



The death enzyme CP14 is a unique papain-like cysteine proteinase with a pronounced S2 subsite selectivity

Melanie Paireder^a, Ulrich Mehofer^a, Stefan Tholen^b, Andreas Porodko^a, Philipp Schähns^a, Daniel Maresch^c, Martin L. Biniossek^b, Renier A.L. van der Hoorn^d, Brigita Lenarcic^e, Marko Novinec^e, Oliver Schilling^{b,f}, Lukas Mach^{a,*}

^a Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, Vienna, Austria

^b Institute of Molecular Medicine and Cell Research, University of Freiburg, Germany

^c Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria

^d The Plant Chemetics Laboratory, Department of Plant Sciences, University of Oxford, United Kingdom

^e Department of Chemistry and Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia

^f BIOS Centre for Biological Signaling Studies, University of Freiburg, Germany

ARTICLE INFO

Article history:

Received 7 April 2016

Received in revised form

25 May 2016

Accepted 26 May 2016

Available online 28 May 2016

Keywords:

Cathepsin

Cysteine protease

Cell death

Degradomics

Plant

ABSTRACT

The cysteine protease CP14 has been identified as a central component of a molecular module regulating programmed cell death in plant embryos. CP14 belongs to a distinct subfamily of papain-like cysteine proteinases of which no representative has been characterized thoroughly to date. However, it has been proposed that CP14 is a cathepsin H-like protease. We have now produced recombinant *Nicotiana benthamiana* CP14 (NbCP14) lacking the C-terminal granulin domain. As typical for papain-like cysteine proteinases, NbCP14 undergoes rapid autocatalytic activation when incubated at low pH. The mature protease is capable of hydrolysing several synthetic endopeptidase substrates, but cathepsin H-like aminopeptidase activity could not be detected. NbCP14 displays a strong preference for aliphatic over aromatic amino acids in the specificity-determining P2 position. This subsite selectivity was also observed upon digestion of proteome-derived peptide libraries. Notably, the specificity profile of NbCP14 differs from that of aleurain-like protease, the *N. benthamiana* orthologue of cathepsin H. We conclude that CP14 is a papain-like cysteine proteinase with unusual enzymatic properties which may prove of central importance for the execution of programmed cell death during plant development.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The tobacco-related plant species *Nicotiana benthamiana* is a popular model organism for the study of plant-pathogen interactions [1]. Substantial evidence has been obtained for a molecular arms race between plant-derived papain-like cysteine proteinases (PLCPs) and antagonistic effectors released by

Abbreviations: ALP, aleurain-like protease; Bz, benzoyl; CHN₂, diazomethyl; CDR, complementarity-determining region; DTT, dithiothreitol; mAb, monoclonal antibody; MCA, 4-methylcoumaryl-7-amide; Nb, *Nicotiana benthamiana*; PICS, Proteomic Identification of Protease Cleavage Sites; PLCP, papain-like cysteine proteinase; Z, benzyloxycarbonyl.

* Corresponding author. Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences, Muthgasse 18, A-1190 Vienna, Austria.

E-mail address: lukas.mach@boku.ac.at (L. Mach).

<http://dx.doi.org/10.1016/j.ab.2016.05.017>

0003-9861/© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

pathogens at the site of infection [2]. Intriguingly, a delicate balance between endogenous PLCPs and cysteine proteinase inhibitors also dictates proper timing of programmed cell death in tobacco embryos [3]. Hence, *N. benthamiana* is a suitable host for further studies on the diverse physiological functions of PLCPs. Furthermore, *N. benthamiana* is increasingly used as expression platform for the production of biotherapeutics [4]. In particular, *N. benthamiana* lines have been established which permit the production of recombinant proteins such as monoclonal antibodies (mAbs) with customized post-translational modifications and thus superior biological activities [5,6]. However, a frequent challenge encountered during the production of mAbs in plants is their proteolytic degradation [7–10]. Recent studies have highlighted that mAb fragmentation in *N. benthamiana* and tobacco involves serine and cysteine proteases including PLCPs [11–14].

Considerable progress is currently made in the biochemical and

structural characterization of PLCPs and their animal counterparts, the cysteine cathepsins [15,16]. Based on a comprehensive phylogenetic analysis, PLCPs have been recently divided into 9 subfamilies. The first six subfamilies comprise cathepsin L-like proteases differing in individual molecular aspects, with two of them (subfamilies 1 and 4) featuring C-terminal extensions. Subfamilies 7–9 are phylogenetically distinct and display homologies to mammalian cathepsins F, H and B, respectively [17]. The development of sophisticated activity-based probes has allowed the selective identification of several PLCPs in different plant species including *N. benthamiana* [18]. However, only two *N. benthamiana* PLCPs have been studied thoroughly on the protein level so far: aleurain-like protease (ALP), assigned to subfamily 8, and cathepsin B, a member of subfamily 9 [19–21]. Another *N. benthamiana* PLCP has recently received substantial attention. This enzyme promotes plant immunity in a similar way as tomato C14 (subfamily 1) and thus was initially named NbC14 [22,23]. However, it later became evident that NbC14 belongs to subfamily 4 [17] and is the orthologue of NtCP14, a protease involved in the spatiotemporal control of programmed cell death [3]. Based on specificity profiling with a small number of synthetic PLCP substrates, it was concluded that the enzymatic properties of NtCP14 resemble those of mammalian cathepsin H [3].

To study the catalytic features of *N. benthamiana* NbC14 (hereafter referred to as NbCP14) in depth, we have now produced recombinant forms of this protease in different heterologous expression systems. This enabled us to conduct a detailed characterization of its enzymatic properties using synthetic substrates as well as proteome-derived peptide libraries. Our results demonstrate that CP14 lacks cathepsin H-like aminopeptidase activity and exhibits a specificity profile distinct of ALP, the plant orthologue of cathepsin H.

2. Materials and methods

2.1. Heterologous expression of NbCP14 in insect cells

The plasmid pTP11 containing the complete open reading frame of NbCP14 [22] was used as a PCR template. Two different constructs were generated: pVTBacHis-FLAG-NbCP14 encodes a truncated version of the proenzyme lacking the C-terminal granulin domain (NbCP14 residues 30–397, preproprotein numbering), whereas both the granulin domain and the proline-rich linker region were removed in the case of pVTBacHis-FLAG-NbCP14 Δ P (NbCP14 residues 30–367). The same forward primer (5'-AAATC-TAGATTTACTACTGATTTTCCAATAC-3') was used in conjunction with different reverse primers, 5'-TCTGGTACCTCAACTTGGTTTTGGAAAAG-3' (NbCP14) or 5'-TCTGGTACCTCAAGAGGCTTCTTTCGTTG-3' (NbCP14 Δ P). The PCR products were cleaved with *Xba*I and *Kpn*I (Fermentas, St. Leon-Rot, Germany) at the underlined sites and ligated into pVTBacHis-FLAG [24] digested with the same enzymes. In the resulting plasmids, the respective NbCP14 sequence is positioned in-frame behind a leader segment consisting of the melittin signal peptide followed by a 4-kDa linker region containing six consecutive histidine residues and the FLAG epitope.

Heterologous production of pro-NbCP14 and pro-NbCP14 Δ P in insect cells was conducted according to previously published procedures [24,25]. Briefly, recombinant baculoviruses were generated by co-transfection of *Spodoptera frugiperda* Sf9 cells with the respective pVTBacHis-FLAG construct and baculoviral DNA. After infection of *S. frugiperda* Sf21 cells, the recombinant proteins were isolated from the culture supernatants by two-step affinity chromatography using Chelating Sepharose (GE Healthcare, Little Chalfont, United Kingdom) and Anti-FLAG M2 Affinity Gel (Sigma-Aldrich, St. Louis, USA).

2.2. Production of NbCP14 in *Escherichia coli*

The sequence encoding NbCP14 residues 30–397 was amplified by PCR with the primer combination 5'-GCAGCTAGCTTTACTACT-GATTTTCCAATACTA-3' (forward) and 5'-GCACCTCGACTCAACTsTGGTTTTGGAAAAG-3' (reverse). The PCR product was then cleaved with *Nhe*I and *Xho*I at the underlined positions and ligated into the corresponding sites of pET-32/28 prior to expression in the *E. coli* strain Rosetta-gami B (DE3) pLysS as described previously [21,26]. The resulting cell pellet was resuspended in 20 mM Hepes (pH 7.4) supplemented with 500 mM NaCl and 20 mM imidazole. After cell lysis by sonication, debris was removed by centrifugation at 16,000g for 15 min at 4 °C. The supernatant was then passed through a column of Chelating Sepharose charged with Ni²⁺ ions and equilibrated in resuspension buffer. After successive washes with 40, 60 and 80 mM imidazole, recombinant pro-NbCP14 was eluted with 250 mM imidazole. Protein-containing eluate fractions were pooled, dialysed against phosphate-buffered saline (20 mM sodium phosphate, pH 7.4, 150 mM NaCl) containing 0.02% (w/v) NaN₃ and then concentrated by ultrafiltration.

2.3. Activity assays

Pro-NbCP14 (0.3 mg/ml) was autocatalytically activated by incubation in 0.1 M sodium acetate (pH 5.0) containing 2.5 mM DTT for 30 min at 37 °C. Stopped assays were performed at 37 °C as outlined previously [27,28], using 10 μ M of the respective peptidyl-MCA substrate (Bachem, Bubendorf, Switzerland; PeptaNova, Sandhausen, Germany) in 0.1 M sodium acetate (pH 5.0), 1% dimethylsulfoxide, 5 mM DTT. The reactions were terminated by addition of an equal volume of 0.1 M monochloroacetic acid, 0.1 M sodium acetate (pH 4.3) prior to analysis by spectrofluorimetry. Active-site titration was done by preincubation with varying concentrations of the irreversible inhibitor E-64 for 60 min at 0 °C [29]. Kinetic measurements with Z-Leu-Arg-MCA or Z-Phe-Arg-MCA as substrate were performed at 25 °C following previously published procedures [26]. The assay buffer was the same as stated above. 5% dimethylsulfoxide was used to enhance the solubility of Z-Phe-Arg-MCA at high substrate concentrations. Reaction rates were determined from the respective progress curves. The kinetic parameters K_m and k_{cat} were derived by non-linear regression analysis using the Henri-Michaelis-Menten equation and GraphPad Prism 5.0 software.

2.4. Specificity profiling with proteome-derived peptide libraries

Specificity profiling of *N. benthamiana* aleurain-like protease (NbALP) and NbCP14 was performed using the Proteomic Identification of Protease Cleavage Sites (PICS) procedure [30–32] or an adaption thereof [33]. Briefly, proteome-derived peptide libraries were prepared by digestion of *E. coli* or HEK293 lysates with trypsin or GluC. These peptide samples (300 μ g) were incubated with the enzyme to be tested (0.6–3 μ g) in assay buffer (NbALP: 0.2 M Mes (pH 5.0), 5 mM DTT; NbCP14: 0.1 M sodium acetate (pH 5.0), 5 mM DTT) for up to 16 h at room temperature. Protease-treated and control samples were then differentially labelled by reductive dimethylation with either light (¹²COH₂) or heavy formaldehyde (¹³COD₂), combined and fractionated by strong-cation exchange chromatography prior to analysis by liquid chromatography-tandem mass spectrometry as described previously [34]. Spectrum to sequence assignment and relative peptide quantitation were performed as reported [35] with the following adaptations: asymmetric precursor mass error of +150 ppm/–0 ppm, fragment ion mass tolerance of 0.1 Da, semi-GluC or semi-trypsin specificity. Semi-specific peptides that were enriched >4-fold in the protease-

treated samples were considered as cleavage products. The corresponding prime and non-prime sequences were reconstructed bioinformatically through database searches and visualized using iceLogo [36].

2.5. Profiling with activity-based probes

Active-site labelling of NbCP14 was performed in 0.1 M sodium acetate (pH 5.5) supplemented with 5 mM DTT and varying concentrations of DCG-04 [37]; Matthew Bogyo, Stanford University, USA), biotin-CA074 [38]; Norbert Schaschke, University of Bielefeld, Germany) or biotin-Leu-Val-Gly-CHN₂ [39]; Gilles Lalmanach, University of Tours, France). After incubation for 30 min at 37 °C, the reactions were stopped by addition of 50 µg/ml E-64 and subsequent treatment for 5 min at 95 °C.

2.6. Digestion of mAbs

The human anti-HIV mAbs 2F5 and 2G12 (Dietmar Katinger, Polymun Scientific GmbH, Klosterneuburg, Austria) were tested for their susceptibility to NbCP14 as outlined earlier [12]. Briefly, 2F5 or 2G12 (200 µg/ml) were treated with NbCP14 (50 µg/ml) in 0.1 M sodium acetate (pH 5.5) containing 2 mM DTT at 37 °C. After incubation for up to 16 h, reactions were stopped by treatment for 5 min at 95 °C.

2.7. SDS-PAGE and western blotting analysis

Samples were denatured for 5 min at 95 °C under reducing conditions and then subjected to 12.5% SDS-PAGE. Separated polypeptides were then either subjected to silver staining or electrophoretically transferred onto Hybond-C nitrocellulose membranes (GE Healthcare). After probing the membranes with monoclonal mouse anti-FLAG M2 (Sigma-Aldrich), bound immunoglobulins were visualized with peroxidase-conjugated goat anti-mouse IgG antibodies (Jackson ImmunoResearch, West Grove, USA) and enhanced chemiluminescence reagents (Bio-Rad, Richmond, USA). mAbs and their heavy-chain fragments were detected with peroxidase-labelled anti-human IgG (γ-chain-specific; Sigma-Aldrich). Streptavidin-peroxidase (Vector Laboratories, Burlingame, USA) was used for the detection of biotinylated proteins on western blots.

2.8. Mass spectrometry

Purified pro-NbCP14 (0.5 mg/ml) was fractionated on a Thermo ProSwift RP-4H column (250 × 0.20 mm) using a Dionex UltiMate 3000 HPLC system (Thermo Scientific, Waltham, USA). After application of the sample (5 µl), elution was performed at 80 °C and a flow rate of 8 µl/min with a gradient of 20–95% solvent B (80% acetonitrile in 0.01% trifluoroacetic acid) in solvent A (0.05% trifluoroacetic acid) over 40 min as follows: 20–65% B (20 min), 65–95% B (20 min). Eluted polypeptides were analysed online on a maXis 4G ETD QTOF mass spectrometer (Bruker, Billerica, USA) equipped with an electrospray ionization source and operated in the positive ion mode (m/z range: 400–3800). The m/z values of the 8 most prominent charge states were used to deduce the molecular mass of recombinant pro-NbCP14.

2.9. Other methods

Recombinant pro-NbALP was produced and activated as described previously [21]. N-terminal sequence analysis of blotted bands was conducted at the Department of Molecular and Biomedical Sciences (Jozef Stefan Institute, Ljubljana, Slovenia) and

the Protein Micro-Analysis Facility (Medical University of Innsbruck, Austria) as outlined earlier [12]. Enzymatic deglycosylation of proteins was performed as reported [25]. Total protein content was determined with the BCA Protein Assay Kit (Pierce, Rockford, USA), using bovine serum albumin as standard.

3. Results

3.1. Heterologous expression of NbCP14

Full-length NbCP14 (505 amino acids; 56.9 kDa; GenBank: KU212214) shares 93% identity with NtCP14 (GenBank: KF113573), and the catalytic domains differ only at 11 positions (5%). Both enzymes belong to PLCP subfamily 4 [17] and are considerably larger than other PLCPs due to the additional presence of a C-terminal granulin domain (108 amino acids), which is connected to the other parts of the protease by a 30-residue proline-rich linker region. The latter segment is preceded by an N-terminal signal sequence (29 amino acids), a propeptide (116 residues) encompassing the ERFNIN motif typical for cathepsin L-like PLCPs [15,16], and the catalytic domain (222 amino acids) with the three active-site residues (Cys¹⁷⁰, His³⁰⁸ and Asn³²⁸; Table S1).

Biochemical information on subfamily 4 enzymes is still sparse, but related PLCPs belonging to subfamily 1 have been characterized in more detail. Since the C-terminal granulin domain of subfamily 1 proteases is dispensable for proteolytic activity [20] but renders the enzymes prone to precipitation [40], we have first generated a truncated NbCP14 precursor (residues 30–397, preproprotein numbering) devoid of the granulin domain (pro-NbCP14) in order to improve the likelihood of proper secretion of the proenzyme into the culture medium. Although pro-NbCP14 could be expressed in insect cells, the amounts of recombinant protein accumulating in the culture supernatant were too low for further characterization. We have therefore generated a second deletion construct (residues 30–367) which additionally lacks the proline-rich region (pro-NbCP14ΔP). This protein was more efficiently secreted than pro-NbCP14 and could be purified to near homogeneity. When purified pro-NbCP14ΔP was subjected to SDS-PAGE analysis, a major 43-kDa band was observed. This apparent molecular mass is in reasonable agreement with its theoretical size (41.9 kDa). However, a 31-kDa polypeptide was also consistently present in purified pro-NbCP14ΔP samples (Fig. 1A). Incubation at low pH in the presence of a reducing agent resulted in the rapid conversion of pro-NbCP14ΔP into the faster migrating species, even in the absence of the processing enhancer dextran sulphate [21]. It is well established

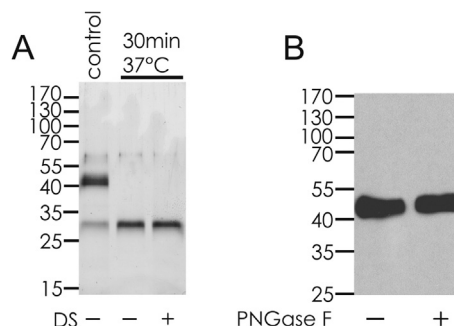


Fig. 1. Characterization of NbCP14 produced in insect cells. (A) Purified pro-NbCP14ΔP was incubated in 0.1 M sodium acetate (pH 5.0), 2 mM cysteine in the absence or presence of 10 µg/ml dextran sulphate (DS) prior to analysis by SDS-PAGE and silver staining. An untreated sample served as control. (B) Pro-NbCP14ΔP was subjected to treatment with peptide N-glycosidase F prior to immunoblotting analysis with anti-FLAG antibodies. The migration positions of prestained molecular mass standards are indicated, with their respective masses expressed in kDa.

that cysteine cathepsins and PLCPs can undergo autocatalytic maturation under acidic conditions [28,41]. Although pro-NbCP14 Δ P contains an *N*-glycosylation sequon within its prodomain, the electrophoretic mobility of the recombinant proenzyme was not increased by treatment with peptide *N*-glycosidase F. This suggests that the potential *N*-glycosylation site of pro-NbCP14 Δ P (Asn⁸⁷) is structurally buried and thus inaccessible for *N*-glycosylation enzymes (Fig. 1B). Furthermore, these results indicate that the discrepancy between the size of the 31-kDa autocatalytic maturation product and its predicted molecular mass (24.4 kDa) is not due to *N*-glycosylation.

Opposite to the baculovirus expression system, pro-NbCP14 containing the proline-rich linker region could be produced in *E. coli* with good yields and purity (3 mg proenzyme per liter bacterial culture). The calculated molecular mass of the recombinant proenzyme (43220.3 Da) could be experimentally verified by liquid chromatography-electrospray ionization-mass spectrometry (43220.1 \pm 1.5 Da), thus confirming proper removal of the initiator methionine and formation of the predicted 4 disulphide bridges (Fig. 2A). However, the recombinant proenzyme migrated in SDS-PAGE gels as a 51-kDa polypeptide, which gave rise to a 36-kDa band upon acid-induced autocatalytic maturation (Fig. 2B). The main N-terminal sequence of the 36-kDa protein as determined by Edman degradation was found to correspond to Ser¹⁴³-Cys-Asp-Val-Pro, accompanied by smaller amounts of Thr¹⁴¹-Thr-Ser-Cys-Asp and Asp¹⁴⁵-Val-Pro-Pro-Ser. Similar to cysteine cathepsins [42], self-processing of NbCP14 thus takes place 1–5 residues upstream of the first amino acid of the protease domain (Val¹⁴⁶; Table S1). The theoretical molecular mass of the main form of the autocatalytically matured enzyme is therefore 27.3 kDa. Hence, the electrophoretic mobility of mature NbCP14 is slower than expected as also noted for its precursor and 31-kDa NbCP14 Δ P. Similar observations have been previously reported for recombinant pro-NtCP14 as well as cathepsins B and L [3,43,44].

3.2. Catalytic features of NbCP14

Activity-based probes are used increasingly to detect PLCPs in complex samples such as plant extracts. We have therefore tested mature NbCP14 for its reactivity with three such probes: DCG-04 [37], biotin-CA074 [38] and biotin-Leu-Val-Gly-CHN₂ [39] (see Fig. S1 for the chemical structures of these compounds). NbCP14 was strongly labelled by DCG-04 in a concentration-dependent manner. The protease also showed a pronounced reaction with biotin-CA074. By contrast, labelling of NbCP14 with biotin-Leu-Val-Gly-CHN₂ was comparatively weak, rendering DCG-04 and biotin-CA074 the preferred reagents for the detection of active forms of NbCP14 in cell and tissue homogenates (Fig. 3).

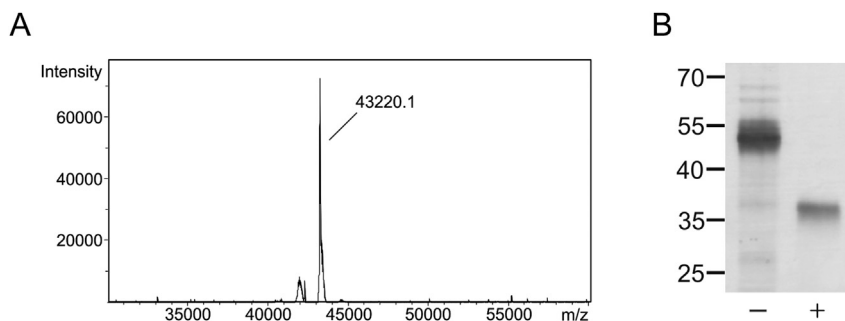


Fig. 2. Characterization of NbCP14 produced in *E. coli*. (A) Purified pro-NbCP14 was subjected to analysis by liquid chromatography-electrospray ionization-mass spectrometry. The theoretical molecular mass of the recombinant protein is 43220.3 Da. (B) Pro-NbCP14 was incubated in 0.1 M sodium acetate (pH 5.0), 2.5 mM DTT for 30 min at 37 °C (+) prior to analysis by SDS-PAGE and silver staining. An untreated sample (–) served as control. The migration positions of prestained molecular mass standards are indicated.

We have then assessed the catalytic properties of mature NbCP14 using a series of synthetic substrates frequently used to monitor the activities of cysteine cathepsins and PLCPs. While the enzyme cleaved a number of endopeptidase substrates, no activity was observed with aminopeptidase substrates such as Arg-MCA (Table 1). In particular, NbCP14 was found to display high hydrolytic activity towards Z-Leu-Arg-MCA. It is of note that Z-Phe-Arg-MCA is cleaved much less efficiently by this enzyme, mainly due to a far slower turnover (Table 2). This indicates that NbCP14 favours leucine over phenylalanine in its specificity-determining S2 subsite.

We have also tested NbCP14 for its capacity to degrade native proteins. For these experiments, the two monoclonal anti-HIV antibodies 2F5 and 2G12 were chosen as substrates. We have previously observed that 2F5 is far more sensitive to PLCPs than 2G12 [12]. Treatment of 2F5 with two other *N. benthamiana* PLCPs, NbALP and cathepsin B, preferentially results in the formation of a discrete 40-kDa polypeptide due to cleavage within the CDR H3 loop of the antibody (Leu-Phe-Gly¹⁰⁸↓Val¹⁰⁹-Pro-Ile) [21]. Entirely different results were obtained for NbCP14. This enzyme completely degraded both mAbs, with only transient accumulation of fragments with apparent masses of 30–40 kDa (Fig. 4). The N-terminus of the 30-kDa 2F5 polypeptide produced by NbCP14 was determined as Thr²⁴⁰, indicative of cleavage at the same site (Lys-Thr-His²³⁹↓Thr²⁴⁰-Cys-Pro; Table S2) as previously reported for papain [45] and human cathepsin L [12].

3.3. Subsite specificity profile of NbCP14

To assess the subsite specificity of NbCP14, we applied the conventional Proteomic Identification of Protease Cleavage Sites (PICS) procedure [30–32] as well as a recently reported adaptation thereof [33]. Preliminary experiments revealed that an enzyme-to-library ratio of 1:100 and an incubation time of 16 h result in the most informative data sets. Standard PICS analysis of tryptic *E. coli* and HEK293 libraries treated with NbCP14 under these conditions yielded 1971 and 1193 unique cleavage sequences, respectively (Fig. 5). The substrate specificity of cysteine cathepsins and PLCPs is largely defined by their S2 pocket [46]. Both sequence logos depict a

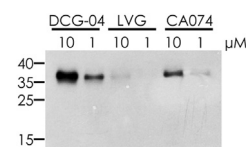


Fig. 3. Profiling of NbCP14 with activity-based probes. Mature NbCP14 was incubated for 30 min at 37 °C with the indicated concentrations of DCG-04, biotin-Leu-Val-Gly-CHN₂ (LVG) or biotin-CA074 (CA074) followed by SDS-PAGE and western blotting using streptavidin-peroxidase for detection. The migration positions of prestained molecular mass standards are indicated.

Table 1

Hydrolysis of peptidyl-MCA substrates by NbCP14. The following substrates were not cleaved (<0.2 mU/mg): Ala-MCA, Arg-MCA, Leu-MCA, Phe-MCA, Gly-Phe-MCA, Z-Arg-MCA, Z-Arg-Arg-MCA, Z-Ala-Ala-Asn-MCA, Z-Gly-Pro-Arg-MCA, acetyl-Tyr-Val-Ala-Asp-MCA, methoxysuccinyl-Ala-Ala-Pro-Val-MCA. All substrates were tested at a final concentration of 10 μ M. mU, nmol per min; Suc, succinyl.

Substrate	Specific activity (mU/mg)	Relative activity (%)
Z-Leu-Arg-MCA	771	100
Boc-Val-Leu-Lys-MCA	158	20.4
Bz-Phe-Val-Arg-MCA	112	14.5
Boc-Val-Arg-MCA	48	6.2
Z-Phe-Arg-MCA	39	5.1
Z-Leu-Leu-Glu-MCA	26	3.3
Z-Val-Val-Arg-MCA	25	3.2
Suc-Leu-Leu-Val-Tyr-MCA	4.4	0.6
Z-Leu-Leu-Leu-MCA	3.8	0.5
Z-His-Glu-Lys-MCA	0.3	<0.1

Table 2

Kinetic parameters of NbCP14.

Enzyme	Substrate	K_m [μ M]	k_{cat} [s^{-1}]	k_{cat}/K_m [$mM^{-1} s^{-1}$]
NbCP14	Z-Leu-Arg-MCA	9.4 ± 1.5	1.7 ± 0.4	181
	Z-Phe-Arg-MCA	23 ± 2	0.05 ± 0.02	2.2
Cat H ^a	Bz-Phe-Val-Arg-MCA	38 ± 8	110 ± 13	2900
	Z-Phe-Arg-MCA	82 ± 26	4.3 ± 1.0	52.5
Papain ^b	Z-Phe-Arg-MCA	420 ± 70	52 ± 6	119
	Asp158Ala ^b	920 ± 70	0.46 ± 0.08	0.5

^a Cat H, recombinant human cathepsin H lacking the mini-chain; data taken from Ref. [55].

^b Asp158Ala, mutant papain (Asp¹⁵⁸ \rightarrow Ala); data taken from Ref. [59].

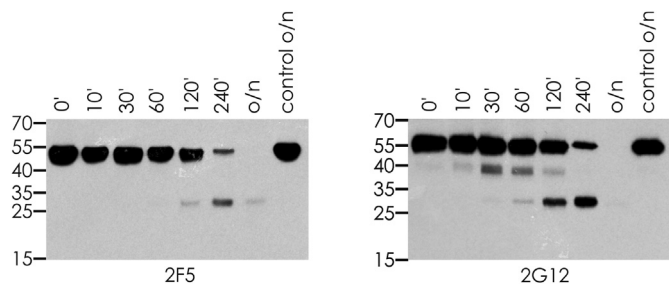


Fig. 4. Processing of anti-HIV mAbs 2F5 and 2G12 by NbCP14. mAbs (200 ng) were incubated with mature NbCP14 (50 ng) at pH 5.5 for the indicated times and then analysed by immunoblotting with antibodies to the heavy chain of human IgG. Untreated mAb was used as control. The migration positions of selected molecular mass standards are indicated. o/n, overnight (16 h).

strong preference of NbCP14 for branched, aliphatic amino acids in P2 (leucine > valine > isoleucine). In P1, NbCP14 accepts small amino acids like glycine and threonine, but also tolerates glutamic acid and glutamine as reported for cathepsins B, L and S [32]. The known preference of PLCPs for basic amino acids in P1, as revealed by positional scanning of combinatorial synthetic substrate libraries [47], cannot be assessed using tryptic peptide libraries. A moderate selectivity of NbCP14 for histidine and hydrophobic amino acids in P3 is consistent with similar findings for cysteine cathepsins and other PLCPs [32,47], but a distinct substrate specificity is not observed for this position. The same applies to P1' where small amino acids as well as aspartic acid/asparagine and glutamic acid/glutamine were found to be weakly enriched. Overall, the subsite profile of NbCP14 as determined by standard PICS argues for P2 as the major specificity factor with only minor contributions from other positions.

The NbCP14 subsite specificity was also assessed with an adapted PICS procedure [33]. This method avoids chemical

modification of primary amines prior to exposure to a protease under investigation and thus also enables the profiling of lysine specificities. *E. coli* libraries generated by treatment with trypsin or GluC were incubated with NbCP14 under the same conditions as above prior to differential labelling of protease-treated and control samples by reductive dimethylation using different stable isotopes of formaldehyde. Upon analysis by liquid chromatography-mass spectrometry, NbCP14 cleavage products could be identified as semi-specific peptides enriched for the corresponding isotopic label. Two independent digests of tryptic *E. coli* libraries with NbCP14 yielded 149 and 207 cleavage sequences, respectively. Consistent with the data acquired by conventional PICS, the preference of NbCP14 for aliphatic amino acids in P2 was pronounced. Glycine and threonine were the most strongly enriched P1 residues (Fig. 6), which is also in good agreement with the results obtained using the

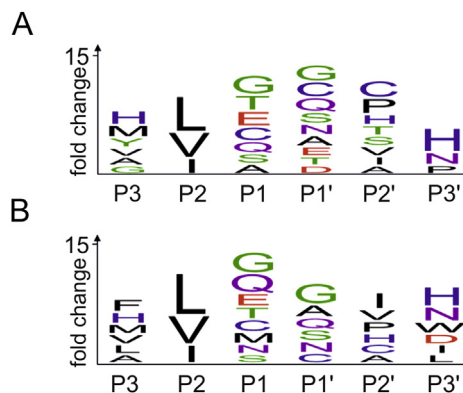


Fig. 5. Specificity profiling of NbCP14 using conventional PICS. Tryptic peptide libraries derived from *E. coli* (A) or HEK293 cells (B) were incubated with mature NbCP14 (protease-to-library ratio: 1:100) for 16 h at room temperature. IceLogos are used to present the identified cleavage sites as cumulative fold-change normalized for the natural abundance of the respective amino acids.

conventional PICS protocol. Glycine has been previously reported to be tolerated well in P1 and P1' of human cathepsins B, L and S. Furthermore, threonine is frequently observed in the P1 position of cysteine cathepsin and PLCP specificity profiles [32,47]. P3 was again found to contribute little to the specificity of NbCP14 with solely histidine exceeding its natural abundance by more than 2-fold at this position. However, the adapted PICS procedure revealed that NbCP14 readily accepts acidic amino acids in P1'.

Many PLCPs accept lysine or arginine residues in the P1 position. Such sequences are rare in tryptic libraries due to the substrate specificity of this enzyme. Therefore, we additionally used the adapted PICS procedure to investigate the action of NbCP14 on a GluC-treated *E. coli* library, obtaining 213 cleavage sequences (Fig. 6). For P2, the same subsite specificity was observed as in the

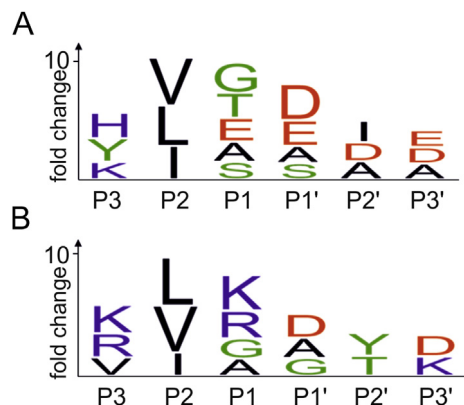


Fig. 6. Specificity profiling of NbCP14 using the adapted PICS procedure. *E. coli* peptide libraries prepared with trypsin (A) or GluC (B) were incubated with mature NbCP14 (protease-to-library ratio: 1:100) for 16 h at room temperature. Semi-specific peptides with a more than 4-fold enrichment were used for reconstruction of the cleavage sites displayed as iceLogos.

aforementioned profiles. However, NbCP14 was found to display a preference for lysine and arginine in P1, which is in good agreement with the data available for other PLCPs [47]. An enrichment of basic amino acids was also observed at P3, whereas small and acidic amino acids dominated in P1'. Taken together, these results demonstrate that the adapted and conventional PICS protocols yield comparable results for the subsite specificity of NbCP14. However, it should be pointed out that the sequence logos obtained with the modified procedure were much easier to interpret than those derived using the standard protocol.

The PLCP most closely related to cathepsin H is ALP. Like its mammalian counterpart, ALP primarily functions as aminopeptidase with moderate endopeptidolytic activity [21,48]. This is attributed to the so-called mini-chain, which partially occupies the active-site cleft and thereby restricts the substrate-binding sites of cathepsin H and related enzymes [49]. Adapted PICS assays of NbALP with tryptic HEK293 libraries led to the identification of 252 cleavage sequences. Strikingly, a large fraction of the detected cleavage events (60%) resulted from the removal of a single residue and were thus due to the aminopeptidase activity of the enzyme. 27% corresponded to the removal of two or three amino acids. Notable differences were observed between the subsite preferences of NbALP and NbCP14. In P1, NbALP readily accepts bulky side chains such as those of phenylalanine, tyrosine and methionine. Serine and threonine are also tolerated in this position (Fig. 7). Similar features have been reported for cathepsin H [50,51]. The S1' pocket of NbALP prefers hydrophobic residues (valine, alanine, isoleucine) as well as tyrosine and aspartic acid. The P2 selectivity of NbALP is not as pronounced as that of NbCP14 and displays an enrichment of small neutral amino acids such as threonine, valine

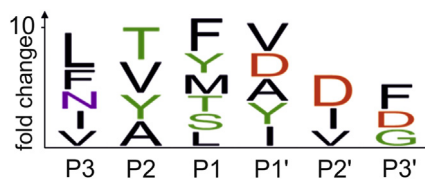


Fig. 7. Specificity profiling of NbALP using the adapted PICS procedure. A tryptic peptide library derived from HEK293 cells was incubated with mature NbALP (protease-to-library ratio: 1:100) for 16 h at room temperature. Semi-specific peptides with a more than 4-fold enrichment were used for reconstruction of the cleavage sites displayed as iceLogo.

and alanine. Taken together, NbALP and NbCP14 clearly differ in their modes of action as well as in their substrate specificities.

4. Discussion

In tobacco, NtCP14 was identified as a key component of a bipartite module controlling the initiation of programmed cell death during the early stages of embryogenesis [3]. In the case of *N. benthamiana*, previous studies have revealed that NbCP14 is required for a powerful response to plant pathogens. This is reminiscent of tomato and potato C14 which also promote plant immunity [22,23]. The latter two enzymes are closely related to *Arabidopsis thaliana* RD21, an extensively studied protease from PLCP subfamily 1 [17,20,40]. However, NbCP14 appears to differ from RD21 in various aspects, thus providing evidence for substantial biochemical variations between subfamilies 1 and 4. For instance, our results unequivocally demonstrate that NbCP14 is capable of autocatalytic activation. By contrast, previous studies have indicated that maturation of recombinant RD21 produced in insect cells requires the addition of plant extracts [40]. Hence, it has been suggested that RD21 activation depends on the action of vacuolar processing enzymes, an unrelated class of cysteine proteases. However, we have recently obtained genetic evidence that vacuolar processing enzymes are dispensable for RD21 activation in *A. thaliana* [20]. Another difference between NbCP14 and RD21 relates to the functional significance of the proline-rich region connecting the catalytic and granulin domains. In the case of RD21, deletion of this spacer element resulted in accumulation of an apparently inactive precursor protein, suggesting that the proline-rich segment has to be present for proper folding of RD21-like proteases [20]. Conversely, recombinant NbCP14 lacking this linker sequence is capable of rapid autocatalytic maturation, thus highlighting remarkable differences between PLCP subfamilies 1 and 4 with respect to their folding competence and cellular activation mechanisms.

Opposite to NbALP and *N. benthamiana* cathepsin B [21], NbCP14 failed to cleave the anti-HIV mAb 2F5 in its CDR H3 loop. Interestingly, the cleavage site of the former enzymes in this region of the antibody (Leu-Phe-Gly↓Val-Pro-Ile) features a phenylalanine in the P2 position. Using synthetic peptide substrates, we observed that NbCP14 displays a pronounced preference for leucine at this location (see Table 1). This unique feature prompted us to test NbCP14 for its reaction with various activity-based probes which are frequently used to map the subsite specificities of cysteine proteinases. The prototypical PLCP probe is DCG-04, whose structure is based on the potent cysteine proteinase inhibitor E-64 [37]. DCG-04 and derivatives thereof have been successfully used to detect PLCPs in plant extracts, including the NbCP14 orthologue XBCP3 from *A. thaliana* [17,52]. Interestingly, all these compounds contain a leucine residue which docks into the affinity-determining S2 subsite of target PLCPs [53]. Hence, the strong labelling of NbCP14 by DCG-04 is in good agreement with the observed substrate specificity of this protease. NbCP14 also gave pronounced signals with another E-64 relative, biotin-CA074. This probe has been originally designed as a selective high-affinity label for cathepsin B [38]. The reaction of biotin-CA074 with NbCP14 indicates that the latter enzyme can tolerate the C-terminal dipeptide extension of the inhibitor in its prime-site pockets. By contrast, the binding of biotin-Leu-Val-Gly-CHN₂ to NbCP14 was relatively weak. This could reflect the preference of NbCP14 for leucine over valine in the P2 position. However, it has been also noted that the biotin moiety of the probe can interfere with its efficient accommodation in the active site of PLCPs due to the absence of a linker segment [39].

For plants, very few proteomics-based studies of protease

specificity have been conducted so far. Recently, the PICS methodology has been used to determine the cleavage-site specificities of matrix metalloproteinases from *A. thaliana* [54]. In this study, we have now used NbCP14 to evaluate an improved PICS protocol [33] for its suitability to assess the substrate preferences of PLCPs. This new procedure circumvents a number of known PICS limitations, such as the lack of normalization and the modification of certain amino acids during library generation. The adapted method omits the dimethylation of primary amines and preserves free thiols in the peptide library allowing the recognition of unmodified N-termini, lysines and cysteines by the test protease. Importantly, the adapted and conventional PICS protocols yielded similar results with respect to the subsite preferences of NbCP14. As expected, the observed cleavage patterns are largely determined by the distinct P2 specificity of the enzyme, which clearly favours aliphatic over aromatic amino acids at this position as previously observed for human cathepsin S [32].

The strong preference of NbCP14 for leucine over phenylalanine in P2 is not only due to enhanced substrate affinity, but at large the result of a major difference in the turnover rate. Similar observations have been made previously for cathepsin H, which favours valine over phenylalanine in P2 like NbALP [55,56] (see Table 2). However, the residues forming the S2 subsite of NbCP14 (Met²¹⁴, Ala²⁸⁰, Leu³⁰⁶, Ala³⁰⁹, Leu³⁵⁸) are inconspicuous when compared with papain [57]. Hence, a different molecular feature shared by NbCP14 and cathepsin H could contribute to the distinct enzymatic properties that these two enzymes have in common. In the proximity of the active-site histidine, cathepsin H contains a small four-residue loop (Lys^{155A}-Thr-Pro-Asp^{155D}) that is not found in other cysteine cathepsins [49]. It has been put forward that this flexible loop could cause partial obstruction of the S2 subsite [55]. Interestingly, NbCP14 also contains such an insertion (Ser³⁰¹-Asn-Pro-Asp³⁰⁴) at the equivalent position.

Opposite to cathepsin H, NbCP14 does not display any aminopeptidase activity. Clearly, this can be attributed to the absence of the mini-chain, a segment of the propeptide which remains attached to mature cathepsin H and its plant orthologues via a disulphide bond not present in NbCP14 [21,55,58]. Another striking difference between NbCP14 and cathepsin H-like enzymes relates to the amino acid preceding the active-site histidine. For papain, it has been shown that this residue (Asp¹⁵⁸) is involved in a hydrogen-bonding network stabilizing the thiolate-imidazolium ion pair required for catalysis [59]. While cathepsin H and NbALP feature an asparagine at this position, Asp¹⁵⁸ is replaced by Ser³⁰⁷ in NbCP14. Site-directed mutagenesis experiments with papain have revealed that substitution of Asp¹⁵⁸ with Asn has little impact on the catalytic efficiency of the protease [60]. However, the enzymatic activity of an Asp¹⁵⁸-Ala variant was found to be severely compromised, which was mainly due to a more than 100-fold decrease in k_{cat} [59] (see Table 2). Hence, Ser³⁰⁷ could at least in part account for the moderate processivity of NbCP14 as compared to papain and cathepsin H. Taken together, the unique molecular features of NbCP14 identified in this study may explain its unusual low-affinity interactions with endogenous cysteine proteinase inhibitors [3] and thus provide a rationale for the non-redundant role of the enzyme during execution of programmed cell death in the plant embryo.

Author contributions

MP performed experiments, analysed data and co-wrote the manuscript. UM, ST, AP, PS, DM and MB performed experiments and analysed data; RH contributed essential material; BL, MN and OS planned experiments and analysed data; LM conceived the study, planned and performed experiments, analysed data and co-wrote

the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

Databases

NbCP14 nucleotide sequence data are available in the DDBJ/EMBL/GenBank databases under the accession number KU212214.

Acknowledgements

The expert technical assistance of Franz Jehl, Bettina Mayer and Barbara Svoboda is gratefully acknowledged. We thank Matthew Bogyo, Dietmar Katinger, Gilles Lalmanach and Norbert Schaschke for reagents. Financial support was provided by the Austrian Science Fund (FWF): project W1224-B09; Deutsche Forschungsgemeinschaft (DFG): SCHI 871/2, SCHI 871/5, SCHI 871/6, GR 1748/6, INST 39/900-1, SFB850 - Project B8); European Research Council (Programme “Ideas” – Call identifier: ERC-2011- StG 282111-ProteaSys); and the Excellence Initiative of the German Federal and State Governments (EXC 294, BIOS).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.abb.2016.05.017>.

References

- [1] M.M. Goodin, D. Zaitlin, R.A. Naidu, S.A. Lommel, *Nicotiana benthamiana*: its history and future as a model for plant-pathogen interactions, *Mol. Plant Microbe Interact.* 21 (2008) 1015–1026.
- [2] A.C. Horger, R.A. van der Hoorn, The structural basis of specific protease-inhibitor interactions at the plant-pathogen interface, *Curr. Opin. Struct. Biol.* 23 (2013) 842–850.
- [3] P. Zhao, X.M. Zhou, L.Y. Zhang, W. Wang, L.G. Ma, L.B. Yang, X.B. Peng, P.V. Bozhkov, M.X. Sun, A bipartite molecular module controls cell death activation in the basal cell lineage of plant embryos, *PLoS Biol.* 11 (2013) e1001655.
- [4] Y.Y. Gleba, D. Tuse, A. Giritch, Plant viral vectors for delivery by Agrobacterium, *Curr. Top. Microbiol. Immunol.* 375 (2014) 155–192.
- [5] R. Strasser, F. Altmann, H. Steinkellner, Controlled glycosylation of plant-produced recombinant proteins, *Curr. Opin. Biotechnol.* 30 (2014) 95–100.
- [6] A. Loos, J.S. Gach, T. Hackl, D. Maresch, T. Henkel, A. Porodko, D. Bui-Minh, W. Sommeregger, G. Wozniak-Knopp, D.N. Forthall, F. Altmann, H. Steinkellner, L. Mach, Glycan modulation and sulfoengineering of anti-HIV-1 monoclonal antibody PG9 in plants, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) 12675–12680.
- [7] B. De Muynck, C. Navarre, M. Boutry, Production of antibodies in plants: status after twenty years, *Plant Biotechnol. J.* 8 (2010) 529–563.
- [8] V.K. Hehle, M.J. Paul, P.M. Drake, J.K. Ma, C.J. van Dolleweerd, Antibody degradation in tobacco plants: a predominantly apoptotic process, *BMC Biotechnol.* 11 (2011) 128.
- [9] C. Goulet, M. Khalf, F. Sainsbury, M.A. D'Aoust, D. Michaud, A protease activity-depleted environment for heterologous proteins migrating towards the leaf cell apoplast, *Plant Biotechnol. J.* 10 (2012) 83–94.
- [10] M. Donini, R. Lombardi, C. Lonoce, M. Di Carli, C. Marusic, V. Morea, P. Di Micco, Antibody proteolysis: a common picture emerging from plants, *Bio-engineered* 6 (2015) 299–302.
- [11] S. Robert, M. Khalf, M.C. Goulet, M.A. D'Aoust, F. Sainsbury, D. Michaud, Protection of recombinant mammalian antibodies from development-dependent proteolysis in leaves of *Nicotiana benthamiana*, *PLoS One* 8 (2013) e70203.
- [12] M. Niemer, U. Mehofer, J.A. Torres Acosta, M. Verdianz, T. Henkel, A. Loos, R. Strasser, D. Maresch, T. Rademacher, H. Steinkellner, L. Mach, The human anti-HIV antibodies 2F5, 2G12, and PG9 differ in their susceptibility to proteolytic degradation: down-regulation of endogenous serine and cysteine proteinase activities could improve antibody production in plant-based expression platforms, *Biotechnol. J.* 9 (2014) 493–500.
- [13] V.K. Hehle, R. Lombardi, C.J. van Dolleweerd, M.J. Paul, P. Di Micco, V. Morea, E. Benvenuto, M. Donini, J.K. Ma, Site-specific proteolytic degradation of IgG monoclonal antibodies expressed in tobacco plants, *Plant Biotechnol. J.* 13 (2015) 235–245.
- [14] K. Duwadi, L. Chen, R. Menassa, S. Dhaubhadel, Identification, characterization

- and down-regulation of cysteine protease genes in tobacco for use in recombinant protein production, *PLoS One* 10 (2015) e0130556.
- [15] M. Martinez, I. Cambra, P. Gonzalez-Melendi, M.E. Santamaria, I. Diaz, C1A cysteine-proteases and their inhibitors in plants, *Physiol. Plant* 145 (2012) 85–94.
- [16] V. Turk, V. Stoka, O. Vasiljeva, M. Renko, T. Sun, B. Turk, D. Turk, Cysteine cathepsins: from structure, function and regulation to new frontiers, *Biochim. Biophys. Acta* 1824 (2012) 68–88.
- [17] K.H. Richau, F. Kaschani, M. Verdoes, T.C. Pansuriya, S. Niessen, K. Stuber, T. Colby, H.S. Overkleef, M. Bogyo, R.A. Van der Hoorn, Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics, *Plant Physiol.* 158 (2012) 1583–1599.
- [18] H. Lu, B. Chandrasekar, J. Oeljeklaus, J.C. Misa-Villamil, Z. Wang, T. Shindo, M. Bogyo, M. Kaiser, R.A. van der Hoorn, Subfamily-specific fluorescent probes for cysteine proteases display dynamic protease activities during seed germination, *Plant Physiol.* 168 (2015) 1462–1475.
- [19] E.M. Gilroy, I. Hein, R. van der Hoorn, P.C. Boevink, E. Venter, H. McLellan, F. Kaffarnik, K. Hrubikova, J. Shaw, M. Holeva, E.C. Lopez, O. Borrás-Hidalgo, L. Pritchard, G.J. Loake, C. Lacomme, P.R. Birch, Involvement of cathepsin B in the plant disease resistance hypersensitive response, *Plant J.* 52 (2007) 1–13.
- [20] C. Gu, M. Shabab, R. Strasser, P.J. Wolters, T. Shindo, M. Niemer, F. Kaschani, L. Mach, R.A. van der Hoorn, Post-translational regulation and trafficking of the granulin-containing protease RD21 of *Arabidopsis thaliana*, *PLoS One* 7 (2012) e32422.
- [21] M. Niemer, U. Mehofer, M. Verdianz, A. Porodko, P. Schahs, D. Kracher, B. Lenarcic, M. Novinec, L. Mach, *Nicotiana benthamiana* cathepsin B displays distinct enzymatic features which differ from its human relative and aleurain-like protease, *Biochimie* 122 (2016) 119–125.
- [22] F. Kaschani, M. Shabab, T. Bozkurt, T. Shindo, S. Schornack, C. Gu, M. Ilyas, J. Win, S. Kamoun, R.A. van der Hoorn, An effector-targeted protease contributes to defense against *Phytophthora infestans* and is under diversifying selection in natural hosts, *Plant Physiol.* 154 (2010) 1794–1804.
- [23] T.O. Bozkurt, S. Schornack, J. Win, T. Shindo, M. Ilyas, R. Oliva, L.M. Cano, A.M. Jones, E. Huitema, R.A. van der Hoorn, S. Kamoun, *Phytophthora infestans* effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 20832–20837.
- [24] E. Kowalewski-Nimmerfall, P. Schahs, D. Maresch, D. Rendic, H. Kramer, L. Mach, *Drosophila melanogaster* cellular repressor of E1A-stimulated genes is a lysosomal protein essential for fly development, *Biochim. Biophys. Acta* 1843 (2014) 2900–2912.
- [25] P. Schähns, P. Weidinger, O.C. Probst, B. Svoboda, J. Stadlmann, H. Beug, T. Waerner, L. Mach, Cellular repressor of E1A-stimulated genes is a bona fide lysosomal protein which undergoes proteolytic maturation during its biosynthesis, *Exp. Cell Res.* 314 (2008) 3036–3047.
- [26] M. Novinec, M. Pavsic, B. Lenarcic, A simple and efficient protocol for the production of recombinant cathepsin V and other cysteine cathepsins in soluble form in *Escherichia coli*, *Protein Expr. Purif.* 82 (2012) 1–5.
- [27] A.J. Barrett, H. Kirschke, Cathepsin B, cathepsin H, and cathepsin L, *Methods Enzymol.* 80 (1981) 535–561.
- [28] L. Mach, J.S. Mort, J. Glossl, Maturation of human procathepsin B. Proenzyme activation and proteolytic processing of the precursor to the mature proteinase, *in vitro*, are primarily unimolecular processes, *J. Biol. Chem.* 269 (1994) 13030–13035.
- [29] C.G. Knight, Active-site titration of peptidases, *Methods Enzymol.* 248 (1995) 85–101.
- [30] O. Schilling, C.M. Overall, Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites, *Nat. Biotechnol.* 26 (2008) 685–694.
- [31] O. Schilling, P.F. Huesgen, O. Barre, U. Auf dem Keller, C.M. Overall, Characterization of the prime and non-prime active site specificities of proteases by proteome-derived peptide libraries and tandem mass spectrometry, *Nat. Protoc.* 6 (2011) 111–120.
- [32] M.L. Biniossek, D.K. Nagler, C. Becker-Pauly, O. Schilling, Proteomic identification of protease cleavage sites characterizes prime and non-prime specificity of cysteine cathepsins B, L, and S, *J. Proteome Res.* 10 (2011) 5363–5373.
- [33] M.L. Biniossek, M. Niemer, K. Maksimchuk, B. Mayer, J. Fuchs, P.F. Huesgen, D.G. McCafferty, B. Turk, G. Fritz, J. Mayer, G. Haecker, L. Mach, O. Schilling, Identification of protease specificity by combining proteome-derived peptide libraries and quantitative proteomics, *Mol. Cell. Proteom.* 15 (2016), <http://dx.doi.org/10.1074/mcp.O115.056671> (in press).
- [34] H. Shahinian, D. Loessner, M.L. Biniossek, J.N. Kizhakkedathu, J.A. Clements, V. Magdolen, O. Schilling, Secretome and degradome profiling shows that kallikrein-related peptidases 4, 5, 6, and 7 induce TGFβ1 signaling in ovarian cancer cells, *Mol. Oncol.* 8 (2014) 68–82.
- [35] S. Guo, W. Skala, V. Magdolen, P. Briza, M.L. Biniossek, O. Schilling, J. Kellermann, H. Brandstetter, P. Goettig, A single glycan at the 99-loop of human kallikrein-related peptidase 2 regulates activation and enzymatic activity, *J. Biol. Chem.* 291 (2016) 593–604.
- [36] N. Colaert, K. Hensens, L. Martens, J. Vandekerckhove, K. Gevaert, Improved visualization of protein consensus sequences by icelogo, *Nat. Methods* 6 (2009) 786–787.
- [37] D. Greenbaum, K.F. Medzihradsky, A. Burlingame, M. Bogyo, Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools, *Chem. Biol.* 7 (2000) 569–581.
- [38] N. Schaschke, I. Assfalg-Machleidt, T. Lassleben, C.P. Sommerhoff, L. Moroder, W. Machleidt, Epoxysuccinyl peptide-derived affinity labels for cathepsin B, *FEBS Lett.* 482 (2000) 91–96.
- [39] G. Lalmanach, R. Mayer, C. Serveau, J. Scharfstein, F. Gauthier, Biotin-labelled peptidyl diazomethane inhibitors derived from the substrate-like sequence of cystatin: targeting of the active site of cruzipain, the major cysteine proteinase of *Trypanosoma cruzi*, *Biochem. J.* 318 (1996) 395–399.
- [40] K. Yamada, R. Matsushima, M. Nishimura, I. Hara-Nishimura, A slow maturation of a cysteine protease with a granulin domain in the vacuoles of senescing *Arabidopsis* leaves, *Plant Physiol.* 127 (2001) 1626–1634.
- [41] T. Vernet, H.E. Khouri, P. Laflamme, D.C. Tessier, R. Musil, B.J. Gour-Salin, A.C. Storer, D.Y. Thomas, Processing of the papain precursor: purification of the zymogen and characterization of its mechanism of processing, *J. Biol. Chem.* 266 (1991) 21451–21457.
- [42] A.D. Rowan, P. Mason, L. Mach, J.S. Mort, Rat procathepsin B. Proteolytic processing to the mature form *in vitro*, *J. Biol. Chem.* 267 (1992) 15993–15999.
- [43] L. Mach, H. Schwihla, K. Stuwe, A.D. Rowan, J.S. Mort, J. Glossl, Activation of procathepsin B in human hepatoma cells: the conversion into the mature enzyme relies on the action of cathepsin B itself, *Biochem. J.* 293 (1993) 437–442.
- [44] P. Sosnowski, D. Turk, Caught in the act: the crystal structure of cleaved cathepsin L bound to the active site of cathepsin L, *FEBS Lett.* 590 (2016) 1253–1261.
- [45] B. Yan, J. Valliere-Douglass, L. Brady, S. Steen, M. Han, D. Pace, S. Elliott, Z. Yates, Y. Han, A. Bolland, W. Wang, D. Pettit, Analysis of post-translational modifications in recombinant monoclonal antibody IgG1 by reversed-phase liquid chromatography/mass spectrometry, *J. Chromatogr. A* 1164 (2007) 153–161.
- [46] D. Turk, G. Guncar, M. Podobnik, B. Turk, Revised definition of substrate binding sites of papain-like cysteine proteases, *Biol. Chem.* 379 (1998) 137–147.
- [47] Y. Choe, F. Leonetti, D.C. Greenbaum, F. Lecaille, M. Bogyo, D. Bromme, J.A. Ellman, C.S. Craik, Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities, *J. Biol. Chem.* 281 (2006) 12824–12832.
- [48] B.C. Holwerda, J.C. Rogers, Purification and characterization of aleurain: a plant thiol protease functionally homologous to mammalian cathepsin H, *Plant Physiol.* 99 (1992) 848–855.
- [49] G. Guncar, M. Podobnik, J. Pungercar, B. Strukelj, V. Turk, D. Turk, Crystal structure of porcine cathepsin H determined at 2.1 Å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function, *Structure* 6 (1998) 51–61.
- [50] T. Takahashi, A.H. Dehdarani, J. Tang, Porcine spleen cathepsin H hydrolyzes oligopeptides solely by aminopeptidase activity, *J. Biol. Chem.* 263 (1988) 10952–10957.
- [51] H. Angliker, P. Wikstrom, H. Kirschke, E. Shaw, The inactivation of the cysteinyl exopeptidases cathepsin H and C by affinity-labelling reagents, *Biochem. J.* 262 (1989) 63–68.
- [52] R.A. van der Hoorn, M.A. Leeuwenburgh, M. Bogyo, M.H. Joosten, S.C. Peck, Activity profiling of papain-like cysteine proteases in plants, *Plant Physiol.* 135 (2004) 1170–1178.
- [53] K.I. Varughese, F.R. Ahmed, P.R. Carey, S. Hasnain, C.P. Huber, A.C. Storer, Crystal structure of a papain-E-64 complex, *Biochemistry* 28 (1989) 1330–1332.
- [54] G. Marino, P.F. Huesgen, U. Eckhard, C.M. Overall, W.P. Schroder, C. Funk, Family-wide characterization of matrix metalloproteinases from *Arabidopsis thaliana* reveals their distinct proteolytic activity and cleavage site specificity, *Biochem. J.* 457 (2014) 335–346.
- [55] O. Vasiljeva, M. Dolinar, V. Turk, B. Turk, Recombinant human cathepsin H lacking the mini chain is an endopeptidase, *Biochemistry* 42 (2003) 13522–13528.
- [56] X.Q. Xin, B. Gunesekeera, R.W. Mason, The specificity and elastolytic activities of bovine cathepsins S and H, *Arch. Biochem. Biophys.* 299 (1992) 334–339.
- [57] H.E. Khouri, T. Vernet, R. Menard, F. Parlati, P. Laflamme, D.C. Tessier, B. Gour-Salin, D.Y. Thomas, A.C. Storer, Engineering of papain: selective alteration of substrate specificity by site-directed mutagenesis, *Biochemistry* 30 (1991) 8929–8936.
- [58] J. Dodt, J. Reichwein, Human cathepsin H: deletion of the mini-chain switches substrate specificity from aminopeptidase to endopeptidase, *Biol. Chem.* 384 (2003) 1327–1332.
- [59] R. Menard, H.E. Khouri, C. Plouffe, P. Laflamme, R. Dupras, T. Vernet, D.C. Tessier, D.Y. Thomas, A.C. Storer, Importance of hydrogen-bonding interactions involving the side chain of Asp158 in the catalytic mechanism of papain, *Biochemistry* 30 (1991) 5531–5538.
- [60] R. Menard, H.E. Khouri, C. Plouffe, R. Dupras, D. Ripoll, T. Vernet, D.C. Tessier, F. Lalberte, D.Y. Thomas, A.C. Storer, A protein engineering study of the role of aspartate 158 in the catalytic mechanism of papain, *Biochemistry* 29 (1990) 6706–6713.