

Orthocaspases are proteolytically active prokaryotic caspase homologues: the case of *Microcystis aeruginosa*

Marina Klemenčič, Marko Novinec and Marko Dolinar*

Department of Chemistry and Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Večna pot 113, SI-1000 Ljubljana, Slovenia.

Summary

Caspases are a family of cysteine-dependent proteases known to be involved in the process of programmed cell death in metazoans. Recently, cyanobacteria were also found to contain caspase-like proteins, but their existence has only been identified *in silico* up to now. Here, we present the first experimental characterisation of a prokaryotic caspase homologue. We have expressed the putative caspase-like gene *MaOC1* from the toxic bloom-forming cyanobacterium *Microcystis aeruginosa* PCC 7806 in *Escherichia coli*. Kinetic characterisation showed that MaOC1 is an endopeptidase with a preference for arginine in the P1 position and a pH optimum of 7.5. MaOC1 exhibited high catalytic rates with the k_{cat}/K_M value for Z-RR-AMC substrate of the order $10^6 \text{ M}^{-1} \text{ s}^{-1}$. In contrast to plant or metazoan caspase-like proteins, whose activity is calcium-dependent or requires dimerisation for activation, MaOC1 was activated by autocatalytic processing after residue Arg219, which separated the catalytic domain and the remaining 55 kDa subunit. The Arg219Ala mutant was resistant to autoprocessing and exhibited no proteolytic activity, confirming that processing of MaOC1 is a prerequisite for its activity. Due to their structural and functional differences to other known caspase-like proteins, we suggest to name these evolutionary primitive proteins orthocaspases.

Introduction

Caspases are cysteine-dependent aspartate-directed proteases and are indispensable for the regulation of metazoan apoptosis, a well-characterised form of programmed

cell death (PCD).¹ They are synthesised as inactive proenzymes and are divided into two types based on their overall structure and activation modes. Executioner caspases are activated by proteolytic separation of the large (p20) and small (p10) subunits, whereas initiator caspases have an N-terminal prodomain that is needed for oligomerisation and recruitment into protein complexes called apoptosomes (Fuentes-Prior and Salvesen, 2004; Riedl and Salvesen, 2007). Caspases are not found in plants, yeasts or protozoa. Instead, two groups of caspase-related cysteine proteases were identified in these organisms and based on domain structure and sequence similarity divided into paracaspases and metacaspases (Uren *et al.*, 2000). Together with caspases, metacaspases and paracaspases are classified as C14 family proteases that belong to the CD clan of cysteine peptidases. A common structural feature of this clan is the typical caspase/hemoglobinase fold (Aravind and Koonin, 2002).

Metacaspases represent the largest sub-family and are further classified into three types: type I members contain an additional proline-rich repeat and zinc-finger motifs in the N-terminal prodomain (Vercammen *et al.*, 2004), whereas type II metacaspases lack such domains and are hallmarked by the presence of an extended linker region between p20-like and p10-like domains. Recently, type III metacaspases have been identified with an unusual rearrangement of domains, having the p10-like domain located at the N-terminus instead of the C-terminus of the protein as seen in other caspase-like proteins (Choi and Berges, 2013). Increasing biochemical evidence reveals that plant, yeast and protozoan metacaspases and metazoan paracaspases do not share the substrate specificity of caspases. Instead, they are cysteine proteases that cleave their substrates after Arg or Lys at the P1 position (Bozhkov *et al.*, 2005; Watanabe and Lam, 2005; 2011; Vercammen *et al.*, 2006; Lee *et al.*, 2007; Moss *et al.*, 2007; Hachmann *et al.*, 2012). The first indication for a basic amino acid

¹Abbreviations used: AMC, 7-amino-4-methylcoumarin; CARD, caspase activation and recruitment domain; DD, death domain; DEVD, Asp-Glu-Val-Glu; FGE-sulfatase, formylglycine-generating sulfatase; FMK, fluoromethylketone; FR, Phe-Arg; GGR, Gly-Gly-Arg; GKR, Gly-Lys-Arg; GRR, Gly-Arg-Arg; GUN4, genomes uncoupled protein 4; LVSR; Leu-Val-Ser-Lys; MaOC, *Microcystis aeruginosa* orthocaspase; MC, metacaspase; PCD, programmed cell death; PRR, prolin-rich repeat; RR, Arg-Arg; Z, benzoyl.

Accepted 25 June, 2015. *For correspondence. E-mail marko.dolinar@fkt.uni-lj.si; Tel. +38614798551; Fax 003862419144.

preference was the determination of the *Arabidopsis thaliana* AtMC9 autoprocessing site (Vercammen *et al.*, 2004) followed by identification of other type I and type II metacaspase cleavage sites, all with basic cleavage site sequences (Sundstrom *et al.*, 2009; Ojha *et al.*, 2010). Substrates with basic Arg or Lys residues at the P1 position were readily hydrolysed by recombinant *Trypanosoma brucei* metacaspase TbMC2 (Moss *et al.*, 2007), *A. thaliana* type I and type II metacaspases AtMC1, AtMC4, AtMC5, AtMC8, AtMC9 (Vercammen *et al.*, 2004; Watanabe and Lam, 2005; 2011; He *et al.*, 2008), as well as immunoprecipitated metacaspases LdMC1 and LdMC2 from *Leishmania donovani* (Lee *et al.*, 2007) and even paracaspases (Hachmann *et al.*, 2012).

This P1 specificity shift still remains elusive, despite new information on phylogeny of the metacaspases, paracaspases and caspases. Emerging evidence suggests that caspase-like proteins evolved from two ancestral forms of metacaspases: type I metacaspases that can be found in lineages from Proteobacteria to plants, and so-called metacaspase-like proteins, widely represented in Eubacteria but absent in plants and green algae (Choi and Berges, 2013). Interestingly, metacaspase-like proteins seem to be lost after primary endosymbiosis with the emergence of type II metacaspases, which are exclusively found in green algae and plants. In contrast to other members of the caspase-like superfamily of proteins, metacaspase-like proteins lack sequence homology to p10 domain, bearing only the core peptide motif of the caspase-hemoglobinase fold. Due to their incomplete caspase-like structure, metacaspase-like proteins were hypothesised to be inactive, which would count for their absence in higher evolutionary lineages (Choi and Berges, 2013).

To address this issue, we have chosen to explore the metacaspase-like proteins of Cyanobacteria, the most likely ancestors of chloroplasts in plants. The typical representative of unicellular cyanobacteria, *Synechocystis* sp. PCC 6803 contains only an inactive metacaspase-like protein as judged from the absence of the catalytic His-Cys dyad. For this reason, metacaspase-like proteins of its most closely related strain based on 16S rRNA, *Microcystis aeruginosa* (Jiang *et al.*, 2010), were chosen for further investigation. Our data suggest that the most structurally conservative cyanobacterial metacaspase-like protein (MaOC1) exhibits proteolytic activities similar to metacaspases and paracaspases. Due to their evolutionary importance and structural distinctiveness we termed metacaspase-like proteins *orthocaspases*.

Results and discussion

Identification of *M. aeruginosa* PCC 7806 orthocaspases

Programmed cell death is a well characterised mechanism with caspases being one of the most widely studied meta-

zoan enzymes. Caspases are synthesised as inactive precursors comprised of a prodomain, a large catalytic p20 and a small p10 domain. More than a decade ago, caspase orthologues were identified based on the presence of the structurally homologous p20 domain with the caspase-hemoglobinase fold (Uren *et al.*, 2000). Identified proteins were termed metacaspases (found in plants, yeasts and protozoa) and paracaspases (found in metazoa and slime moulds). With increasing numbers of available genomes, new evidence emerges on the presence of caspase-like proteins in Euryarchaeota and even Proteobacteria, which complies well with Koonin's almost 20 year old idea of a bacterial origin of the PCD (Aravind *et al.*, 2001). Interestingly, type II metacaspases can exclusively be found in plants and green algae, whereas an ancient form of type I metacaspases (without prodomain) has been identified in β -Proteobacteria, δ -Proteobacteria, Actinobacteria and Nitrospirae (Choi and Berges, 2013). α -Proteobacteria, on the other hand, together with Cyanobacteria and Archaea contain a well conserved p20-like domain but uniformly lack the p10-like domain (Fig. 1A). These proteins were termed metacaspase-like proteins (Choi and Berges, 2013). Due to their incomplete caspase-like structure, they have been hypothesised most likely to be non-functional or at least exhibit considerably different substrate specificity. Based on their demonstrated functionality herein and together with their suggested evolutionary importance, we termed the metacaspase-like proteins *orthocaspases*.

Characterisation of *M. aeruginosa* orthocaspase sequences

Cyanobacteria, an ancient group of photoautotrophic bacteria, contain a rich pool of putative orthocaspases. The distribution of these genes in Cyanobacteria is an integrated function of genome size and ecological habitat: most unicellular cyanobacteria contain none or only one gene, whereas filamentous cyanobacteria may contain up to 12 caspase homologues (Jiang *et al.*, 2010). The architecture of most of the orthocaspases is simple, comprising only the p20-like domain (Jiang *et al.*, 2010), yet additional domains known to be involved in signal recognition were also identified (Asplund-Samuelsson *et al.*, 2012). In the colony-forming cyanobacterium *M. aeruginosa* PCC 7806, five putative genes coding for caspase-like proteins have been identified thus far (Frangeul *et al.*, 2008). Performing a homology DELTA-BLAST search, we discovered an additional sequence, leading to a total of six *M. aeruginosa* PCC 7806 orthocaspases. We designated the members from MaOC1 to MaOC6 in accordance to the nomenclature established for *A. thaliana* metacaspases (Tsiatsiani *et al.*, 2011), and the alignment of their amino acid sequences is shown in Fig. 1B. In addition to the catalytic domain, MaOC3 and MaOC5 contain

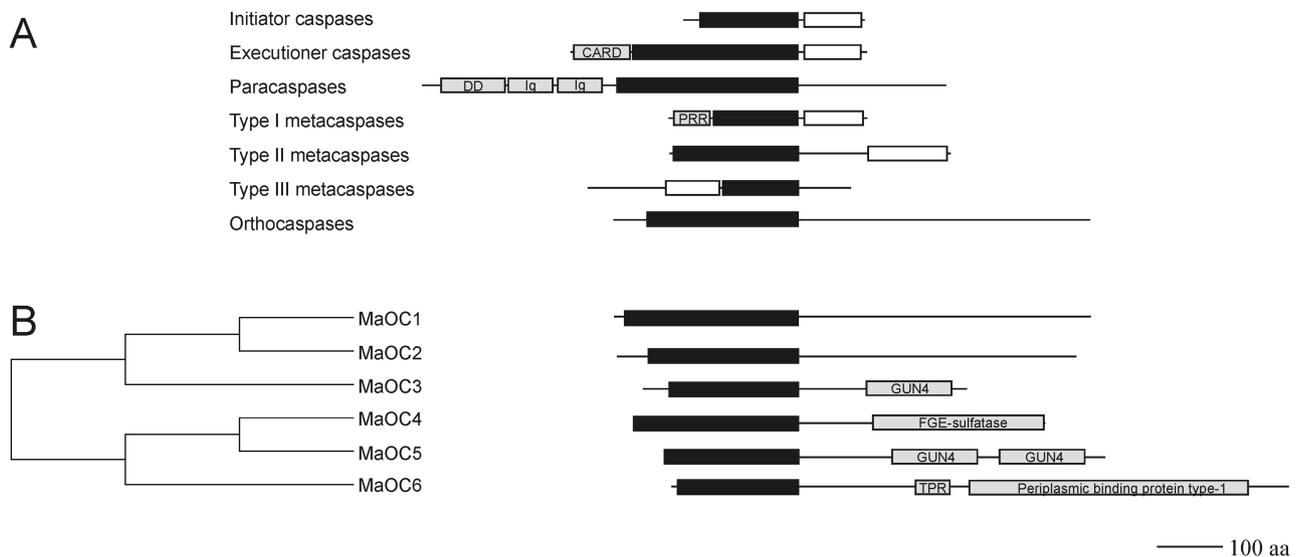


Fig. 1. Domain organisation of orthocaspases and other typical caspase-like protein superfamily members.

A. Domain organisation of caspases and paracaspases in metazoa, metacaspases in plants and diatoms and orthocaspases in bacteria. Domains were identified using InterPro protein sequence analysis and classification tool (Mitchell *et al.*, 2015). In addition to p20 (black rectangle) and p10 (white rectangle) various recruitment domains are found in executioner caspases (CARD), paracaspases (DD) and type I metacaspases (PRR). The following representatives are shown: human caspase-9 as an executioner caspase, human caspase-3 as an initiator caspase, human MALT-1 as a paracaspase, *A. thaliana* AtMC1 as type I metacaspase, *A. thaliana* AtMC4 as type II metacaspase, *T. pseudonana* metacaspase as type III metacaspase and *Synechocystis* PCC 6803 caspase-like protein as a representative of orthocaspases.

B. Unrooted phylogenetic tree of *M. aeruginosa* PCC 7806 orthocaspase family. The tree is based on a PROMALS alignment of p20-like sequences from the six identified MaOCs and was constructed using the Neighbor-Joining method of the MEGA software (version 6). The catalytic p20-like domain is drawn as a black rectangle while additional domains are drawn in gray. CARD, caspase activation and recruitment domain; DD, death domain; PRR, proline-rich repeat; GUN4, genomes uncoupled protein 4; FGE-sulfatase, formylglycine-generating sulfatase. NCBI accession numbers: caspase-9, 842; caspase-3, 836; MALT-1, NG_033893; AtMC1, AEE27396; AtMC4, AEE36232; TpMC, 7450918; SyMC, 499175590; MaOC1, CAO88281; MaOC2, CAO87907; MaOC3, CAO88826; MaOC4, CAO87362; MaOC5, CAO86659; MaOC6, CAO86810.

sequences homologous to the GUN4 domain that is proposed to participate in plastid-to-nucleus signalling in *A. thaliana* by regulating magnesium-protoporphyrin IX transfer (Larkin *et al.*, 2003). A sequence homologous to sulfatase-modifying factor enzyme 1 is present at the C-terminus of MaOC5. As this is a domain found in several eukaryotic proteins where it is required for post-translational sulfatase modification (Landgrebe *et al.*, 2003), its function in *M. aeruginosa* cannot easily be deduced. MaOC6 is the only member that is most likely a transmembrane protein with an extracellular or periplasmic receptor domain at the C-terminus.

Neither MaOC1 nor MaOC2 contain identifiable domains other than the p20-like domain. The hallmark of this domain is the conserved His/Cys dyad. Consensus regions surrounding the catalytic residues in *M. aeruginosa* are H(Y/F)SGHG and D(A/S)C(H/R) and are highly conserved in orthocaspases (Jiang *et al.*, 2010) (Fig. 2) as well as metacaspases (Vercammen *et al.*, 2004). Interestingly, MaOC2, like many other cyanobacterial orthocaspases and metacaspases TbMC1 and TbMC4 (Szallies

et al., 2002), encodes Tyr at the position of His, and Ser in place of the catalytic Cys, which renders it catalytically inactive. For further characterisation we have chosen the MaOC1, with the most conserved structure containing only the p20-like domain.

Overexpression of MaOC1 results in autocatalytic processing

The cleavage of caspases, which results in separation of the catalytic p20 domain and the p10 domain followed by dimerisation, is a prerequisite for their activity (Earnshaw *et al.*, 1999). Similarly, plant type I metacaspases (AtMC2) and yeast metacaspases (Yca1) were shown to undergo proteolytic processing upon overproduction in bacteria, but it was unclear whether cleavage was necessary for activation (Watanabe and Lam, 2005; Wong *et al.*, 2012). Metacaspase AtMC4 was shown to generate two fragments of approximately 20 and 10 kDa (Watanabe and Lam, 2011), and autoprocessing of AtMC9 was proven to be a prerequisite for its activity (Vercammen *et al.*, 2004).

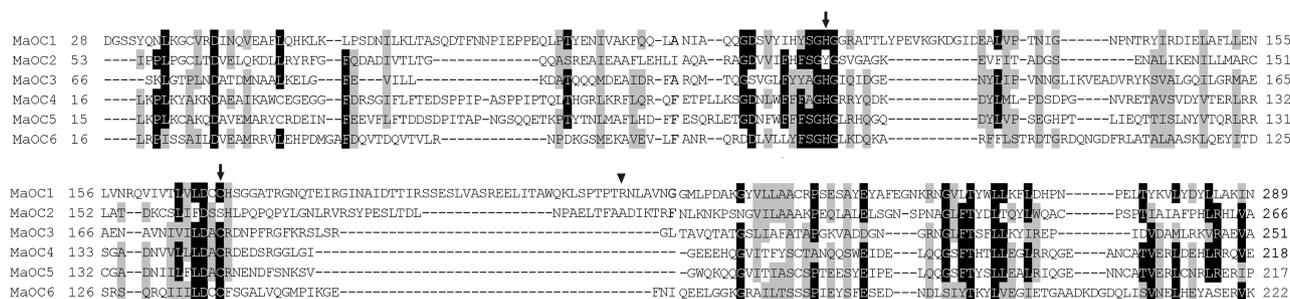


Fig. 2. Multiple sequence alignment of *M. aeruginosa* orthocaspases. Identical residues are shaded black and similar amino acids are shaded gray with 60% threshold for shading. The putative catalytic His and Cys residues are indicated by arrows and the P1 Arg involved in autocatalytic processing is represented by a triangle. The sequence alignment was performed using PROMALS (Pei and Grishin, 2007).

Conversely, protozoan metacaspase TbMC2 required no autoprocessing for its activity (Moss *et al.*, 2007), whereas LdMC1 and NdMC2 do not undergo proteolytic processing (Lee *et al.*, 2007). When we expressed and affinity purified MaOC1 in *Escherichia coli*, no band corresponding to the full-length protein (approximate molecular mass of 78 kDa) could be detected. Instead, two bands with apparent molecular masses of about 55 kDa and 25 kDa were clearly visible (Fig. 3). Although various chromatographic methods were used in attempt to separate the processed fragments (ionic, hydrophobic, size-exclusion), both proteolytic fragments were co-purified in the same fraction. Sequence alignment of MaOCs reveals that MaOC1 as well as MaOC6 contain two adjacent Cys residues (Cys¹⁶⁹ and Cys¹⁷⁰) that could potentially form a catalytic dyad with His in the active site. To test which of the Cys residues is required for the catalytic activity, Cys¹⁶⁹ and Cys¹⁷⁰, together with His¹¹⁰, were individually substituted with alanines. SDS-PAGE analysis revealed that MaOC1^{C170A} and MaOC1^{H110A} did not undergo processing, thus proving that Cys¹⁷⁰ represents the catalytic cysteine (Fig. 2). The equivalent of Cys¹⁷⁰ in *T. brucei* (Cys²¹³), which also harbours neighbouring Cys residue at

position 212, has similarly been identified as the catalytic Cys (Moss *et al.*, 2007). The MaOC1^{C169A} mutant on the other hand was able to process autocatalytically, but exhibited a distinct cleavage pattern, suggesting its possible role in MaOC1 substrate co-ordination.

MaOC1 cleaves after P1-Arg substrates

As bacterial overproduction of the MaOC1 resulted in autoprocessing into an active form (see below), we characterised the cleavage products. N-termini of the cleaved fragments were consistent with cleavage after residues Arg⁵ and Arg²¹⁹ (Fig. 3). This clearly suggests that orthocaspases, in analogy to executioner caspases and plant metacaspases, separate the catalytic p20-like domain from the rest of the polypeptide chain. While cleavage sites in type I and type II metacaspases are highly conserved (Lam and Zhang, 2012), the cleavage in MaOC1 occurs in a long linker region with no homology in other MaOC proteins (Fig. 2). The nature of the autoprocessing suggests that prokaryotic orthocaspases share the substrate specificity of metacaspases and paracaspases, which in contrast to caspases (with aspartate preference), cleave their substrates after positively charged residues. Various synthetic oligopeptides with Arg at P1 site were efficiently cleaved by MaOC1, with a requirement of a minimum length of two residues, as no cleavage was observed using the substrate H-R-AMC (Fig 4A). Expectedly, MaOC1 did not cleave the classical caspase substrate Ac-DEVD-AMC. Optimal substrates with Arg at P1 site were cleaved with k_{cat}/K_M values of up to $10^6 \text{ M}^{-1} \text{ s}^{-1}$ (Table 1), which is one or two orders of magnitude higher than the specificity constants determined for optimal substrates of metacaspase AtMC9 or paracaspase MALT-1, respectively (Vercammen *et al.*, 2006; Hachmann *et al.*, 2012), rendering MaOC1 the most efficient member of the caspase-like superfamily of proteins. In addition to Arg specificity at P1 position, metacaspases and paracaspases cleave their substrates

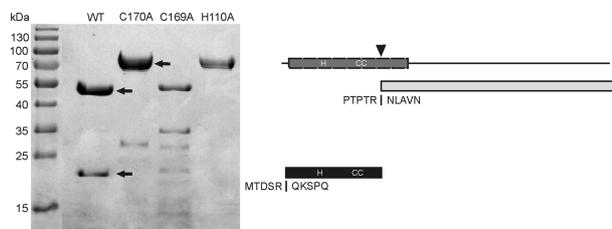


Fig. 3. Analysis of autoprocessing of MaOC1 expressed in *E. coli*. Five micrograms of purified C-terminally His-tagged MaOC1 and its mutant forms were separated on 12% SDS-PAGE and stained with Coomassie brilliant blue. The detected wild-type autoprocessing fragments with the determined N-terminal amino acid sequences are represented schematically on the right. The location of P1 Arg involved in autocatalytic processing is represented by a triangle. *Dark gray*, InterPro (Mitchell *et al.*, 2015) annotated caspase C14 domain; *black*, p20-like domain; *light gray*, C-terminal domain.

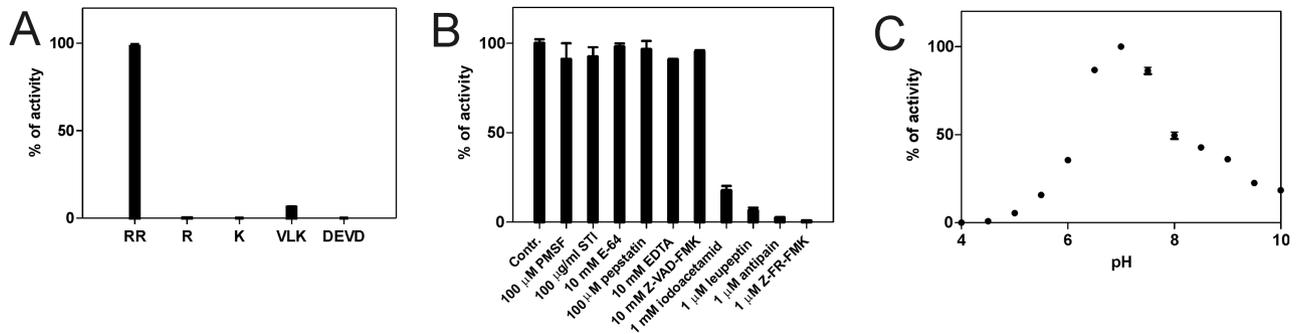


Fig. 4. Enzymatic properties of MaOC1.

A. MaOC1 substrate specificity was determined using various fluorogenic peptides: Z-RR-AMC (*RR*), R-AMC (*R*), K-AMC (*K*), Boc-VLK-AMC (*VLK*) or Ac-DEVD-AMC (*DEVD*). Initial hydrolysis rates were determined and are presented as percentage of the maximal rate using Z-RR as a substrate.

B. Effect of different protease inhibitors on MaOC1 was determined with Z-RR-AMC as a substrate. Relative activity is expressed as percentage of the activity of the enzyme without any inhibitor.

C. Initial velocities using Z-RR-AMC as a substrate were determined in buffers with different pH as explained in *Materials and methods* section. Relative activity is expressed as percentage of the activity at optimal pH. All assays were performed using 10 μ M substrate and 1 nM enzyme concentrations in 20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM DTT at 25°C. All experiments were done as three independent assays. The analysis was performed with GraphPad Prism 5.0 Software. Mean values with standard deviation are shown.

readily after Lys residues. MaOC1, on the other hand, exhibited only weak catalytic efficiencies using VLK as a substrate, due to a very low k_{cat} and moderately higher K_{M} value in comparison with substrates with Arg at position P1. Consistently, orthocaspase activity is completely abolished in the presence of two arginal reversible inhibitors leupeptin and antipain (Fig. 4B), which were also proven to be potent inhibitors of *A. thaliana*, *L. donovani* and *T. brucei* metacaspases (Bozhkov *et al.*, 2005; Watanabe and Lam, 2005; Moss *et al.*, 2007). Activity was neither observed in the presence of the inhibitory substrate analogue Z-FR-FMK, confirming the specificity for basic P1 residues in orthocaspase substrates.

Biochemical analysis demonstrated that MaOC1 has a pH optimum at pH 7, with substantial activity also at higher pH values (Fig. 4C). This is in accordance with previously published data of various type I or type II meta-

caspases as well as paracaspases that indicates that they all prefer neutral to slightly basic pH for their activity. The only exception is AtMC9 with a pH optimum of pH 5.5 (Vercammen *et al.*, 2004; Lee *et al.*, 2007; Moss *et al.*, 2007; Hachmann *et al.*, 2012).

MaOC1 is activated by autoprocessing

Members of the caspase superfamily of proteins are synthesised as inactive zymogens and can be further classified on the basis of their mode of activation: while proteolytic processing is required for the activation of mammalian caspases as well as type II metacaspases, the activity of type I metacaspases does not depend on autoprocessing (Fuentes-Prior and Salvesen, 2004; McLuskey *et al.*, 2012). In addition to autoprocessing mechanisms, the activity of type I as well as most of the type II metacaspases is calcium regulated (Adams, 2003; Wong *et al.*, 2012). With the exception of calcium-independent AtMC9 (Zhang and Lam, 2011), type II metacaspases are activated in the presence of mM concentrations of calcium, whereas μ M concentrations are sufficient for activation of type I metacaspases (Madeo *et al.*, 2002). Paracaspases, on the other hand, are activated by dimerisation without cleavage (Hachmann *et al.*, 2012). To determine the mode of activation of the orthocaspase MaOC1, different experiments were set up to test each of the possibilities. As shown in Fig. 5A, Ca^{2+} concentrations ranging from low micromolar to high millimolar had no effect on the rate of MaOC1 hydrolysis. As no significant change in the activity was observed neither in presence of 10 mM EDTA, we concluded that calcium has no effect on the hydrolytic activity of MaOC1. Size exclusion chromatography followed by immediate measurement of enzyme activity was

Table 1. Catalytic efficiencies of MaOC1.

Substrate	k_{cat} (s^{-1})	K_{M} (μM)	$k_{\text{cat}}/K_{\text{M}}$ ($\text{M}^{-1}\cdot\text{s}^{-1}$)
Z-LR-AMC	8.58	34.83 (\pm 6.35)	2.44×10^5
Z-FR-AMC	8.72	13.52 (\pm 1.68)	2.50×10^5
Z-RR-AMC	8.53	4.98 (\pm 0.47)	1.71×10^6
Boc-GKR-AMC	9.65	82.46 (\pm 4.61)	1.16×10^5
Boc-GRR-AMC	6.94	27.3 (\pm 2.61)	2.52×10^5
Z-VVR-AMC	15.30	9.14 (\pm 1.03)	1.67×10^6
Boc-VLK-AMC	0.61	51.74 (\pm 7.39)	1.14×10^4

Assays were performed in 20 mM HEPES (pH 7.5) containing 150 mM NaCl and 5 mM DTT at 25°C using 1 nM enzyme with variable peptidyl-AMC concentrations (0–250 μM). Active enzyme concentrations were determined by active site titration with the irreversible inhibitor Z-FR-FMK. Kinetic constants were determined performing non-linear regression analysis using Graphpad Prism 5.0 Software.

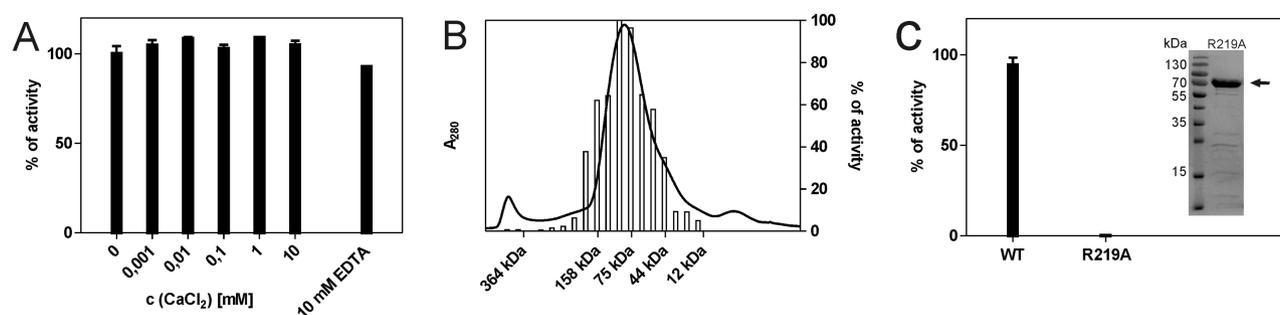


Fig. 5. Activation mode of MaOC1.

A. Effect of calcium concentrations on MaOC1 was determined by performing the measurements with calcium concentrations ranging from μM to mM concentrations or 10 mM EDTA without calcium. Relative activity is expressed as percentage of the activity of the enzyme in absence of calcium.

B. Size-exclusion chromatography followed by immediate measurement of enzyme activity was performed to determine the possible oligomeric state of the enzyme. Purified MaOC1 was applied to a Superdex 200 size-exclusion chromatography (SEC) column connected to ÄKTA FPLC system and 0.5 ml fractions were collected for enzymatic analysis.

C. Proteolytic activities of the wild type MaOC1 and its R219A mutant were measured and relative activity is expressed as percentage of the activity of the wild type. All assays were performed using 10 μM substrate and 1 nM enzyme concentrations in 20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM DTT at 25°C. All experiments were done as three independent assays. Statistical analysis was performed with GraphPad Prism 5.0 Software. Mean values with standard deviation are shown. Panel on the right shows purified MaOC1 R219A mutant separated on 12% SDS-PAGE and stained with Coomassie brilliant blue.

performed to test the possibility of activation by dimerisation. The majority of the protein eluted as a monomer with an elution volume corresponding to the mass of full-length MaOC1 and maximal hydrolytic activity corresponded to the maximal elution peak (Fig. 5B). To rule out the possibility of a preferred monomer state in diluted solution, we used sodium citrate, a kosmotropic salt that decreases the entropy of the system, thereby allowing ordering and oligomerisation processes to take place as was shown for paracaspase MALT-1 (Hachmann *et al.*, 2012). We assayed the activity of MaOC1 in buffers with increasing sodium citrate concentration and no increase in reaction rates has been observed (data not shown). Finally, we confirmed that proteolytic cleavage drives the activation of the MaOC1 by constructing a mutant form of MaOC1, where we replaced the P1 Arg at the cleavage site with Ala (MaOC1^{R219A} mutant). SDS-PAGE analysis shown in Fig. 5C demonstrates that the mutant remained uncleaved and no proteolytic activity could be detected. This clearly indicates that it could not autoprocess, confirming that the autoprocessing of MaOC1 occurs specifically at Arg²¹⁹ and that orthocaspases strongly depend on autoprocessing for their enzymatic activity.

Conclusion

In this work we present the identification and first characterisation of orthocaspases, evolutionarily primitive prokaryotic caspase homologues that, in contrast to other members of the caspase-like superfamily, contain a well-conserved catalytic p20-like domain but lack sequence homology with the small p10-like domain. Despite this structural uniqueness, our results demonstrate that ortho-

caspases are members of the C14 family of proteins belonging to the CD clan of proteases. In caspase-1, three conserved positively charged residues (Arg¹⁷⁹, Gln²⁸³ and Arg³⁴¹) form a basic S1 pocket that underlines the Asp-P1 preference for clan CD proteases (Fuentes-Prior and Salvesen, 2004). The alignment of the caspase sequences with metacaspases and paracaspases reveals striking structural conservation as paracaspase MALT-1 uses acidic residues at the same location to form an acidic S1 pocket (Yu *et al.*, 2011). In addition to highly conserved Asp³⁶⁵ and Asp⁴⁶² (according to MALT-1 residue numbering) in all members of the caspase-like family, Glu⁵⁰⁰ is involved in formation of acidic S1 pocket in MALT-1. Recently, Asp³⁵³ in *Nicotiana tabacum* L. metacaspase (NtMC1), corresponding to Glu⁵⁰⁰ in MALT-1 or Asp²⁶⁶ in AtMC9, was identified as the residue in p10 domain that may contribute to the formation of the substrate-binding pocket (Acosta-Maspons *et al.*, 2014). Although this negatively charged residue is conserved in MaOC1 (Glu²⁴⁷) (Fig. S1), the region surrounding it shows little sequence similarity to otherwise well conserved consensus sequence defining the p10 domain (SGCXDXQTSADV). This sequence forms the 280-loop in TbMC2, which in the presence of calcium undergoes conformational changes and stabilises substrates in the active site (McLuskey *et al.*, 2012). As no change in activity is observed in the presence or absence of calcium in MaOC1, the catalytic role of Glu²⁴⁷ remains to be determined.

The MaOC1 preference for cleavage after basic amino acid residues is common to caspase-like proteins characterised up to now. In contrast to type I metacaspases that readily cleave after Arg or Lys residues, type II metacaspases AtMC4 and AtMC9 exhibit somehow lower catalytic

activities towards substrates with Lys at P1 positions (Vercammen *et al.*, 2004; Watanabe and Lam, 2005), a feature observed also in cyanobacterial orthocaspase MaOC1. In the absence of the three-dimensional structure, we are unable to give explanation for this characteristic. Despite extensive efforts, no adequate computational model of the MaOC1 p20 domain could be constructed due to very low overall similarity of the p20 domain to known metacaspase or paracaspase protein structures (as evident from Fig. S1).

This study suggests that measurements of caspase-like activity using caspase-specific substrates (harbouring Asp at P1) should not be used for indication of orthocaspase catalytic activities in prokaryotes. Instead, we encourage the use of the substrates with positively charged Arg or Lys residues at P1 position, as it has already been suggested by Tsiatsiani for measurement of metacaspase activities (Tsiatsiani *et al.*, 2011). Intriguingly, elevated caspase-3-like activity was observed in metabolic stress-induced *Xanthomonas* (Wadhawan *et al.*, 2014) as well as in *E. coli* MG1655 (Wadhawan *et al.*, 2013), and upregulated levels of caspase-like genes were reported to be paralleled by caspase activity rates in *Trichodesmium* (Bar-Zeev *et al.*, 2013). In light of our results, these data suggest that a yet unidentified class of enzymes but not orthocaspases are most likely responsible for the detected caspase-like activities in prokaryotic organisms. It has been speculated that cryptic caspase-like proteases might be activated downstream of orthocaspases/metacaspases (Tsiatsiani *et al.*, 2011), but as caspase-3-like activity in *M. aeruginosa* was not found to be inhibited in the presence of the synthetic caspase inhibitor Z-VAD-FMK (Ding *et al.*, 2012), additional data are needed to clarify this observation.

Our results support the hypothesis of Choi and Berges (2013), who proposed that type I metacaspases and metacaspase-like proteins are two ancestral forms of metacaspases that evolved into type II and type III metacaspases, in plants and protozoa respectively. Their evolutionary position is further illustrated by the most simplistic mode of activation without the requirement of additional cofactors. Alignment of metacaspases and orthocaspases reveals that not all Asp residues that were in TbMC2 shown to be involved in calcium co-ordination are conserved in MaOC1 (Fig. S1). Little is known about the importance of calcium as a signalling molecule in prokaryotes and even less about its role in cyanobacteria, but the hypothetical concept stating that first forms of calcium signalling occurred immediately after the primary endosymbiosis (Blackstone, 2014) could explain the absence of calcium-coordinating mechanisms in orthocaspases.

Despite increased data on identification of potent cyanobacterial peptidase inhibitors, we report the first biochemical characterisation of a cyanobacterial cysteine

peptidase. Orthocaspases with their high catalytic efficiencies and strict substrate specificities are therefore compelling proteases not only for detailed kinetic studies but also for elucidation of their potent role in prokaryotic programmed cell death.

Experimental procedures

Identification of *M. aeruginosa* caspase homologue

Microcystis aeruginosa PCC 7806 caspase-like proteins were identified performing the DELTA-BLAST search of the non-redundant protein sequence database using the sequence of previously identified caspase-like protein in *Synechocystis* sp. PCC 6803 (Jiang *et al.*, 2010). Structure analyses of the obtained sequences were performed using the SMART tool relying on profile-hidden Markov models (Schultz *et al.*, 1998). Domains were identified using protein sequence analysis and classification tool InterPro (Mitchell *et al.*, 2015). The identified caspase-like sequences were aligned using PROMALS (Alberts *et al.*, 1994) with default alignment parameters. Unrooted phylogenetic tree based on aligned sequences was constructed using Neighbour-Joining method of the MEGA software (version 6) (Tamura *et al.*, 2013).

Construction of plasmids

The coding sequence for *M. aeruginosa* orthocaspase 1 full-length protein (MaOC1) was amplified from axenic *M. aeruginosa* PCC 7806 cell lysate by PCR using the primer pair MaOC1F and MaOC1R (Table S1). The amplified sequence was ligated into the bacterial expression vector pET-28b(+) in frame with a C-terminal hexahistidine tag using *Nco*I and *Xho*I restriction sites. This construct was subsequently used as a template for preparation of MaOC1 mutants (H110A, C169A, C170A or R219A) by site-directed mutagenesis. All mutations were introduced using the QuickChange Site-directed mutagenesis Kit (Stratagene) with primers listed in Table S1, according to the manufacturer's instructions. Correct ligation, incorporation of mutations and in-frame insertion of the fragments were verified by DNA sequencing.

Protein expression and purification

Escherichia coli BL21(DE3) bacteria were transformed with the expression plasmids and grown in shaker cultures at 37 °C in overexpression medium containing 30 µg ml⁻¹ kanamycin for 6 h (Studier, 2005). After lowering the temperature to 16 °C, cultures were left shaking for additional 18 h overnight. The cell pellet collected from 400 ml of bacterial culture was resuspended in 20 ml of resuspension buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole) and sonified 3 × 6 min on ice. Following centrifugation at 30 000 × g for 10 min to remove insoluble debris, the supernatant was applied to a Ni-NTA Superflow Cartridge (Qiagen) connected to ÄKTA FPLC system, washed with the resuspension buffer and eluted in the same buffer, but containing 250 mM imidazole. The peak fractions were collected, dialysed against the

buffer containing 50 mM HEPES (pH 7.5) and 150 mM NaCl, concentrated to approximately 5 mg ml⁻¹ using an Amicon filtration unit (Millipore Corp.) equipped with a 10 kDa exclusion membrane, aliquoted and stored at -80°C.

Size-exclusion chromatography

Purified proteins were applied to a Superdex 200 size-exclusion chromatography (SEC) column connected to ÄKTA FPLC system. The column was equilibrated in 20 mM HEPES pH 7.5, 150 mM NaCl. The flow rate was 0.5 ml min⁻¹ and 0.5 ml fractions were collected for analysis.

N-terminal sequencing of protein samples

For N-terminal sequencing, proteins were resolved on a 12% polyacrylamide gel followed by Western blotting onto a PVDF membrane. The membrane was briefly stained with Coomassie brilliant blue R-250, destained and washed with water. Bands of interest were excised and analysed by N-terminal sequencing on a Procise protein sequencing system 492A (PE Applied Biosystems).

Kinetic assays

Fluorogenic substrates R-AMC, K-AMC, Boc-GRR-AMC, Boc-GKR-AMC, Boc-VLK-AMC and Ac-DEVD-AMC were obtained from Peptanova, Z-RR-AMC, Z-FR-AMC, Z-LR-AMC and Z-VVR-AMC that were from Bachem. The inhibitors were from Sigma-Aldrich, except for caspase inhibitor Z-VAD-FMK (InvivoGen) and Z-FR-FMK (Bachem). The activity of the enzyme was monitored by measuring the release of fluorescent group AMC (7-amino-4-methylcoumarin) from the substrates at excitation and emission wavelengths of 383 nm and 455 nm, respectively, in a Perkin-Elmer LS50B spectrofluorimeter. Assays were performed in 20 mM HEPES (pH 7.5) containing 150 mM NaCl and 5 mM DTT at 25°C. Michaelis-Menten constants (k_{cat} and K_M) were determined by non-linear regression analysis using Graphpad Prism 5.0 Software (GraphPad Software, Inc., 7825 Fay Avenue, Suite 230, La Jolla, CA 92037 USA). For the determination of the pH profile of MaOC1, the buffers used were 100 mM acetate (pH 4–pH 5.5), 100 mM phosphate (pH 6–pH 7.5), 100 mM Tris (pH 8–pH 9) or 100 mM carbonate (pH 9.5–pH 10). All assays were performed in 2 ml disposable acrylic cuvettes with 4 μM substrate and 1 nM enzyme unless otherwise stated. All measurements were run repeated in three independent experiments.

Acknowledgements

This project has in part received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No. 308518, CyanoFactory.

References

- Acosta-Maspons, A., Sepulveda-Garcia, E., Sanchez-Baldoquin, L., Marrero-Gutierrez, J., Pons, T., Rocha-Sosa, M., and Gonzalez, 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.
- Adams, J.M. (2003) Ways of dying: multiple pathways to apoptosis. *Genes Dev* **17**: 2481–2495.
- Alberts, G.F., Hsu, D.K., Peifley, K.A., and Winkles, J.A. (1994) Differential regulation of acidic and basic fibroblast growth factor gene expression in fibroblast growth factor-treated rat aortic smooth muscle cells. *Circ Res* **75**: 261–267.
- Aravind, L., and Koonin, E.V. (2002) Classification of the caspase-hemoglobinase fold: detection of new families and implications for the origin of the eukaryotic separins. *Proteins* **46**: 355–367.
- Aravind, L., V., Dixit, M., and Koonin, E.V. (2001) Apoptotic molecular machinery: vastly increased complexity in vertebrates revealed by genome comparisons. *Science* **291**: 1279–1284.
- Asplund-Samuelsson, J., Bergman, B., and Larsson, J. (2012) Prokaryotic caspase homologs: phylogenetic patterns and functional characteristics reveal considerable diversity. *PLoS ONE* **7**: e49888.
- Bar-Zeev, E., Avishay, I., Bidle, K.D., and Berman-Frank, I. (2013) Programmed cell death in the marine cyanobacterium *Trichodesmium* mediates carbon and nitrogen export. *ISME J* **7**: 2340–2348.
- Blackstone, N.W. (2014) The impact of mitochondrial endosymbiosis on the evolution of calcium signaling. *Cell Calcium* **57**: 133–139.
- Bozhkov, P.V., Suarez, M.F., Filonova, L.H., Daniel, G., Zamyatnin, A.A., Jr, Rodriguez-Nieto, S., et al. (2005) Cysteine protease mclI-Pa executes programmed cell death during plant embryogenesis. *Proc Natl Acad Sci USA* **102**: 14463–14468.
- 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.
- Ding, Y., Gan, N., Li, J., Sedmak, B., and Song, L. (2012) Hydrogen peroxide induces apoptotic-like cell death in *Microcystis aeruginosa* (Chroococcales, Cyanobacteria) in a dose-dependent manner. *Phycologia* **51**: 567–575.
- Earnshaw, W.C., Martins, L.M., and Kaufmann, S.H. (1999) Mammalian caspases: structure, activation, substrates, and functions during apoptosis. *Annu Rev Biochem* **68**: 383–424.
- Franguel, L., Quillardet, P., Castets, A.M., Humbert, J.F., Matthijs, H.C., Cortez, D., et al. (2008) Highly plastic genome of *Microcystis aeruginosa* PCC 7806, a ubiquitous toxic freshwater cyanobacterium. *BMC Genomics* **9**: 274.
- Fuentes-Prior, P., and Salvesen, G.S. (2004) The protein structures that shape caspase activity, specificity, activation and inhibition. *Biochem J* **384**: 201–232.
- Hachmann, J., Snipas, S.J., van Raam, B.J., Cancino, E.M., Houlihan, E.J., Poreba, M., et al. (2012) Mechanism and specificity of the human paracaspase MALT1. *Biochem J* **443**: 287–295.
- He, R., Drury, G.E., Rotari, V.I., Gordon, A., Willer, M., Farzaneh, T., et al. (2008) Metacaspase-8 modulates programmed cell death induced by ultraviolet light and H₂O₂ in *Arabidopsis*. *J Biol Chem* **283**: 774–783.
- Jiang, Q., Qin, S., and Wu, Q.Y. (2010) Genome-wide com-

- parative analysis of metacaspases in unicellular and filamentous cyanobacteria. *BMC Genomics* **11**: 198.
- Lam, E., and Zhang, Y. (2012) Regulating the reapers: activating metacaspases for programmed cell death. *Trends Plant Sci* **17**: 487–494.
- Landgrebe, J., Dierks, T., Schmidt, B., and von Figura, K. (2003) The human SUMF1 gene, required for posttranslational sulfatase modification, defines a new gene family which is conserved from pro- to eukaryotes. *Gene* **316**: 47–56.
- Larkin, R.M., Alonso, J.M., Ecker, J.R., and Chory, J. (2003) GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Science* **299**: 902–906.
- Lee, N., Gannavaram, S., Selvapandiyan, A., and Debrabant, A. (2007) Characterization of metacaspases with trypsin-like activity and their putative role in programmed cell death in the protozoan parasite *Leishmania*. *Eukaryot Cell* **6**: 1745–1757.
- 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.
- Madeo, F., Herker, E., Maldener, C., Wissing, S., Lachelt, S., Herlan, M., et al. (2002) A caspase-related protease regulates apoptosis in yeast. *Mol Cell* **9**: 911–917.
- Mitchell, A., Chang, H.Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., et al. (2015) The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res* **43**: D213–D221.
- 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.
- Ojha, M., Cattaneo, A., Hugh, S., Pawlowski, J., and Cox, J.A. (2010) Structure, expression and function of *Allomyces arbuscula* CDP II (metacaspase) gene. *Gene* **457**: 25–34.
- Pei, J., and Grishin, N.V. (2007) PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* **23**: 802–808.
- Riedl, S.J., and Salvesen, G.S. (2007) The apoptosome: signalling platform of cell death. *Nat Rev Mol Cell Biol* **8**: 405–413.
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci USA* **95**: 5857–5864.
- Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**: 207–234.
- Sundstrom, J.F., Vaculova, A., Smertenko, A.P., Savenkov, E.I., Golovko, A., Minina, E., et al. (2009) Tudor staphylococcal nuclease is an evolutionarily conserved component of the programmed cell death degradome. *Nat Cell Biol* **11**: 1347–1354.
- Szallies, A., Kubata, B.K., and Duzsenko, M. (2002) A metacaspase of *Trypanosoma brucei* causes loss of respiration competence and clonal death in the yeast *Saccharomyces cerevisiae*. *FEBS Lett* **517**: 144–150.
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**: 2725–2729.
- Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zaviyalov, 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.
- Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., et al. (2004) Type II metacaspases Atmc4 and Atmc9 of *Arabidopsis thaliana* cleave substrates after arginine and lysine. *J Biol Chem* **279**: 45329–45336.
- Vercammen, D., Belenghi, B., van de Cotte, B., Beunens, T., Gavigan, J.A., De Rycke, R., et al. (2006) Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for metacaspase 9. *J Mol Biol* **364**: 625–636.
- Wadhawan, S., Gautam, S., and Sharma, A. (2013) A component of gamma-radiation-induced cell death in *E. coli* is programmed and interlinked with activation of caspase-3 and SOS response. *Arch Microbiol* **195**: 545–557.
- Wadhawan, S., Gautam, S., and Sharma, A. (2014) Involvement of proline oxidase (PutA) in programmed cell death of *Xanthomonas*. *PLoS ONE* **9**: e96423.
- Watanabe, N., and Lam, E. (2005) Two *Arabidopsis* metacaspases AtMCP1b and AtMCP2b are arginine/lysine-specific cysteine proteases and activate apoptosis-like cell death in yeast. *J Biol Chem* **280**: 14691–14699.
- Watanabe, N., and Lam, E. (2011) Calcium-dependent activation and autolysis of *Arabidopsis* metacaspase 2d. *J Biol Chem* **286**: 10027–10040.
- Wong, A.H., Yan, C., and Shi, Y. (2012) Crystal structure of the yeast metacaspase Yca1. *J Biol Chem* **287**: 29251–29259.
- 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.
- Zhang, Y., and Lam, E. (2011) Sheathing the swords of death: post-translational modulation of plant metacaspases. *Plant Signal Behav* **6**: 2051–2056.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.