PNAS

Structure–function study of a Ca²⁺-independent metacaspase involved in lateral root emergence

Simon Stael[a,b,c,1] (iD), Igor Sabljić[c,1] (iD), Dominique Audenaert[d,e], Thilde Andersson[f] (iD), Liana Tsiatsiani[a,b] (iD), Robert P. Kumpf[2] (iD), Andreu Vidal-Albalat[f], Cecilia Lindgren[f] (iD), Dominique Vercammen[a,b] (iD), Silke Jacques[a,b], Long Nguyen[d,e], Maria Njo[a,b], Álvaro D. Fernández-Fernández[a,b] (iD), Tine Beunens[a,b], Evy Timmerman[g,h], Kris Gevaert[g,h] (iD), Marc Van Montagu[a,b,3] (iD), Jerry Ståhlberg[c] (iD), Peter V. Bozhkov[c], Anna Linusson[f] (iD), Tom Beeckman[a,b], and Frank Van Breusegem[a,b,3] (iD)

Metacaspases are part of an evolutionarily broad family of multifunctional cysteine proteases, involved in disease and normal development. As the structure–function relationship of metacaspases remains poorly understood, we solved the X-ray crystal structure of an *Arabidopsis thaliana* type II metacaspase (AtMCA-IIf) belonging to a particular subgroup not requiring calcium ions for activation. To study metacaspase activity in plants, we developed an in vitro chemical screen to identify small molecule metacaspase inhibitors and found several hits with a minimal thioxodihydropyrimidine-dione structure, of which some are specific AtMCA-IIf inhibitors. We provide mechanistic insight into the basis of inhibition by the TDP-containing compounds through molecular docking onto the AtMCA-IIf crystal structure. Finally, a TDP-containing compound (TDP6) effectively hampered lateral root emergence in vivo, probably through inhibition of metacaspases specifically expressed in the endodermal cells overlying developing lateral root primordia. In the future, the small compound inhibitors and crystal structure of AtMCA-IIf can be used to study metacaspases in other species, such as important human pathogens, including those causing neglected diseases.

AtMCA-IIf crystal structure | cysteine protease | lateral root development | metacaspase | small chemical inhibitor

Metacaspases are part of the C14 family of cysteine-dependent proteases together with caspases and paracaspases (1). Metacaspases are conserved throughout plants, fungi, protists, and bacteria, whereas caspases and paracaspases are present in mammals. The interest in metacaspases was initially sparked by their sequence similarities, entailing a common hemoglobinase structural fold and a conserved histidine/cysteine dyad in the catalytic domain, with the well-studied mammalian caspases and hence a suspected functional involvement in programmed cell death (PCD) outside metazoa (2, 3). Two decades later, it becomes clear that metacaspases are a diversified clade of multifunctional proteins that diverge from caspases in protein structure, substrate preference, and functionalities (4–9). Unlike caspases that cleave their substrate proteins after aspartate, metacaspases cleave exclusively after the basic amino acids arginine and lysine. Metacaspases are classified according to their domain organization: Type-I metacaspases have an N-terminal prodomain that is absent in type-II metacaspases, and the latter have a longer linker connecting the p20 and p10 regions. Type III metacaspases were more recently identified in genomes of phytoplanktonic protists and have a rearrangement of the p10 region preceding the p20, but otherwise have a similar activity profile to type I and type II metacaspases (10, 11). The conserved p20 and p10 regions together build up the core protease domain. They were designated p20 and p10 because activated enzymes typically show two bands at 20 kDa and 10 kDa, respectively, on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gels. Most metacaspases require calcium ions (Ca²⁺) and a neutral pH for activity, with the notable exception of a subgroup of type II metacaspases present only in vascular plants that function optimally at mildly acidic pH (around pH 5.5) and do not require Ca²⁺ (4). Apart from their role in pathogen-induced PCD (in plants also called the hypersensitive response) (12, 13), metacaspases function in senescence, aging and protein aggregate clearing (14–16), clearance of cell corpses (17), wound-induced damage-associated molecular pattern signaling (18, 19), and developmental cell death events (17).

To date, two protein crystal structures of type I MCAs are published from *Saccharomyces cerevisiae* and *Trypanosoma brucei* (TbMCA-Ib) (20, 21), and one type II calcium-dependent MCA from *Arabidopsis thaliana* (AtMCA-IIa/AtMC4) (22). The core of all three structures is a typical caspase/hemoglobinase alpha–beta–alpha sandwich fold (23). Differences exist within the type-specific regions: In the type I MCA structures, the N-terminal part of the protein binds in the active site, which explains why these enzymes are not fully active before

Significance

Metacaspases are widely expressed throughout all kingdoms of life, except mammals. A detailed understanding of their protease activity has been hampered by a paucity of structural information and small molecule inhibitors. We solved the protein structure of an atypical plant metacaspase and, from a chemical library of 10,000 compounds, identified an unknown class of small molecule metacaspase inhibitors. These inhibitors disrupted lateral root emergence in plants, effectively overcoming genetic redundancy. The structural data can aid the understanding of activation mechanism for metacaspases and related proteases, including caspases and paracaspases in mammals, and the inhibitors can form the basis for development of tools and drugs to target metacaspases.

removing the N-terminal prodomain. However, it is not yet understood how Ca^{2+} activates the protease since the proposed Ca^{2+}-binding sites are far away from the active site. Furthermore, the cysteine and histidine residues that make up the catalytic dyad are too far apart and, in a conformation where the catalytic histidine cannot act as base responsible for cysteine deprotonation, which is otherwise usually proposed as mechanism for other cysteine proteases like papains. This conundrum presents itself also in the caspases for which the catalytic mechanism is not completely explained (24). In the AtMCA-IIa structure, part of the linker region binds in the active site, which explains the need for autocleavage in this region for activation (22). The resolved structures of AtMCA-IIa do not contain Ca^{2+}, therefore it remains unclear where Ca^{2+} binds in type II MCAs, and how it triggers activation. Similar to the conundrum in type I MCAs, AtMCA-IIa structures also show an inactive form of the enzyme with the catalytic residues too far apart. In addition, for the Ca^{2+}-independent metacaspases, a representative crystal structure is lacking, which further impedes the understanding of metacaspase activation mechanisms.

Lateral roots make up a significant part of the plant root system and ensure optimal nutrient absorption, water uptake, and soil anchorage of plants. Their developmental trajectory is remarkable, as they originate from precursor cells deep within the main root and have to push through the terminally differentiated outer cell layers of the root. Lateral roots originate from pericycle cells that acquire lateral root founder cell identity and divide asymmetrically, a process that is tightly regulated by the plant hormone auxin. Under the control of auxin, these founder cells develop into a lateral root primordium (LRP), which then grows through the overlying endodermis and cortex until it eventually breaks through the epidermis and emerges from the primary root. Evidence has been provided that a tight integration of chemical and mechanical signaling between the LRP and the surrounding tissue is essential for its proper development while the overlying layers have to accommodate the outgrowth of lateral roots (25). In Arabidopsis, the pericycle-overlying endodermal cells represent a major obstacle for the developing LRP because of the presence of their lignified Casparian strips in the primary cell wall. Both endodermal cell elimination and accommodation due to cell shape changes have been reported to contribute to the outgrowth of primordia through the endodermis (25, 26). The reported specific spatiotemporal transcriptional induction of *AtMCA-IIf*/*AtMC9* in the endodermis in front of an initiation LRP, suggested its involvement in PCD of the endodermis to clear the way of the developing primordium. However, most likely due to functional redundancy, single *atmca-IIf* mutants do not experience problems during lateral root emergence (26). The Arabidopsis genome contains three type I and six type II MCAs (2, 4, 5).

Here, we took a chemical biology approach to address the role of metacaspases in lateral root formation. First, we solved the crystal structure of AtMCA-IIf, the sole Ca^{2+}-independent metacaspase of Arabidopsis. We then developed an in vitro chemical screen and identified a series of small molecule metacaspase inhibitors with a minimal thioxodihydropyrimidine-dione (TDP) structure, some of which could specifically inhibit activity of AtMCA-IIf. Finally, we show that an AtMCA-IIf-specific TDP-containing inhibitor (TDP6) was effective at suppressing LRP emergence in vivo.

Results

The Crystal Structure of AtMCA-IIf C147A Mutant Is the First Structure of a Ca^{2+}-Independent Metacaspase.
Recombinant AtMCA-IIf wild-type and catalytically inactive C147A mutant,

with N-terminal His-tag (4), were expressed in *E. coli*. No crystals were obtained of the wild-type enzyme, presumably due to autocleavage, but the C147A mutant did crystallize, and the structure was solved and refined at 1.95 Å resolution and R/R$_{free}$ values of 0.17/0.21. At position 271, we found an unintended Phe to Tyr mutation, but it does not seem to induce any structural disturbances. Twenty amino acid residues are not seen in the structure—the six first N-terminal residues, seven in the first gap (102-VKSAHPF-108), and seven in a second gap (166-SSNISPA-172). The structure of AtMCA-IIf C147A is shown in comparison with the Ca^{2+}-dependent type-II (AtMCA-IIa C139A) and type-I (TbMCA-Ib C213G) metacaspases (Fig. 1). All are built from a core protease domain comprising the p20 (with catalytic His-Cys dyad) and p10 regions. For details and crystallography statistics, see *SI Appendix, Text* and Table S1.

Structure Comparison of AtMCA-IIf Reveals an Overall Similarity with AtMCA-IIa.
As expected, AtMCA-IIf is more similar to AtMCA-IIa than to TbMCA-Ib, both in terms of primary sequence organization (Fig. 1A), overall structure (Fig. 1B), and autoinhibitory motif (AIM) (Fig. 1C). In the type-I TbMCA-Ib, the active site is blocked by the N-terminal prodomain, which is absent in the type-II proteases (Fig. 1A). Instead, in the region between p20 and p10, they contain an AIM followed by a domain of unknown function (UNK). The AIM embedded in the active site differs in sequence (Fig. 1C). In AtMCA-IIf, Arg183 (180-ITSRALP-186) is a key residue that is cleaved upon activation, whereas Lys225 (222-AKDKSLP-228) is the corresponding residue in AtMCA-IIa (4, 28). In the following, residue numbers refer to AtMCA-IIf unless otherwise indicated. The linker (residues 165 to 179) is much shorter in AtMCA-IIf with 15 residues (of which six are missing in the structure), compared to around 70 in AtMCA-IIa, of which 60 are missing from the current structure. The UNK domain (residues 187 to 241), not present in type-I, is also shorter in AtMCA-IIf. It contains only helices, four in AtMCA-IIf and eight in AtMCA-IIa, with 55 and 105 residues, respectively (Fig. 1B).

Binding of the AIM in the Active Site Resembles AtMCA-IIf Substrate Preference.
At the catalytic center, well-defined and unambiguous electron density (*SI Appendix,* Fig. S1) shows the binding of the AIM with its backbone carbonyl of Arg183 placed between the catalytic dyad residues (His95 and Cys147Ala; Fig. 1C). Thus, with the catalytic residues on opposite sides of the peptide bond to be cleaved, the histidine will not be in direct contact with and cannot act as activating proton acceptor to the catalytic cysteine (Cys147 in AtMCA-IIf wild type). Instead, His95 is suitably positioned for protonation of the peptide nitrogen of the scissile bond, suggesting a role in the first, acylation step of the reaction (*SI Appendix,* Fig. S2A).

Whereas the AIM is quite different in TbMCA-Ib, it is similar in the type-II enzymes. The sidechain of Arg183 (and Lys225 in AtMCA-IIa) is entirely buried in the S$_1$ pocket (Fig. 1C). In the N-terminal direction from Arg183, four residues (179-TITS-182) form a beta strand, bound antiparallel with the 258 to 261 β-strand in p10. Their sidechains are exposed and no well-defined pockets are distinguished, except for a predominantly hydrophobic S$_4$ pocket, suggesting a preference for a hydrophobic residue at the P4 position of substrates (*SI Appendix,* Fig. S3). In the C-terminal direction, no S$_1$' pocket can be distinguished and all residues in the vicinity are the same in AtMCA-IIf and AtMCA-IIa (*SI Appendix,* Fig. S2B). There is a clear hydrophobic S$_2$' pocket though, harboring a leucine residue from the AIM in both
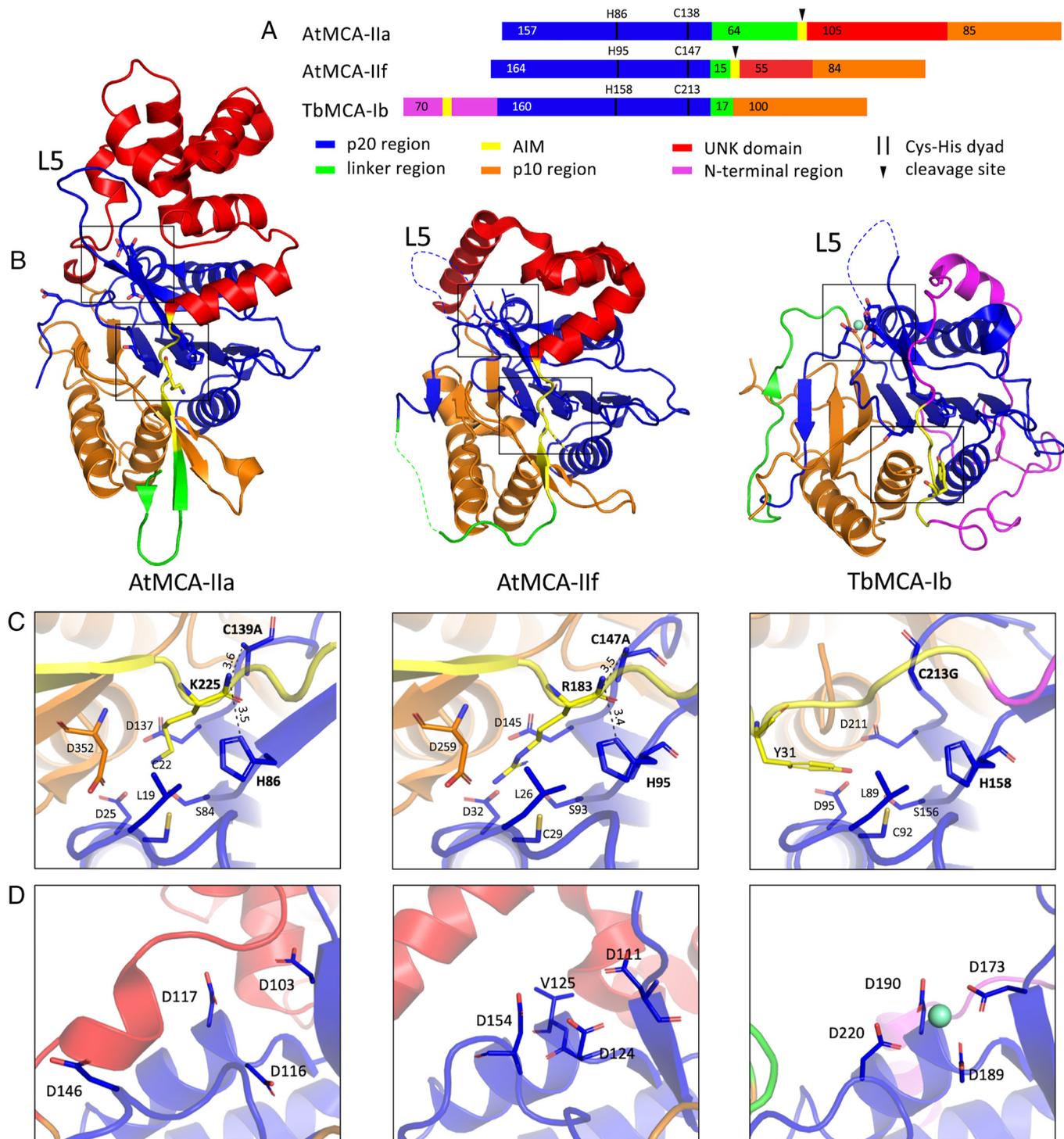
Fig. 1. Comparison of the obtained AtMCA-IIf C147A crystal structure to available MCA structures. (*A*) Scheme of the protein sequences drawn to scale and colored by region (color key in figure). Regions p20 and p10 were designated as described (27). The positions of the catalytic dyad His and Cys residues are indicated as well as the number of amino acids in each region. (*B*) Overall structures of AtMCAII-a (PDB 6W8S; C139A mutation), AtMCA-IIf (PDB 8A53; C147A mutation), and TbMCA-Ib (PDB 4AFP; C213G mutation) color-coded by region. Selected residues are shown in stick representation, at the catalytic center (Lower box), and at the putative calcium-binding site (Upper box). The position of loop L5 is indicated, which is visible in AtMCA-IIa but not in the two other structures. (*C*) Catalytic center with autoinhibitory motifs bound. (*D*) Putative calcium-binding site, with Sm^{3+} bound in the TbMCA-Ib structure (*Right*).

structures, where several residues are conserved. Also, the first helix of the UNK domain (187 to 197) is oriented slightly different between the two AtMCAs, which may influence the S$_2'$ pocket formation. Next, there is no clear S$_3'$ pocket. A proline sidechain points out in solution in both structures, after which the peptide chain turns about 90° and continues into the first α-helix of the UNK domain.

To assess substrate specificity, a positional proteomics approach was used to screen the root proteome of Arabidopsis plants for proteins that are differentially cleaved in AtMCA-IIf loss- or gain-of-function mutants by N-terminal combined fractional diagonal chromatography (COFRADIC) (29) (*SI Appendix,* Fig. S4*A*). Identification of so-called neo-N-terminal peptides revealed a sub-fraction of proteins in the Arabidopsis proteome that are

potentially cleaved by AtMCA-IIf and the cleavage sites in those proteins (*SI Appendix* Fig. S4C and Dataset S1). Enriched amino acid sequences within the P4 to P4′ position spanning the cleavage site (P4-P3-P2-P1↓P1′-P2′-P3′-P4′) (30) of the Arg/Lys-cleaved proteins display a cleavage signature that is similar to the one previously described for AtMCA-IIf in protein substrates from 2-d-old seedlings (31) (*SI Appendix,* Fig. S4E). Based on the crystal structure, we were now able to match the substrate cleavage specificity to the active site pockets of AtMCA-IIf. Folded proteins were used for the in vitro and in vivo COFRADIC analyses, so that an enrichment of charged amino acid residues on the surface of substrates can be expected. However, no clear explanation is at hand for the particular enrichment of basic amino acid residues in P3, neither for acidic side chains in P1′ position, as clear pockets were missing from the AtMCA-IIf structure and no known technical issues with COFRADIC can explain the observed signature. Convincingly, mostly hydrophobic amino acids were found at the P2′ position that can fit the hydrophobic S2′ pocket. To a lesser extent, there was a fit for the hydrophobic S4 pocket, where only in the previous study by Tsiatsiani et al. there was an enrichment for Ala in P4 (as well as Glu and Val; *SI Appendix,* Fig. S4E). Altogether, an AtMCA-IIf substrate cleavage signature emerged wherein P3 and P1 are occupied by basic, P1′ by acidic, and P4 and P2′ by hydrophobic amino acid residues. This is reflected at least in the active site pockets S1 (acidic) and S2′ (hydrophobic) from the newly acquired AtMCA-IIf crystal structure.

Absence of Calcium-Binding Sites in AtMCA-IIf. To assess the calcium independency of AtMCA-IIf, we compared the proposed Ca^{2+}-binding sites between the available 3D crystal structures (Fig. 1D). TbMCA-Ib has a Sm^{3+} ion bound as a Ca^{2+} surrogate, and this site is almost identical in both type I structures (20, 21). The Sm^{3+} ion is coordinated by four Asp residues and two water molecules, while in AtMCA-IIf, only three Asp residues are conserved (the fourth is replaced by valine), which could be connected to the AtMCA-IIf Ca^{2+}-independency (Fig. 1D). In

AtMCA-IIa, an alternative Ca^{2+}-binding site has been proposed (22), composed of four negatively charged amino acids in a row (96-EDDD-99) at the turn of loop L5 that is embedded in a positively charged pocket of the UNK domain (Fig. 1B). Negatively charged residues are conserved here in both type I and II of Arabidopsis, except for AtMCA-IIf, which is also lacking the part of the UNK domain that supports loop L5 in AtMCA-IIa (*SI Appendix,* Fig. S5). The corresponding loop of AtMCA-IIf is disordered and not visible in the crystal structure (102-VKSAHPF-108) and does not contain a similar cluster of negatively charged residues. These structural differences could explain why AtMCA-IIf is not activated by Ca^{2+}.

A Chemical Screen Finds Metacaspase Small Molecule Inhibitors with a Common Thioxodihydropyrimidinedione (TDP) Scaffold. To overcome potential metacaspase functional redundancy and phenotypic plasticity as a result of stable metacaspase mutations, we aimed for a chemical approach to inhibit the enzymatic activity in situ. Therefore, we developed a high-throughput adaptation of an in vitro biochemical assay based on cleavage of a fluorescently labeled VRPR-AMC tetrapeptide by recombinant AtMCA-IIf (4, 32). The miniaturized assay was used to screen a chemical library of 10,000 small organic compounds at a concentration of 10 μM to identify inhibitors of AtMCA-IIf (Fig. 2A). The screen was performed twice against the 10,000 compounds and together led to 177 compounds that showed at least 75% inhibition. These 177 compounds were then selected for a third screen. Together, 81 compounds were identified that reduced the metacaspase activity on average with at least 75% in all three screens (Dataset S2). Investigation of the chemical diversity of the inhibitors by performing a clustering based on the compounds' chemical structures revealed an enrichment for hits containing a TDP scaffold (Fig. 2B and Dataset S2). This TDP substructure was found in 30 (37%) of the hit compounds, compared to 79 (0.8%) present in the entire screening library, which made us to prioritize the TDP-containing hits for further investigation.
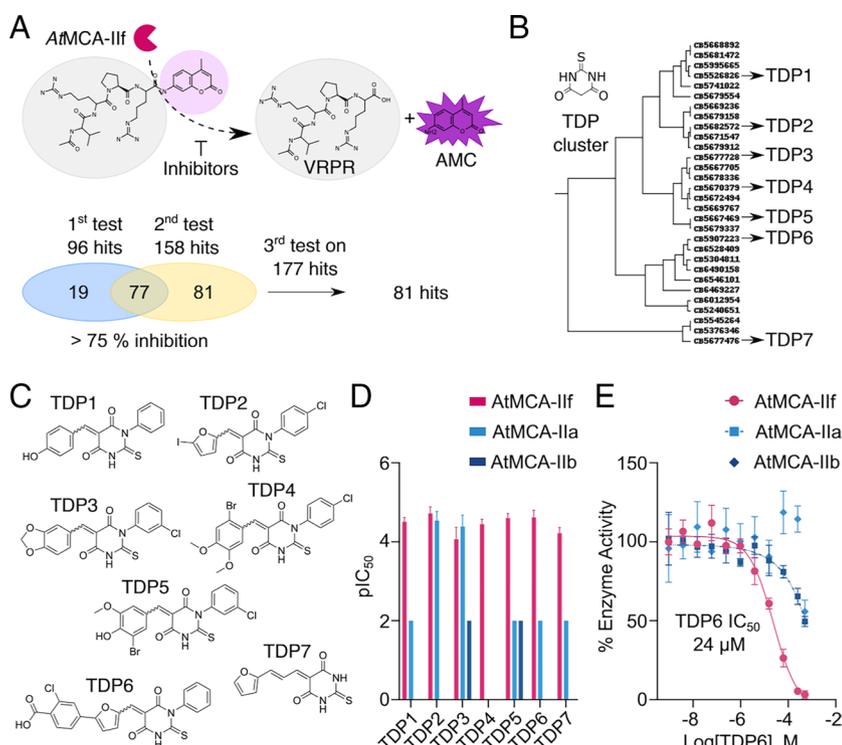
Fig. 2. A chemical screen reveals a cluster of AtMCA-IIf protease inhibitors that share a TDP substructure. (A) Schematic overview of the chemical screen against AtMCA-IIf. The full screen (10.000 compounds) was performed twice resulting in the identification of 96 and 158 compounds with an inhibitory effect larger than 75%. The resulting 177 compounds were retested and 81 compounds were retained with an average inhibition larger than 75% in the three tests. (B) Structure-based clustering of chemical compounds with 75% AtMCA-IIf inhibitory effect. From 81 final hits, 30 compounds contain a thioxodihydropyrimidinedione (TDP) substructure. (C) Chemical structures of the seven representative compounds from each TDP subcluster based on their inhibitory effect against the VRPR-ase activity of AtMCA-IIf. (D) Inhibitory activity of the TDP compounds against three metacaspases expressed as pIC_{50}. Compounds that showed some inhibitory activity ($IC_{50} > 200$ μM) at the highest tested concentrations are set to $pIC_{50} = 2$ to illustrate the inhibitory difference between the enzymes. (E) Dose–response curve for inhibitory activity for TDP6 against three metacaspases, obtained from the activity-based in vitro assay.

Dose-Response Analysis of TDP-Containing Compounds to Probe Selectivity toward AtMCA-IIf, AtMCA-IIa, and AtMCA-IIb. Based upon their inhibitory activity and chemical diversity, seven compounds containing the TDP scaffold were selected for a full dose–response analysis (TDP1 to TDP7; Fig. 2*C*; and *SI Appendix,* Fig. S6 *A* and Table S2). In addition to AtMCA-IIf, the inhibitory activities of the compounds were evaluated on AtMCA-IIb/AtMC5, because of its overlapping gene expression pattern in roots (*SI Appendix,* Fig. S9), and AtMCA-IIa, the most abundant and relatively well-studied Arabidopsis metacaspase (13, 18, 22, 27, 28). Measurements were performed on commercially available compounds, of which the chemical structures had been confirmed by NMR characterization (Dataset S3). The compounds all carry different aromatic or heteroaromatic moieties linked to the TDP scaffold via a carbon–carbon double bond, generating two possible isomers, depending on the substitution pattern (E and Z). Before testing, the ratio between the E and Z isomers was determined by ^{1}H-NMR to be 1:1 for all TDPs, except TDP7 that only was found in the E configuration (Dataset S3). Thus, six out of seven TDPs are a mixture of two compounds with the same chemical formula (diastereomeric). The full dose–response analysis of the TDPs toward AtMCA-IIf corresponded well with the screening data, as all hits showed dose-dependent inhibition toward the protease, with IC_{50} values ranging from 19 to 86 μM (Fig. 2*D* and *SI Appendix,* Fig. S6*A* and Table S2). A comparison of the inhibition profiles between the proteases showed that the majority of the TDPs selectively inhibited AtMCA-IIf. Only two of the compounds (TDP2 and TDP3) inhibited the activity of AtMCA-IIa in a dose-dependent manner, exhibiting an inhibitory effect similar as for AtMCA-IIf, while additional four showed some activity at the highest concentrations. The compounds showed even less inhibitory activity against AtMCA-IIb; only two (TDP3 and TDP5) where active at the highest tested concentrations. Investigation of the inhibition profile toward AtMCA-IIf, revealed

a steep dose-dependency (Hillslope >1) for some of the TDPs. This might be related to the diastereomeric mixture of the TDP compounds, where the two compounds possible can interact differently with the protease causing the steepness of the curve. We developed an orthogonal in vitro assay to probe inhibition of metacaspase activities based on the cleavage of a substrate protein, PROPEP1 (18, 19, 22). TDP6 was selected as a representative of the inhibitors with the lowest IC_{50} of the compounds that are selective toward AtMCA-IIf in the VRPR-AMC cleavage assay. Increased cleavage of a recombinant GST-TEV-PROPEP1 fusion protein by AtMCA-IIf correlated with decreased concentrations of TDP6 in the reaction mixture (*SI Appendix,* Fig. S7 *A* and *B*), while this correlation was obscure for AtMCA-IIa and AtMCA-IIb (*SI Appendix,* Fig. S7 *C* and *D*). In conclusion, several TDP-containing small molecules were found to inhibit AtMCA-IIf enzymatic activity in vitro, without inhibiting the two other AtMCAs of the same type.

Molecular Docking Simulations to the Crystal Structure of AtMCA-IIf Suggest Binding of TDP Compounds to the Active Site Pocket. To understand how the TDPs could interact with AtMCA-IIf and inhibit the catalysis, we performed molecular docking to the active site of AtMCA-IIf of TDP6, as selective inhibitor (Fig. 2 *D* and *E*), and TDP2 and TDP3, which inhibit both AtMCA-IIf and AtMCA-IIa (Fig. 3 and *SI Appendix,* Fig. S8). The AIM was excluded from AtMCA-IIf to expose the active site prior to docking. Since the TDPs were tested experimentally as a mixture of isomers, both the E- and Z-configurations of TDP2, TDP3, and TDP6 were considered. The molecular docking revealed multiple poses of the TDP inhibitors due to the E- and Z-configurations and the apparent symmetric nature of the TDP scaffold with its phenyl and benzylidene substituents. Two molecular interaction types appeared to be the driving force for the generated binding poses in the active site of
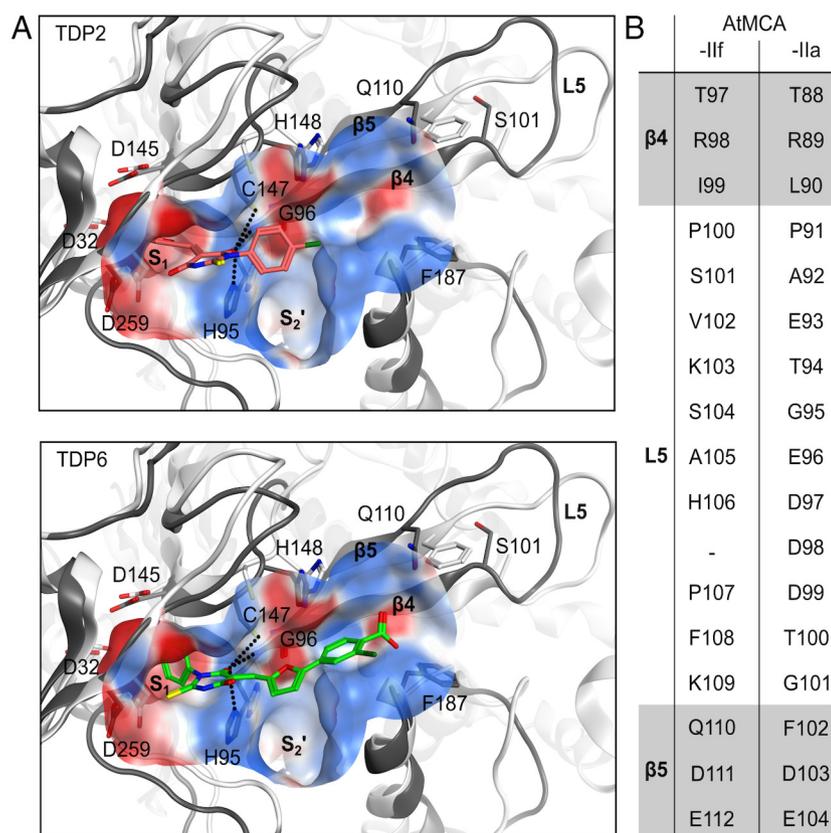
Fig. 3. Potential binding modes of TDP6 based on docking to AtMCA-IIf. (*A*) The tertiary structure and selected residues of AtMCA-IIf (PDB 8A53) are displayed in dark gray and AtMCA-IIa (PDB: 6W8S) in white. One of the aromatic substituents of the thiobarbituric acid scaffold of the E-configuration of TDP2 and TDP6 is placed in the S1 pocket of AtMCA-IIf, and the second aromatic substituent is placed along the β4 sheet. Possible hydrogen bonding interactions to the carbonyl oxygens of the thiobarbituric acid moiety are shown with dotted lines. The electrostatic surface of the active site is displayed. The L5 loop of AtMCA-IIf has been modeled for visualization purposes. Residue numbers for AtMCA-IIf are shown. (*B*) Comparison of the sequence of the L5 loop of AtMCA-IIf and AtMCA-IIa.

	AtMCA -IIf	AtMCA -IIa
β4	T97	T88
	R98	R89
	I99	L90
L5	P100	P91
	S101	A92
	V102	E93
	K103	T94
	S104	G95
	A105	E96
	H106	D97
	-	D98
	P107	D99
	F108	T100
	K109	G101
β5	Q110	F102
	D111	D103
	E112	E104

AtMCA-IIf: i) an aromatic moiety of the inhibitors (the phenyl or the benzylidene substituent) projected into the S1 pocket of the protein for electrostatic and dispersion interactions and ii) hydrogen bonding interactions between a carbonyl oxygen of the TDP scaffold and one or more –NH of the protein, (the –NH+ of His95, and the backbone –NH's of Gly96 and Cys147). Note that in the crystal structure where the AIM is included, i) corresponds to the pocket where the side chain of Arg183 is bound and ii) is the presumed oxy-anion hole where the peptide carbonyl oxygen of Arg183 is located. In the molecular dockings, it was the type of substituents on the phenyl or the benzylidene moieties together with the E- and Z-configurations that determined which one of the aromatic moieties was projected into the S1 pocket and thus which of the two carbonyls of the TDP scaffold that formed the hydrogen bonding interactions. In addition to the key interactions that were observed for all docked compounds, the TDP inhibitors with E-configuration placed the other aromatic substituent along the β4 sheet (Fig. 3A), while the docking poses for the Z-isomers varied more. In many docking poses, the other aromatic moiety of the Z-isomers was projected out to the solvent in the opposite direction of the β4 sheet, or in a few poses in the S2′ pocket where Leu185 of the AIM is located. The multiple alternatives of binding poses for this class of inhibitors might explain its unpronounced structure–activity relationship. Inspection of the docking poses showed that the larger benzylidene substituent of E-TDP6 had a more extensive interaction surface with the β4 sheet directed toward the L5 loop, that connects β4 with β5, compared to the nonselective TDP2 and TDP3 (Fig. 3A and *SI Appendix,* Fig. S8). Interestingly, while the amino acid residues forming the β4 and β5 sheets are highly similar between the two proteases, there are large differences in the L5 loop. The ten amino acid residue Pro100-Lys109 loop of AtMCA-IIf shows no sequence similarity, except for Pro91 in the L5 loop of AtMCA-IIa (Fig. 3B). Instead, there are significant differences that likely influence the dynamics of the loops and thus also the character of the β-sheet surfaces. The loop of AtMCA-IIa comprises two glycines (Gly95 and Gly101), while the loop of AtMCA-IIf instead has an additional proline (Pro107). Thus, we propose that the loss of activity for TDP6 toward AtMCA-IIa, while TDP2 and TDP3 maintain inhibition, was because the β4-sheet of AtMCA-IIa could not accommodate the larger substituent of TDP6 in a similar manner as for AtMCA-IIf due to differences in the dynamic pattern of the β-sheet surfaces.

AtMCA-IIf Spatiotemporal Expression Pattern Suggests a Role in Lateral Root Emergence. As demonstrated earlier (26), *AtMCA-IIf* is expressed in the endodermis in cells overlying developing lateral root primordia (LRP; *SI Appendix,* Fig. S9A), which suggests a role in the emergence of lateral roots through the endodermis (Fig. 4A). Furthermore, we studied the spatiotemporal expression of the other Arabidopsis metacaspase genes in roots using promoter GUS-GFP reporter lines. Similar to *AtMCA-IIf*, *AtMCA-IIb* is expressed in endodermal cells overlying developing primordia (*SI Appendix,* Fig. S9B). Besides *AtMCA-IIc/AtMC6*, all metacaspase genes were expressed in the root, although with distinctive expression levels and spatial patterns (*SI Appendix,* Fig. S10). The expression pattern of *AtMCA-IIf* was investigated in more detail with a translational promoter AtMCA-IIf-GFP fusion construct confirming expression in the endodermis during the first stages of lateral root development (Fig. 4B). As previously reported (17), AtMCA-IIf was also expressed early on in the xylem strands (adjacent to the stage II LRP, Fig. 4B, and Movie S1), which disappears in the later stages of LRP development during which the xylem loses the GFP signal (Stage V LRP, Fig. 4B, and Movie S1). Next to the tight engulfment of the endodermal cell(s) surrounding the LRP, a striking partitioning of the AtMCA-IIf -GFP fusion protein in the endodermis to the side of the cell bordering the developing lateral root could be discerned (Fig. 4C). LRP emergence in T-DNA knockout lines of *atmca-IIb*, *atmca-IIf*, and the double *atmca-IIb atmca-IIf* mutant were measured (*SI Appendix,* Fig. S11 *A* and *B*). Surprisingly, no significant differences were observed between wild-type plants and single or double knockout lines in the ratio of nonemerged lateral roots (*SI Appendix,* Fig. S11C). This confirms previous findings that phenotypes in single or double mutants have no effect on
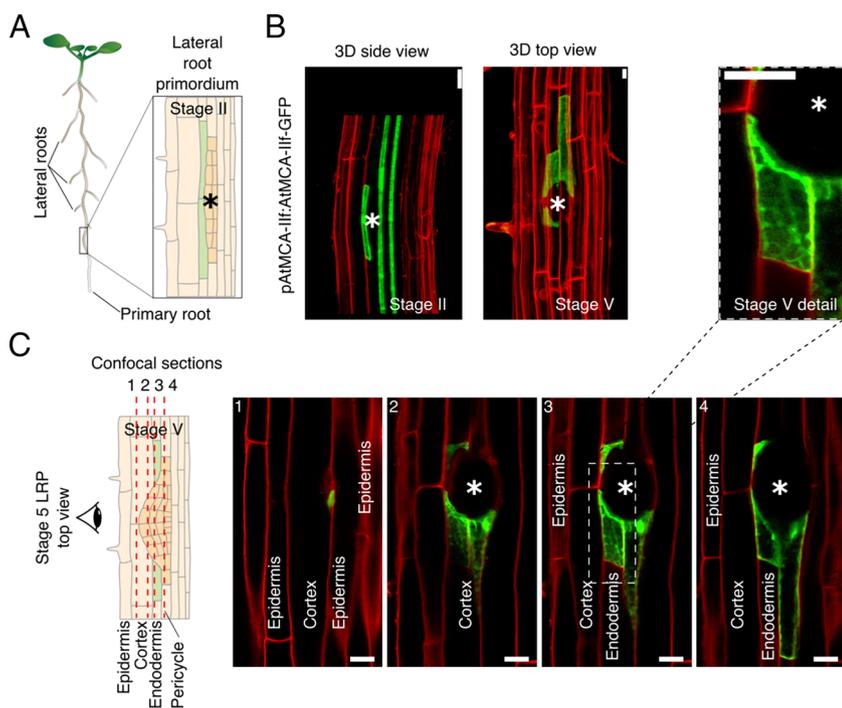
Fig. 4. AtMCA-IIf is specifically expressed in endodermal cells overlying the LRP in Arabidopsis. (*A*) Schematic overview of the site of lateral root development in the primary Arabidopsis root. (*B*) 3D reconstruction of pAtMCA-IIf:AtMCA-IIf-GFP fluorescence from confocal sections of stage 2 to 3 and stage 5 lateral root primordia (LRP). The primordium is indicated with an asterisk (*). Cell walls are stained with PI (red). (*C*) Confocal sections through the tissue layers overlying the stage 5 LRP. (Scale bar is 20 μm.)

LRP emergence possibly because of functional redundancy of metacaspases (26).

Inhibition of Metacaspase Activity by TDP6 Suppresses Lateral Root Emergence. To assess the phenotypic effect of metacaspase inhibition on LRP development, 8-d-old Arabidopsis seedlings were analyzed after a 5-d growth period in media supplemented with 50 μM of the inhibitors TDP1-7. Several compounds decreased the number of lateral roots, with TDP3 and TDP6 having the most pronounced effect on lateral root density (Fig. 5A). Unlike TDP3, TDP6 did not have a significant effect ($P < 0.05$) on primary root length, indicating that TDP6 was a more specific inhibitor of lateral root development (Fig. 5 *A* and *B* and *SI Appendix,* Fig. S12). Interestingly, TDP2, which showed a high inhibitory activity toward both AtMCA-IIf and AtMCA-IIa (Fig. 2D and *SI Appendix,* Table S2), affected lateral root development only to a minor degree. This suggests that factors unaccounted for in an in vitro assay, such as chemical stability and tissue uptake, are critical for the in vivo effect of a chemical probe.

A detailed assessment of the TDP3 and TDP6 effect on lateral root emergence showed that the proportion of nonemerged primordia in treated seedlings was on average 20% higher compared to the untreated control (Fig. 5C). LRP go through various stages of development (33), which can be conveniently scored by a root bending assay in which seedlings are turned 90° after which the root tips grow toward the shifted vector of gravity, resulting in a bend in the root and the development of a LRP in the bend. In this assay, primordia emergence appeared to have slowed down as they accumulate primarily at stage 5 in TDP6-treated seedlings (Fig. 5D). To confirm these findings by genetic means, we reduced type II metacaspase gene expression with an amiRNA construct directed against *AtMCA-IIa, b, c, d* and *-IIf* in an *AtMCA-IIe/AtMC8* T-DNA background line (*SI Appendix,* Fig. S13 *A* and *B*). Similar to TDP3 and TDP6 treatment of wild-type seedlings (Fig. 5C), two independent knockdown lines (KD1 and KD2) contained more nonemerged LRP than wild type (*SI Appendix,* Fig. S13C), although the effect was more pronounced with the TDP compounds. Detailed analysis of LRP stages showed a significantly increased number of LRP ($P < 0.05$) that have not yet protruded the endodermis (stages 1 to 3; *SI Appendix,* Fig. S12D), suggesting that the endodermis is the restraining tissue for lateral root emergence when type II metacaspase levels are down-regulated. These findings are recapitulated in a root bending assay of the KD1 and KD2 lines (Fig. 5D). Unlike with the knockdown lines, inhibition of lateral root emergence can be reverted by transfer of seedlings from TDP6 to control medium (*SI Appendix,* Fig. S14 *A* and *B*). Importantly, a reversal was observed in the part of the root that was previously exposed to TDP6 (*SI Appendix,* Fig. S14C), underscoring the importance of TDP6 as a versatile tool to tune lateral root emergence. In conclusion, where genetic ablation of AtMCA-IIf in previous and current research failed to uncover an effect on lateral root development, application of the small molecule AtMCA-IIf inhibitor TDP6 reveals a positive effect of metacaspase on LRP emergence.

Discussion

The AtMCA-IIf Crystal Structure Has Led to an Improved Understanding of Metacaspase Activity. We report here the crystal structure of a Ca^{2+}-independent type II metacaspase, AtMCA-IIf. While improving the understanding of metacaspase activity, nevertheless, this study raised further questions about the catalytic mechanism of the cysteine protease clan CD that contains, among others, metacaspases, paracaspases such as mucosa-associated
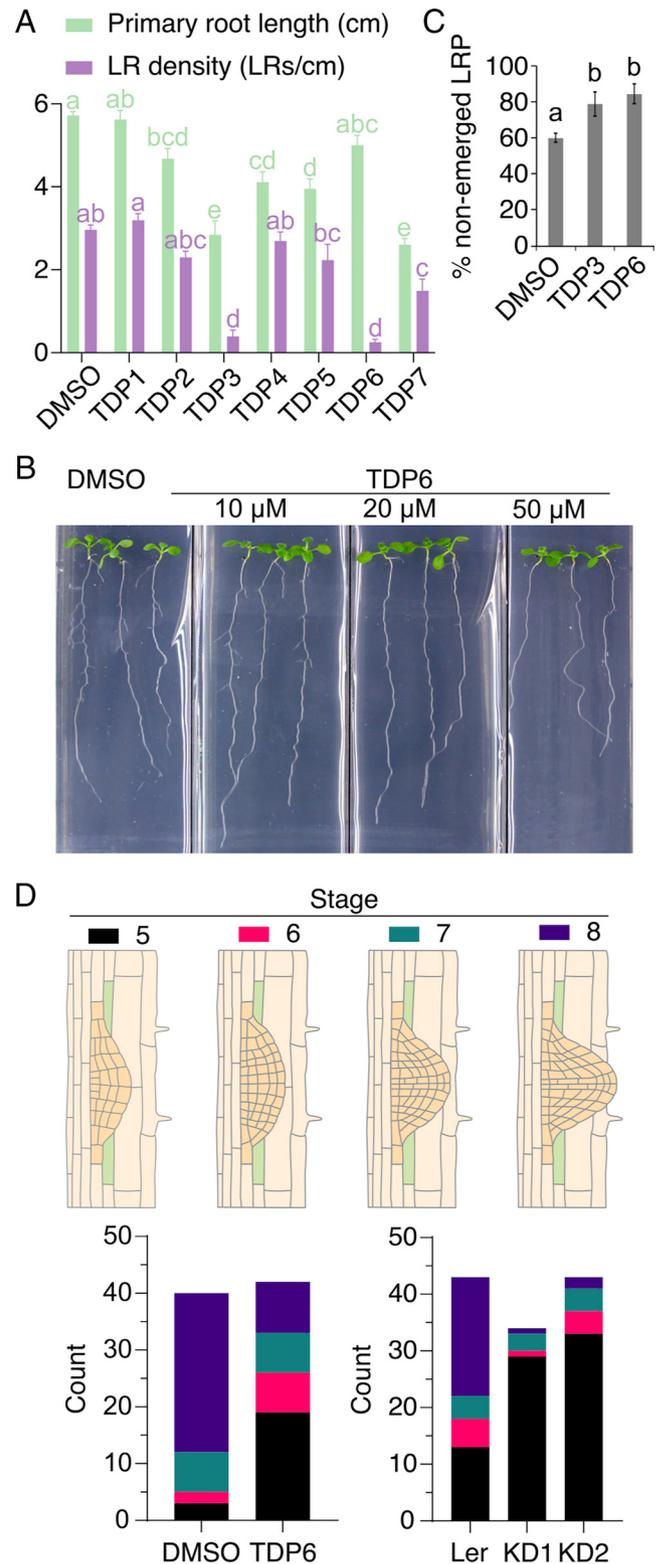
Fig. 5. TDP6 inhibits lateral root emergence. (*A*) The effect of the selected TDP-containing inhibitors on primary root length and lateral root density compared to control (DMSO). Error bars represent SEM, treatments were compared separately for primary root length and lateral root density by one-way ANOVA post hoc Tukey's test ($P < 0.05$). (*B*) Visual phenotype of TDP6 treatment on lateral root growth. (*C*) The number of nonemerged lateral root primordia (LRP) is increased by 20% on average by TDP3 and TDP6. Error bars represent SEM; treatments were compared with one-way ANOVA post hoc Tukey's test ($P < 0.0001$). (*D*) Stages of LRP development scored in a root bending assay for TDP6 and two metacaspase II family amiRNA knockdown lines (KD1 and KD2) compared to their controls (DMSO and wild-type Landsberg *erecta* seedlings, respectively).

lymphoid tissue lymphoma translocation protein 1 (MALT1), caspases, legumain, and separase. The MEROPS database states that the His-Cys catalytic dyad acts with a mechanism distinct from that in clan CA (papain and homologs) (1, 34). Yet, several later studies have used the mechanism of papain to describe the role of the His-Cys dyad in clan CD. In papain, the Cys and His residues are positioned in direct contact with each other, on the same side of the peptide bond to be cleaved, where His acts as proton acceptor to the cysteine and the deprotonated, negatively charged Cys makes a nucleophilic attack at the peptide carbonyl carbon. In clan CD on the other hand, the His and Cys residues are positioned on opposite sides of the scissile peptide bond, making it unlikely that the His residue can act as a proton acceptor for the catalytic Cys residue. It may be argued that the catalytic center may undergo substantial conformational changes during the reaction, but such changes have not been reported for any clan CD proteases. Like in clan CA, it is well established that the catalytic cysteine forms a covalent bond with the peptide carbonyl in the acyl-enzyme intermediate. However, QM/MM simulations of the reaction mechanism for human legumain indicated that the cysteine needs to be in its protonated, neutral state for a productive nucleophilic attack and rupture of the scissile peptide bond (35). It is thus uncertain how the His residue participates in the first acylation step. In the second hydrolysis step, the role of the His residue is presumably to deprotonate and activate the catalytic water molecule and shuttle the proton to the carbonyl oxygen of the acyl–enzyme intermediate.

For AtMCA-IIa, no major structural differences were noted upon treatment of wild-type crystals with Ca^{2+}, apart from the disappearance of electron density for Lys255 (presumably due to cleavage) (22). Also, when a wide range of clan CD protein structures, inactive mutants, and zymogen forms as well as activated enzymes with trapped covalent intermediates, including metacaspases, caspases, legumains, and butelases (*SI Appendix,* Fig. S15) are superposed, the positions of the His-Cys dyad-bearing loops are very similar, indicating that zymogen activation does not involve large movements of the His-Cys dyad residues for recruitment of the catalytic machinery. Rather, the machinery already seems to be in place for peptide bond cleavage. The complementarity of the AIM sequence to the active site suggests that the AtMCA-IIa and AtMCA-IIf structures may be close to the conformation of the Michaelis complex, i.e. how the substrate is bound for cleavage. If this is the case, why is the AIM not cleaved then in the zymogen state of the enzymes, or rather, why is it so slow?

Actually, cleavage of the AIM does seem to happen at low rates also in the absence of Ca^{2+} or low pH as indicated by i) the fact that both recombinant AtMCA-IIa and AtMCA-IIf are partially autoprocessed when produced in *E. coli* (4), and ii) the presence of a p20 band on western blots against AtMCA-IIa and AtMCA-IIf from plant protein extracts in steady state (13, 18, 36). However, strong and rigid binding of the AIM sequence on both sides of the catalytic center may hold both the peptide carbonyl and nitrogen atoms so firmly in place that they are less likely to undergo the conformational changes required for bond cleavage. The activation energy barrier will thus be very high and the reaction rate very low. Furthermore, if the peptide bond is cleaved, the new N-terminal amino group may block the access for a water molecule to attack the acyl-enzyme intermediate and complete the hydrolysis, while the amino group itself will be well positioned for nucleophilic attack and re-ligation of the peptide bond. This is analogous to the function of bovine pancreatic trypsin inhibitor (BPTI), addressed in a QM/MD simulation study of BPTI interactions with trypsin (37) (for details, see *SI Appendix, Text*), as well as the inhibition of legumain by the reactive center loop of

human cystatin E (hCE), which is slowly hydrolyzed and re-ligated by legumain in a pH-dependent manner (38). Additionally, the AIM in type II metacaspases is covalently linked to the rest of the protein. If hydrolytic cleavage occurs and the newly formed ends dissociate from the active site, their effective concentrations will be very high and probably rebind and re-ligate. Stronger binding of the noncleaved substrate than after cleavage will affect the equilibrium of the reaction and favor noncleaved over cleaved AIM sequence.

The conservation of an LPL/F motif in the AIM sequence across type II metacaspases suggests an important role (*SI Appendix,* Fig. S5). The first Leu of this motif, represented by Ile and Met in some homologs, is deeply embedded in the S2′ pocket. Interestingly, structurally related *A. thaliana* legumains also harbor a hydrophobic S2′ pocket that preferentially binds Leu residues, and which is shown to be important for ligation/transpeptidation reactions (39), pointing toward a putative role in metacaspases to confer latency by proenzyme re-ligation. The following Pro residue is strictly conserved and is partially exposed at the S3′ site (*SI Appendix,* Fig. S2 *A* and *B*). After the Pro residue, the peptide chain makes a sharp turn into the first alpha helix of the UNK domain, and the N-terminal end of the helix is anchored by a conserved hydrophobic residue (L/I/Y/F). We hypothesize that the binding of this motif is crucial for autoinhibition and that the function of the UNK domain is to keep this region firmly bound at the active site to prevent cleavage of the AIM.

We also speculate that the activation mechanism involves increased dynamics of the UNK domain, at least of the first alpha helix, leading to larger conformational freedom of the AIM so that it is more readily cleaved. In AtMCA-IIf, an additional cleavage site was found in the UNK domain in the degradome study (31) that may play a role here. The sequence cleaved is LFGR(216) DAGLKF. If AtMCA-IIf is cleaved both at Arg183 of the AIM and at Arg216 of the UNK domain, the resulting 33-residue peptide could thus be free to detach and diffuse away from the activated enzyme. In AtMCA-IIa, a similar cleavage site has not yet been identified, but there are several Lys and Arg residues in the UNK domain. For example, cleavage at Lys249 in UNK and at Lys225 in the AIM would generate a 24-residue peptide comprising the 1st helix and the following loop.

We also note that the UNK domain in AtMCA-IIa contains an unusually large number of charged residues, salt bridges, and positive residues that are buried within the domain. Buried or partially buried Lys/Arg residues make salt bridges at both proposed Ca^{2+}-binding sites mentioned above. Ca^{2+} binding would disrupt these salt bridges and thus destabilize the structure of the UNK domain, which may play a key role in the Ca^{2+} activation mechanism. In the UNK of AtMCA-IIf on the other hand, there are few Lys/Arg, none of which are buried, and few salt bridges, which may further help explain why this enzyme is Ca^{2+} independent. However, future studies are needed to understand if, and in that case how, a low pH destabilizes the UNK domain of AtMCA-IIf and binding of the AIM at the catalytic center.

A Group of Small Molecules for In Vivo Inhibition of Metacaspase Activity. To overcome genetic and functional redundancy and to develop a tunable inhibition strategy, we performed a target based chemical screen for small molecules that inhibit type II metacaspases. Small molecule inhibitors are often used as a means to inhibit therapeutic targets in a clinical setting, including proteases (40). In contrast, few examples exist in the literature of small molecules that modulate plant protease activity in vivo. Mainly those can be found that are substrate mimetics: a peptide based on the substrate cleavage signature and modified with a "warhead" that covalently

binds and inactivates the protease active site during catalysis (41). For metacaspases, the most used substrate mimetic inhibitor is z-VRPR-fmk, which was derived from a preferred tetrapeptide substrate as determined by scanning of a combinatorial tetrapeptide library for AtMCA-IIf activity (32) and has a fluoromethyl ketone (fmk) warhead. Not only is z-VRPR-fmk used in enzymatic and in vivo assays in plants (18, 42, 43), but also in vivo against human paracaspase MALT1. However, in human studies z-VRPR-fmk suffers from poor cell permeability, probably due to the two Arg residues (44), and biological toxicity (45, 46). More recently, new activity-based probes with an acyloxymethylketone warhead were developed based on the AtMCA-IIf tetrapeptide library study and COFRADIC analysis (31, 32), but in vivo efficacy was not tested (47). Substrate mimetics and derivatives were also designed for inhibition of protozoan metacaspases (48). Few studies of small molecule inhibitors in *Trypanosoma* sp. and *Plasmodium* sp. are available, and even fewer that are not peptide or substrate-like, that would contain more diversity in their chemical structure. Nonetheless, through rational design based on the observation that z-FA-fmk can inhibit *Plasmodium falciparum* metacaspase-2 (PfMCA-2), a nonpeptidyl molecule was synthesized, designated SS-5 that induces cell death in *P. falciparum* (49). More recently, two small molecules derived from SS-5, called C532 and C533, were found to effectively block transmission of *P. falciparum* and *P. berghei*, reducing the parasite burden in a mosquito host (50). Bear in mind that protozoa exclusively have type I metacaspases (5). In this study, we screened against 10,000 chemicals thereby enlarging the chemical space of potential type II metacaspase inhibitors. The uncovered small molecule inhibitors in this study and their derivatives could prove to be active against medical important type I metacaspase in the future. However, given that highly similar TDP-containing small molecules can have very different effects in vivo (for instance, on LRP emergence), it is hard to predict which small molecules will be effective without further detailed studies in vivo.

There is an interesting discrepancy between the seemingly specific inhibition of AtMCA-IIf and LRP emergence by TDP6 on the one hand, and lack of LRP phenotype in single *atmca-IIf* and double *atmca-IIb atmca-IIf* mutants on the other hand. Several plausible explanations exist: i) TDP6 prefers to inhibit AtMCA-IIf and thereby is more specific than TDP2 or TDP3, but can inhibit AtMCA-IIa to a lesser extent (Fig. 2*E* and *SI Appendix*, Fig. S6*A* and Table S2). A fine balance likely exists in vivo between inhibition of AtMCA-IIf and Ca^{2+}-dependent MCAs, similar to the effect of reducing type II *MCA* gene expression in the KD1 and KD2 amiRNA lines (Fig. 5*D*). ii) Genetic and functional redundancy probably occurs in the stable knockout mutants. Other MCAs could assume the role of AtMCA-IIf and AtMCA-IIb in the single and double mutants, because of the pronounced genetic redundancy in Arabidopsis that contains six type II MCAs. Alternatively, completely different proteases altogether might take over, leading to functional redundancy. A precedent was described for the cysteine protease SENESCENCE-ASSOCIATED GENE 12 (SAG12), where the activity of an aspartic protease APOPLASTIC, EDS1-DEPENDENT 1 (AED1) is highly up-regulated in a *sag12* mutant, leading to a lack of phenotype in the *sag12* mutant during senescence (51). Genetic and functional redundancy is less likely to occur during short treatments with chemical inhibitors. iii) There is growing evidence that zymogen and active forms of the same MCA protein may each have a distinct biochemical role. One example is yeast MCA1. While proteolytically active MCA1 executes cell death, its zymogen has a chaperone activity counteracting build-up of protein aggregates during replicative aging (14, 16). It is tempting to speculate that Arabidopsis MCAs can likewise carry different biochemical activities depending on whether they are present as zymogens or mature proteases. This could explain why genetic knock-out of MCAs and pharmacological inhibition of its protease activity do not furnish the same phenotype.

MCA-Dependent Substrate Cleavages as a Potential Driver of Cell Separation and Heightened Immune Response in Lateral Root Emergence. Lateral root emergence requires an interplay between the emerging LRP and the overlying tissues, including endodermal cells (52, 53). Auxin signaling, mechanical feedback, regulated cell death, cytoskeleton dynamics, and symplastic isolation are involved in this tightly regulated process of new organ development (26, 54–57). Genes expressed specifically in the endodermal cells adjacent to LRPs, such as *MYB36* and *PLDP5* (54, 58), can control the process of lateral root emergence. Here, we demonstrated that metacaspases with a defined spatiotemporal endodermal expression pattern are necessary for a proficient outgrowth of LRPs. Endodermal cells are differentiated cells containing two impermeable polymers in their cell walls: lignin and suberin. Both the lignified Casparian strips and suberized lamellae are critical to form a tightly controlled bidirectional barrier between the environment (e.g. soil nutrients, water, microbiota) and the plant vasculature (59). Cell separation is required to allow for the developing LRP to protrude through the endodermal cell layer, while at the same time, remodeling of the lignin and suberin barriers is required to regulate nutrient and water flow, and likely to prevent colonization of the root interior by pathogens (25, 60, 61). The substrate specificity studies with a COFRADIC analysis on *atmca-IIf* single mutants in whole Arabidopsis roots revealed potential coexpressed substrate proteins of AtMCA-IIf related to cell wall remodeling, including glycosyl hydrolases of the GH family 1 (AT3G09260), 32 (AT1G12240), 38 (AT3G26720), and peroxidases (AT4G30170, AT1G05240, AT5G17820; Dataset S1).

Alternatively, AtMCA-IIf could be important for execution of cell death or clearance of dead cells (17), as cell death of the endodermal cells overlying the LRP was found to be important for lateral root emergence (26). As not all proximal endodermal cells die during passage of the developing lateral root (26), it is possible that a mixture of cell death and cell wall remodeling of endodermal cells takes place during LRP development. Furthermore, cell wall remodeling and cell death are not necessarily mutually exclusive events, as AtMCA-IIf can be active postmortem for example in the clearance of cellular remnants in "dead" xylem cells (17). In the future, a detailed spatiotemporal analysis of cell wall remodeling-mediated cell separation and PCD in the root endodermis should facilitate our mechanistic comprehension of the role of AtMCA-IIf in lateral root emergence.

Besides their potential roles in cell wall remodeling, we can also speculate on a defensive role for metacaspases during LRP emergence. Physical damage-induced or temporary developmental breaches of root tissue integrity might create the opportunity for soil pathogens to invade the host plant. Zhou et al. noted that cortical cells in the vicinity of LRPs have a heightened responsiveness to bacterial elicitors such as flg22 (61). We and others have previously shown that upon wounding of plants, the calcium-dependent metacaspase AtMCA-IIa releases the immunomodulatory peptide Pep1 from its precursor PROPEP1 (18, 19). PROPEPs are a family of precursor proteins from which small signaling peptides, Peps, are derived and typically are perceived by the leucine-rich repeat receptor-like kinases (LRR-RLKs), PEP RECEPTOR 1 and 2 (PEPR1 and PEPR2), and its coreceptor BRI1 ASSOCIATED RECEPTOR KINASE 1 (BAK1) (62–65). Recently, Pep7 was found to be perceived by another LRR-RLK, SUCROSE-INDUCED RECEPTOR KINASE 1 (SIRK1), and its coreceptor QIAN SHOU KINASE 1 (QSK1) (66). Strikingly, the *sirk1* and *propep7* mutants have a delayed LRP emergence, whereas

application of synthetic Pep7 peptide speeds up LRP emergence. The mature Pep7 peptide was identified previously from root tip protein extracts through COFRADIC analysis (67) and PROPEP7 can be cleaved by AtMCA-IIa in vitro (18) and in Arabidopsis protoplasts (19). Therefore, it is possible that AtMCA-IIf and potentially Ca^{2+}-dependent type II MCAs exert their effect on LRP emergence through cleavage and release of Pep7. Depending on the presence of PEPR1 and SIRK1 in the vicinity of the LRP, Pep7 could have an effect both on immunity and development, respectively. Nevertheless, the molecular mechanism behind the effect of Pep7 on LRP emergence remains unclear and metacaspases likely mediate LRP emergence through cleavage of multiple different substrate proteins. Certainly, further studies are required to tease apart the potential developmental and immune-related effects of metacaspase substrate cleavage in LRP emergence.

Conclusion

In summary, we report here on the structure of the calcium-independent metacaspase AtMCA-IIf and provide a rationale for the inhibitory action of a class of TDP-containing small molecules against metacaspases. Importantly, some of the TDP-containing small molecules (such as TDP6) have a specific effect in plants on LRP emergence where, due to the specific expression pattern of AtMCA-IIf and AtMCA-IIb in the endodermal cells overlying the LRP, metacaspases can control LRP emergence. Furthermore, the structure of AtMCA-IIf contributes to a better understanding of the activation mechanism of calcium-independent metacaspases and other clan CD proteases, including paracaspases, caspases, legumain, and separase. The TDP-containing small molecule inhibitors provide an extra tool to study metacaspase function in diverse plant species and can initiate the development of drugs against neglected diseases caused by *Trypanosoma* and *Plasmodium* where metacaspases—absent in humans—are considered as suitable therapeutic targets.

Materials and Methods

Plant Lines. T-DNA insertion lines of AtMCA-IIb (SAIL-284-C06) and AtMCA-IIf (GK-540H06) were crossed. Type-II metacaspase silenced lines were generated by means of the artificial microRNA (amiRNA) technology (knockdown lines KD1 and KD2). Transcription and translational fusion lines for expression analysis were generated. Construction and genotyping of these lines, is described in more detail in *SI Appendix, Materials and Methods*.

Tissue Expression Analysis. GUS staining was performed on 7-d-old plants and is described in more detail in *SI Appendix, Materials and Methods*. Fluorescence microscopy of the pAtMCA:GFP:GUS fusion reporter lines was performed with a confocal microscope 100 M and software package LSM 510 version 3.2 (Zeiss, Jena, Germany). Excitation was performed with a 488-nm argon laser. Emission fluorescence for GFP was captured via a 500 to 550-nm band-pass filter, and for propidium iodide (PI) via a 585-nm long-pass filter. Fluorescence microscopy of the pAtMCA-IIf:AtMCA-IIf-GFP translational fusion lines was performed on a Zeiss LSM 710 confocal scanning microscope, excitation with a 488-nm argon laser and appropriate emission windows for GFP and PI stain.

Preparation and Purification of Native Recombinant Metacaspases. Recombinant AtMCA-IIa and AtMCA-IIf were produced as previously detailed (4) and are described in more detail in *SI Appendix, Materials and Methods*.

Crystallization, Data Collection, and Structure Determination of Inactive AtMCA-IIf C147A. Crystallization screening was performed using a mosquito-LCP robot (SPT Labtech) in three-drop 96-well protein crystallography plate

(Corning). Crystals were grown in a RockImager 1,000 system (Formulatrix) at 4 and 20 °C. Four commercial screens were used: JCSG+ (Qiagen), Index, Crystal Screen HT, and PEGSuite (Hampton research) and are described in more detail in *SI Appendix, Materials and Methods*.

N-Terminal COFRADIC Analysis. Roots from 3-wk-old Arabidopsis Columbia-0 (Col-0), AtMCA-IIf T-DNA insertion (GK-506H04-019739) and AtMCA-IIf overexpression lines (32) were used. Protein extracts were modified for N-terminal COFRADIC analysis as described (29), with few modifications. More details about proteome extraction, N-terminal labeling, and peptide identification are described in *SI Appendix, Materials and Methods*.

Screening of AtMCA-IIf Inhibitors. High-throughput screens of 10,000 compounds comprising the DIVERSet™ compound library of Chembridge were carried out in 384-well plates (Genetix X7001), and the reactions were performed in 50 μL final volume and are described in more detail in *SI Appendix, Materials and Methods*.

In Silico Analyses. IceLogo (68) analyses were performed via the web application (http://iomics.ugent.be/icelogoserver/logo.html) using the Arabidopsis UniProt proteome as reference. Hierarchical clustering of the small-molecule metacaspase inhibitors was done via the open access software ChemMine (http://chemmine.ucr.edu/) based on structural and physicochemical similarities of the compounds using the Tanimoto similarity coefficient.

In Vitro HIT Verification and Full Dose–Response Validation. The enzymatic activity of recombinant AtMCA-IIa, AtMCA-IIb and, AtMCA-IIf in the presence of the TDP substances were determined in vitro by following the cleavage of the fluorogenic metacaspase substrate Val-Arg-Pro-Arg-7-amino-4-methyl Coumarin (VRPR-AMC). The kinetic measurements were performed on a BioTek Synergy™ H4 hybrid microplate reader, monitoring the change in fluorescence light at Ex355/Em 460 nm. An isothiouronium salt, purchased from Sigma Aldrich, was used as a positive control toward AtMCA-IIa, AtMCA-IIb, and AtMCA-IIf (*SI Appendix,* Fig. *S6B*). Assay measurements on AtMCA-IIa and AtMCA-IIb are described in more detail in *SI Appendix, Materials and Methods*.

Compound Characterization by NMR. Selected TDP-containing compounds were reordered from Chembridge (TDP1 to TDP7; Dataset S2). ¹H NMR spectra of TDP1 to TDP7 were recorded on a Bruker 400 MHz Avance III spectrometer at 400 MHz (¹H) at 25 °C. Chemical shifts are reported in ppm, and δ values were referenced to the residual solvent signal of DMSO-d_6 (2.50 ppm) as an internal standard (Dataset S3).

Inhibition of GST-TEV-PROPEP1 Cleavage. An orthogonal in vitro assay based on the cleavage of a recombinant MCA substrate protein was developed as alternative for the VRPR-AMC based activity assay and is described in more detail in *SI Appendix, Materials and Methods*.

Molecular Docking Simulations. The X-ray crystal structure of *At*MCA-IIf was prepared for docking using the protein preparation wizard implemented in Maestro (69) and is described in more detail in *SI Appendix, Materials and Methods*.

RNA Extraction, Reverse-Transcription PCR and Real-Time Quantitative PCR. Total RNA was extracted from leaves as previously described (70) and is described in more detail in *SI Appendix, Materials and Methods*.

Plant Treatments and Growth Conditions. Surface-sterilized seeds were grown on vertical plates containing half-strength Murashige and Skoog (½MS) salts, 1% (w/v) sucrose, 0.8% (w/v) plant tissue culture agar (LabM, Bury, UK), 0.5 mg/l nicotinic acid, 0.5 mg/l pyridoxin, and 1 mg/l thiamin. The seeds were stratified at 4 °C for two nights in the dark and then left to germinate in continuous light (80 to 100 μmol m^{-2}s^{-1}) at 21 °C. Three days after germination, seedlings were transferred to ½MS agar plates containing 50 μM compounds and then 5 d later their root growth and phenotypical appearance were monitored. Staining of LRP and developmental staging was performed 10 d after germination as described (33). The TDP6 to DMSO transfer experiment is described in more detail in *SI Appendix, Materials and Methods*.

Root-Bending Assay and LRP Staging. Surface-sterilized Arabidopsis seeds of Col-0, L*er*, and AtMCA-II amiRNA knockdown lines KD1 and KD1, were sown on ½MS plates containing 1% (w/v) sucrose, 0.01% (w/v) myo-inositol and 0.05% (w/v) MES-KOH, pH 5.7. The plates were stratified for 2 d and transferred to the growth chamber to grow vertically under continuous light at 22 °C. After 48 h, seedlings were transferred to ½MS plates containing 25 μM TDP6 or an equivalent volume of DMSO as control. After 24 h, plates were rotated 90° for 48 h. The seedlings were then fixed with FAA for 30 min and were rehydrated for 20 min each in 70%, 50%, 30%, and 10% ethanol and infiltrated overnight in 50% glycerol. The fixation and infiltration steps were taking place on ice and all solutions were kept cold. The seedlings were mounted in 50% glycerol under glass coverslips. Images of LRP were collected with an AxioCam MR3 (Carl Zeiss, Göttingen, Germany). The number and stage of LRP were registered according to (33).

Data, Materials, and Software Availability. All data porting this study are included in the main text and *SI Appendix*. Coordinates and structure factors for AtMCA-IIf have been deposited in the Protein Data Bank under accession number 8A53 (71). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD040376.

Author affiliations: [a]Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium; [b]Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium; [c]Department of Molecular Sciences, Uppsala BioCenter, Swedish University of Agricultural Sciences and Linnean Center for Plant Biology, 75007 Uppsala, Sweden; [d]VIB Screening Core, VIB, 9052 Ghent, Belgium; [e]Centre for Bioassay Development and Screening, Ghent University, 9000 Ghent, Belgium; [f]Department of Chemistry, Umeå University, 90187 Umeå, Sweden; [g]Department of Biomolecular Medicine, Ghent University, 9052 Ghent, Belgium; and [h]Center for Medical Biotechnology, VIB, 9052 Ghent, Belgium

1. N. D. Rawlings *et al.*, The *MEROPS* database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **46**, D624–D632 (2018).
2. A. G. Uren *et al.*, 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 (2000).
3. W. G. van Doorn *et al.*, Morphological classification of plant cell deaths. *Cell Death Differ.* **18**, 1241–1246 (2011).
4. D. Vercammen *et al.*, Type II metacaspases Atmc4 and Atmc9 of *Arabidopsis thaliana* cleave substrates after arginine and lysine. *J. Biol. Chem.* **279**, 45329–45336 (2004).
5. L. Tsiatsiani *et al.*, Metacaspases. *Cell Death Differ.* **18**, 1279–1288 (2011).
6. E. Lam, Y. Zhang, Regulating the reapers: Activating metacaspases for programmed cell death. *Trends Plant Sci.* **17**, 487–494 (2012).
7. E. A. Minina, N. S. Coll, H. Tuominen, P. V. Bozhkov, Metacaspases versus caspases in development and cell fate regulation. *Cell Death Differ.* **24**, 1314–1325 (2017).
8. E. A. Minina *et al.*, Classification and nomenclature of metacaspases and paracaspases: No more confusion with caspases. *Mol. Cell* **77**, 927–929 (2020).
9. M. Klemenčič, C. Funk, Evolution and structural diversity of metacaspases. *J. Exp. Bot.* **70**, 2039–2047 (2019).
10. C. J. Choi, J. A. Berges, New types of metacaspases in phytoplankton reveal diverse origins of cell death proteases. *Cell Death Dis.* **4**, e490 (2013).
11. M. Klemenčič, C. Funk, Type III metacaspases: Calcium-dependent activity proposes new function for the p10 domain. *New Phytol.* **218**, 1179–1191 (2018).
12. N. S. Coll *et al.*, *Arabidopsis* type I metacaspases control cell death. *Science* **330**, 1393–1397 (2010).
13. N. Watanabe, E. Lam, Arabidopsis metacaspase 2d is a positive mediator of cell death induced during biotic and abiotic stresses. *Plant J.* **66**, 969–982 (2011).
14. R. E. C. Lee, S. Brunette, L. G. Puente, L. A. Megeney, Metacaspase Yca1 is required for clearance of insoluble protein aggregates. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13348–13353 (2010).
15. N. S. Coll *et al.*, The plant metacaspase AtMC1 in pathogen-triggered programmed cell death and aging: Functional linkage with autophagy. *Cell Death Differ.* **21**, 1399–1408 (2014).
16. S. M. Hill, X. Hao, B. Liu, T. Nyström, Life-span extension by a metacaspase in the yeast *Saccharomyces cerevisiae*. *Science* **344**, 1389–1392 (2014).
17. B. Bollhöner *et al.*, *Post mortem* function of AtMC9 in xylem vessel elements. *New Phytol.* **200**, 498–510 (2013).
18. T. Hander *et al.*, Damage on plants activates Ca^{2+}-dependent metacaspases for release of immunomodulatory peptides. *Science* **363**, eaar7486 (2019).
19. W. Shen, J. Liu, J.-F. Li, Type-II metacaspases mediate the processing of plant elicitor peptides in *Arabidopsis*. *Mol. Plant* **12**, 1524–1533 (2019).
20. K. McLuskey *et al.*, Crystal structure of a *Trypanosoma brucei* metacaspase. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7469–7474 (2012).
21. A.H.-H. Wong, C. Yan, Y. Shi, Crystal structure of the yeast metacaspase Yca1. *J. Biol. Chem.* **287**, 29251–29259 (2012).
22. P. Zhu *et al.*, Structural basis for Ca^{2+}-dependent activation of a plant metacaspase. *Nat. Commun.* **11**, 2249 (2020).
23. L. Aravind, E. V. Koonin, Classification of the caspase–hemoglobinase fold: Detection of new families and implications for the origin of the eukaryotic separins. *Proteins* **46**, 355–367 (2002).
24. C. A. Ramos-Guzmán, K. Zinovjev, I. Tuñón, Modeling caspase-1 inhibition: Implications for catalytic mechanism and drug design. *Eur. J. Med. Chem.* **169**, 159–167 (2019).
25. J. E. M. Vermeer *et al.*, A spatial accommodation by neighboring cells is required for organ initiation in *Arabidopsis*. *Science* **343**, 178–183 (2014).
26. S. Escamez *et al.*, Cell death in cells overlying lateral root primordia facilitates organ growth in *Arabidopsis*. *Curr. Biol.* **30**, 455–464 (2020).
27. J. Fortin, E. Lam, Domain swap between two type-II metacaspases defines key elements for their biochemical properties. *Plant J.* **96**, 921–936 (2018).
28. N. Watanabe, E. Lam, Calcium-dependent activation and autolysis of Arabidopsis metacaspase 2d. *J. Biol. Chem.* **286**, 10027–10040 (2011).
29. A. Staes *et al.*, Selecting protein N-terminal peptides by combined fractional diagonal chromatography. *Nat. Protoc.* **6**, 1130–1141 (2011).
30. I. Schechter, A. Berger, On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162 (1967).
31. L. Tsiatsiani *et al.*, The *Arabidopsis* METACASPASE9 degradome. *Plant Cell* **25**, 2831–2847 (2013).
32. D. Vercammen *et al.*, Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for metacaspase 9. *J. Mol. Biol.* **364**, 625–636 (2006).
33. J. E. Malamy, P. N. Benfey, Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. *Development* **124**, 33–44 (1997).
34. L. Polgár, "Catalytic mechanisms of cysteine peptidases" in *Handbook of Proteolytic Enzymes*, N. D. Rawlings, G. Salvesen, Eds. (Academic Press, Amsterdam, ed. 3, 2013), **vol. 2**, pp.1773–1784.
35. B. Elsässer *et al.*, Distinct roles of catalytic cysteine and histidine in the protease and ligase mechanisms of human legumain as revealed by DFT-based QM/MM simulations. *ACS Catalysis* **7**, 5585–5593 (2017).
36. B. Belenghi *et al.*, Metacaspase activity of *Arabidopsis thaliana* is regulated by *S*-nitrosylation of a critical cysteine residue. *J. Biol. Chem.* **282**, 1352–1358 (2007).
37. M. Peräkylä, P. A. Kollman, Why does trypsin cleave BPTI so slowly? *J. Am. Chem. Soc.* **122**, 3436–3444 (2000).
38. E. Dall, J. C. Fegg, P. Briza, H. Brandstetter, Structure and mechanism of an aspartimide-dependent peptide ligase in human legumain. *Angew. Chem. Int. Ed.* **54**, 2917–2921 (2015).
39. E. Dall *et al.*, Structural and functional studies of *Arabidopsis thaliana* legumain beta reveal isoform specific mechanisms of activation and substrate recognition. *J. Biol. Chem.* **295**, 13047–13064 (2020).
40. Q. Wu, W. Qian, X. Sun, S. Jiang, Small-molecule inhibitors, immune checkpoint inhibitors, and more: FDA-approved novel therapeutic drugs for solid tumors from 1991 to 2021. *J. Hematol. Oncol.* **15**, 143 (2022).
41. K. Morimoto, R. A. L. van der Hoorn, The increasing impact of activity-based protein profiling in plant science. *Plant Cell Physiol.* **57**, 446–461 (2016).
42. E. Berenguer *et al.*, Suppression of metacaspase- and autophagy-dependent cell death improves stress-induced microspore embryogenesis in *Brassica napus*. *Plant Cell Physiol.* **61**, 2097–2110 (2020).
43. S. Graff van Creveld *et al.*, Biochemical characterization of a novel redox-regulated metacaspase in a marine diatom. *Front. Microbiol.* **12**, 688199 (2021).
44. L. Fontán *et al.*, Specific covalent inhibition of MALT1 paracaspase suppresses B cell lymphoma growth. *J. Clin. Invest.* **128**, 4397–4412 (2018).
45. A. Demeyer, J. Staal, R. Beyaert, Targeting MALT1 proteolytic activity in immunity, inflammation and disease: Good or bad? *Trends Mol. Med.* **22**, 135–150 (2016).
46. I. Hamp, T. J. O'Neill, O. Plettenburg, D. Krappmann, A patent review of MALT1 inhibitors (2013-present). *Expert Opin. Ther. Pat.* **31**, 1079–1096 (2021).
47. V. Štrancar *et al.*, Activity-based probes trap early active intermediates during metacaspase activation. *iScience* **25**, 105247 (2022).
48. M. Berg *et al.*, Design and evaluation of *Trypanosoma brucei* metacaspase inhibitors. *Bioorg. Med. Chem. Lett.* **20**, 2001–2006 (2010).
49. Vandana *et al.*, A nonpeptidyl molecule modulates apoptosis-like cell death by inhibiting *P. falciparum* metacaspase-2. *Biochem. J.* **477**, 1323–1344 (2020).
50. V. Kumari *et al.*, Dissecting the role of *Plasmodium* metacaspase-2 in malaria gametogenesis and sporogony. *Emerg. Microbes Infect.* **11**, 938–955 (2022).
51. M. James *et al.*, SAG12, a major cysteine protease involved in nitrogen allocation during senescence for seed production in *Arabidopsis thaliana*. *Plant Cell Physiol.* **59**, 2052–2063 (2018).
52. J. Banda *et al.*, Lateral root formation in *Arabidopsis*: A well-ordered LRexit. *Trends Plant Sci.* **24**, 826–839 (2019).
53. R. Sager, M. Bennett, J.-Y. Lee, A tale of two domains pushing lateral roots. *Trends Plant Sci.* **26**, 770–779 (2021).
54. R. Sager *et al.*, Auxin-dependent control of a plasmodesmal regulator creates a negative feedback loop modulating lateral root emergence. *Nat. Commun.* **11**, 364 (2020).

55. G. L. Richter, G. B. Monshausen, A. Krol, S. Gilroy, Mechanical stimuli modulate lateral root organogenesis. *Plant Physiol.* **151**, 1855–1866 (2009).

56. P. Marhavý *et al.*, Targeted cell elimination reveals an auxin-guided biphasic mode of lateral root initiation. *Genes Dev.* **30**, 471–483 (2016).

57. A. Vilches Barro *et al.*, Cytoskeleton dynamics are necessary for early events of lateral root initiation in *Arabidopsis*. *Curr. Biol.* **29**, 2443–2454 (2019).

58. M. Fernández-Marcos *et al.*, Control of Arabidopsis lateral root primordium boundaries by MYB36. *New Phytol.* **213**, 105–112 (2017).

59. V. G. Doblas, N. Geldner, M. Barberon, The endodermis, a tightly controlled barrier for nutrients. *Curr. Opin. Plant Biol.* **39**, 136–143 (2017).

60. R. Ursache *et al.*, GDSL-domain proteins have key roles in suberin polymerization and degradation. *Nat. Plants* **7**, 353–364 (2021).

61. F. Zhou *et al.*, Co-incidence of damage and microbial patterns controls localized immune responses in roots. *Cell* **180**, 440–453 (2020).

62. A. Huffaker, G. Pearce, C. A. Ryan, An endogenous peptide signal in *Arabidopsis* activates components of the innate immune response. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10098–10103 (2006).

63. E. Krol *et al.*, Perception of the *Arabidopsis* danger signal peptide 1 involves the pattern recognition receptor *At*PEPR1 and its close homologue *At*PEPR2. *J. Biol. Chem.* **285**, 13471–13479 (2010).

64. Y. Yamaguchi, G. Pearce, C. A. Ryan, The cell surface leucine-rich repeat receptor for AtPep1, an endogenous peptide elicitor in *Arabidopsis*, is functional in transgenic tobacco cells. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10104–10109 (2006).

65. Y. Yamaguchi, A. Huffaker, A. C. Bryan, F. E. Tax, C. A. Ryan, PEPR2 is a second receptor for the Pep1 and Pep2 peptides and contributes to defense responses in *Arabidopsis*. *Plant Cell* **22**, 508–522 (2010).

66. J. Wang *et al.*, PEP7 acts as a peptide ligand for the receptor kinase SIRK1 to regulate aquaporin-mediated water influx and lateral root growth. *Mol. Plant* **15**, 1615–1631 (2022).

67. C. Liu *et al.*, The proteolytic landscape of an Arabidopsis separase-deficient mutant reveals novel substrates associated with plant development. bioRxiv [Preprint] (2017). https://doi.org/10.1101/140962. Accessed 15 December 2022.

68. N. Colaert, K. Helsens, L. Martens, J. Vandekerckhove, K. Gevaert, Improved visualization of protein consensus sequences by iceLogo. *Nat. Methods* **6**, 786–787 (2009).

69. M. G. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aid. Mol. Des.* **27**, 221–234 (2013).

70. L. Bentsink, J. Jowett, C. J. Hanhart, M. Koornneef, Cloning of *DOG1*, a quantitative trait locus controlling seed dormancy in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17042–17047 (2006).

71. I. Sabljić, S. Stael, J. Ståhlberg, P. V. Bozhkov, Crystal structure of AtMCA-IIf C147A (metacaspase 9) from Arabidopsis thaliana. *Protein Data Bank*. https://www.rcsb.org/structure/unreleased/8A535. Deposited 14 June 2022.