Heterologous expression and characterization of a novel serine protease from *Daphnia magna*: A possible role in susceptibility to toxic cyanobacteria

Jacqueline Lange[1]*, Fatih Demir[2], Pitter F. Huesgen[2], Ulrich Baumann[3], Eric von Elert[1] and Christian Pichlo[3]

From the [1]Department of Aquatic Chemical Ecology, University of Cologne, 50674 Cologne, Germany; [2]Central Institute for Engineering, Electronics and Analytics, ZEA-3, Forschungszentrum Jülich, 52425 Jülich, Germany; [3]Department of Biochemistry, University of Cologne, 50674 Cologne, Germany

Running title: *Heterologous expression of D. magna serine protease CT448*

*To whom correspondence should be addressed: Jacqueline Lange: Department of Aquatic Chemical Ecology, University of Cologne, 50674 Cologne, Germany; j.lange@uni-koeln.de; Tel.: 0049 221 470 8233

The final version of this peer-reviewed author manuscript has been published by the Journal Aquatic Toxicology and is available at https://doi.org/10.1016/j.aquatox.2018.09.013

ABSTRACT

Mass developments of toxin-producing cyanobacteria are frequently observed in freshwater ecosystems due to eutrophication and global warming. These mass developments can partly be attributed to cyanobacterial toxins, such as protease inhibitors (PIs), which inhibit digestive serine proteases of *Daphnia*, the major herbivore of phytoplankton and cyanobacteria. To date, mechanisms of this inhibition in the gut of the crustacean *Daphnia magna* are not known. Here, we characterize a single serine protease, chymotrypsin 448 (CT448), which is present in the gut of the crustacean *D. magna*. Sequence alignments with human serine proteases revealed that CT448 has a putative N-terminal pro-peptide which is extended compared to the mammalian homologs and within this pro-peptide two N-linked glycosylation motifs were found. CT448 was heterologously expressed in *Sf*21 insect cells using a baculovirus expression system for optimized protein production and secretion into the medium. The protein was purified via a one-step affinity chromatography, which resulted in a protein yield of 3.45 mg/l medium. The inactive precursor (zymogen) could be activated by tryptic digestion. This is the first example of a recombinant expression of an active crustacean serine protease, which functions in the gut of a *Daphnia*. Proteomic identification of protease cleavage sites (PICS) and hydrolysation of various synthetic substrates showed that CT448 is a chymotrypsin-like elastase. In this study, we confirm that CT448 is a target of cyanobacterial protease inhibitors. Local evolutionary modifications of CT448 might render this proteolytic enzyme less susceptible against cyanobacterial secondary metabolites and might improve the fitness of Daphnia during cyanobacterial blooms.

Keywords: Protease inhibitors, chymotrypsin-like elastase, PICS, *Sf*21, serine protease, recombinant protein expression

1. Introduction

Cyanobacterial blooms have increased in lakes and ponds due to nutrient input and rising temperatures and are predicted to further increase in frequency (Downing et al., 2001; Paerl and Huisman, 2008). Cyanobacteria produce secondary metabolites that are harmful to humans and, in some cases, cause death of livestock when exposed to water with high cyanobacterial concentrations (Cheung et al., 2013). Therefore, cyanobacterial mass developments are a major issue of concern in lake management (Carmichael et al., 2001). Multiple strategies have been conducted to control harmful cyanobacteria blooms, *e.g.* the reduction of nutrient input or phytoplankton by herbicides, but also manipulation of the food chains has been proposed (Shapiro et al., 1975; Hansson et al., 1998). This manipulation addresses the concept of increased size and abundances of herbivorous zooplankton by the manipulation of higher trophic levels (Shapiro et al., 1975; Wright and Shapiro, 1984) in order to reduce cyanobacterial biomass.

During these blooms, cyanobacteria represent a major part of the phytoplankton and thus of the diet of the non-selective filter feeder *Daphnia* although. Cyanobacteria are of particularly low food quality for *Daphnia* (Martin-Creuzburg et al., 2008; Von Elert et al., 2003; Von Elert and Wolffrom, 2001). And more importantly cyanobacterial filaments interfere with the filtering apparatus of *Daphnia* (DeMott et al., 2001; Gliwicz and Lampert, 1990) resulting in a decreased fitness of the grazer. The biomass of *Daphnia* is important for the effective suppression of cyanobacterial biomass (Wright and Shapiro, 1984; Chislock et al., 2013) and crucial for successful food chain manipulation (Leibold, 1989). The prevalence of cyanobacteria in lakes and ponds leads to the well-known low abundances of *Daphnia* in summer (Ghadouani et al., 2003; Sommer et al., 1986; Threlkeld, 1979), which, to a large degree, is caused by toxic and inhibitory cyanobacterial metabolites (Wilson et al., 2006). Many cyanobacterial species produce these toxic secondary metabolites with negative effects on the fitness of *Daphnia* (Gademann and Portmann, 2008; Lürling, 2003; Rohrlack et al., 2001; Sivonen and Jones, 1999) and several studies tried to understand the toxicological impact of the well-studied secondary metabolite microcystin on *Daphnia* and to elucidate molecular mechanisms behind the inhibition of *Daphnia* (Dittmann and Wiegand, 2006; Sadler and Elert, 2014a, 2014b). Among the huge variety of biologically active secondary metabolites

(Gademann and Portmann, 2008), we focus here on protease inhibitors (PIs), as they are more frequently found in surface blooms of cyanobacteria (Agrawal et al., 2001) than the well-studied protein phosphatase inhibitor microcystin.

These cyanobacterial PIs specifically inhibit digestive proteases in the gut of *Daphnia* and thus reduce growth of this herbivore (Lürling, 2003; Schwarzenberger et al., 2010). Von Elert et al. (2004) demonstrated that two classes of serine proteases, chymotrypsin and trypsin, account for 80% of the total digestive protease activity in the gut of *Daphnia* and a strain of *Microcystis aeruginosa* was shown to produce secondary metabolites, which specifically inhibit chymotrypsins in the gut homogenate of *Daphia magna* (Von Elert et al., 2012). Moreover, two chymotrypsin inhibitors, nostopeptin BN920 and cyanopeptolin 954 (CP954) were identified in this cyanobacterium (Von Elert et al., 2005). To date there is no detailed understanding of the interaction of cyanobacterial PIs with digestive proteases in the gut of *Daphnia.* Thus, studies on the protein-inhibitor-interaction are crucial with respect to the control of cyanobacterial blooms.

In nature, *Daphnia* clones differ in their tolerance to PIs depending on their coexistence with cyanobacteria in their habitat, which points at local adaption of *Daphnia* to cyanobacterial serine PIs (Schwarzenberger et al., 2017). In response to exposure to dietary cyanobacterial protease inhibitors, *D. magna* has been shown to respond by the expression of compensatory isoforms, and by upregulated expression of dietary trypsins and chymotrypsins (Schwarzenberger et al., 2012; Schwarzenberger et al., 2010). These adaptations are further, epigenetically transferred to the offspring (Schwarzenberger and Elert, 2013). Using LC-MS/MS based proteomics, three chymotrypsin-like and trypsin-like peptidases were identified in the gut of *D. magna*, as well as the equivalent genes (Schwarzenberger et al., 2010) . Furthermore, the copy number of the most relevant chymotrypsin-like gene, i.e. *ct448,* varied among different *Daphnia* clones. Nevertheless, these variations in gene copy numbers were not related to tolerance against PIs (Schwarzenberger et al., 2017), which strongly suggests that adaptation to cyanobacterial PIs in *Daphnia* is not due to adjusted gene expression levels of *ct448* but rather due to changes of the protein CT448 itself.

Chymotrypsin and trypsin belong to the S1 family of serine proteases that are of major importance for the digestion of proteins in humans and *D. magna*. These proteases are synthesized as inactive zymogens that require activation by proteolytic cleavage. Chymotrypsin and trypsin mainly recognize their substrates by the nature of the residue N-terminal to the scissile bond (S1 site). While chymotrypsin mainly cleaves behind large hydrophobic residues, trypsin cleaves behind positively charged amino acid residues. Elastases which are also chymotrypsin-like have specificity for small hydrophobic amino acids residues. Genes encoding for digestive proteases in *D. magna* have been identified. However, theses digestive proteases have only been characterized in gut homogenates of *Daphnia* and not in isolation.

In the current work, we investigated a novel serine protease expressed in *D. magna*. For the first time, a functional serine protease of *Daphnia* was recombinantly expressed using a baculovirus-insect cell expression system and purified by Strep-tag affinity chromatography. The purified protein was proteolytically activated and its specificity was characterized by Proteomic Identification of protease Cleavage Sites (PICS) revealing that the protein has a elastase-type specificity. Furthermore, we tested its ability to hydrolyze different synthetic substrates and the susceptibility to cyanobacterial inhibitors. Taken together, we provide a novel approach to study *Daphnia* gut proteases at the molecular level with the ultimate aim of providing new approaches for the management of cyanobacterial blooms.

2. Results

2.1 In silico analysis of CT448

Sequence similarity searches using blast reveals that CT448 shares around 33 % sequence identity to the catalytic domain of human elastase-2a, a pancreatic chymotrypsin-like elastase (Fig. 1). Further multiple sequence alignments with human trypsin, chymotrypsin and elastase 1, 2 and 2A revealed that CT448 has a putative N-terminal pro-peptide with two predicted glycosylation site. The pro-peptide is extended compared to the mammalian homologs. Mammalian pancreatic trypsins and chymotrypsin are activated by a cleavage at the C-terminal site of the positively charged residue of the pro-peptide, leading to neo-N-termini starting with isoleucine or valine. In line with that, also in CT448 an arginine (R77) at

the very C-terminus of the putative pro-peptide could be found, followed by two isoleucines, suggesting that CT448 is activated in an analog fashion as its homologs in mammals. An analysis of the residues involved in S1 pocket formation was conducted revealing that CT448 shares no strict conservation to any of the other protease at relevant sites (colored in red) and thus permitting to make accurate prediction about the specificity of this protease.

2.2 Heterologous expression and purification of *D. magna* serine protease in *Sf*21 cells

We cloned the CT448 gene of *D. magna* into the pFL transfer vector of the Multibac system in order to generate afterwards a baculovirus for the heterologous expression of CT448 in in *Sf*21 cells. For an efficient secretion, the putative signal-peptide of CT448 was replaced by a melitin signal peptide. In addition, a twin-Strep-tag was fused to the C-terminus for straightforward one-step purification. For the recombinant expression of CT448, *Sf*21 cells were infected with the baculovirus resulting in a proliferation arrest approximately 24 h after addition. This time point was termed the day of proliferation arrest (dpa). At different time points after dpa, samples of the culture were taken for analysis. Cells were separated from the medium by centrifugation to obtain cell-free medium (M). Subsequently, the cells of the pellet were lysed in order to obtain supernatant and pellet (S+P) and, after centrifugation, supernatant of lysate (S). Western blot analysis of these samples revealed that CT448 became first detectable 48 h after dpa within cells lysates (Fig. 2A). After 72 h, a significant amount of the protein had been secreted into the medium. This time point was set as the optimal day for harvest of the culture medium, as 24 hours later, CT448 was no longer detectable in the medium.

The cell-free supernatant, containing the protein of interest, was loaded onto a column for affinity chromatography. Figure 2B (upper panel) shows the individual steps of purification with a StrepTactin resin. The Coomassie-stained SDS-polyacrylamid gel displayed that washing of the column removed impurities, which were present in the medium. After the elution from the column (eluate) and after concentration, distinct bands with a molecular weight of around 35 kDa were detected and are consistent with the expected mass of 35.7 kDa for the pro-enzyme of CT448 (containing its pro-peptide). Western

blot analysis (Fig. 2B, lower panel) and tryptic mass fingerprint confirmed that all species around 35 kDa were CT448. Further, MS-analysis of tryptic-digested protein revealed that all CT448 species had a complete twin-strep-tag (Fig. S1), suggesting that the distinct species have different N-termini due to an inaccurate cleavage of the signal peptide or N-terminal degradation. The mature enzyme (without pro-peptide (trypsin treated)) would have a calculated molecular mass of 29.1 kDa and the pro-peptide a mass of 6.6 kDa. For both, we could not observe a corresponding band during SDS-PAGE analysis. A stability control, incubation of the eluate at 4°C for 4 days (conc. eluate aged), showed a broader pattern of bands (35-37 kDa), which indicated protein degradation

Overall, the purification via a twin-strep-tag and StrepTactin resin followed by concentration of the eluate seemed to be an adequate procedure to purify CT448 from *Sf21* medium for subsequent assays. The average yield from 1 liter medium was 3.45 mg of protein.

2.3 CT448 – A chymotrypsin-like elastase

With purified recombinant CT448 protein in hands, we proceeded to determine its sequence specificity with proteome-derived peptide libraries using the Proteomic Identification of protease Cleavage Sites (PICS) approach (Biniossek et al., 2016; Schilling and Overall, 2008). In this approach a trypsin-digested E.coli proteome was used as peptide library, allowing to determine the specificity of non-tryptic proteases. The library was incubated with non-activated CT448 (-Tryp), trypsin-activated CT448 (+Tryp), or buffer control. Afterwards peptides which altered in abundance in the different samples were identified and annotated by a MS-based proteomic analysis from which a specificity profile was obtained. This profile showed that non-activated CT448 preferentially cleave at sites with Pro at P2 and Val/Ile at P1 for CT448 (Fig. 3A), resembling the specificity of neutrophil elastase (Biniossek et al., 2016). Activation with trypsin did not alter substrate specificity significantly (Fig. 3B). To validate the PICS results, we performed CT448 enzyme activity assays with synthetic peptide substrates, two specific for elastase and one specific for chymotrypsin. Trypsin-activated CT448 (+ Tryp) showed a significantly higher activity for the elastase substrates N-Succinyl-alanine-alanine-alanine-para-nitroanilide (AAApNA, 16.5 ± 2.24

mU/mg) and N-Succinyl-alanine-alanine-valine-para-nitroanilide (AAVpNA, 5.07 ± 0.17 mU/mg), than for the chymotrypsin-specific substrate N-Succinyl-alanine-alanine-proline-phenylalanine-para-nitroanilide (AAPFpNA, 0.045 ± 0.008 mU/mg) (one way ANOVA: $F=113.299$; $p < 0.001$; Tukey $p < 0.05$, Fig. 3C). CT448 without prior activation (- Tryp) showed a similar preference for the elastase substrates AAApNA (0.37 ± 0.23 mU/mg) and AAVpNA (0.177 ± 0.002 mU/mg) although the activity was strongly reduced compared to the trypsin-activated CT448 (Fig. 3C). Only very little hydrolysis of the chymotrypsin substrate AAPFpNA (0.032 ± 0.002 mU/mg) close to background levels was detected. Trypsin, which was used for activation of CT448, did not hydrolyze any of these substrates (data not shown).

As the specificity of CT448 was similar to that of elastases, we compared the CT448 of *D. magna* to the well-studied porcine pancreatic elastase (Fig. 4A). When using AAApNA as substrate, similar activities of trypsin-activated CT448 (CT448 + Tryp) and porcine elastase were detected. Instead, CT448 - Tryp was only active after a lag phase of 6-7 hours, which suggested that auto-activation over several hours was required, before proteolytic activation became detectable. To exclude contaminating protease activity we used 1 mM and 10 mM of EDTA for metalloprotease inhibition, 1 mM and 10 mM of iodoacetamide for cysteine protease inhibition and 50 µM chymostatin as positive control. As expected, the activity of CT448 was only affected by chymostatin showing that the main activity was derived from a serine protease (Fig. S2).

Digestive proteases of *Daphnia* are often exposed to the well-studied cyanobacterial PIs cyanopeptolin 954 (CP954) and nostopeptin 920 (BN920). We therefore tested CT448 activity in the presence of these two PIs and observed complete inhibition of CT448 by both inhibitors (Fig. 4B). Interestingly, BN920 and CP954 also reduced the activity of porcine elastase to a minimum. Here, CP954 seemed to have a stronger impact than BN920. This demonstrates that these two cyanobacterial PIs, which have previously been reported as chymotrypsin inhibitors (Von Elert et al., 2005; Schwarzenberger et al., 2010), act more broadly as inhibitors of mammalian and crustacean chymotrypsin-like elastases.

3. Discussion

193 Among the major digestive serine proteases in *Daphnia*, there is the chymotrypsin-like protease
194 CT448, which has shown to be of major relevance for the interaction of *Daphnia* with cyanobacterial PIs.
195 Evidence for the importance of CT448 is based on the upregulation of *ct448* gene expression and on an
196 increase in proteolytic activity in *D. magna* clones when being fed with PI producing cyanobacteria
197 (Schwarzenberger et al., 2012). Further, it has been shown that the tolerance of *D. magna* clones to
198 cyanobacteria that contain chymotrypsin inhibitors is correlated to the tolerance of *D. magna's* gut
199 proteases to in-vitro inhibition by these inhibitors (Gademann and Portmann, 2008). However, the
200 molecular mechanism of the increased tolerance to PIs remained unclear. So far, proteases of *D. magna*
201 could not be analyzed, so that no specific characteristics of individual digestive enzymes could be
202 accomplished. Here, we were able to express a specific gut protease with the here described method in
203 adequate amounts (3.45 mg/L) for subsequent analyses. Nevertheless, that we used the highly specific
204 Twin-Strep-tag for purification, several protein bands with molecular masses ranging from 35-37 kDa in
205 the purified fractions (eluate, conc. eluate and conc. eluate aged, Fig. 2B) were observed. Western blot
206 and MS analysis confirmed that all species were CT448. As the Twin-Strep-tag was intact, the distinct
207 species were most likely the result of N-terminal truncations or post-translational modifications such as
208 glycosylation. This, however, requires further investigations.

209 The digestive protease CT448 is annotated as a chymotrypsin (Schwarzenberger et al., 2010). Here,
210 we showed by PICS analysis and by using synthetic substrates that CT448 preferentially cleaves behind
211 small aliphatic residues Ala, Val and Ile meaning that the protease is a elastase, rather than a classic
212 chymotrypsin. Further will the elucidation of the CT448's specificity allow the development and use of
213 more specific substrates and inhibitors to considerably enhance the sensitivity of assays. It remains to be
214 determined if other digestive chymotrypsin-like proteases in *Daphnia* also have an elastase-like
215 specificity, which now can be elucidated with our approach. Interestingly, the study demonstrated that
216 CP954 and BN920, originally described as chymotrypsin inhibitors, also act as elastase inhibitors. For
217 CP954 and BN920, IC_{50} values of 4.5 nM and 3.1 nM for the inhibition of bovine chymotrypsin have
218 been reported (Von Elert et al., 2005), which classified them as the most potent inhibitors containing 3-

amino-6-hydroxy-2-piperidone (Ahp). Similar to CP954 and BN920, scyptolin A isolated from the cyanobacterium *Scytonema hofmanni* PCC7110 was demonstrated to inhibit elastase by occupying parts of the active center. Four N-terminal amino acid residues of scyptolin A bind at subsites S1 through S4 and prevent hydrolysis (Matern et al., 2001; Matern et al., 2003). Like scyptolin A, CP954 and BN920 belong to the class of cyanopeptolines and share similar structures, we suggest that these cyanopeptolines share similar specificities and binding properties with respect to target proteases. To test, whether inhibition of CP954 or BN920 is comparable to the binding of scyptolin A to elastase, we would suggest co-crystallization of CP954 or BN920 with elastase for structure analysis and ultimately with CT448. In several cases *Daphnia* have been shown to evolutionarily adapt to the presence of cyanobacteria (DeMott et al., 1991; Hairston Jr et al., 1999; Sarnelle and Wilson, 2005). The *D. magna* clone used in this study, was isolated from a lake where *D. magna* coexists with cyanobacteria that produce PIs (Schwarzenberger et al., 2013). We expect that CT448 has adapted to these natural cyanobacterial PIs in a way that rendered CT448 less susceptible to these natural inhibitors.

In the future, the approach presented here will be used to compare different *D. magna* clones with regard to their tolerance to cyanobacterial PIs: Local adaptation of *Daphnia* to toxic cyanobacteria has been demonstrated (Sarnelle and Wilson, 2005), and understanding the molecular basis of this adaptation might be helpful for the management of cyanobacterial blooms. In high densities *Daphnia* have the capability to suppress cyanobacteria before they become dominant (Christoffersen et al., 1993). Chislock et al. (2013) demonstrated that *Daphnia* suppressed a toxic cyanobacteria dominated bloom within a few weeks. In this enclosure experiment 96% of the phytoplankton was composed by cyanobacteria and grazing of *Daphnia* reduced this cyanobacterial biomass by 76%. Further, it has been shown that *Daphnia* are able to suppresses even an established bloom of cyanobacteria (Sarnelle, 2007).

As *Daphnia* is able to adapt to toxic secondary metabolites of cyanobacteria and is capable of suppressing cyanobacterial mass formation, we suggest that *Daphnia* clones with a high tolerance to serine protease inhibitors might facilitate the suppression of cyanobacterial blooms. Thus, adapted clones might be used as a tool for the management of cyanobacteria in lakes and ponds in the future.

4. Conclusion

In this study we have provided an approach for recombinant expression of *D. magna* proteases and have shown that CT448 of *D. magna* encodes a chymotrypsin-like elastase with specificity for alanine residues. This will facilitate comparison of the sensitivity of digestion proteases from different *Daphnia* clones to cyanobacterial PIs and thus help to clarify their contribution to adaptation.

5. Experimental procedures

5.1 Protease substrates and inhibitors

The synthetic substrates N-Succinyl-alanine-alanine-alanine-para-nitroanilide (AAApNA), N-Succinyl-alanine-alanine-valine-para-nitroanilide (AAVpNA), N-Succinyl-alanine-alanine-proline-phenylalanine-para-nitroanilide (AAPFpNA) and porcine pancreatic elastase were purchased from Sigma-Aldrich (St. Louis, USA). The cyanobacterial protease inhibitors, cyanopeptolin 954 (CP954) and nostopeptin 920 (BN920), were isolated from the *Microcystis aeruginosa* strain NIVA Cya43 via HPLC according to von Elert et al. (2005). Stock solutions of the substrates were prepared in DMSO.

5.2 Generation of recombinant baculovirus

The full length sequence of *ct448* of *D. magna* clone Sweden May17, collected in May 2010 in Lake Bysjön (situated in Southern Scania, Sweden: N 55.675399 E 13.545805) was optimized for insect cell expression (GeneArt, Regensburg, Germany). Additional sequences were introduced so that the protein harbors an N-terminal melittin signal peptide and C-terminal twin-strep-tag. This gene construct was cloned into pFL MultiBac transfer plasmid (kind gift of Imre Berger) using BamHI and XhoI as restriction sites. Subsequently, chemical-competent *E. coli* DH10EMBacY cells (kind gift of Imre Berger) were transformed with the cloned transfer vector for *in vivo* Tn7 transposition generating a bacmid containing the expression cassette of CT448 (according to (Berger et al., 2004; Bieniossek et al., 2008; Fitzgerald et al., 2006). Putative recombinant pFLCT448 bacmids were selected by white/blue screening.

Isolated bacmid DNA was used to transfect *Spodoptera frugiperda Sf*21 cells (Vaughn et al., 1977) using X-tremeGENE HP (Roche, Basel, Switzerland) resulting in the recombinant virus (CT448-EmbacY).

5.3 Baculovirus expression and protein purification

*Sf*21 cells were grown in suspension in Insect-Xpress medium (Lonza, Basel, Switzerland) at 27°C. To analyze for recombinant protein production, cells were cultured in densities of 10^6 cells/ml and infected with CT448-EmbacY. Samples were taken to determine the time when cells stopped proliferating (day of proliferation arrest, dpa), usually 24 h post infection. Medium and cells were separated and analyzed for recombinant protein production. Cells were lysed in PBS and the cell extract was separated from insoluble material by centrifugation (11 000*g*, 5 min). Proteins were separated in a 15% SDS polyacrylamide gel. The presence of recombinant protein in either medium (M), supernatant of the lysate (S) or supernatant and pellet (S+P) was detected using a horseradish peroxidase conjugated StrepTactin (StrepTactin-HRP, IBA, Goettingen, Germany) against the twin-strep-tag, in a standard Western-blot assay.

For large-scale protein expression and purification, approximately 5 10^5 cells/ml were infected with CT448 EMbacY. Cells were harvested at the time point of highest protein secretion into the medium as the CT448 was secreted into the medium. Medium was collected, cells were removed by centrifugation (200*g*, 45 min) and prepared for affinity chromatography by adjusting the pH to 8.0 and by blocking of biotin with BioLock (IBA, Goettingen, Germany). After a second round of centrifugation (200*g*, 45 min), protein was purified by Strep-tag affinity chromatography using 1.5 ml Strep-Tactin XT Superflow resin (IBA, Goettingen, Germany) according to the user manual, except of exchanging manufacturer's Buffer B and BXT with a buffer containing 25 mM Tris pH 8.0 and 250 mM NaCl. For protein elution 50 mM biotin was supplemented into the washing buffer. Eluted protein was concentrated by centrifugal ultrafiltration. After the cells had been harvested, all steps were performed at 4°C.

5.4 Proteomic Identification of protease Cleavage Sites (PICS)

Proteome-derived peptide libraries were generated by tryptic digest of *E.coli* K12 lysates as described (Chen et al., 2017). The peptide library was split in three aliquots that were treated with recombinant CT448 with and without prior activation by trypsin or with a buffer control for 1.5 h at RT. After incubation, peptides were triplex stable isotope labelled by reductive dimethylation (Boersema et al., 2009). Isotopically light formaldehyde (CH_2O) and sodium cyanoborohydride ($NaBH_3CN$) was used to label control-treated library peptides, deuterated formaldehyde (CD_2O) and $NaBH_3CN$ were used to label the assay with recombinant CT448 without prior activation and heavy formaldehyde ($^{13}CD_2O$) and sodium cyanoborodeuteride ($NaBD_3CN$) were used to label samples after incubation with trypsin-activated CT448. Dimethylation was performed for 2 h at RT and quenched by addition of 100 mM Tris-HCl pH 7.5 for 1 h. Subsequently samples were mixed, desalted and purified by C18 StageTips (Rappsilber et al., 2007).

1 µg peptides were separated using a nano-HPLC system (Ultimate 3000 nano-RSLC, Thermo) operated in a two-column setup (Acclaim PepMap 100 C18, ID 75 µm, trap column length 2 cm, particle size 3 µm, analytical column length 50 cm, particle size 2 µm, Thermo) coupled online to a high resolution Q-TOF mass spectrometer (ImpactII, Bruker) as described (Rinschen et al., 2017). Peptides were eluted with a binary gradient from 5-35% B for 90 min (A: H_2O + 0.1% FA, B: ACN + 0.1% FA), followed by washing and re-equilibration steps to a total runtime of 2 h per sample. The Bruker HyStar Software (v3.2) was used to acquire line-mode MS spectra in a mass range from 200-1750 m/z at an acquisition rate of 4 Hz. For each MS spectrum, the Top17 most intense ions were selected for fragmentation and an exclusion window of 40s was applied.

Peptides were identified and quantified from the acquired MS spectra using the MaxQuant software package, v1.6.0.16 (Tyanova et al., 2016) and a UniProt *E.coli* K12 proteome library (downloaded Nov 2015). Trypsin was set as digestion enzyme for semi-specific searches (e.g. only one side of the peptide was required to match the trypsin specificity). Label multiplicity was set to three, considering light dimethylation (+ 28.0313 Da), medium dimethylation (+ 32.0564 Da) and heavy dimethylation (+

36.0757) as peptide N-terminal and lysine labels. Carbamidomethylation of cysteine residues (+ 57.0215 Da) was set as fixed modification, methionine oxidation (+ 15.9949 Da) and protein N-terminal acetylation (+ 42.0106 Da) was considered as variable modifications. PSM false discovery rate was set to 0.01.

Identified peptides that showed at least a fourfold increase in intensity after protease treatment compared to the control treatment or were exclusively present in the protease-treated condition were considered as putative cleavage products. An in-house developed Perl script was used to remove putative library peptides (trypsin specificity on both sides of the identified peptide) and to reconstruct the full cleavage windows from the identified cleavage products as described (Schilling and Overall, 2008). Aligned validated cleavage windows were visualized as icelogos (Colaert et al., 2009), displaying site-specific differential amino acid abundance calculated as per cent difference compared to the *E.coli* K12 proteome as reference set (p-value 0.05).

5.5 In-gel digest of recombinant CT448

Coomassie-stained bands in preparations of recombinant CT448 were subjected to in-gel digestion with trypsin as described (Demir et al., 2013). Mass spectrometry analysis was performed as noted above and the sequence of recombinant CT448 was appended to contaminant database for peptide identification.

5.6 CT448 activity assays

Activity measurements of CT448 were performed with trypsin activation (+Tryp) and without (-Tryp) immediately after protein purification. CT448 was activated with recombinant trypsin (Sigma-Aldrich, St. Louis, USA) in a 1:20 mass ratio to CT448. For investigation of the substrate specificity, 100 µM of AAApNA, AAVpNA or AAPFpNA were incubated with 25 µg or 10 µg, of each activated (+Tryp) or non-activated (-Tryp), of purified CT448 in a final volume of 150 µl of 100 mM potassium phosphate buffer pH 8.0. Activity was measured as the increase in absorbance at 380 nm within the first 12 h at

27°C in a Biotek Synergy H4 plate reader (Biotek, Winooski, USA). Further, the specific proteolytic activity (mU/mg) of CT448 was calculated from the steepest parts of the curves.

In order to compare CT448 with a mammalian serine protease, we additionally tested the activity of 25 µg of porcine pancreatic elastase (CAS Number: 39445-21-1, Sigma-Aldrich, St. Louis, USA) on 100 µM AAApNA under the same conditions as described above. The effect of the isolated cyanobacterial protease inhibitors cyanopeptolin 954 (CP954) or nostopeptin 920 (BN920) on the activity of CT448 and of porcine elastase was measured. Thus, 20 µM of both inhibitors were separately added to trypsin-activated and non-activated CT448 or porcine elastase. All measurements were carried out in triplicates.

5.7 Statistical analyses

The statistics were conducted with the program Sigmaplot 11.0 (Systat Software GmbH, Erkrath, Germany). The data were analyzed via one-way ANOVA and a post-hoc analysis (Tukey HSD). The level of significance was $p < 0.05$.

5.8 Data availability

The mass spectrometry proteomics data for the PICS experiment have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaíno et al, 2015) partner repository with the dataset identifier PXD010978**.**

Acknowledgement: We acknowledge Katja Preuß for her laboratory support regarding the isolation of the cyanobacterial inhibitors. This study was supported by the German Science Foundation, DFG, with a grant to EvE (EL 179/11-1).

Declaration of interest: none.

AuthorsContribution: JL, CP, EvE designed the study. JL and FD performed the experiments. JL expressed and purified the protein. JL, CP, FD, PFH analyzed the data. EvE, UB, PFH supervised the work. JL wrote the paper with input from all authors. All authors have approved the final article.

.

References

Agrawal, M.K., Bagchi, D., Bagchi, S.N., 2001. Acute inhibition of protease and suppression of growth in zooplankter, *Moina macrocopa*, by *Microcystis* blooms collected in Central India. Hydrobiologia (464), 37–44.

Berger, I., Fitzgerald, D.J., Richmond, T.J., 2004. Baculovirus expression system for heterologous multiprotein complexes. Nature biotechnology 22 (12), 1583.

Bieniossek, C., Richmond, T.J., Berger, I., 2008. MultiBac: multigene baculovirus- based eukaryotic protein complex production. Current protocols in protein science 51 (1), 5.20. 1-5.20. 26.

Biniossek, M.L., Niemer, M., Maksimchuk, K., Mayer, B., Fuchs, J., Huesgen, P.F., McCafferty, D.G., Turk, B., Fritz, G., Mayer, J., 2016. Identification of protease specificity by combining proteome-derived peptide libraries and quantitative proteomics. Molecular & Cellular Proteomics, mc O115. 056671.

Boersema, P.J., Raijmakers, R., Lemeer, S., Mohammed, S., Heck, A., JR, 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nature protocols 4 (4), 484.

Carmichael, W.W., Azevedo, S.M., An, J.S., Molica, R.J., Jochimsen, E.M., Lau, S., Rinehart, K.L., Shaw, G.R., Eaglesham, G.K., 2001. Human fatalities from cyanobacteria: chemical and biological evidence for cyanotoxins. Environmental health perspectives 109 (7), 663.

Chen, C.-y., Mayer, B., Schilling, O., 2017. Profiling of protease cleavage sites by proteome-derived peptide libraries and quantitative proteomics, in: Protein Terminal Profiling. Springer, pp. 197–204.

Cheung, M.Y., Liang, S., Lee, J., 2013. Toxin-producing cyanobacteria in freshwater: a review of the problems, impact on drinking water safety, and efforts for protecting public health. Journal of Microbiology 51 (1), 1–10.

Chislock, M.F., Sarnelle, O., Jernigan, L.M., Wilson, A.E., 2013. Do high concentrations of microcystin prevent *Daphnia* control of phytoplankton? Water research 47 (6), 1961–1970.

Christoffersen, K., Riemann, B., Klysner, A., Søndergaard, M., 1993. Potential role of fish predation and natural populations of zooplankton in structuring a plankton community in eutrophic lake water. Limnol. Oceanogr. 38 (3), 561–573.

Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., Gevaert, K., 2009. Improved visualization of protein consensus sequences by iceLogo. Nature methods 6 (11), 786.

Demir, F., Horntrich, C., Blachutzik, J.O., Scherzer, S., Reinders, Y., Kierszniowska, S., Schulze, W.X., Harms, G.S., Hedrich, R., Geiger, D., 2013. Arabidopsis nanodomain-delimited ABA signaling pathway regulates the anion channel SLAH3. Proceedings of the National Academy of Sciences, 201211667.

DeMott, W.R., Gulati, R.D., van Donk, E., 2001. Daphnia food limitation in three hypereutrophic Dutch lakes: Evidence for exclusion of large- bodied species by interfering filaments of cyanobacteria. Limnol. Oceanogr. 46 (8), 2054–2060.

DeMott, W.R., Zhang, Q.-X., Carmichael, W.W., 1991. Effects of toxic cyanobacteria and purified toxins on the survival and feeding of a copepod and three species of *Daphnia*. Limnol. Oceanogr. 36 (7), 1346–1357. 10.4319/lo.1991.36.7.1346.

Dittmann, E., Wiegand, C., 2006. Cyanobacterial toxins–occurrence, biosynthesis and impact on human affairs. Molecular nutrition & food research 50 (1), 7–17.

Downing, J.A., Watson, S.B., McCauley, E., 2001. Predicting cyanobacteria dominance in lakes. Canadian journal of fisheries and aquatic sciences 58 (10), 1905–1908.

Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., Berger, I., 2006. Protein complex expression by using multigene baculoviral vectors. Nature methods 3 (12), 1021.

Gademann, K., Portmann, C., 2008. Secondary metabolites from cyanobacteria: complex structures and powerful bioactivities. Current Organic Chemistry 12 (4), 326–341.

Ghadouani, A., Pinel- Alloul, B., Prepas, E.E., 2003. Effects of experimentally induced cyanobacterial blooms on crustacean zooplankton communities. Freshwater Biology 48 (2), 363–381.

Gliwicz, Z.M., Lampert, W., 1990. Food thresholds in Daphnia species in the absence and presence of blue- green filaments. Ecology 71 (2), 691–702.

Hairston Jr, N.G., Lampert, W., Cáceres, C.E., Holtmeier, C.L., Weider, L.J., Gaedke, U., Fischer, J.M., Fox, J.A., Post, D.M., 1999. Lake ecosystems: rapid evolution revealed by dormant eggs. Nature 401 (6752), 446.

Hansson, L.-A., Annadotter, H., Bergman, E., Hamrin, S.F., Jeppesen, E., Kairesalo, T., Luokkanen, E., Nilsson, P.-Å., Søndergaard, M., Strand, J., 1998. Biomanipulation as an application of food-chain theory: constraints, synthesis, and recommendations for temperate lakes. Ecosystems 1 (6), 558–574.

Leibold, M.A., 1989. Resource edibility and the effects of predators and productivity on the outcome of trophic interactions. The American Naturalist 134 (6), 922–949.

Lürling, M., 2003. *Daphnia* growth on microcystin- producing and microcystin- free *Microcystis aeruginosa* in different mixtures with the green alga *Scenedesmus obliquus*. Limnol. Oceanogr. 48 (6), 2214–2220.

Martin-Creuzburg, D., Elert, E. von, Hoffmann, K.H., 2008. Nutritional constraints at the cyanobacteria—Daphnia magna interface: the role of sterols. Limnol. Oceanogr. 53 (2), 456–468.

Matern, U., Oberer, L., Falchetto, R.A., Erhard, M., König, W.A., Herdman, M., Weckesser, J., 2001. Scyptolin A and B, cyclic depsipeptides from axenic cultures of *Scytonema hofmanni* PCC 7110. Phytochemistry 58 (7), 1087–1095. 10.1016/S0031-9422(01)00400-9.

Matern, U., Schleberger, C., Jelakovic, S., Weckesser, J., Schulz, G.E., 2003. Binding Structure of Elastase Inhibitor Scyptolin A. Chemistry & Biology 10 (10), 997–1001. 10.1016/j.chembiol.2003.10.001.

Paerl, H.W., Huisman, J., 2008. Blooms like it hot. Science 320 (5872), 57–58.

Rappsilber, J., Mann, M., Ishihama, Y., 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature protocols 2 (8), 1896.

Rinschen, M.M., Hoppe, A.-K., Grahammer, F., Kann, M., Völker, L.A., Schurek, E.-M., Binz, J., Höhne, M., Demir, F., Malisic, M., 2017. N-degradomic analysis reveals a proteolytic network processing the podocyte cytoskeleton. J Am Soc Nephrol 28 (10), 2867–2878.

Rohrlack, T., Dittmann, E., Börner, T., Christoffersen, K., 2001. Effects of cell-bound microcystins on survival and feeding of *Daphnia spp*. Applied and environmental microbiology 67 (8), 3523–3529. 10.1128/AEM.67.8.3523-3529.2001.

Sadler, T., Elert, E. von, 2014a. Dietary exposure of *Daphnia* to microcystins: no in vivo relevance of biotransformation. Aquatic toxicology (Amsterdam, Netherlands) 150, 73–82. 10.1016/j.aquatox.2014.02.017.

Sadler, T., Elert, E. von, 2014b. Physiological interaction of *Daphnia* and *Microcystis* with regard to cyanobacterial secondary metabolites. Aquatic toxicology (Amsterdam, Netherlands) 156, 96–105. 10.1016/j.aquatox.2014.08.003.

Sarnelle, O., 2007. Initial conditions mediate the interaction between *Daphnia* and bloom- forming cyanobacteria. Limnol. Oceanogr. 52 (5), 2120–2127.

Sarnelle, O., Wilson, A.E., 2005. Local adaptation of *Daphnia pulicaria* to toxic cyanobacteria. Limnol. Oceanogr. 50 (5), 1565–1570. 10.4319/lo.2005.50.5.1565.

Schilling, O., Overall, C.M., 2008. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. Nature biotechnology 26 (6), 685.

Schwarzenberger, A., D'Hondt, S., Vyverman, W., Elert, E. von, 2013. Seasonal succession of cyanobacterial protease inhibitors and *Daphnia magna* genotypes in a eutrophic Swedish lake. Aquat Sci 75 (3), 433–445. 10.1007/s00027-013-0290-y.

Schwarzenberger, A., Elert, E. von, 2013. Cyanobacterial protease inhibitors lead to maternal transfer of increased protease gene expression in *Daphnia*. Oecologia 172 (1), 11–20. 10.1007/s00442-012-2479-5.

Schwarzenberger, A., Keith, N.R., Jackson, C.E., Elert, E. von, 2017. Copy number variation of a protease gene of *Daphnia*: Its role in population tolerance. J. Exp. Zool. 327 (2-3), 119–126. 10.1002/jez.2077.

Schwarzenberger, A., Kuster, C.J., Elert, E. von, 2012. Molecular mechanisms of tolerance to cyanobacterial protease inhibitors revealed by clonal differences in *Daphnia magna*. Molecular ecology 21 (19), 4898–4911. 10.1111/j.1365-294X.2012.05753.x.

Schwarzenberger, A., Zitt, A., Kroth, P., Mueller, S., Elert, E. von, 2010. Gene expression and activity of digestive proteases in *Daphnia*: effects of cyanobacterial protease inhibitors. BMC physiology 10, 6. 10.1186/1472-6793-10-6.

Shapiro, J., Lamarra, V.A., Lynch, M., 1975. Biomanipulation: an ecosystem approach to lake restoration.

Sivonen, K., Jones, G., 1999. Cyanobacterial toxins. Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management 1, 43–112.

Sommer, U., Gliwicz, Z.M., Lampert, W., Duncan, A., 1986. The PEG-model of seasonal succession of planktonic events in fresh waters. Arch. Hydrobiol 106 (4), 433–471.

Threlkeld, S.T., 1979. The midsummer dynamics of two *Daphnia* species in Wintergreen Lake, Michigan. Ecology 60 (1), 165–179.

Tyanova, S., Temu, T., Cox, J., 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nature protocols 11 (12), 2301.

Vaughn, J.L., Goodwin, R.H., Tompkins, G.J., McCawley, P., 1977. The establishment of two cell lines from the insect *Spodoptera frugiperda* (lepidoptera; noctuidae). In vitro 13 (4), 213–217.

Vizcaíno JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu QW, Wang R, Hermjakob H (2016). 2016 update of the PRIDE database and related tools. Nucleic Acids Res 44(D1): D447-D456

Von Elert, E., Agrawal, M.K., Gebauer, C., Jaensch, H., Bauer, U., Zitt, A., 2004. Protease activity in gut of *Daphnia magna*: evidence for trypsin and chymotrypsin enzymes. Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology 137 (3), 287–296. 10.1016/j.cbpc.2003.11.008.

Von Elert, E., Martin-Creuzburg, D., Le Coz, J.R., 2003. Absence of sterols constrains carbon transfer between cyanobacteria and a freshwater herbivore (Daphnia galeata). Proceedings of the Royal Society of London B: Biological Sciences 270 (1520), 1209–1214.

Von Elert, E., Oberer, L., Merkel, P., Huhn, T., Blom, J.F., 2005. Cyanopeptolin 954, a chlorine-containing chymotrypsin inhibitor of *Microcystis aeruginosa* NIVA Cya 43. Journal of natural products 68 (9), 1324–1327. 10.1021/np050079r.

Von Elert, E., Wolffrom, T., 2001. Supplementation of cyanobacterial food with polyunsaturated fatty acids does not improve growth of Daphnia. Limnol. Oceanogr. 46 (6), 1552–1558.

Von Elert, E., Zitt, A., Schwarzenberger, A., 2012. Inducible tolerance to dietary protease inhibitors in *Daphnia magna*. The Journal of experimental biology 215 (Pt 12), 2051–2059. 10.1242/jeb.068742.

Wilson, A.E., Sarnelle, O., Tillmanns, A.R., 2006. Effects of cyanobacterial toxicity and morphology on the population growth of freshwater zooplankton: Meta- analyses of laboratory experiments. Limnol. Oceanogr. 51 (4), 1915–1924.

Wright, D.I., Shapiro, J., 1984. Nutrient reduction by biomanipulation: An unexpected phenomenon and its possible cause: With 4 figures and 1 table in the text. Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen 22 (1), 518–524.

6. Figures (Fig. 1; Fig. 3A+3B please in color)

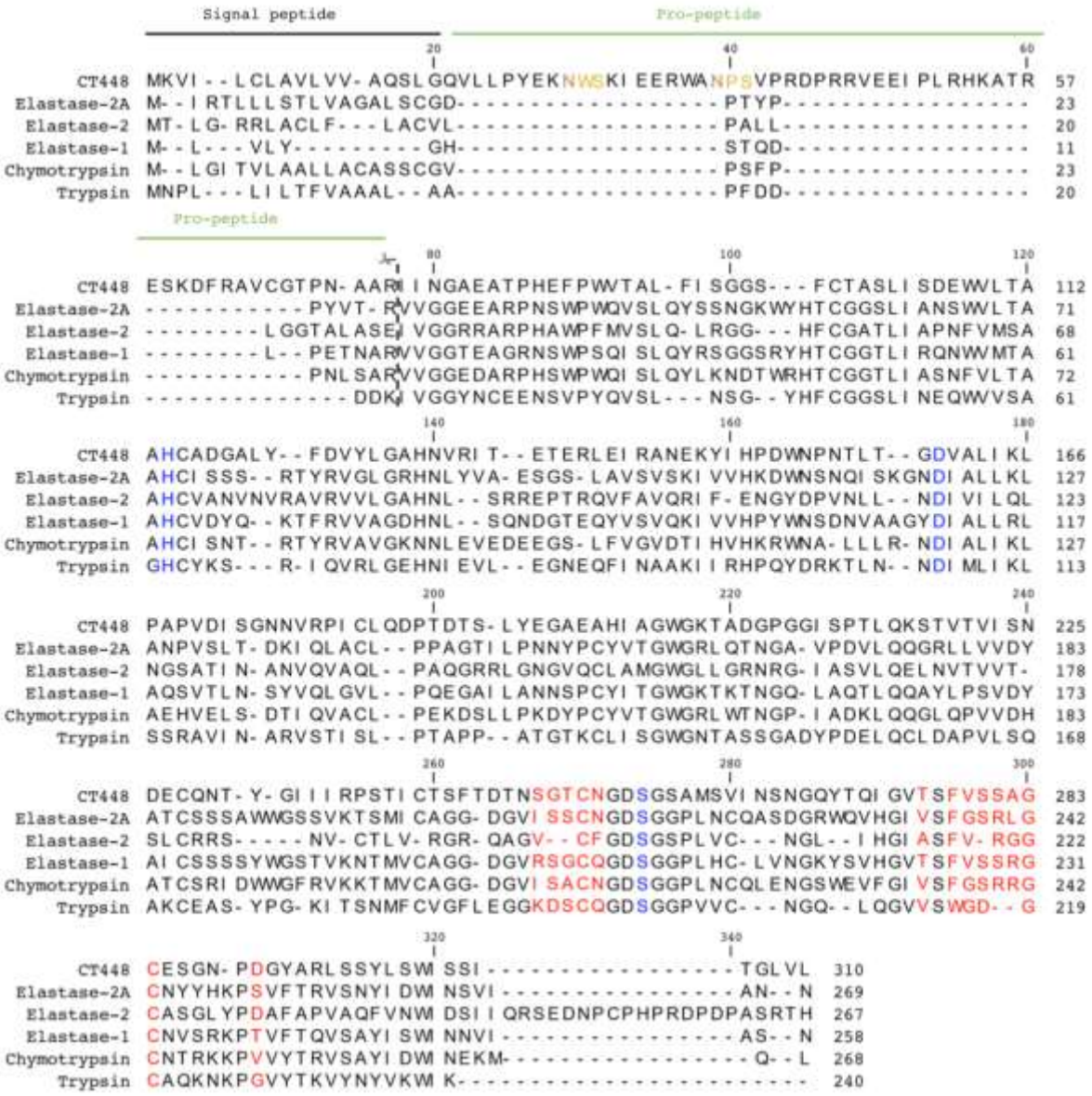
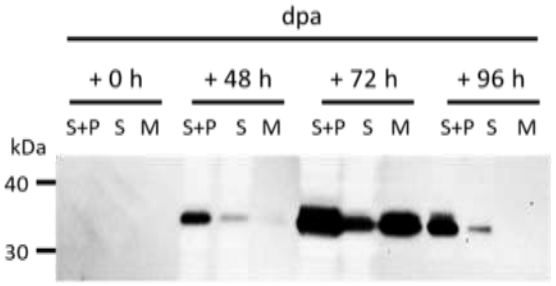


Figure 1: Multiple alignment of crustacean CT448 with human serine proteases. Important regions are highlighted with different colors: The pro-peptide in green, amino acids of the catalytic triad (H122, D174 and S275) in blue and residues relevant for substrate binding in red. The putative glycosylated asparagines are in dark yellow, while residues of the N-glycosylation consensus sequence are in yellow. Glycosylation sites were predicted using the NetNGlyc 1.0 Server. Scissor icon indicates the position for trypsin cleavage and activation.

A



B

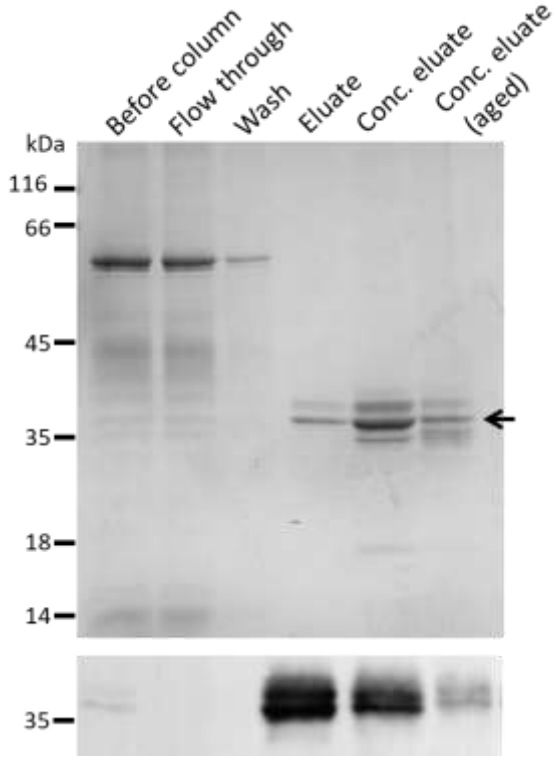
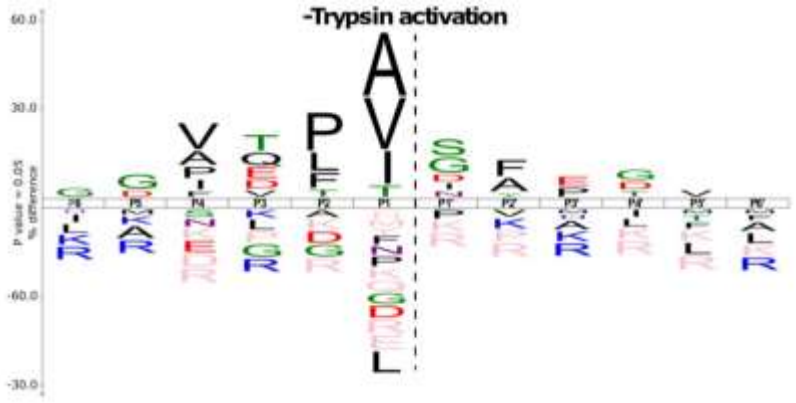


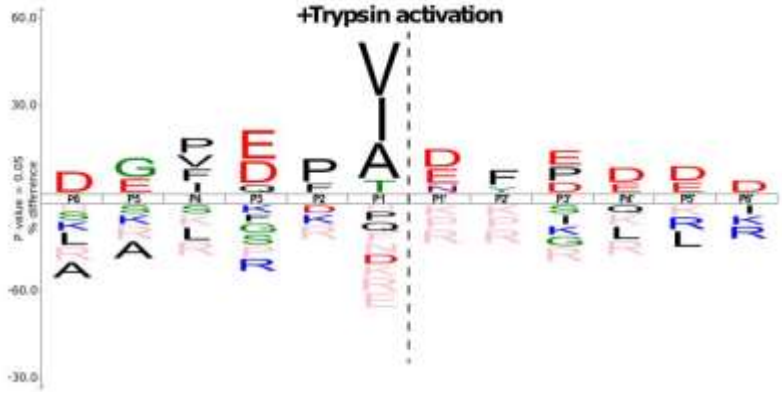
Figure 2: Analysis of the expression and purification of C-terminally tagged CT448 by western blotting and SDS-PAGE. A) The protein was expressed in *Sf*21 cells after infection with a baculovirus that lead to an arrest of proliferation. CT448 was detected at the day of proliferation arrest (dpa) and at different time intervals, thereafter in supernatant and pellet (S+P), in supernatant of the lysate (S) and in the medium (M). Using StrepTactin-HRP conjugate, CT448 was first detected 48 h after dpa and CT448 levels increased until 72 h after dpa. These levels decreased 96 h after dpa and CT448 was not longer detectable

in the medium. B) SDS-PAGE analysis from samples taken during purification. A commassie-stained Polyacrylmide gel is shown in the upper panel while a western blot analysis is shown in the lower panel. CT448 was secreted into the medium and was, after separation from the *Sf*21 cells (before column), purified by affinity chromatography with thoroughly washing of the column (flow through and wash). The eluted protein (eluate) was concentrated in an ultrafiltration centrifugal device (conc. eluate) for subsequent enzymatic assays. Different species of pure CT448 (highlighted by an arrow) with a molecular weight of around 35 kDa were obtained. CT448 was detected using the StrepTactin-HRP conjugate (lower panel) and revealed CT448-specific bands at around 35 kDa which faded upon aging.

A


-Trypsin activation
60.0
30.0
-60.0
-30.0
P value = 0.05
% difference

B


+Trypsin activation
60.0
30.0
-60.0
-30.0
P value = 0.05
% difference

C

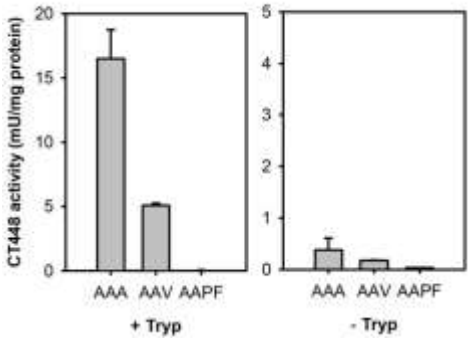
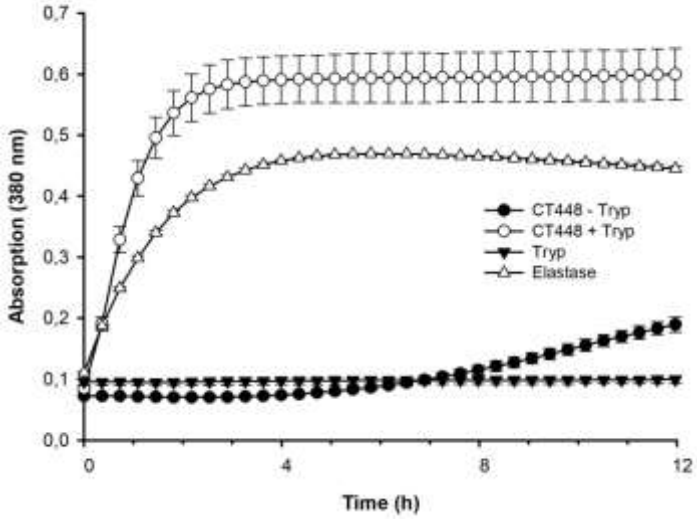

CT448 activity (mU/mg protein)
20
15
10
5
0
AAA AAV AAPF
+ Tryp
5
4
3
2
1
0
AAA AAV AAPF
- Tryp

Figure 3: Characterization of CT448 protease specificity by Proteomic Characterization of Cleavage Sites (PICS) and specific enzymatic activity on different synthetic substrates. A) For PICS analyses, an *E. coli* K12 total proteome-derived peptide library, prepared by tryptic-digest, was incubated with recombinant CT448 protease. IceLogos of 351 aligned cleavage sites of CT448 and B) 181 aligned CT448 cleavages sites of trypsin-activated CT448 display the favored (above horizontal line) and disfavored (below line) amino acids at each position, expressed as difference against the expected amino acid abundance in the *E.coli* K12 proteome (p-value 0.05). The dashed line indicates the cleavage site. Color coding of amino acids depict the different physical-chemical properties: black for hydrophobic, green for hydrophilic, red for acidic and blue for alkaline amino acids. The height of one amino acid in a stack (P1-6; P1'-6') reflects its frequency. C) Further, synthetic substrates consisting of different peptides coupled to C-terminal *para*-nitroanilide were subjected to identical concentrations to CT448 (- Tryp) and with trypsin-activated CT448 (+ Tryp). N-Succinyl-alanine-alanine-proline-phenylalanine-para-nitroanilide (AAPFpNA), N-Succinyl-alanine-alanine-valine-para-nitroanilide (AAVpNA) and N-Succinyl-alanine-alanine-alanine-para-nitroanilide (AAApNA) were used as substrate with a concentration of 100 µM. Protease activity was monitored by detecting the release of p-nitroaniline by measuring the absorption at 380 nm. Only AAApNA and AAVpNA showed high specific activity of trypsin-activated CT448. Depicted is the mean specific enzymatic activity (n = 3, ± SD). Note different scaling of y-axes.

A



B

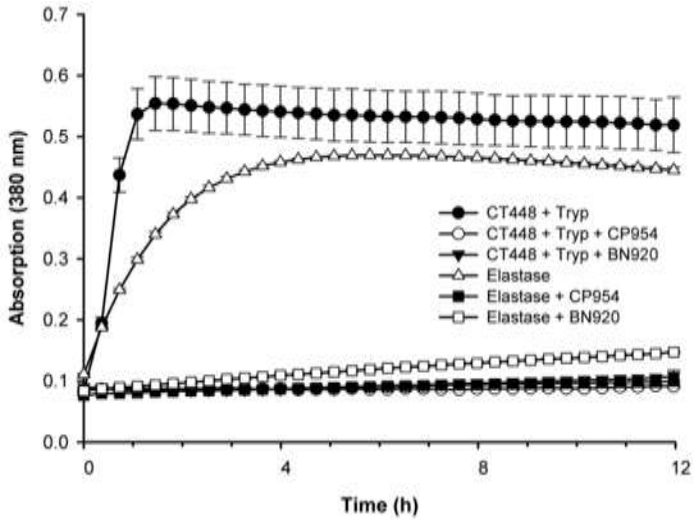


Figure 4: Hydrolysis of a synthetic substrate in the absence/presence of cyanobacterial PIs by heterologously expressed CT448 of *D. magna* and of porcine elastase. A) Comparison of *D.magna* CT448 to porcine elastase on AAApNA. Non-activated (CT448 – Tryp) and trypsin-activated CT448 (CT448 + Tryp), porcine elastase and trypsin were tested, the latter did not hydrolyze the elastase specific substrate. B) Addition of 20 µM of the cyanobacterial serine protease inhibitors cyanopeptolin 954 (CP954) or nostopeptin 920 (BN920), which had been isolated from the cyanobacterium *Microcystis*

aeruginosa NIVA Cya43, to *D. magna* CT448 and porcine elastase. Both serine proteases were completely inhibited with CP954 or BN920. The measurements were carried out at 27°C for 12 hours. Depicted is the mean absorption$_{380nm}$ (n = 3, ± SD).