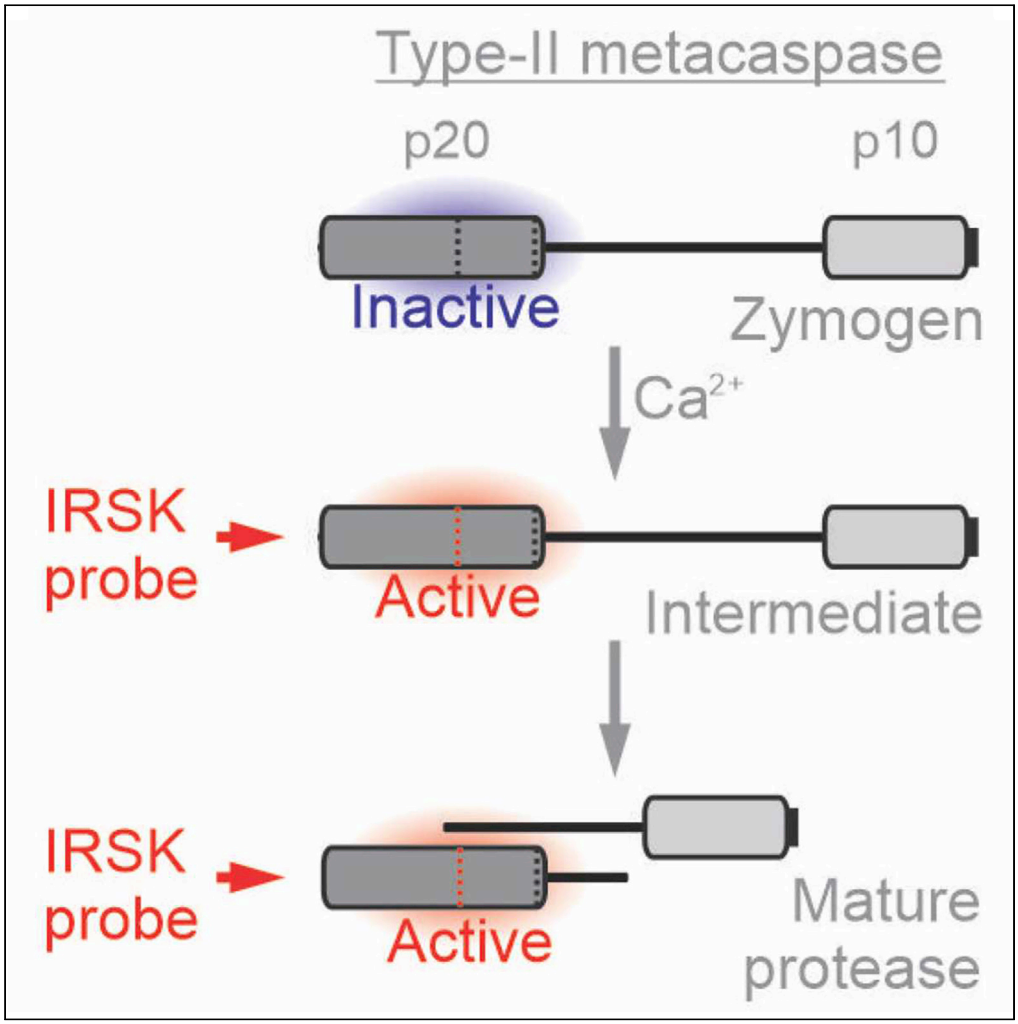
iScience

CellPress
OPEN ACCESS

Article

Activity-based probes trap early active intermediates during metacaspase activation



Vida Štrancar, Katarina P. van Midden, Daniel Krahn, ..., Christopher J. Schofield, Marina Klemenčič, Renier A.L. van der Hoorn

renier.vanderhoorn@biology.ox.ac.uk

Highlights

Four tetrapeptide acyloxymethylketone designed to label metacaspases

All four probes label purified metacaspases representing classes I, II, and III

Labeling and autocatalytic processing requires calcium

Highest affinity IRSK probe traps early intermediates during activation

Štrancar et al., iScience 25, 105247
November 18, 2022 © 2022
https://doi.org/10.1016/j.isci.2022.105247

iScience

CellPress
OPEN ACCESS

Article

Activity-based probes trap early active intermediates during metacaspase activation

Vida Štrancar,[1,8] Katarina P. van Midden,[1,8] Daniel Krahn,[2,3,7] Kyoko Morimoto,[2] Marko Novinec,[1] Christiane Funk,[4] Simon Stael,[5,6] Christopher J. Schofield,[3] Marina Klemenčič,[1,4] and Renier A.L. van der Hoorn[2,9,*]

SUMMARY

Metacaspases are essential cysteine proteases present in plants, fungi, and protists that are regulated by calcium binding and proteolytic maturation through mechanisms not yet understood. Here, we developed and validated activity-based probes for the three main metacaspase types, and used them to study calcium-mediated activation of metacaspases from their precursors *in vitro*. By combining substrate-inspired tetrapeptide probes containing an acyloxymethylketone (AOMK) reactive group, with purified representatives of type-I, type-II, and type-III metacaspases, we were able to demonstrate that labeling of mature metacaspases is strictly dependent on calcium. The probe with the highest affinity for all metacaspases also labels higher molecular weight proteoforms of all three metacaspases only in the presence of calcium, displaying the active, unprocessed metacaspase intermediates. Our data suggest that metacaspase activation proceeds through previously unknown active intermediates that are formed upon calcium binding, before precursor processing.

INTRODUCTION

Metacaspases are multifunctional proteases that are essential for the normal physiology of non-metazoan organisms (Tsiatsiani et al., 2011). Metacaspases are nucleophilic cysteine proteases that, together with orthocaspases and paracaspases, form the C14B protease family, which are structurally homologous to the well-characterized caspases of the C14A family (Minina et al., 2020; Uren et al., 2000). Whereas C14A proteases cleave after acidic Asp residues, C14B proteases cleave after basic Arg/Lys residues. Consequently, these proteases have distinct substrates and require different tools to study. Family C14 proteases all contain a catalytic dyad of His and Cys residues located in the catalytic p20 domain, named after its approximate molecular weight. In addition to the p20 catalytic domain, metacaspases also contain a regulatory p10 domain that harbors the amino acid residues involved in the coordination of calcium ions, which activate these proteases (Klemenčič and Funk, 2018). This p10 domain is absent in orthocaspases, which do not require calcium ions for activity (Klemenčič et al., 2015), and paracaspases, which are activated by dimerization (Yu et al., 2011).

Based on the position of their p20 and p10 domains metacaspases are further classified into type-I (p20-p10), type-II (p20-linker-p10), and type-III metacaspases (p10-p20). Only type-I metacaspases are present in eukaryotes as well as in prokaryotes (archaea and bacteria) (Klemenčič et al., 2019). Eukaryotic type-I metacaspases can contain an additional *N*-terminal pro-domain (Choi and Berges, 2013); those in land plants sometimes contain a zinc finger domain (zf) within the *N*-terminal pro-domain, a motif that is also found in the negative cell-death regulator LSD1 (Lesion Simulating Disease-1, Lam and Zhang, 2012). For activity, type-I metacaspases require micromolar concentrations of calcium ions (Gilio et al., 2017; Machado et al., 2013; van Midden et al., 2021) as well as usually neutral or slightly acidic pH. Type-II metacaspases are present only in green lineage photosynthetic organisms. In Chlorophyta, they are present in Chlamydomonadales and have up to now been identified in all Embriophyta (Klemenčič and Fun k, 2019). Type-II metacaspases do not contain an inhibitory prodomain. Instead, they have a linker region connecting the p20 and p10 domains that obstructs the active site and requires cleavage at a conserved Arg/Lys residue to activate the protease (Fortin and Lam, 2018; Vercammen et al., 2004; Zhu et al., 2020). Type-II metacaspases also require higher calcium ion concentrations for activity (millimolar range) and have

[1]Department of Chemistry and Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, Slovenia

[2]The Plant Chemetics Laboratory, Department of Plant Sciences, University of Oxford, South Park Road, Oxford OX1 3RB, UK

[3]Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute or Antimicrobial Research, University of Oxford, Mansfield Road, Oxford OX1 3TA, UK

[4]Department of Chemistry, Umeå University, Umeå, Sweden

[5]Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium

[6]VIB-UGent Center for Plant Systems Biology, Ghent, Belgium

[7]Present address: Leibniz-Institut für Analytische Wissenschaften –ISAS –e.V. Dortmund, Germany

[8]These authors contributed equally

[9]Lead contact

*Correspondence: renier.vanderhoorn@biology.ox.ac.uk

https://doi.org/10.1016/j.isci.2022.105247

optimal activity at basic pH. Type-III metacaspases were discovered most recently and are present only in unicellular eukaryotes that emerged by secondary endosymbiosis (Choi and Berges, 2013). Their catalytic properties are similar to those of type-I metacaspases, despite the p20-p10 domain swap (Graff van Creveld et al., 2021; Klemenčič and Funk, 2018). Type-III metacaspases can also contain an *N*-terminal prodomain, and its removal increases their proteolytic efficiency (Klemenčič and Funk, 2018). *Pt*MCA-III from the model diatom *Phaeodactylum tricornutum* is redox-regulated (Graff van Creveld et al., 2021).

The existence of various processed intermediate proteoforms of metacaspases, and the regulation of their activity by calcium ions, causes challenges in the identification of active metacaspase proteoforms (Hander et al., 2019). To address this problem, we aimed to develop activity-based probes for metacaspases that covalently react with the active site in an activity-dependent manner, thereby distinguishing active metacaspases from inactive proteoforms. Activity-based probes carrying acyloxymethylketone (AOMK) warheads linked to tetrapeptides have been very useful to display the active proteoforms of other family C14 proteases. AOMK probes with the DEVD tetrapeptides have been used to discover that caspases are activated via early intermediates (Berger et al., 2006), and AOMK probes with tetrapeptides LRSR and LVSR have been used to characterize the MALT1 paracaspase (Eitelhuber et al., 2015; Hachmann et al., 2015). Most of these probes are not suitable for metacaspases because these enzymes are selective for peptides carrying a Lys or Arg residue at the position just before the cleavable bond (P1 position) and prefer a basic residue at the P3 position (Vercammen et al., 2006), as supported by proteomic analysis of substrates of type-II metacaspase *At*MCA-IIf (*At*MC9) of *Arabidopsis thaliana* (Tsiatsiani et al., 2013).

Earlier work with purified *At*MCA-IIa (*At*MCA2d/*At*MC4) revealed that this type-II metacaspase can be labeled with a biotinylated tripeptide chloromethylketone Bio-FPR-cmk (Watanabe and Lam, 2011), which lacks a basic residue at the P3 position as it was originally designed to label serine protease thrombin (Williams et al., 1989). Here, we developed fluorescent high-affinity tetrapeptidyl probes that can also label type-I and type-II metacaspases, and used these probes to study the activation mechanism of purified metacaspases representing the three different types in the presence of calcium ions.

RESULTS AND DISCUSSION

Choice and synthesis of AOMK-based tetrapeptidyl probes

To synthesize activity-based probes for metacaspases, we took advantage of the Cys-reactive acyloxymethylketone (AOMK) warhead because this functional group enables the inclusion of a K as the P1 residue. Although metacaspases prefer P1 = K or R, we could only synthesize probes with P1 = K because of complications in R-AOMK synthesis caused by low yield and partial degradation of the probe upon elution from the column. To increase the affinity for metacaspases further, the probes included a tripeptide targeting sequence, based on substrate preferences that were reported at the time of probe synthesis. IISK, ILSK, and IRSK were the three top substrates for *At*MCA-IIf based on the tetrapeptide library study (Vercammen et al., 2006). As IISK and ILSK are rather similar in their P3 position (I vs. L), we chose to produce probes with IISK and IRSK tetrapeptides. In addition, EKTK and EKAK were the preferred amino acid sequences preceding the cleavage sites in potential *At*MCA-IIf substrate proteins based on an *N*-terminomics study (Tsiatsiani et al., 2013). These four tetrapeptidyl probes were synthesized by Solid Phase Peptide Synthesis (SPPS) and carry an alkyne minitag (Figure 1A). Labeling metacaspases with AOMK probes will result in the formation of a stable thioether bond between the probe and the catalytic Cys residue (Figure 1B). Probe-labeled proteins carry an alkyne minitag that can be coupled to azide-picolyl-Cy5 fluorophore via click chemistry (Figure 1C), facilitating the detection of fluorescent conjugates on protein gels.

GtMCA-I, CrMCA-II, and GtMCA-III are canonical representatives of the three metacaspase types

To determine if the synthesized probes can label metacaspases, we first selected one representative of each of the three (I/II/III) metacaspase types from unicellular algae, which we use for studies on programmed cell death. The main criterion for selection was the possibility to express these proteases recombinantly in *E. coli* as soluble, full-length *C*-terminally His-tagged proteins. For this reason, we selected metacaspase representatives from two unicellular photosynthetic organisms: the model green alga *Chlamydomonas reinhardtii* and a secondary endosymbiont *Guillardia theta*, which was amongst Cryptophyta the first organism with sequenced nuclear genome (Curtis et al., 2012). From the latter, we amplified genes for the type-I metacaspase (*Gt*MCA-I) and the previously characterized *Gt*MCA-III (before *Gt*MC2; Klemenčič and Funk, 2018). As Cryptophyta lack the type II metacaspases, we have chosen to use the

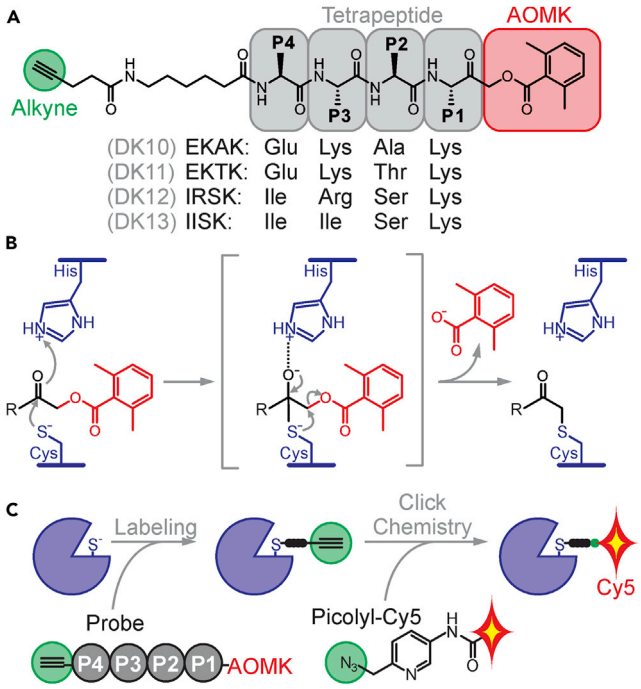


Figure 1. AOMK-based tetrapeptidyl probes to target metacaspases

(A) Structures of the four AOMK probes designed for metacaspases. In addition to the common alkyne minitag (green) and the *C*-terminal AOMK warhead (red), all probes contain a tetrapeptide (grey) with a Lys residue at the P1 position and various residues at the P2, P3, and P4 positions, as summarized below the structure. The P1 residue in a substrate locates just before the cleavable bond. DK numbers are alternative probe names

(B) AOMK labeling mechanism. AOMK probes covalently label Cys protease (blue) by forming a covalent thiomethyl ketone bond with the active site Cys

(C) Two-step labeling of metacaspases. In the first step, the active site Cys of metacaspases is covalently labeled with the AOMK probe, resulting in an alkyne-labeled protease. In the second step, the alkyne is coupled to an azide-picolyl-Cy5 fluorophore via click chemistry, under denaturing conditions

only type II metacaspase present in *C. reinhardtii*, *Cr*MCA-II (Figure 2A). All proteins could be expressed as soluble proteins in *E. coli* with apparent molecular weights (MWs) corresponding to their predicted MWs (Figure 2B, see Figure S1 for full gels with markers).

As *Gt*MCA-I and *Cr*MCA-II have not yet been previously biochemically characterized, we first determined their optimal proteolytic conditions by monitoring their activity at different pH values in the presence of calcium (5 mM) and subsequently determined their activity at various calcium concentrations at their optimal pH. Although *Gt*MCA-III had been previously characterized (Klemenčič and Funk, 2018), this protease was included in parallel. For all three metacaspase types, the fluorescent dipeptide Z-FR-AMC was chosen as a substrate, as it had a comparable affinity with all three metacaspase types in comparison with Z-RR-AMC (Figure S2).

*Gt*MCA-I exhibited the highest proteolytic activity at neutral to slightly acidic pH (optimum at pH 6.5, Figure 3A), which is similar to other characterized type-I metacaspases (He et al., 2008; Tsiatsiani et al., 2011). Furthermore, *Gt*MCA-I's proteolytic activity is strongly calcium-dependent, as no fluorescent product was detected at calcium concentrations lower than 10 μM (Figure 3B). The optimal calcium concentration was reached at a low millimolar range and noticeably decreased above 30 mM. This phenomenon has also been observed for three *Schizophyllum commune* type-I metacaspases, where calcium concentrations higher than 25 mM result in a remarkable decrease in proteolytic activity (Leang et al., 2019), and for *Tb*MCA-Ib (originally *Tb*MCA2) of *Trypanosoma brucei*, where activity increases sharply up to ~500 μM of $CaCl_2$ but decreases above these concentrations (Machado et al., 2013).

The activation of *Gt*MCA-I at pH 6.5- and 5-mM calcium is associated with processing, resulting in a protein with slightly reduced apparent MW, as detected on the SDS gel (Figure 3C, original images in Figure S3).

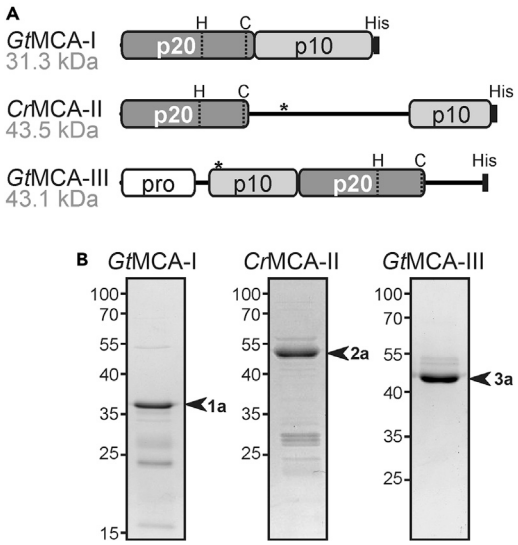


Figure 2. The three metacaspase types purified as soluble proteins

(A) Schematic domain organization of type-I metacaspase *Gt*MCA-I, type-II metacaspase *Cr*MCA-II, and type-III metacaspase *Gt*MCA-III used in this study. The p20-like domain (dark gray) contains the catalytic His and Cys residues. Also indicated are the regulatory p10 domain (light gray), the *C*-terminal His-tag (black), the *N*-terminal prodomain (white), putative cleavage sites (*), and the predicted MW of the unprocessed proteins.

(B) Purified metacaspases, separated on a 12% SDS-PAGE gel under reducing conditions. *C*-terminally His-tagged metacaspases were expressed in *E. coli* and purified by IMAC. The arrows show the positions of full-length proteins. Left: apparent MW of marker proteins in kDa.

*Tb*MCA-Ib is also processed in the presence of calcium at two lysine residues (Lys55 and Lys268), which are located in the *N*-terminal domain, and the so-called p280-loop within the p10 domain, respectively (McLuskey et al., 2012). However, as *Gt*MCA-I, unlike *Tb*MCA-Ib, lacks the *N*-terminal domain, we presumed that the cleavage occurred at the *C*-terminus of the protein. This was further confirmed using anti-His antibodies, which failed to immunostain the calcium-activated *Gt*MCA-I (Figures S4 and S5). As the cleaved *Tb*MCA-Ib fragments are still connected by non-covalent bonds and there are no data that such autoprocessing occurs within the p20-p10 core of other type-I metacaspases *in vitro* or *in vivo* (Coll et al., 2010, 2014; Laverrière et al., 2012; Moss et al., 2007; van Midden et al., 2021), we hypothesize that the observed *C*-terminal processing of *Gt*MCA-I has no physiological relevance.

The proteolytic nature of *Cr*MCA-II is very different from that of *Gr*MCA-I. *Cr*MCA-II has maximal proteolytic activity at alkaline pH, with an optimum pH of 8.5 (Figure 3B). At optimal pH, *Cr*MCA-II's activity is negligible at calcium concentrations below 1 mM, but *Cr*MCA-II is fully active in the presence of 10-mM calcium, as well as at higher calcium concentrations (Figure 3B). This behavior is typical of metacaspases similar to *At*MCA-IIa (originally *At*MC4) of *Arabidopsis thaliana*. These type-IIa-like metacaspases are present in all representatives of Viridiplantae, from Chlorophytae to Angiosperms (Klemenčič and Funk, 2019) and contain a large linker region between the catalytic p20 domain and the regulatory p10 domain, which is also present in *Cr*MCA-II (Figure 2A). This linker crosses the S1 pocket of the active site and thus acts as an autoinhibitory domain (Zhu et al., 2020). These type-IIa-like metacaspases are activated only after cleavage at a conserved Lys or Arg residue within the linker region. However, unlike type-I and type-III metacaspases, type-IIa-like metacaspases undergo fast autoprocessing once they are activated (Vercammen et al., 2004; Watanabe and Lam, 2011; Zhu et al., 2020). Likewise, we observed cleavage of the 55-kDa full-length *Cr*MCA-II into two major fragments migrating at 25 and 30 kDa when resolved by SDS-PAGE (Figure 3C).

The activity of *Gt*MCA-III is optimal at slightly acidic pH (Figure 3A), similar to the closely related type-I metacaspases (Choi and Berges, 2013). However, *Gt*MCA-III's activity requires high concentrations of calcium (1 mM), and in contrast to type-I metacaspases, its activity is not inhibited at higher calcium concentrations (Figure 3B). Similarly, also *Pt*MCA-IIIc requires high calcium concentrations for optimal proteolytic activity, and proteolytic activity is not inhibited at high calcium concentrations (Graff van Creveld et al., 2021; Klemenčič and Funk, 2018). *Gt*MCA-III is activated by cleaving the 45-kDa pro-enzyme into a 37-kDa protease as observed by SDS-PAGE (Figure 3C). These molecular weights are consistent with the removal of the *N*-terminal prodomain, whereas the remaining p10-p20 core structure does not undergo further processing (Klemenčič and Funk, 2018). An additional weak signal at 24 kDa also was observed in many experiments, which might represent the p20 domain, released upon processing between the p10 and p20 domains. But because this conversion is incomplete and can be suppressed by adding EDTA in the sample buffer before SDS-PAGE, the processing releasing this 24-kDa signal might be an artifact of sample preparation.

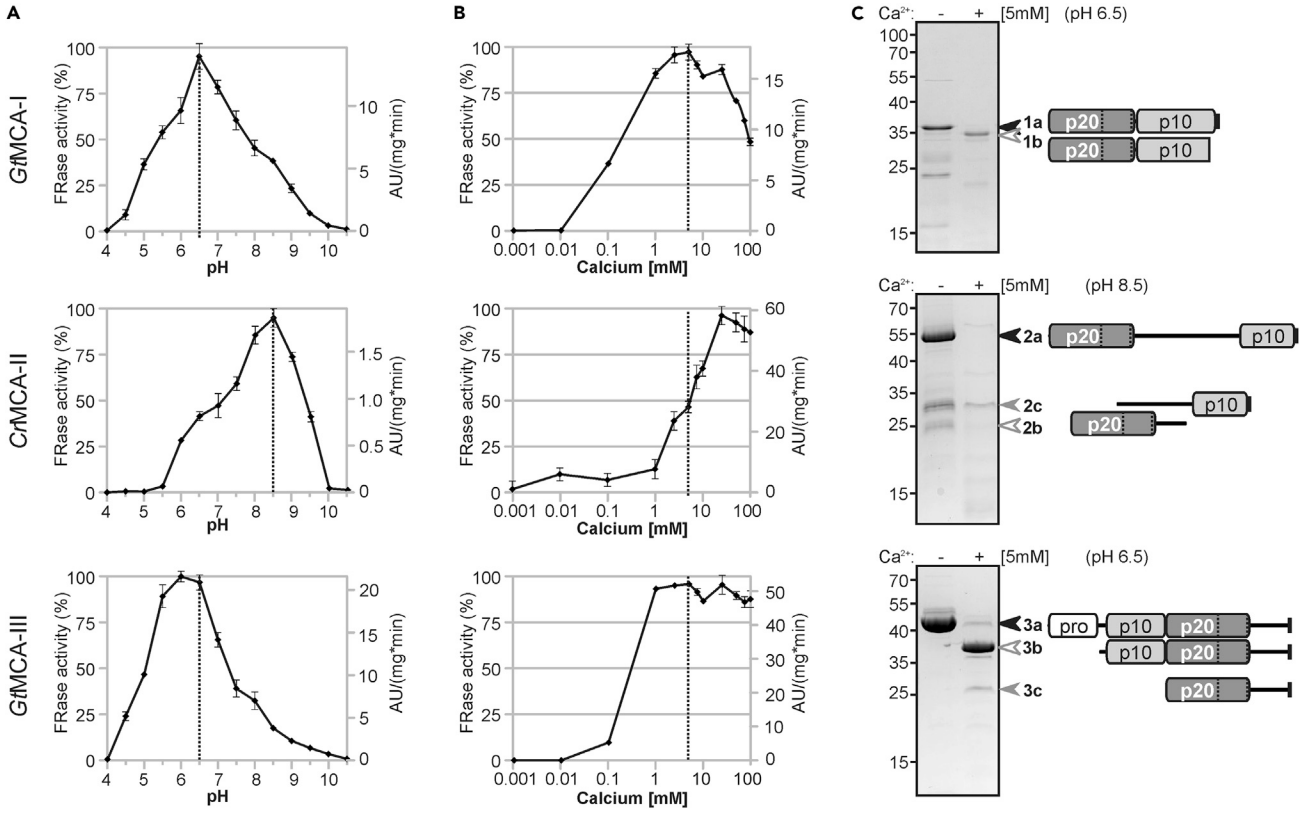


Figure 3. The three metacaspase types have different catalytic preferences

(A and B) Proteolytic activities were measured using Z-FR-AMC as the substrate at various pH in the presence of 5-mM $CaCl_2$ (A) or different calcium concentrations (B) at pH 6.5 (*Gt*MCA-I and *Gt*MCA-III) or pH 8.5 (*Cr*MCA-II). The dotted line represents the selected conditions used in this study. Both the relative activity (% of maximum witin experiment) and the total fluorescence (in arbitrary units (AU) per mg and per minute) were shown. Error bars represent standard deviations of n=3 replicates.

(C) Metacaspases are processed upon activation. Purified metacaspases were incubated with and without 5-mM $CaCl_2$ at 22°C for 2 h in buffers at the indicated pH values. Proteins were separated on 15% SDS-PAGE gels and stained with Coomassie. Proproteases, mature proteases, and the p20 domain of *Gt*MCA-III are indicated with black, white, and grey arrowheads, respectively. The presumed domain organizations of corresponding proteins are shown on the right. Left: apparent MW of marker proteins in kDa.

The three metacaspase types are proteolytically active in distinct proteoforms

To analyze labeling patterns of the three metacaspases with the four probes, metacaspases were activated at their optimal pH and calcium conditions, incubated with the four probes, coupled to azide-picolyl-Cy3 via click chemistry, and separated by SDS PAGE. All four probes successfully labeled the three metacaspases but uncovered different labeling patterns (Figure 4, originals and replicates are shown as Figures S6–S12). In the case of *Gt*MCA-I, two distinct bands were observed: a higher one with MW of 37 kDa, corresponding to a full-length protein, and the other with MW of 35 kDa, corresponding to the *C*-terminally processed *Gt*MCA-I (Figure 4A). The fact that both *Gt*MCA-I proteoforms were labeled by the probes indicates that *C*-terminal processing does not substantially affect *Gt*MCA-I activity. Interestingly, the labeling pattern of the various probes differed for *Gt*MCA-I, as seen by Coomassie stain. As the probes inactivate the metacaspase when they covalently bind to the Cys of the active site, the differences between the probes are likely caused by their different affinities for *Gt*MCA-I. This implies that *C*-terminal processing occurs intramolecularly and that IRSK is able to trap full-length *Gt*MCA-I before its *C*-terminus is processed.

Labeling of *Cr*MCA-II results in a single signal at MW of 25 kDa, corresponding to the processed catalytic p20 domain (Figure 4B). Coomassie staining identified also the 55-kDa full-length protein, but this proteoform was not fluorescently labeled with any of the probes. Activation of type-II metacaspases requires processing within the linker region to separate the p20 and p10 domains and to allow for conformational changes to the p20 and p10 domains to activate the protease (Zhu et al., 2020). The p20 and p10 domains

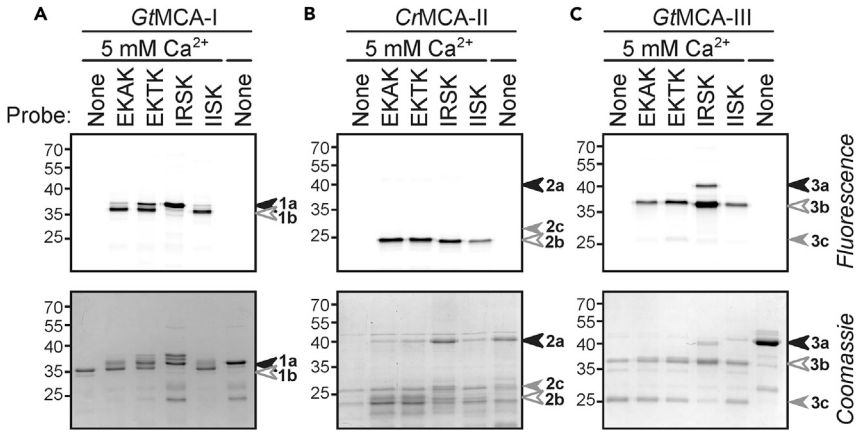


Figure 4. All four AOMK-based probes label all three metacaspase types

Purified metacaspases *Gt*MCA-1 (A), *Cr*MCA-II (B) and *Gt*MCA-III (C) were pre-incubated with and without 5-mM calcium ions in 150-mM NaCl and 5-mM DTT at pH 6.5 (20-mM MES for *Gt*MCA-I and *Gt*MCA-III) or pH 8.5 (20-mM HEPES for *Cr*MCA-II), and then incubated with and without 2-μM AOMK probes for 1 h. Samples were separated on SDS-PAGE gels under reducing conditions, scanned for fluorescence (top panels) and stained with Coomassie (bottom panels). Full-length proforms and mature proteases are indicated with black and white arrowheads, respectively. The grey arrowhead indicates the weakly labeled p20 domain of *Gt*MCA-III. Left: apparent MW of marker proteins in kDa.

remain in a complex upon processing because the catalytic activity of the p20 domain is still regulated by calcium binding to the p10 domain (Acosta-Maspons et al., 2014; Klemenčič and Funk, 2018; Zhu et al., 2020). Interestingly, we detect the accumulation of the full-length Coomassie-stained proteins in the presence of the probes, especially in the presence of IRSK, but this full-length proform is not fluorescently labeled (Figure 4B). The accumulation of the unlabeled proform in the presence of the probes indicates that inhibition of activated *Cr*MCA-II by the probes can block the processing of full-length *Cr*MCA-II. This implies that full-length *Cr*MCA-II is activated *in trans* by mature *Cr*MCA-II. Activation of type-II metacaspases by *in trans* cleavage is consistent with the observation that catalytically inactive *At*MCA-IIa can be cleaved by active *At*MCA-IIa, although this was less efficient compared with autocatalytic cleavage (Watanabe and Lam, 2011). This earlier study also demonstrated that the precursor of *At*MCA-IIa can be labeled with Bio-FPR-cmk, but only at high probe concentrations (100 μM; Watanabe and Lam, 2011). In contrast, we did not detect much precursor labeling of *Cr*MCA-II with 2-μM IRSK-AOMK, whilst the presence of the probe blocks precursor processing (Figure 4B). This mechanism is distinct from the activation of *Gt*MCA-I and -III, for which precursors are labeled with 2 μM IRSK-AOMK.

Labeling of *Gt*MCA-III resulted in a fluorescent signal at 37 kDa, consistent with all probes, corresponding to the mature protease (Figure 4C). Notably, labeling with IRSK caused an additional fluorescent signal at 42 kDa, which corresponded to the full-length protein. This observation was confirmed in the Coomassie-stained gel, only for the IRSK-labeled sample. Labeling of the full-length metacaspase confirmed a previous hypothesis that type-III metacaspases do not require removal of the *N*-terminal prodomain for their activity (Klemenčič and Funk, 2018). These data indicate that IRSK traps an early intermediate of *Gt*MCA-III activation. Interestingly, we also detect a stronger fluorescent signal at 37 kDa for IRSK when compared with the other three probes, and this fluorescent signal corresponds to a higher accumulation of this mature protease on the Coomassie-stained gel. The accumulation of the mature protease indicates that IRSK labeling stabilizes the mature protease by preventing its further processing to produce a 24-kDa protein that we detect both by fluorescence and by Coomassie analyses in all incubated samples, though less for IRSK. This observation implies that *Gt*MCA-III also exists as a processed proteoform, presumably cleaved between the p10 and p20 domains and that IRSK can prevent this cleavage.

Labeling of metacaspases is activity-dependent

To investigate if probes also label the metacaspase precursors, we performed labeling experiments on metacaspase precursors in the presence or absence of EDTA or calcium ions. In the absence of calcium ions, the metacaspases remain in their precursor proteoform, even without adding EDTA (Figure 5, original images in Figure S13–S15). No labeling was observed of these precursors with the IISK probe, whereas the

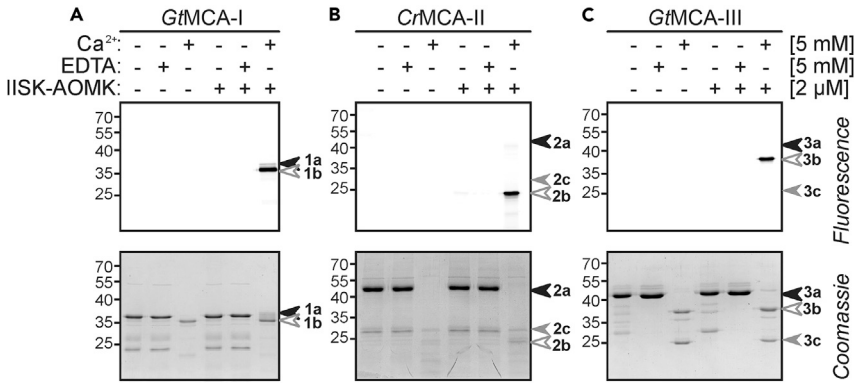


Figure 5. IRSK probe does not label metacaspase proforms

Purified metacaspases *Gt*MCA-1 (A), *Cr*MCA-II (B) and *Gt*MCA-III (C) were pre-incubated with and without 5-mM calcium ions or 5-mM EDTA at pH 6.5 (*Gt*MCA-I and *Gt*MCA-III) or pH 8.5 (*Cr*MCA-II) and then incubated with and without 2-µM IISK-AOMK for 1 h. Samples were separated on the SDS-PAGE under reducing conditions, scanned for fluorescence (top panels) and stained with Coomassie (bottom panels). Full-length proforms and mature proteases are indicated with black and white arrowheads, respectively. Left: apparent MW of marker proteins in kDa.

mature metacaspases, generated by the presence of calcium ions, were labeled (Figure 5). These data demonstrate that the IISK probe does not label metacaspase precursors, consistent with not being active proteases.

To test if labeling of mature metacaspase proteases is activity-dependent, we generated mature metacaspases by adding calcium ions, and then added an excess of EDTA to chelate calcium ions. Incubation of these samples with the IRSK probe did not result in substantial labeling, even though the mature proteases accumulate on the protein gel (Figure 6A, longer exposure shown in Figure S16). By contrast, strong labeling was observed when the calcium ions were not chelated with EDTA (Figure 6B, original images in Figure S17). These data demonstrate that labeling mainly occurs in the presence of calcium ions, indicating that calcium is required not only for autocatalytic activation but also for the activity of the mature p20-p10 enzyme.

Calcium-dependent labeling is consistent with the properties of these metacaspases. All three metacaspases contain eight conserved Asp residues that form two calcium-binding sites. A presumed high-affinity calcium-binding site is located in the p20 domain, whereas the second, presumed low-affinity calcium-binding site is formed by two Asp residues in the p20 and two Asp residues in the p10 domain (Figure S8) (Acosta-Maspons et al., 2014; Klemenčič and Funk, 2018; Machado et al., 2013;). The K_D values for the low-affinity binding site for *Tb*MCA-Ib (Machado et al., 2013) and *Gt*MCA-III (Klemenčič and Funk, 2018) are both in the low millimolar range. Calcium ions bound to both sites are thought to cause conformational changes that bring the p20 and p10 domains together, forming a functional activity site (van Midden et al., 2021). In the case of metacaspases that lack any peptide sequence covering the active site (e.g. *Gt*MCA-I) this renders the protease instantly active. The active site of the other metacaspases is blocked by the prodomain (*Gt*MCA-III) or the linker region (*Cr*MCA-II), and these sequences are proteolytically removed before the protease can process substrates. However, the fact that the IRSK probe labels an unprocessed, active intermediate of *Gt*MCA-II indicates that calcium-binding already causes dissociation of the inhibitory motif before it is proteolytically removed. This two-step activation mechanism is similar to that described for caspases (Berger et al., 2006).

IRSK probe is the best inhibitor for all three metacaspases

The ability of the IRSK probe to trap early intermediates in the activation of *Gt*MCA-I and *Gt*MCA-III and to cause the accumulation of zymogen precursor in *Cr*MCA-II indicates that this probe has a higher affinity. We therefore determined the inactivation constants (k_{ina}) of the four probes on the three metacaspases. In the absence of the probes, hydrolysis of the substrate resulted in linearly increasing fluorescence for all three metacaspases, suggesting a linear increase in product formation. However, in the presence of probes, product curves were hyperbolic in shape, consistent with a slow loss of enzyme activity owing to inactivation by the AOMK-based probes. This is characteristic of slow inactivation by the AOMK-based

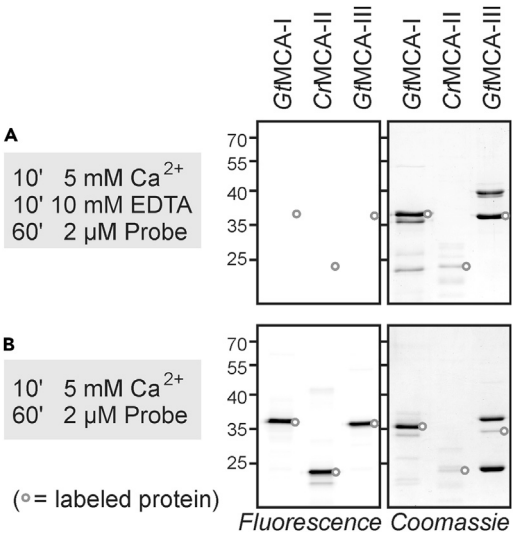


Figure 6. Labeling of mature metacaspases requires calcium ions

(A and B) Purified metacaspases were pre-incubated with 5-mM calcium ions for 10 min and then labeled for 1 h with 2 μM IISK-AOMK after incubation with (A) or without (B) 10-mM EDTA for 10 min. Alkyne-labeled proteins were coupled to the fluorophore via click chemistry. The samples were separated on the SDS-PAGE under reducing conditions, scanned for fluorescence (left panels) and stained with Coomassie (right panels). Mature metacaspases are indicated with grey circles. The fluorescent images are from one gel and one exposure. Left: apparent MW of marker proteins in kDa.

probes (Baici et al., 2009). The binding mechanism and inactivation rate was determined for each combination of probe and metacaspase and the corresponding plots of the pseudo-first-order rate constants determined from each progress curve versus probe concentration were plotted (Figure S19). The hyperbolic shape of the graph using *Gt*MCA-I suggests that the probe binding most likely follows a mechanism with two distinguishable steps, first forming a reversible non-covalent complex, followed by irreversible covalent bond formation resulting in inactivation. The equilibrium constant for the formation of the complex (Ki) and the first-order rate constant for the inhibition of the enzyme (k4) are shown in Figure S19. For *Cr*MCA-II and *Gt*MCA-III, the formation of a covalent bond between the enzyme and the probes was faster and occurred in one detectable binding step. From these datasets, we were able to determine the inactivation constants (k_{ina}) of each probe for each metacaspase. The higher the k_{ina} value, the more efficient probe for a chosen metacaspase representative.

For all three metacaspases, the IRSK probe showed the best inhibition properties with k_{ina} values of 3.27×10^5 M^{-1} s^{-1} for *Gt*MCA-I; 16.90×10^5 M^{-1} s^{-1} for *Cr*MCA-II; and 2.50×10^5 M^{-1} s^{-1} for *Gt*MCA-III (Figure 7). For comparison, Z-RFK-AOMK inhibits cathepsin L with k_{ina} 8.75×10^5 M^{-1} s^{-1} (Torkar et al., 2013) and Z-FR-AOMK inhibits cathepsin B with k_{ina} 1.1×10^5 M^{-1} s^{-1} (Yoon et al., 2021). Our IRSK probe has a similar or better inhibition constant and therefore emerges as a very powerful probe to study metacaspases of all three types and especially type-II metacaspases.

The other three AOMK probes were approximately 10-fold less effective in inhibiting the metacaspases when compared with the IRSK probe, with the exception of the IISK probe, which is a stronger inhibitor of *Cr*MCA-II when compared with the EKAK and EKTK probes. This suggests that different metacaspases exhibit distinct affinities for residues in the P2-P4 positions. Generally, all probes showed the highest inhibition of *Cr*MCA-II in comparison with the other two metacaspases. This higher binding affinity is consistent with the fact that the tetrapeptide sequences of the probes were determined based on the specificity scan of type-II metacaspase *At*MCA-IIf (Vercammen et al., 2006).

Both p20 and p10 domains bind the tetrapeptide to the metacaspase active site

To explain the observed high inactivation constant by the IRSK probe, we performed a molecular docking analysis of the IRSK tetrapeptide with a metacaspase. The MALT-1 paracaspase with the bound VRPR inhibitor was used as a modeling template (Yu et al., 2011), because it is the only available structure of a C14B protease in a catalytically active conformation. *Gt*MCA-I was chosen for modeling because it was the only of the three metacaspases that could be modeled reliably. This structure is very similar to that of yeast metacaspase Yca1, which does not contain a bound inhibitor (Wong et al., 2012). In all non-metazoan caspase homologues, the specificity pocket, which attracts positively charged P1 amino acid residues, is formed by one Ser and two highly conserved Asp residues (Figure 8A). These three residues are also present in *Gt*MCA-I (D26, S84, and D136) and the other two metacaspases (Figure S18) and form a deep, negatively charged S1 pocket, which explains the robust labeling of probes having P1 = Lys.

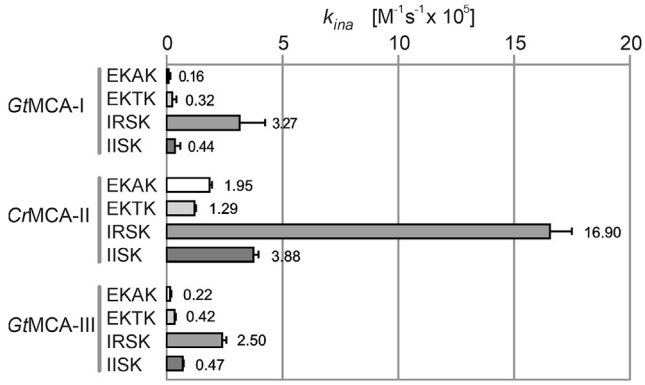


Figure 7. IRSK-AOMK is the best tested inhibitor for all three metacaspase types

Graphical comparison of k_{ina} values for the four probes for each of the three metacaspase representatives is shown, with mean values written next to the bars. The mean values and their standard deviation values were calculated from the plots shown in Figure S4.

IRSK has a 4-fold higher k_{ina} value compared with IISK, which must be caused by the positively charged P3 = Arg residue. This Arg residue can bind tightly in the S3 pocket, which is formed by the negatively charged Asp183 and Asp185 residues in the 280-loop of the p10 domain (Figure 8B). These residues are also present in *Cr*MCA-II and *Gt*MCA-III (Figure S18). The observation that metacaspases prefer P3 = Arg was also demonstrated by inhibition or cleavage of type-I and type-II metacaspases by Val-Arg-Pro-Arg (VRPR) containing inhibitors or substrates, respectively (McLuskey et al., 2012; Machado et al., 2013; Vercammen et al., 2006).

The EKAK and EKTK probes are less efficient in inhibiting metacaspases when compared with the IISK and IRSK probes, which is presumably caused by the fact that the P4 = Ile residue in the tetrapeptide fits better in the uncharged S4 pocket in the p10 domain, formed by Ser182, Gly184, and Val186 (Figure 8B). However, these residues are different in the other two tested metacaspases (Figure S19). *Cr*MCA-II, for instance, has Ala, Ala, and Gly at these positions, which might explain the improved labeling of *Cr*MCA-II by IISK.

The residue at the P2 position seems to be less relevant for metacaspases because we do not see distinctive differences between the EKAK and EKTK probes that carry P2 = Ala and Thr, respectively. The IRSK and IISK probes carry P2 = Ser but their high affinity is rather explained by having P4 = Ile. The less relevant role of

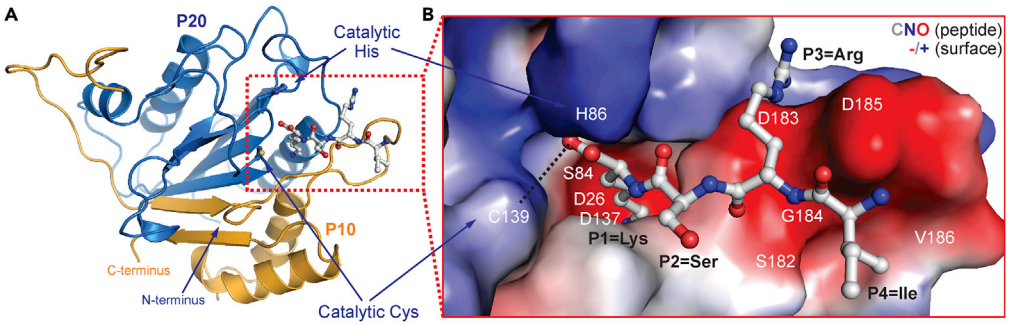


Figure 8. Simulation-based model of *Gt*MCA-I bound to tetrapeptide IRSK

(A) Overall structure of *Gt*MCA-I with the P10 (orange) and P20 (blue) domains as cartoon and the tetrapeptide (grey) in the stick mode.

(B) Enlarged substrate-binding groove of *Gt*MCA-I in electrostatic surface presentation and the tetrapeptide (grey) in the stick mode. The acidic P1- and P3-binding pockets contain aspartates (white letters) and the P4-binding pocket is hydrophobic. Labeling with IRSK-AOMK would result in a covalent bond (dashed line) between the catalytic C139 and the *C*-terminus of the peptide. The homology model of *Gt*MCA-I was built with AlphaFold. The tetrapeptide IRSK was docked into the active site with AutoDock Vina and the complex was equilibrated in a short 10-ns molecular dynamics simulation using NAMD3.

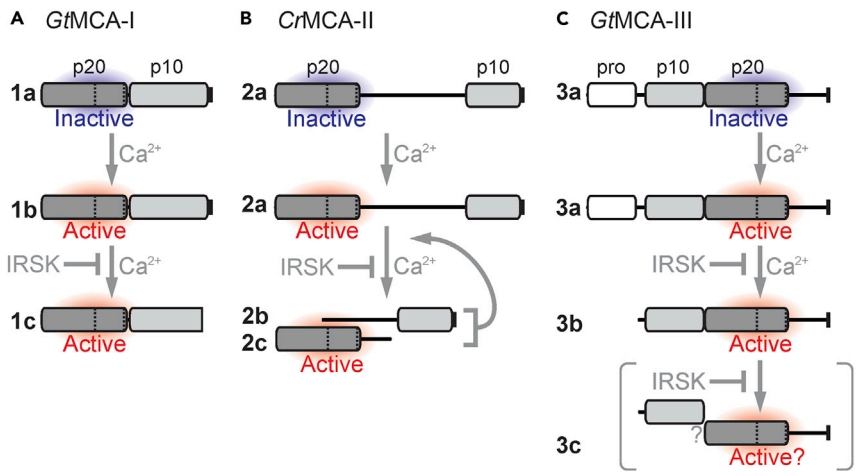


Figure 9. Model for the activation of type-I (A), type-II (B), and type-III (C) metacaspases

The proteoforms that can be labeled with the AOMK probes are highlighted in red. The zymogens cannot be labeled and have blue highlights. The IRSK probe traps intermediates because they can prevent autocatalytic processing by quickly inhibiting the metacaspases. The formation of the 24-kDa proteoform of *Gt*MCA-III (3c) might be an artifact of sample preparation.

the residue at the P2 position is consistent with the modelling, which indicates that this P2 residue points away from the enzyme (Figure 8B).

Whilst our AOMK probes label representatives of all three metacaspase classes, there are several potential limitations of these probes. First, it remains to be tested if these probes also label metacaspases that may have a very different substrate specificity. Second, substrate specificity and metacaspase maturation might be influenced by pH, which was fixed in our experiments but can be an important parameter influencing metacaspase labeling. Third, some of the processing events we observed in these metacaspases, such as *C*-terminal processing of *Gt*MCA-I, might not be a class-specific feature but rather be specific to the studied enzyme. Finally, we have only shown that these probes specifically label the active proteoforms of purified metacaspases. Profiling endogenous metacaspases is a distinct challenge that we aim to address in future work.

Conclusion

We have established activity-dependent labeling of representatives of the three major types of metacaspases and discovered that the fastest labeling tetrapeptide (IRSK) traps active intermediate proteoforms of all metacaspases, uncovering striking differences between activation mechanisms of the three metacaspase types (Figure 9). Type-I metacaspase *Gt*MCA-I is activated by calcium at pH 6.5 and processes itself at the *C*-terminus, which is without a known relevance (Figure 9A). Type-II metacaspase *Cr*MCA-II is activated by calcium at pH 8.5 by cleaving the linker region *in trans*, by its mature proteoform (Figure 9B). And type-III metacaspase *Gt*MCA-III is activated by calcium at pH 6.5 and can already be active when still carrying its prodomain. Removal of the prodomain of *Gt*MCA-III results in a mature protease that is processed further, presumably by cleaving between the p10 and p20 domains (Figure 9C). The high-affinity binding of the tetrapeptide is explained by the complementary nature of the substrate-binding pockets in docking experiments.

Limitations of the study

We have only reported the labeling of purified proteins here. Although no labeling was observed in the absence of calcium and inactive proteoforms, the reactivity of these probes in extracts or living cells remains to be tested.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105247.

ACKNOWLEDGMENTS

The authors acknowledge the financial support from the Slovenian Research Agency core funding No. P1-0179 and research project No. J4-2550; the Research Foundation-Flanders grant FWO14/PDO/166 (to SS); and the European Research Council ERC-AdG-2020 101019324 'ExtraImmune' (to RH).

AUTHOR CONTRIBUTIONS

M.K. and R.v.d.H. conceived the project; S.S., D.K., and R.H. designed the probes; D.K. synthesized the probes in the CJS laboratory; K.M. performed initial labeling studies; V.S., K.v.M., and M.K. performed labeling experiments on recombinant metacaspases, cloned in the CF laboratory; M.F. generated the IRSK-bound GtMCA-I model; M.K. and R.H. wrote the manuscript with input from all authors. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 10, 2022
Revised: August 11, 2022
Accepted: September 27, 2022
Published: November 18, 2022

REFERENCES

Acosta-Maspons, A., Sepúlveda-García, E., Sánchez-Baldoquín, L., Marrero-Gutiérrez, J., Pons, T., Rocha-Sosa, M., and González, L. (2014). Two aspartate residues at the putative p10 subunit of a type II metacaspase from *Nicotiana tabacum* L. may contribute to the substrate-binding pocket. Planta *239*, 147–160.

Baici, A., Schenker, P., Wächter, M., and Rüedi, P. (2009). 3-Fluoro-2, 4-dioxa-3-phosphadecalins as inhibitors of acetylcholinesterase. A reappraisal of kinetic mechanisms and diagnostic methods. Chem. Biodivers. *6*, 261–282.

Berger, A.B., Witte, M.D., Denault, J.B., Sadaghiani, A.M., Sexton, K.M., Salvesen, G.S., and Bogyo, M. (2006). Identification of early intermediates of caspase activation using selective inhibitors and activity-based probes. Mol Cell *23*, 509–521. https://doi.org/10.1016/j.molcel.2006.06.021.

Cha, S. (1975). Tight-binding inhibitors—I: kinetic behavior. Biochem. Pharmacol. *24*, 2177–2185.

Choi, C.J., and Berges, J.A. (2013). New types of metacaspases in phytoplankton reveal diverse origins of cell death proteases. Cell Death Dis. *4*, e490.

Coll, N.S., Smidler, A., Puigvert, M., Popa, C., Valls, M., and Dangl, J.L. (2014). The plant metacaspase AtMC1 in pathogen-triggered programmed cell death and aging: functional linkage with autophagy. Cell Death Differ. *21*, 1399–1408.

Coll, N.S., Vercammen, D., Smidler, A., Clover, C., Van Breusegem, F., Dangl, J.L., and Epple, P. (2010). Arabidopsis Type I metacaspases control cell death. Science *330*, 1393–1397.

Curtis, B.A., Tanifuji, G., Burki, F., Gruber, A., Irimia, M., Maruyama, S., Arias, M.C., Ball, S.G., Gile, G.H., Hirakawa, Y., et al. (2012). Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature *492*, 59–65.

Eitelhuber, A.C., Vosyka, O., Nagel, D., Bognar, M., Lenze, D., Lammens, K., Schlauderer, F., Hlahla, D., Hopfner, K.P., Lenz, G., et al. (2015). Activity-based probes for detection of active MALT1 paracaspase in immune cells and lymphomas. Chem. Biol. *22*, 129–138. https://doi.org/10.1016/j.chembiol.2014.10.021.

Fortin, J., and Lam, E. (2018). Domain swap between two type-II metacaspases defines key elements for their biochemical properties. Plant J. *96*, 921–936.

Gilio, J.M., Marcondes, M.F., Ferrari, D., Juliano, M.A., Juliano, L., Oliveira, V., and Machado, M.F.M. (2017). Processing of metacaspase 2 from *Trypanosoma brucei* (*Tb*MCA2) broadens its substrate specificity. Biochim. Biophys. Acta. Proteins Proteom. *1865*, 388–394.

Graff van Creveld, S., Ben-Dor, S., Mizrachi, A., Alcolombri, U., Hopes, A., Mock, T., Rosenwasser, S., and Vardi, A. (2021). Biochemical characterization of a novel redox-regulated metacaspase in a marine diatom. Front. Microbiol. *12*, 2578.

Hachmann, J., Edgington-Mitchell, L.E., Poreba, M., Sanman, L.E., Drag, M., Bogyo, M., and Salvesen, G.S. (2015). Probes to monitor activity of the paracaspase MALT1. Chem Biol *22*, 139–147. https://doi.org/10.1016/j.chembiol.2014.11.011.

Hander, T., Fernández-Fernández, Á.D., Kumpf, R.P., Willems, P., Schatowitz, H., Rombaut, D., Staes, A., Nolf, J., Pottie, R., Yao, P., et al. (2019). Damage on plants activates Ca2+-dependent metacaspases for release of immunomodulatory peptides. Science *363*, eaar7486.

He, R., Drury, G.E., Rotari, V.I., Gordon, A., Willer, M., Farzaneh, T., Woltering, E.J., and Gallois, P. (2008). Metacaspase-8 modulates programmed cell death induced by ultraviolet light and H_2O_2 in Arabidopsis. J. Biol. Chem. *283*, 774–783.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589.

Kato, D., Boatright, K.M., Berger, A.B., Nazif, T., Blum, G., Ryan, C., Chehade, K.A.H., Salvesen, G.S., and Bogyo, M. (2005). Activity-based probes that target diverse cysteine protease families. Nat. Chem. Biol. *1*, 33–38.

Klemenčič, M., Asplund-Samuelsson, J., Dolinar, M., and Funk, C. (2019). Phylogenetic distribution and diversity of bacterial pseudo-orthocaspases underline their putative role in photosynthesis. Front. Plant Sci. *10*, 293.

Klemenčič, M., and Funk, C. (2018). Type III metacaspases: calcium-dependent activity proposes new function for the p10 domain. New Phytol. *218*, 1179–1191.

Klemenčič, M., and Funk, C. (2019). Evolution and structural diversity of metacaspases. J. Exp. Bot. *70*, 2039–2047.

Klemenčič, M., Novinec, M., and Dolinar, M. (2015). Orthocaspases are proteolytically active prokaryotic caspase homologues: the case of *Microcystis aeruginosa*. Mol. Microbiol. *98*, 142–150.

Lam, E., and Zhang, Y. (2012). Regulating the reapers: activating metacaspases for programmed cell death. Trends Plant Sci. *17*, 487–494.

Laverrière, M., Cazzulo, J.J., and Alvarez, V.E. (2012). Antagonic activities of *Trypanosoma cruzi* metacaspases affect the balance between cell proliferation, death and differentiation. Cell Death Differ. *19*, 1358–1369.

Leang, L., McDonald, M.C., Mineo, C.R., Jones, B., Barker, T., Gagliardi, C., and Fox, K.M. (2019). Identification and characterization of *Schizophyllum commune* type I metacaspases. Biochem. Biophys. Rep. *20*, 100706.

Machado, M.F.M., Marcondes, M.F., Juliano, M.A., McLuskey, K., Mottram, J.C., Moss, C.X., Juliano, L., and Oliveira, V. (2013). Substrate specificity and the effect of calcium on *Trypanosoma brucei* metacaspase 2. FEBS J. *280*, 2608–2621.

McLuskey, K., Rudolf, J., Proto, W.R., Isaacs, N.W., Coombs, G.H., Moss, C.X., and Mottram, J.C. (2012). Crystal structure of a Trypanosoma brucei metacaspase. Proc. Natl. Acad. Sci. USA *109*, 7469–7474.

Minina, E.A., Staal, J., Alvarez, V.E., Berges, J.A., Berman-Frank, I., Beyaert, R., Bidle, K.D., Bornancin, F., Casanova, M., Cazzulo, J.J., et al. (2020). Classification and nomenclature of metacaspases and paracaspases: no more confusion with caspases. Mol. Cell *77*, 927–929.

Morrison, J.F., and Walsh, C.T. (1988). The behavior and significance of slow-binding enzyme inhibitors. Adv. Enzymol. Relat. Areas Mol. Biol. *61*, 201–301.

Morrison, J.F. (1982). The slow-binding and slow, tight-binding inhibition of enzyme-catalysed reactions. Trends Biochem. Sci. *7*, 102–105.

Moss, C.X., Westrop, G.D., Juliano, L., Coombs, G.H., and Mottram, J.C. (2007). Metacaspase 2 of *Trypanosoma brucei* is a calcium-dependent cysteine peptidase active without processing. FEBS Lett. *581*, 5635–5639.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera - a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Phillips, J.C., Hardy, D.J., Maia, J.D.C., Stone, J.E., Ribeiro, J.V., Bernardi, R.C., Buch, R., Fiorin, G., Hénin, J., Jiang, W., et al. (2020). Scalable molecular dynamics on CPU and GPU architectures with NAMD. J. Chem. Phys. *153*, 044130.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. *41*, 207–234.

Torkar, A., Lenarčič, B., Lah, T., Dive, V., and Devel, L. (2013). Identification of new peptide amides as selective cathepsin L inhibitors: the first step towards selective irreversible inhibitors? *Bioorg*. Med. Chem. Lett. *23*, 2968–2973.

Trott, O., and Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. *31*, 455–461.

Tsiatsiani, L., Timmerman, E., De Bock, P.J., Vercammen, D., Stael, S., van de Cotte, B., Staes, A., Goethals, M., Beunens, T., Van Damme, P., et al. (2013). The Arabidopsis metacaspase9 degradome. Plant Cell *25*, 2831–2847. https://doi.org/10.1105/tpc.113.115287.

Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lam, E., and Bozhkov, P.V. (2011). Metacaspases. Cell Death Differ. *18*, 1279–1288.

Uren, A.G., O'Rourke, K., Aravind, L.A., Pisabarro, M.T., Seshagiri, S., Koonin, E.V., and Dixit, V.M. (2000). Identification of paracaspases and metacaspases: two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. Mol. Cell *6*, 961–967.

van Midden, K.P., Peric, T., and Klemenčič, M. (2021). Plant type I metacaspases are proteolytically active proteases despite their hydrophobic nature. FEBS Lett. *595*, 2237–2247.

Vercammen, D., Belenghi, B., van de Cotte, B., Beunens, T., Gavigan, J.-A., De Rycke, R., Brackenier, A., Inzé, D., Harris, J.L., and Van Breusegem, F. (2006). Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for metacaspase 9. J. Mol. Biol. *364*, 625–636.

Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenberghe, I., Van Beeumen, J., Inzé, D., and Van Breusegem, F. (2004). Type II metacaspases Atmc4 and Atmc9 of *Arabidopsis thaliana* cleave substrates after arginine and lysine. J. Biol. Chem. *279*, 45329–45336.

Watanabe, N., and Lam, E. (2011). Calcium-dependent activation and autolysis of Arabidopsis metacaspase 2d. J. Biol. Chem. *286*, 10027–10040.

Williams, E.B., Krishnaswamy, S., and Mann, K.G. (1989). Zymogen/enzyme discrimination using peptide chloromethyl ketones. J. Biol. Chem. *264*, 7536–7545.

Wong, A.H.H., Yan, C., and Shi, Y. (2012). Crystal structure of the yeast metacaspase Yca1. J. Biol. Chem. *287*, 29251–29259.

Yoon, M.C., Solania, A., Jiang, Z., Christy, M.P., Podvin, S., Mosier, C., Lietz, C.B., Ito, G., Gerwick, W.H., Wolan, D.W., et al. (2021). Selective neutral pH inhibitor of Cathepsin B designed based on cleavage preferences at cytosolic and lysosomal pH conditions. ACS Chem. Biol. *16*, 1628–1643.

Yu, J.W., Jeffrey, P.D., Ha, J.Y., Yang, X., and Shi, Y. (2011). Crystal structure of the mucosa-associated lymphoid tissue lymphoma translocation 1 (MALT1) paracaspase region. Proc. Natl. Acad. Sci. USA *108*, 21004–21009.

Zhu, P., Yu, X.-H., Wang, C., Zhang, Q., Liu, W., McSweeney, S., Shanklin, J., Lam, E., and Liu, Q. (2020). Structural basis for Ca^{2+}-dependent activation of a plant metacaspase. Nat. Commun. *11*, 2249.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal Anti-6X His tag® antibody	abcam	Cat#ab9108
Goat Anti-Rabbit IgG H&L (HRP)	abcam	Cat#ab205718
Bacterial and virus strains		
Escherichia Coli BL21 [DE3] Competent cells	Thermo Scientific™	Cat#EC0114
Chemicals, peptides, and recombinant proteins		
Fmoc-L-Lys(Boc)-OH	Iris Biotech	Cat#FAA1125
polystyrene-Me-NH2 resin	Iris Biotech	Cat#BR-1000C
dichloromethane	VWR	Cat#L13089.AP
methanol	VWR	Cat#20834.291
dimethylformamide	VWR	Cat#210585000
acetonitrile	VWR	Cat#K981-1L
Fmoc-AHX-OH	Fluorochem	Cat#045380
4-pentynoic acid	Fluorochem	Cat#164204
1-hydroxybenzotriazole hydrate	Fluorochem	Cat#M02875
N,N-diisopropylcarbodiimide	Fluorochem	Cat#132050
2,6-dimethylbenzoic acid	Alfer Aesar	Cat#A15014
diethylamine	ABCR	Cat#AB182463
trifluoroacetic acid	ABCR	Cat#AB212435
syringe reactors	Multi Syn Tech	Cat#V050PE063
EKTK-AOMK	this paper	N/A
EKAK-AOMK	this paper	N/A
IRSK-AOMK	this paper	N/A
IISK-AOMK	this paper	N/A
recombinant *Gt*MCA-I	this paper	N/A
recombinant *Cr*MCA-II	this paper	N/A
recombinant *Gt*MCA-III	Klemenčič M, Funk C. 2018. Type III metacaspases: calcium-dependent activity proposes new function for the p10 domain. *New Phytol.* 218, 1179–1191.	N/A
Z-FR-AMC	PeptaNova	Cat#3095-v
Z-RR-AMC	PeptaNova	Cat#3123-v
Na_2HPO_4	Merck	Cat#106586
KH_2PO_4	Sigma-Aldrich	Cat#60218
NH_4Cl	Fluka	Cat#11209
Na_2SO_4	Fluka	Cat#71962
$MgSO_4$	Sigma-Aldrich	Cat#13143
Glycerol	Sigma-Aldrich	Cat#G6279
D(+)-Glucose	Sigma-Aldrich	Cat#16325
α-lactose	Sigma-Aldrich	Cat#L3625
Kanamycin Sulfate	Fisher Bioreagents	Cat#BP906-5

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HEPES Free Acid	Carl Roth	Cat#HN78.3
NaCl	GramMol	Cat#P148590
imidazole	Sigma-Aldrich	Cat#56749
Sodium Acetate	Sigma-Aldrich	Cat#S7545
MES Hydrate	Sigma-Aldrich	Cat#M8250
TRIS (Tris(Hydroxymethyl)aminomethane Hydrochloride)	Fluka	Cat#93349
CAPS (3-(Cyclohexylamino)-1-propanesulfonic acid)	Sigma-Aldrich	Cat#C2632
Calcium Chloride-2-Hydrate	Sigma-Aldrich	Cat#31307
EDTA (Ethylenediaminetetraacetic Acid disodium salt)	Sigma-Aldrich	Cat#E-5134
DTT (dithiothreitol)	Thermo Scientific	Cat#R0861
DMSO (Dimethyl Sulfoxide)	Sigma-Aldrich	Cat#D5879
acetone	Sigma-Aldrich	Cat#32201
picolyl-Cy5	Click Chemistry Tools	Cat#1177-1
Copper(II) Sulfate Pentahydrate	Merck	Cat#02790
TCEP (Tris(2-carboxyethyl)phosphine)	Sigma-Aldrich	Cat#C4706
TBTA (Tris((1-benzyl-4-triazolyl)methyl)amin)	Sigma-Aldrich	Cat#678937
Oligonucleotides		
*Gt*MCA-I-NcoI-F: 5′- catgccatgg ccggcaggaagaaagctcttctgatc-3′	Integrated DNA Technologies	N/A
*Gt*MCA-I-XhoI-R: 5′-atccctcgag cttggggcgaatcttcctcctc-3′	Integrated DNA Technologies	N/A
Recombinant DNA		
codon-optimized *Cr*MCA-II gene (gBlock fragment)	Integrated DNA Technologies	N/A
*Gt*MCA-III cDNA	Klemenčič M, Funk C. 2018. Type III metacaspases: calcium-dependent activity proposes new function for the p10 domain. *New Phytol.* 218, 1179–1191.	N/A
plasmid pET-28b(+)	Novagen	Cat#69865
Software and algorithms		
GraphPad Prism 9.1.1 software for Windows	GraphPad Software	https://www.graphpad.com/
UCSF Chimera	Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612.	https://www.cgl.ucsf.edu/chimera/
AlphaFold	Jumper J, Evans R, Pritzel A et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589.	https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb#scrollTo=kOblAo-xetgx
AutoDock Vina	Trott O, Olson AJ. 2010. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461.	https://vina.scripps.edu/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Renier van der Hoorn (renier.vanderhoorn@biology.ox.ac.uk).

Materials availability

Plasmids and probes generated in this study are available from the lead contact on request.

Data and code availability

Data: All data reported in this paper will be shared by the lead contact upon request.

Code: This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Escherichia coli BL21(DE3) bacteria were grown in shaker cultures at 37°C in auto-expression medium [1% N-Z-amine AS, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.5% glycerol (54 mM), 0.05% glucose (2.8 mM), 0.2% α-lactose (5.6 mM) (Studier, 2005)], containing 50 μg/mL kanamycin.

METHOD DETAILS

Chemicals and solvents for synthesis of AOMK-based probes

Chemicals and materials were purchased from the following vendors. Fmoc-L-Lys(Boc)-OH (FAA1125, Iris Biotech); polystyrene-Me-NH2 resin (100–200 mesh; BR-1000C, Iris Biotech); dichloromethane (L13089.AP, VWR); methanol (20834.291, VWR); dimethylformamide (210585000, VWR); acetonitrile (K981-1L, VWR); Fmoc-AHX-OH (045380, Fluorochem); 4-pentynoic acid (164204, Fluorochem); 1-hydroxybenzotriazole hydrate (M02875, Fluorochem); N,N-diisopropylcarbodiimide (132050, Fluorochem); 2,6-dimethylbenzoic acid (A15014, Alfer Aesar); diethylamine (AB182463, ABCR), trifluoroacetic acid (AB212435, ABCR) and syringe reactors (V050PE063, Multi Syn Tech).All purifications were carried out using a Jasco preparative reversed-phase high-performance liquid chromatography (HPLC) system, equipped with a RP-C18-column and an applied flow rate of 20 mL/min. Peak detection was carried out at 210 nm. High-resolution mass spectrometry (HRMS) analyses were performed at a Bruker μTOF ESI-Spectrometer using the direct flow-injection method.

Synthesis of AOMK-based probes

The synthesis of all metacaspase probes was carried out according to a general SPPS procedure, following Fmoc-strategy. Couplings were conducted in a syringe reactor utilizing the corresponding Fmoc-amino acids (also including Fmoc-AHX-OH and 4-pentynoic acid) (4 equivalent), HOBt (4 equivalent), DIC (4 equivalent), at room temperature and a reaction time of 45 min per coupling reaction.

The coupling solution was discarded after each coupling step and the resin was washed alternating with 3× DMF, 3× MeOH and 3× DCM. Fmoc-deprotection after each coupling step was carried out by agitating the resin in a solution of 5% diethylamine in DMF for 15 min (2×, changing cleavage solution in between). The cleavage solution was discarded, and the resin was washed alternating with 3× DMF, 3× MeOH and 3× DCM. Fmoc-Lys(Boc)-AOMK was synthesised and loaded onto freshly prepared semicarbazide resin according to a literature procedure (Kato et al., 2005). Resin loading was carried out by agitating the semicarbazide resin with a solution of Fmoc-Lys(Boc)-AOMK in DMF at 50°C for 3 h in a syringe reactor. The loading solution was discarded and the resin was washed alternating with 3× DMF and 3× DCM. The resultant resin loading was determined at 1.3 mmol/g. After completion of the synthesis, the crude probes were cleaved off the resin (5% H_2O in TFA, room temperature, agitation, 1.5 h) and purified by reversed phase HPLC (H_2O:ACN, 0.1% TFA; gradient: 3 → 80%).

EKTK: Yield: 21.7 mg (25.7 μmol), HRMS (ESI): m/z = calculated for $C_{42}H_{66}N_7O_{11}^+$ $[M + H]^+$ 844.48141, found 844.48074. EKAK: Yield: 4.02 mg (5.2 μmol), HRMS (ESI): m/z = calcd for $C_{41}H_{64}N_7O_{10}^+$ $[M + H]^+$ 814.47092,

found 814.47083. IRSK: Yield: 2.26 mg (2.7 μmol), HRMS (ESI): m/z = calcd for $C_{42}H_{68}N_9O_9^+$ $[M + H]^+$ 842.51345, found 842.51404. IISK: Yield: 0.92 mg (2.7 μmol), HRMS (ESI): m/z = calcd for $C_{42}H_{67}NO_6^+$ $[M + H]^+$ 799.49640, found 799.49616.

Cloning of the metacaspase constructs

To obtain cDNA which was used for amplification of *Gt*MCA-I and *Gt*MCA-III (before *Gt*MC1 and *Gt*MC2, respectively), we isolated total RNA from 200 mL of *G. theta* cells (CCMP2712) in the late exponential phase as explained before (Klemencic and Funk, 2018). *Gt*MCA-I was amplified using the primers introducing the *Nco*I and *Xho*I restriction sites at 5′ and 3 of the construct, respectively, (5′- catgccatggccggcaggaa-gaaagctcttctgatc-3′ and 5′-atccctcgagcttggggcgaatcttcctcctc-3′)). The gene encoding for full-length *Cr*MCA-II, codon-optimized for expression in *E. coli*, was ordered as a gBlock fragment at IDT (Coralville, Iowa, USA) also with the *Nco*I and *Xho*I restriction sites at the 5′ and 3′ of the gene, respectively. Both amplicons were ligated into the bacterial expression vector pET-28b(+) using *Nco*I and *Xho*I restriction sites. Correct ligation and nucleotide sequences were verified by DNA sequencing. Construction of the *Gt*MCA-III gene has been reported (Klemenčič and Funk, 2018). Table S1 contains nucleotide sequences of all three used metacaspases.

Expression and purification of proteins

Escherichia coli BL21(DE3) bacteria were transformed with the expression plasmids and grown in shaker cultures at 37°C in auto-expression medium [1% N-Z-amine AS, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.5% glycerol (54 mM), 0.05% glucose (2.8 mM), 0.2% α-lactose (5.6 mM) (Studier, 2005)], containing 50 μg/mL kanamycin. After 6 h, the flasks were cooled down to 16°C and left shaking for an additional 18 h (Studier, 2005). The cell pellet collected from 400 mL of bacterial culture was resuspended in 20 mL of resuspension buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole) and sonicated 5 × 30 s (80% power) on ice. Following centrifugation at 20,000 × g for 30 min to remove insoluble debris, the supernatant was applied to a HisTrap™ FF column (Cytiva), connected to an ÄKTA FPLC system. After the wash with the resuspension buffer, bound proteins were eluted in the same buffer, but containing 300 mM imidazole. The peak fractions were collected (~5 mL), diluted to 100 mL of ion-exchange binding buffer (20 mM HEPES pH 8.2), and subsequently loaded onto a HiTrap ANX FF column (GE Healthcare Life Sciences, Marlborough, Massachusetts, USA). Specifically, bound proteins were eluted with binding buffer containing 1 M NaCl. The peak fractions were collected, concentrated to approximately 5 mg/mL using an Amicon filtration unit (Millipore Corp., Temecula, California, USA) equipped with a 10 kDa exclusion membrane, and applied to a Superdex 75 size-exclusion chromatography column (GE Healthcare Life Sciences, Marlborough, Massachusetts, USA) connected to an ÄKTA FPLC system. The column was equilibrated in 20 mM HEPES pH 7.5, 500 mM NaCl, and a flow rate of 0.5 mL/min was used for the separation of proteins. Selected fractions containing the protein of choice were kept at −80°C until use.

Kinetic assays

Protease activity was measured by monitoring the release of the fluorescent group AMC (7-amino-4-methylcoumarin) from Z-FR-AMC (from PeptaNova, Sandhausen, Germany) at excitation and emission wavelengths of 383 nm and 455 nm, respectively, using a *Infinite M200 Pro* spectrofluorimeter (*Tecan Trading, Männedorf, Switzerland*). For analysis of pH profiles, the buffers used were: 100 mM acetate (pH 4.0–pH 5.5), 100 mM MES (pH 6.0–pH 6.5), 100 mM HEPES (pH 7.0–pH 8.0), 100 mM Tris (pH 8.5–pH 9.0) and 100 mM CAPS (pH 9.5–pH 11.0). All assays were performed with 10 nM enzyme and 5 μM Z-FR-AMC or Z-RR-AMC in a buffer containing 150 mM NaCl, 5 mM $CaCl_2$ and 5 mM dithiothreitol (DTT). The enzyme concentration is based on protein assuming purity. To determine the calcium-dependency profiles, assays were performed in 20 mM MES, pH 6.5, 150 mM NaCl, 5 mM DTT for *Gt*MCA-I and *Gt*MCA-III or 20 mM HEPES, pH 8.5, 150 mM NaCl, 5 mM DTT for *Cr*MCA-II at respective calcium concentrations. All data were analyzed using GraphPad Prism 9.1.1 (GraphPad Software Inc., La Jolla, CA, USA).

Determination of K_M values

For determination of the Michaelis-Menten constant (K_M) fluorogenic substrates Z-FR-AMC and Z-RR-AMC (both from PeptaNova, Sandhausen, Germany) were used. Assays were performed in 20 mM MES, pH 6.5, 150 mM NaCl, 5 mM DTT, 10 mM $CaCl_2$ for *Gt*MCA-I (final enzyme concentration 95 nM), 20 mM HEPES, pH 8.5, 150 mM NaCl, 5 mM DTT, 10 mM $CaCl_2$ for *Cr*MCA-II (final enzyme concentration 6 nM) and 20 mM

MES, pH 6.5, 150 mM NaCl, 5 mM DTT, 10 mM $CaCl_2$ for *Gt*MCA-III (final enzyme concentration of 35 nM). All assays were performed in 2 mL disposable acrylic cuvettes. All measurements were repeated in three independent experiments. Michaelis-Menten constant (K_M) was determined by non-linear regression analysis using GraphPad Prism 9.1.1 Software.

Inhibition kinetics measurements

To determine the inactivation parameters of each of the four probes against the representatives of the three metacaspase types, product concentration curves of the hydrolysis of the fluorogenic substrate Z-RR-AMC were recorded in the presence of increasing concentrations of each probe. Stock solutions of the probes (100 μM) were prepared in DMSO and stored at −20°C prior to use. Time-dependent release of fluorescent product AMC was followed using a Luminiscence Spectrometer LS50B (Perkin Elmer) at 25°C, using λ_{ex} of 383 nm and λ_{em} of 455 nm with slit width of 5 nm. Fluorescence of the reaction mixtures was measured in a 5 × 5 mm quartz cuvette (Hellma, Müllheim, Germany). Reaction mixtures consisted of 20 mM MES, pH 6.5, 150 mM NaCl, 5 mM DTT, 5 mM $CaCl_2$, 5 μM Z-RR-AMC for *Gt*MCA-I (final enzyme concentration 28 nM), 20 mM HEPES, pH 8.5, 150 mM NaCl, 5 mM DTT, 10 mM $CaCl_2$, 25 μM Z-RR-AMC for *Cr*MCA-II (final enzyme concentration 8 nM) and 20 mM MES, pH 6.5, 150 mM NaCl, 5 mM DTT, 5 mM $CaCl_2$, 6 μM Z-RR-AMC for *Gt*MCA-III (final enzyme concentration of 12 nM). Final concentrations of probes in reaction mixtures were ranging from 5 to 500 nM. All reactions were initiated by the addition of enzyme to mixtures of substrate and inhibitor in the respective buffer.

Inhibition kinetics data analysis

The time-dependent inhibition of metacaspase activities was evaluated by progress curve analysis using GraphPad Prism 9.1.1 software. Progress curves of product formation were obtained in the absence and the presence of probes as described above. Progress curves were fitted with a modified version of the equation used by Cha (Cha (1975) and Morrison (Morrison (1982) for the description of slow-binding inhibitors:

$$[AMC] = \frac{v_0}{k'}\left(1 - e^{-k't}\right) + d \qquad \text{(Equation 1)}$$

where k is a pseudo-first order rate constant which describes the bending of the curve and is linked to the mechanism of inactivation, as described below, v_0 is the initial reaction rate and d is the signal displacement along the y-axis (Morrison and Walsh, 1988).

Subsequently, plots of the calculated values of k versus probe concentration were generated and fitted with Equation 2 or 3, depending on the shape of the regression curve (linear of hyperbolic, respectively), which reflects the mechanism of metacaspase activity inactivation in one or two steps, as depicted in Figure S20. Corresponding rate constants were determined from the fits and expressed as the overall rate constant of inactivation k_{ina}. For one-step inactivation, k_{ina} was determined from the slope of the linear regression curve using Equation 2. For two-step inactivation, k_{ina} was determined as the ratio of k_4/K_i, obtained by fitting Equation 3 to the data (Morrison and Walsh, 1988).

$$k' = \frac{k_{ina}}{1+\sigma}[I];\ \sigma = \frac{[S]}{K_m} \qquad \text{(Equation 2)}$$

$$k' = \frac{k_4\frac{[I]}{K_i}}{1+\sigma+\frac{[I]}{K_i}};\ \sigma = \frac{[S]}{K_m} \qquad \text{(Equation 3)}$$

By comparing constants of inactivation (k_{ina}), we were able to assess the efficiency of the probe.

Metacaspase labelling

Two labelling buffers were used: 20 mM MES, pH 6.5, 150 mM NaCl for *Gt*MCA-I and *Gt*MCA-III and 20 mM HEPES, pH 8.5, 150 mM NaCl for *Cr*MCA-II. Labelling reaction mixtures were set by mixing 5 μg of a recombinant protease and a probe (final concentration 2 μM in DMSO) in 50 μL of the appropriate buffer. In control reactions (NPC, no probe control), only DMSO was added instead of the probe. Only after 5 min incubation at room temperature, $CaCl_2$ was added to the reaction mixture to a final 5 mM concentration. Reaction mixtures were left incubating by gently shaking at room temperature for 1 h and stopped by addition of 4 volumes of ice-cold acetone followed by vortexing and centrifugation for 2 min at 13,000 g at room

temperature. The supernatant was discarded and pellets were dissolved by incubation first for 30 min at 37°C followed by 10 min at 90°C in 1% SDS in PBS. Click reaction was performed by addition of 2.5 μM picolyl-Cy5 (in DMSO), 1 mM $CuSO_4$, 2 mM TCEP (Tris(2-carboxyethyl)phosphine) and 0.1 mM TBTA (Tris((1-benzyl-4-triazolyl)methyl)amin). Reactions were left shaking in dark for 1 h at room temperature and stopped by addition of 4 volumes of ice-cold acetone as before. The supernatant was discarded and 25 μL of dH_2O and 25 μL of 2× loading buffer with DTT was added to the pellet, which was left to dissolve for 30 min at 37°C. 10 μL were loaded into the wells of 15% SDS-PAGE gel, which was left running at 150 V for approximately 2.5 h. Gels were briefly washed with dH_2O and scanned with a ChemiDoc MP Imaging System using settings for Cy5 fluorescence detection (BioRad, Hercules, California, United States). The same gels were subsequently stained with Coomassie Brilliant Blue, destained and photographed using the same imager.

Molecular modeling and docking

The homology model of *Gt*MCA-I was built with AlphaFold (Jumper et al., 2021) MALT-1 with bound Z-VRPR-FMK as a template (PDB ID: 3UOA) (Yu et al., 2011). The tetrapeptide IRSK was modeled with UCSF Chimera (Pettersen et al., 2004) and then docked into the active site of *Gt*MCA-I with AutoDock Vina (Trott and Olson, 2010) using fully flexible ligand (IRSK tetrapeptide) and rigid receptor (*Gt*MCA-I). The best binding mode was selected based on the comparison of the obtained binding mode with the crystal structure of the human paracaspase MALT1 with bound tetrapeptide-based irreversible inhibitor (Yu et al., 2011). The obtained complex was further optimized by a 10 ns molecular dynamics simulation using NAMD3 (Phillips et al., 2020) which was sufficient to equilibrate the system. The simulation was performed at 298 K in explicit water solvent. The system was ionized and neutralized by addition of sodium and chloride ions to a final concentration of 0.15 M NaCl. The system was first minimized in a 1000 step gradient and then simulated for 10 ns using periodic boundary conditions.

QUANTIFICATION AND STATISTICAL ANALYSIS

Activity assays shown in Figures 3A, 3B, 7, and S2 are based on flurescence measurements, shown as the mean with standard deviation of n = 3 replicates. Michaelis-Menten constant was determined by non-linear regression analysis using GraphPad Prism 9.1.1 Software.