose a structural model of *C.p.*TryIII. Since all trypsins known to date show a high degree of conservation on the structural level [26–29], this is an appropriate method to gain basic structural information. The X-ray structure of a trypsin from the crayfish *P. leptodactylus* [10] (PDB code 2F91) was used as a template since it represents the most homologous trypsin sequence available so far.

As expected, the generated model of *C.p.*TryIII shows the typical overall backbone fold of trypsins including the amount and position of the secondary structures. The insertions of crustacean trypsins are solely located within three loops (Fig. 4A). Most interesting is the inversed distribution pattern of potential autolysis sites compared to vertebrate trypsins (Fig. 4B), which clearly explains the origin of the observed autolysis fragments. Although the total number of Arg/Lys (11) residues remains almost conserved, only the sites Lys82 and Lys87 are located within the critical N-terminal domain of *C.p.*TryIII, which are, however, suggested to be protected by the structure itself [15,16]. The highly sensitive sites of mammalian trypsins Lys61 and Arg117 [13,14] as well as another five N-terminal Arg/Lys residues are absent in *C.p.*TryIII. In contrast, the crab trypsin contains six additional Arg/Lys residues compared to rat trypsin within the C-terminal domain, which are mostly located in loop structures. However, the secondary structures of the C-terminal part of trypsin molecules are highly cross-linked by various disulphide bonds, which conserve the general structural integrity and, thereby, the activity even if partial autolysis occurs within the loops [30]. The two missing disulphide bonds of *C.p.*TryIII are supposed to be counterbalanced by electrostatic interactions as well as main-chain hydrogen bonds, as reported for the homologous crayfish trypsin [10].

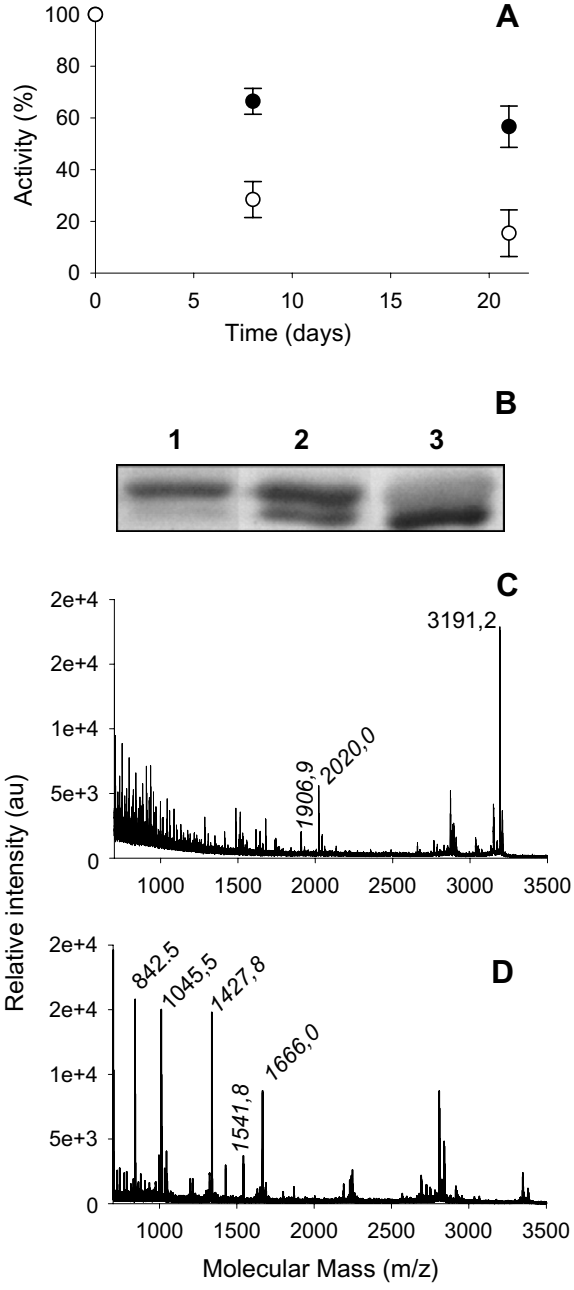

Fig. 3. Activity and autoproteolytic degradation of *C.p.*TryIII compared to porcine trypsin. (A) Activities of *C.p.*TryIII (filled circles) and porcine trypsin (open circles) after incubation at RT up to 21 days. (B) SDS–PAGE analysis of C-terminal autolytic truncation of *C.p.*TryIII (26–25 kDa) immediately after purification (1), after 2 weeks (2), and after 4 weeks (3) of storage at 4 °C. The appearance of autoproteolytic peptides from *C.p.*TryIII (C) and porcine trypsin (D) was also tested by MALDI-TOF/TOF analysis after 3 h of incubation at room temperature. Unambiguously identified peptides are denoted by regular font numbers, peptides identified by sequence tags in a *de novo* sequencing approach by italic font numbers.

Effect of calcium

Trypsins from vertebrates are significantly stabilized by Ca^{2+} ions [26,27]. However, in *C.p.*TryIII the presence of Ca^{2+}-ions (10 mM) neither affected the activity nor the stability up to

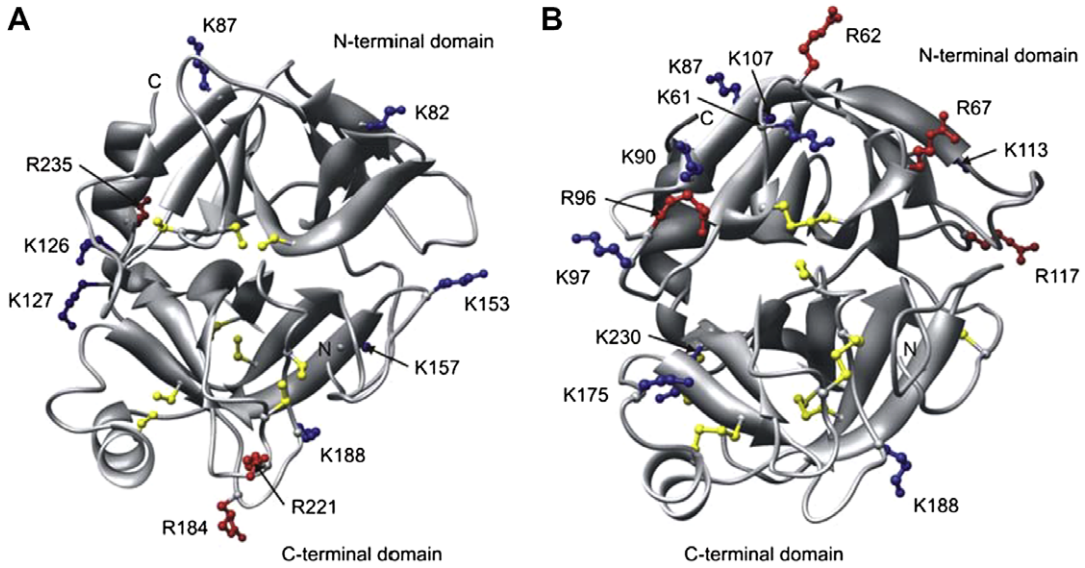

Fig. 4. Computer generated model of the three-dimensional structure of *C.p.*TryIII (A) compared to the three-dimensional structure of *Rattus norvegicus* trypsin (B) (PDB code 1HJ8). Arginine (red), lysine (blue), and cysteine (yellow) residues are displayed in 'ball-and-stick' representation. All Arg/Lys residues are labelled according to the common chymotrypsin numbering. (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

30 °C (data not shown). This different behaviour of Ca^{2+} also depends on structural features. In bovine trypsins Ca^{2+} is coordinated by residues Glu70 and Glu80. Since the Ca^{2+} site is placed within the 'self-destruction' segment Lys61 to Arg117, it is supposed that Ca^{2+} forms an autolysis-resistant conformation of this polypeptide chain [13]. The Ca^{2+}-binding site is also conserved in the sequence of crustacean trypsins including *C.p.*TryIII [4,7–9]. However, there are no accessible autolysis sites in the N-terminal domain, which need to be stabilized by Ca^{2+} coordination. Thus, a significant effect of Ca^{2+} was not observed.

Conclusion

In this study we proposed a molecular mechanism that clearly explains the enhanced autolysis stability of a novel crab trypsin. In contrast to earlier suggestions [4,5,7], the key factor of resistance is supposed to be the absence of the N-terminal 'self-destruction' segment rather than missing cleavage sites within the C-terminal part of the *C.p.*TryIII molecules. Due to the conservation of the required structural features within crustacean trypsins, a general validation for this class of enzymes is suggested. Autolysis is a limiting factor for specific application of proteases in laboratory and in biotechnological processes. Therefore, crustacean trypsins are interesting candidates for techniques that require proteases characterised by long-term stability and minimal autolysis.

Acknowledgments

We thank the crews of R/V Aade and R/V Uthörn for sampling crabs and A.-K. Langer as well as M. Risch for excellent laboratory assistance.

References

[1] E. DeVillez, K. Buschlen, Survey of a tryptic digestive enzyme in various species of crustacean, Comp. Biochem. Physiol. 21 (1967) 541–546.
[2] K.A. Walsh, Trypsinogens and trypsins of various species, Methods Enzymol. 19 (1970) 41–63.
[3] L. Hellgren, B. Karlstam, V. Mohr, J. Vincent, Krill enzymes. A new concept for efficient debridement of necrotic ulcers, Int. J. Dermatol. 30 (1991) 102–103.
[4] R. Zwilling, G. Pfleiderer, H.H. Sonneborn, V. Kraft, I. Stucky, The evolution of endopeptidases—V. Common and different traits of bovine and crayfish trypsin, Comp. Biochem. Physiol. 28 (1969) 1275–1287.
[5] K. Titani, T. Sasagawa, R.G. Woodbury, L.H. Ericsson, H. Dörsam, M. Kraemer, H. Neurath, R. Zwilling, Amino acid sequence of crayfish (*Astacus fluviatilis*) trypsin I_f, Biochemistry 22 (1983) 1459–1465.
[6] P. Hernández-Cortés, L. Cerenius, F. García-Carreño, K. Söderhäll, Trypsin from *Pacifastacus leniusculus* hepatopancreas: purification and cDNA cloning of the synthesized zymogen, Biol. Chem. 380 (1999) 499–501.
[7] B.J. Gates, J. Travis, Isolation and comparative properties of shrimp trypsin, Biochemistry 8 (1969) 4483–4489.
[8] B. Klein, G. Le Moullac, D. Sellos, A. Van Wormhoudt, Molecular cloning and sequencing of trypsin cDNAs from *Penaeus vannamei* (Crustacea, Decapoda): use in assessing gene expression during the moult cycle, Int. J. Biochem. Cell Biol. 28 (1996) 551–563.
[9] G.N. Rudenskaya, Y.A. Kislitsin, D.V. Rebrikov, Collagenolytic serine protease PC and trypsin PC from king crab *Paralithodes camtschaticus*: cDNA cloning and primary structure of the enzymes, BMC Struct. Biol. 4 (2004) 2.
[10] K. Fodor, V. Harmat, C. Hetényi, J. Kardos, J. Antal, A. Perczel, A. Patthy, G. Katona, L. Gráf, Extended intermolecular interactions in a serine protease–canonical inhibitor complex account for strong and highly specific inhibition, J. Mol. Biol. 350 (2005) 156–169.
[11] J.J. Perrona, C.A. Tsu, C.S. Craik, R.J. Fletterick, Crystal structure of an ecotin–collagenase complex suggests a model for recognition and cleavage of the collagen triple helix, Biochemistry 36 (1997) 5381–5392.
[12] R.L. Smith, E. Shaw, Pseudotrypsin. A modified bovine trypsin produced by limited autodigestion, J. Biol. Chem. 244 (1969) 4704–4712.
[13] E. Várallyay, G. Pal, A. Patthy, L. Szilagyi, L. Graf, Two mutations in rat trypsin confer resistance against autolysis, Biochem. Biophys. Res. Commun. 243 (1998) 56–60.
[14] B. Szenthe, C. Frost, L. Szilagyi, A. Patthy, R. Naude, L. Graf, Cloning and expression of ostrich trypsinogen: an avian trypsin with a highly sensitive autolysis site, Biochim. Biophys. Acta 1748 (2005) 35–42.
[15] Z. Kukor, M. Tóth, G. Pál, M. Sahin-Tóth, Arg-117 is the reactive site of an inhibitory surface loop that controls spontaneous zymogen activation, J. Biol. Chem. 277 (2002) 6111–6117.
[16] S. Maroux, P. Desnuelle, On some autolyzed derivatives of bovine trypsin, Biochim. Biophys. Acta 181 (1969) 59–72.
[17] R. Saborowski, G. Sahling, M.A. Navarrete del Toro, I. Walter, F.L. García-Carreño, Stability and effects of organic solvents on endopeptidases from the gastric fluid of the marine crab *Cancer pagurus*, J. Mol. Catal. B Enzym. 30 (2004) 109–118.
[18] B.F. Erlanger, N. Kokowsky, W. Cohen, The preparation and properties of two new chromogenic substrates of trypsin, Arch. Biochem. Biophys. 95 (1961) 271–278.
[19] A. Fontana, Modification of tryptophan with BNPS skatole (2-(2-nitrophenylsulfenyl)-3-methyl-3-bromoindolenine), Methods Enzymol. 25 (1972) 419–423.
[20] A. Shevchenko, M. Wilm, O. Vorm, O.N. Jensen, A.V. Podtelejnikov, G. Neubauer, P. Mortensen, M. Mann, A strategy for identifying gel separated proteins in sequence databases by MS alone, Biochem. Soc. Trans. 24 (1996) 893–896.

[21] D. Benndorf, U. Balcke, H. Harms, M. von Bergen, Functional metaproteome analysis of protein extracts from contaminated soil and ground water, ISMI J. (2007) 1–11.
[22] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[23] J.D. Thompson, D.G. Higgins, T.J. Gibson, ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[24] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: an automated protein homology-modeling server, Nucleic Acids Res. 31 (2003) 3381–3385.
[25] A. Roussel, C. Cambillau, TURBO-FRODO, A General-Purpose Molecular-Modelling Environment, AFMB-CNRS (2001), 13402 Marseille CEDEX 20, France.
[26] W. Bode, P. Schwager, The refined crystal structure of bovine betatrypsin at 1.8 A resolution. II. Crystallographic refinement, calcium binding site, benzamidine binding site and active site at pH 7.0, J. Mol. Biol. 98 (1975) 693–717.
[27] L.T. Delbaere, W.L. Hutcheon, M.N. James, W.E. Thiessen, Tertiary structural differences between microbial serine proteases and pancreatic serine enzymes, Nature 257 (1975) 758–763.
[28] W. Bode, Z. Chen, K. Bartels, C. Kutzbach, G. Schmidt-Kastner, H. Bartunik, Refined 2 A X-ray crystal structure of porcine pancreatic kallikrein A, a specific trypsin-like serine proteinase. Crystallization, structure determination, crystallographic refinement, structure and its comparison with bovine trypsin, J. Mol. Biol. 164 (1983) 237–282.
[29] A.M. Lesk, W.D. Fordham, Conservation and variability in the structures of serine proteinases of the chymotrypsin family, J. Mol. Biol. 258 (1996) 501–537.
[30] P.M. Harrison, M.J. Sternberg, The disulphide beta-cross: from cystine geometry and clustering to classification of small disulphide-rich protein folds, J. Mol. Biol. 264 (1996) 603–623.