

RESEARCH ARTICLE

An Unusual Member of the Papain Superfamily: Mapping the Catalytic Cleft of the *Marasmius oreades* agglutinin (MOA) with a Caspase Inhibitor

Gabriele Cordara^{1,2*}, André van Eerde¹, Elin M. Grahn¹, Harry C. Winter³, Irwin J. Goldstein³, Ute Krengel^{1*}

1 Department of Chemistry, University of Oslo, Oslo, Norway, **2** Department of Biochemistry, Institute for Cancer Research, The Norwegian Radium Hospital, Oslo, Norway, **3** Department of Biological Chemistry, Medical School, University of Michigan, Ann Arbor, Michigan, United States of America

* gabriele.cordara@kjemi.uio.no (GC); ute.krengel@kjemi.uio.no (UK)



OPEN ACCESS

Citation: Cordara G, van Eerde A, Grahn EM, Winter HC, Goldstein IJ, Krengel U (2016) An Unusual Member of the Papain Superfamily: Mapping the Catalytic Cleft of the *Marasmius oreades* agglutinin (MOA) with a Caspase Inhibitor. PLoS ONE 11(2): e0149407. doi:10.1371/journal.pone.0149407

Editor: Eugene A. Permyakov, Russian Academy of Sciences, Institute for Biological Instrumentation, RUSSIAN FEDERATION

Received: November 13, 2015

Accepted: February 1, 2016

Published: February 22, 2016

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: The atomic coordinates and structure factors have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Databank (<http://www.rcsb.org/>) under PDB IDs 5D61, 5D62 and 5D63.

Funding: The work was funded by the University of Oslo and carried out as part of the GlycoNor consortium. This work was supported in part by the Norwegian Research Council (grant number: 216625; www.forskningsradet.no) and BioStruct-X (www.biostruct-x.org/).

Abstract

Papain-like cysteine proteases (PLCPs) constitute the largest group of thiol-based protein degrading enzymes and are characterized by a highly conserved fold. They are found in bacteria, viruses, plants and animals and involved in a number of physiological and pathological processes, parasitic infections and host defense, making them interesting targets for drug design. The *Marasmius oreades* agglutinin (MOA) is a blood group B-specific fungal chimerolectin with calcium-dependent proteolytic activity. The proteolytic domain of MOA presents a unique structural arrangement, yet mimicking the main structural elements in known PLCPs. Here we present the X-ray crystal structure of MOA in complex with Z-VAD-fmk, an irreversible caspase inhibitor known to cross-react with PLCPs. The structural data allow modeling of the substrate binding geometry and mapping of the fundamental enzyme-substrate interactions. The new information consolidates MOA as a new, yet strongly atypical member of the papain superfamily. The reported complex is the first published structure of a PLCP in complex with the well characterized caspase inhibitor Z-VAD-fmk.

Introduction

Cysteine proteases are involved in a number of physiological and pathological processes. They are classified in families based on sequence and fold conservation and further grouped into superfamilies or clans [1]. There are currently 77 recognized families of cysteine proteases in the MEROPS database, divided into 13 clans of C (Cys) and P (mixed) types, 48 of which have been structurally characterized [2].

Proteases belonging to clan CA are referred to as *papain-like cysteine proteases* (PLCPs, EC 3.4.22) and take their namesake from papain, the superfamily holotype. All PLCPs have the same fold, composed of two subdomains, the L(left)- and R(right)-domain, named after their

[biostruct-x.eu](https://doi.org/10.1371/journal.pone.0149407)). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

position in the standard view (Fig 1). They feature a Cys-His catalytic dyad and a conserved enzyme-substrate interaction geometry [3, 4]. These enzymes constitute the cysteine protease superfamily with the largest number of members [5]. PLCPs are found in bacteria, viruses, plants, and animals [6]. They are involved in a number of physiological and pathological processes, including antigen presentation [6], cancer, inherited diseases, parasitic infections [7] and host defense [8, 9]. Their role in pathology makes them suitable targets for drug design [10]; moreover, some of the enzymes can be exploited in integrated pest management [9, 11]. PLCPs are also represented among the fungal taxa. While recognized members include homologues of animal proteases such as bleomycin hydrolase [8], deubiquitinating enzymes [12] and calpain [13], little is known about fungal-specific families of proteases carrying the papain fold.

The *Marasmius oreades* agglutinin (MOA) is a 293 amino acid, homodimeric, histo-blood-group-B specific chimerolectin extracted from the fruiting bodies of the common fairy ring mushroom [17]. Recent literature suggests a role for MOA and related proteins as active players in fungal defense against external threats [18–20]. Each MOA protomer is composed of two domains, accounting for the lectin's sugar-binding and proteolytic functions, respectively [21]. The proteolytic activity is associated to the C-terminal $\alpha+\beta$ dimerization domain, which closely resembles the consensus papain fold (r.m.s.d.: 2.2 Å), including the Cys/His catalytic dyad (Cys215, His257) (Fig 1) [18, 22]. The papain-like L- and R-domain partitioning is conserved across the MOA dimer, with the L-domain borrowing structural elements from the other protomer. In contrast to other known PLCPs, the proteolytic domain of MOA carries a binuclear calcium binding site [14]. Calcium binding leads to an active site rearrangement essential for catalysis [14, 18, 22].

X-ray crystal structures of PLCP-inhibitor complexes have historically been fundamental to gain a better understanding of the active site structure, the enzyme-substrate interaction geometry and the subtle differences determining substrate specificity [23]. As with other cysteine-dependent enzymes, the proteolytic activity of MOA can be inhibited by thiol-modifying agents (e.g., *N*-ethyl maleimide, iodoacetamide) and by specific thiol-reactive compounds.

Z-VAD-fmk (Fig 2) belongs to a family of irreversible substrate-mimetic ketone inhibitors. Originally designed as an inhibitor for the aspartate-specific ICE/caspase-1 cysteine protease [24, 25], Z-VAD-fmk was subsequently found to efficiently inhibit the proteolytic activity of other enzymes in the caspase family [26]. Later publications have shown cross-reactivity between caspase-specific inhibitors and unrelated thiol-dependent enzymes, including peptide: *N*-glycanases (PNGases) [27] and PLCPs [28], Z-VAD-fmk has been successfully used to assess the physiological role of caspases [29]. In spite of available crystal structures of Z-VAD-fmk in complex with different caspases and a PNGase [30, 31], there is currently no structure of the inhibitor in complex with any representative of the PLCP class of proteases.

The unique features of the PLCP domain of MOA and its nature of fungal-specific papain-like protease provide an opportunity to gain insights into the fungal branch of the papain superfamily. Here, we report X-ray crystal structures of MOA in complex with the Z-VAD-fmk inhibitor. Our work sheds light on the substrate binding geometry at the active site of MOA and identifies the positions of the different binding subsites in the protein's catalytic cleft. Furthermore, the structural data argue for the correct placement of MOA in the cysteine protease family tree as the representative of a novel fungal-specific PLCP subfamily.

Materials and Methods

Expression and purification

An IPTG-inducible pT7 vector (MOApT7-LO) containing the cDNA for wild-type MOA (described in [22]) was expressed in *E. coli* strain BL21 (DE3). Bacteria were grown at 37°C in

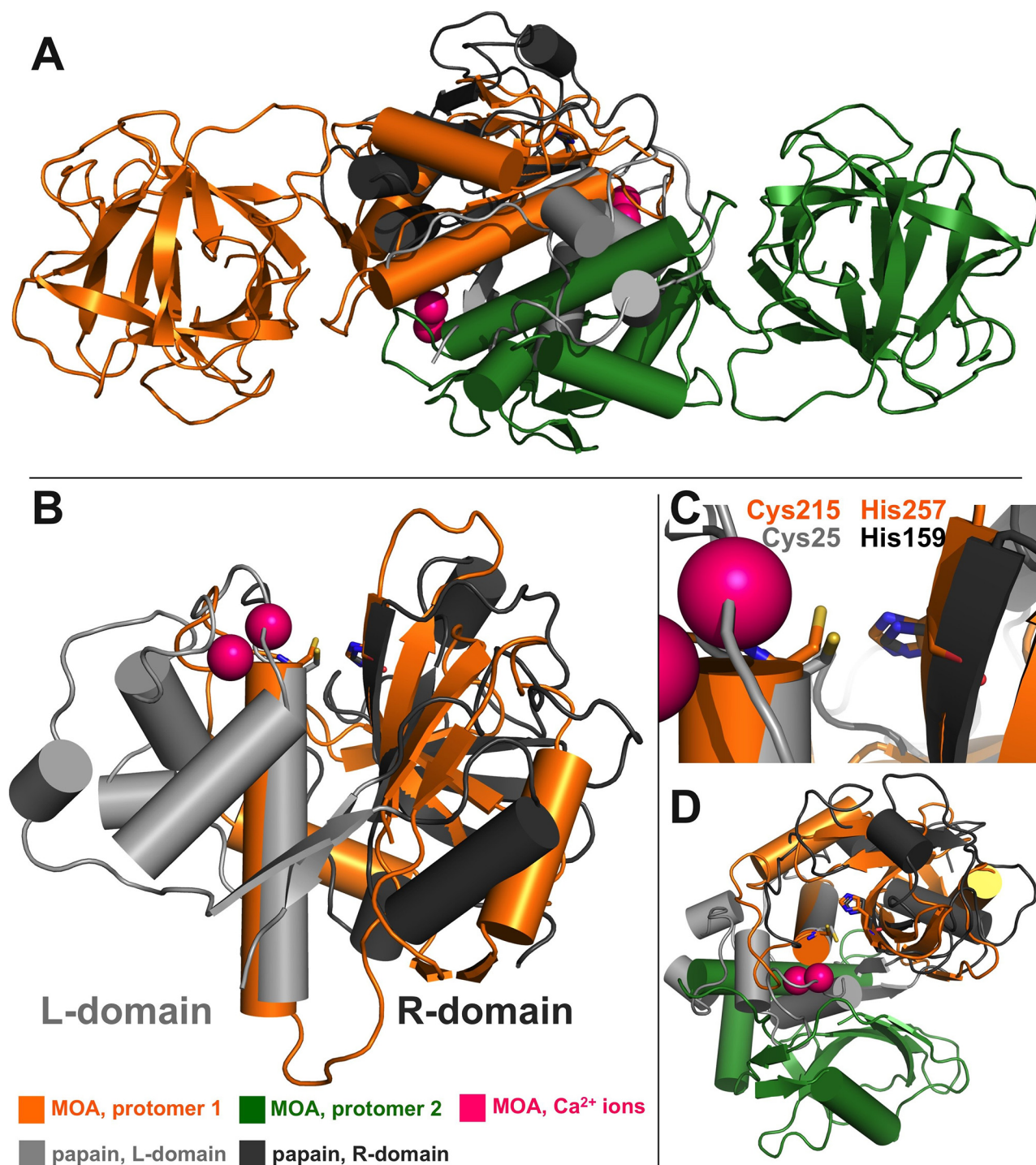


Fig 1. Comparison of MOA and papain PLCP domains. (A) The dimerization domain of MOA is a good structural match of papain and papain-like cysteine proteases. (B,C) This is clearly visible from the structural superposition of MOA (PDB ID: 3EF2 [14]) and papain (PDB ID: 1CVZ [15]), aligned according to the *standard view*, first described by Heinemann *et al.* [16]. The catalytic Cys-His dyads are indicated. In the figure, the L(left)- and R(right)-domain of papain are represented in different colors. (D) The fold conservation between the two enzymes is partially lost in the L-domain, where most structural elements of papain are replaced by the MOA dimerization interface.

doi:10.1371/journal.pone.0149407.g001

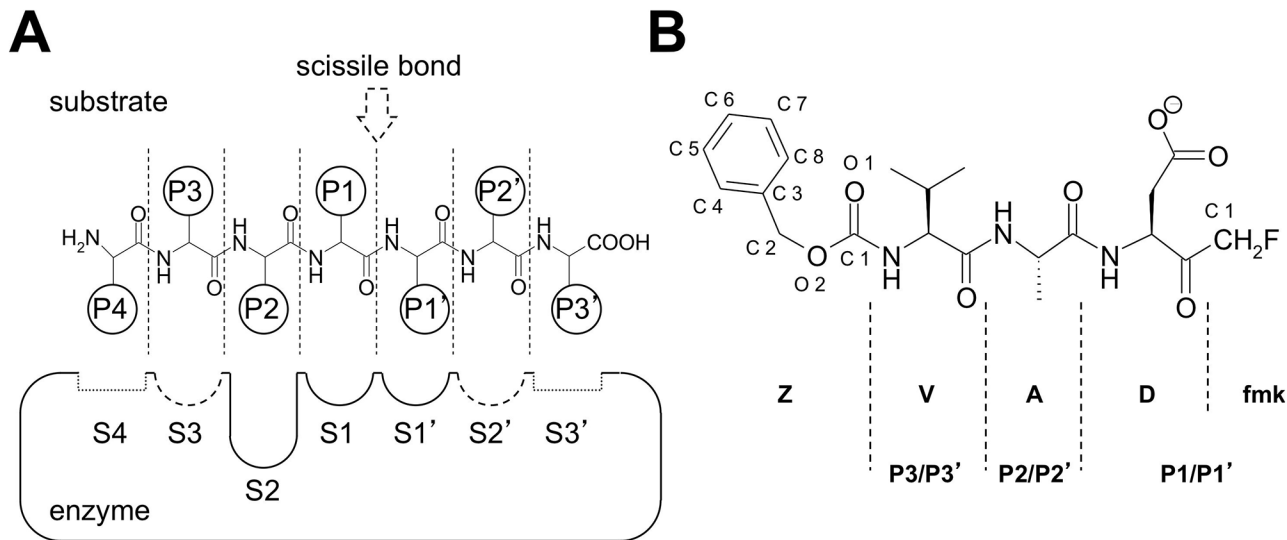


Fig 2. Z-VAD-fmk inhibitor and interaction with PLCP active site. (A) PLCP active site, as mapped by Schechter & Berger [32] and revised by Turk *et al.* [4]; figure adapted from [33]. In a simplified representation of a generic PLCP substrate, the residues on the N-terminal side of the scissile bond are defined as P1-P4 moieties, counting outwards, while those on the C-terminal side are defined as P1'-P3'. Following this description, the scissile bond lies between positions P1 and P1'. The binding subsites on the enzyme are numbered S1-S4 (*unprimed* subsites) and S1'-S3' (*primed* subsites), depending on the substrate position that they interact with. The S2 binding site is represented as a deeper well to highlight its character of substrate binding pocket. Subsites S3 and S2' are drawn as dotted lines to represent their nature of "binding areas". Sites S4 and S3' are represented as shallow grooves to stress their very low conservation among enzymes of the PLCP superfamily. (B) The Z-VAD-fmk molecule is a substrate-mimetic Val-Ala-Asp tripeptide inhibitor carrying a thiol-reactive fluoromethylketone (fmk) on the carboxy-terminus and a capping benzoyl carbonyl (Z) moiety on its N-terminus.

doi:10.1371/journal.pone.0149407.g002

LB medium until the early log phase, induced using 1 mM IPTG and grown at 18°C for 24 hours. Cells were harvested by centrifugation (5000 rcf, 15 min), washed once with a buffer containing 50 mM Tris pH 8.0 and 0.15 M NaCl and stored at -80°C overnight before lysis. After thawing, the bacterial pellets were resuspended in a lysis buffer containing 50 mM Tris pH 8.0, 0.15 M NaCl, 2 mM EDTA, 1x concentrated complete protease inhibitor cocktail EDTA free (Roche Diagnostics Ltd), 1 µl/ml Benzonase nuclease (Thermo Scientific) and 4 mg/ml hen egg white lysozyme. After incubation on a shaker for two hours at RT, the insoluble fraction was removed by two rounds of centrifugation (20000 rcf, 45 min).

The purification protocol for MOA takes advantage of the residual affinity of the sugar binding domain for galactose: as the capture step, the clarified cell lysate was passed through a D-Gal-sepharose affinity column (Thermo Scientific), followed by extensive washing with 20 mM Tris pH 8.0 buffer and elution of the protein using a 1.0 M D-Gal single step gradient. Protein fractions were pooled and concentrated using a 10000 MWCO PES membrane (Vivaspin, Sartorius AG), followed by overnight dialysis against 20 mM acetate pH 4.5, 2 mM EDTA, 2 mM DTT using 7000 MWCO Snakeskin dialysis tubing (Thermo Scientific). Further purification was carried out by cation exchange on a HiTrap SP XL column (GE Healthcare Life Sciences), using a loading buffer containing 20 mM acetate pH 4.5, 2 mM EDTA, 2 mM DTT and eluting the protein with a single step 1.0 M NaCl gradient. Final polishing of the protein preparation was carried out by concentrating the protein to a volume of 500 µl using 10000 MWCO PES membrane concentrator tubes followed by size-exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare Life Sciences) and a buffer containing 20 mM acetate pH 5.0, 2 mM EDTA, 0.2 M D-Gal, 0.15 M NaCl and 2 mM DTT. The fractions containing the purified protein were pooled, concentrated to a final protein concentration of 15–20 mg/ml using concentrator tubes with a 10000 MWCO PES membrane (Vivaspin,

Sartorius AG) and underwent three rounds of buffer exchange against 20 mM acetate pH 5.0, 2 mM EDTA, 2 mM DTT.

Crystallization

Before crystallization, MOA proteolytic activity was tested as described in [22]. MOA crystals grew from a solution containing the purified protein at a concentration of 5 mg/ml, pre-mixed directly before the experiments with the Gal α 1,3(Fuc α 1,2)Gal trisaccharide (Dextra; 1:20 MOA:sugar molar ratio) and the Z-VAD-fmk inhibitor (Sigma-Aldrich). Z-VAD-direct, Z-VAD-inverted and Z-VAD-dual crystals were obtained from 1:3, 1:20 and 1:30 molar ratios (MOA:inhibitor), respectively. The formulation of the crystallization mixture differed among the three crystal forms with respect to precipitant concentration and absence or the presence of DMSO. The crystallization solutions had the following composition: Z-VAD-direct: 0.1 M imidazole pH 8.0, 12% PEG 8000, 0.2 M calcium acetate; Z-VAD-inverted: 0.1 M imidazole pH 8.0, 10% PEG 8000, 5% DMSO, 0.2 M calcium acetate; Z-VAD-dual: 0.1 M imidazole pH 8.0, 16% PEG 8000, 5% DMSO, 0.2 M calcium acetate. The crystals grew as trigonal prisms with dimensions of 0.1 mm \times 0.1 mm \times 0.2 mm. Fully grown crystals were cryoprotected in mother liquor supplemented with 15% ethylene glycol and flash-frozen in liquid nitrogen for data collection. The protein consistently crystallized in space group $P6_322$, with cell parameters of approximately $a = 121$ Å, $b = 121$ Å, and $c = 100$ Å.

Data collection, processing, scaling and structure determination

Final diffraction data were collected at beamlines ID29 and ID23-2 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). The images were processed and scaled using XDS [34]. A summary of the data collection and scaling statistics is given in Table 1. All the structures were solved by molecular replacement with the software PHASER [35], using a modified model of the calcium-bound structure of MOA as search model (PDB ID: 3EF2 [14], lacking the Pro54-Val56 loop and with residues showing flexible or generally poorly defined side chains mutated to Ala). The PHASER solution identified a single MOA protomer in the asymmetric unit; and $mFo-DFc$ maps showed well-defined, positive electron density peaks for the binuclear metal binding site, the three sugar binding sites and the putative active site cleft.

Model building and refinement

Real space refinement of the PHASER output was carried out using Coot [38] by first removing ill-defined side chains and the loop Ile53-Asn55, and subsequently adding missing structural elements in a step-wise fashion as the quality of the electron density map improved. Model building was alternated with refinement cycles using REFMAC5 [39]. The model was completed by first adding the metal ions, then the sugar ligands, the water molecules and small ligands present either in the mother liquor (chloride ions, DMSO) or the cryoprotecting solution (ethylene glycol). The Z-VAD-fmk inhibitor was modeled at the end of the refinement process, when the difference electron density map allowed the unambiguous tracing of the inhibitor molecule in the MOA active site. Ligand occupancy was determined by minimizing the residual difference electron density for the inhibitor molecule and taking the B-factors of nearby interacting atoms into account.

The final model contains residues 2–293, with no electron density accountable for the first methionine residue, suggesting that it might be cleaved off during protein synthesis. The three sugar binding sites of MOA show a different preference for the anomeric form of the reducing end galactose of the Gal α 1,3(Fuc α 1,2)Gal trisaccharide: while the sugar binding α site (residues 20–48) showed full occupancy for β -D-Gal, the β site (residues 72–100) also contains a minor

Table 1. Data collection and refinement statistics.

	ZVAD-direct	ZVAD-dual	ZVAD-inverted
A. Data collection			
Beamline	ESRF ID29	ESRF ID23-2	ESRF ID23-2
Wavelength (Å)	0.9763	0.8726	0.8726
Collected data (deg.)	45	67	50
Oscillation range (deg.)	0.1	0.4	0.5
Space group	<i>P</i> 6 ₃ 22	<i>P</i> 6 ₃ 22	<i>P</i> 6 ₃ 22
Cell parameters: a, b, c (Å)	120.6 120.6 100.1	121.1 121.1 100.0	121.0 121.0 100.0
Resolution (Å) ^a	46.31–1.60 (1.63–1.60)	46.42–1.65 (1.68–1.65)	46.41–1.70 (1.73–1.70)
<i>R</i> _{merge} (%) ^{ab}	6.0 (67.8)	12.1 (71.1)	9.0 (61.0)
<i>R</i> _{meas} ^{ac}	6.7 (76.4)	12.9 (80.6)	10.2 (75.9)
<i>R</i> _{p.i.m.} ^{ad}	3.0 (34.4)	4.6 (36.2)	4.6 (43.5)
CC _{1/2}	99.9 (77.4)	99.6 (70.3)	99.7 (64.4)
Mean <i>I</i> / σ(<i>I</i>) ^a	15.1 (2.3)	11.2 (1.6)	12.0 (1.4)
Completeness (%) ^a	99.6 (97.1)	93.3 (64.7)	98.5 (83.2)
Multiplicity ^a	4.8 (4.8)	7.2 (4.0)	4.7 (2.6)
No. reflections (unique)	275191 (56827)	355074 (48993)	221657 (47195)
B. Refinement			
Resolution (Å)	46.31–1.60	46.42–1.65	46.41–1.70
<i>R</i> _{work} / <i>R</i> _{free} (%) ^e	16.5 / 18.8	16.9 / 19.0	17.7 / 20.3
Macromolecules / a.s.u.	1	1	1
<i>No. atoms</i>			
Protein	2361	2390	2386
Water	262	277	229
Ligands	130	147	134
<i>B-factor (Å²)</i>			
Protein	18.6	13.2	16.8
Water	28.6	24.0	27.0
Ligands	25.3	20.6	26.3
<i>r.m.s.d. from ideal values</i>			
Bond lengths (Å)	0.02	0.02	0.02
Bond angles (deg.)	2.1	2.0	2.0
<i>Ramachandran plot</i>			
Core region (%)	97.4	97.1	97.4
Outliers (%)	0.0	0.0	0.0
PDB ID	5D61	5D62	5D63

^aValues in parentheses refer to highest resolution shell

^b $R_{\text{merge}} = \sum_{\mathbf{h}} \sum_j |I_{\mathbf{h}j} - \langle I_{\mathbf{h}} \rangle| / \sum_{\mathbf{h}} \sum_j I_{\mathbf{h}j}$, where $\langle I_{\mathbf{h}} \rangle$ is the mean intensity of symmetry-related reflections $I_{\mathbf{h}}$

^c $R_{\text{meas}} = \sum_{\mathbf{h}} [N_{\mathbf{h}} / (N_{\mathbf{h}} - 1)]^{1/2} \sum_j |I_{\mathbf{h}j} - \langle I_{\mathbf{h}} \rangle| / \sum_{\mathbf{h}} \sum_j I_{\mathbf{h}j}$, where N is the redundancy of reflection \mathbf{h} [36]

^d $R_{\text{p.i.m.}} = \sum_{\mathbf{h}} [1 / (N_{\mathbf{h}} - 1)]^{1/2} \sum_j |I_{\mathbf{h}j} - \langle I_{\mathbf{h}} \rangle| / \sum_{\mathbf{h}} \sum_j I_{\mathbf{h}j}$ [37]

^e R_{free} was calculated from 5% of randomly selected data for each data set

doi:10.1371/journal.pone.0149407.t001

fraction of the α-D-Gal anomer (unmodeled), and the γ site (residues 123–151) exclusively carries the α-D-Gal anomer. The residual density in the catalytic cleft differs among the three data sets, defining the presence of Z-VAD-fmk in two alternative orientations, referred to as ‘direct’ and ‘inverted’. In the ‘ZVAD-direct’ and ‘ZVAD-inverted’ structures, the inhibitor molecule was modeled at full occupancy in one of these two orientations; additional very low electron

density for the opposite inhibitor orientation present in the ‘ZVAD-inverted’ data set was not modeled. In the ‘ZVAD-dual’ structure, occupancies for the two inhibitor conformations were refined to 0.6/0.4 (direct/inverted). Validation of the model was carried out using Coot [38], the MolProbity server (<http://molprobity.biochem.duke.edu>) [40], and phenix.validate [41]; r. m.s.d. values were calculated using the PDBeFold server (<http://www.ebi.ac.uk/msd-srv/ssm/>) [42]. Refinement statistics are summarized in Table 1. Electron density maps were mostly calculated using the program FFT, which is part of the CCP4 software suite for macromolecular crystallography [43], except for simulated annealing composite OMIT maps (not shown), which were calculated with phenix.refine [44]. All the figures displaying structural data were generated with PyMOL Molecular Graphics System, version 1.5.0.4 (Schrödinger LLC).

Results and Discussion

MOA in complex with the Z-VAD-fmk inhibitor

The structure of MOA in complex with calcium, the irreversible caspase inhibitor Z-VAD-fmk and the branched histo-blood-group B trisaccharide (Gal α 1,3(Fuc β 1,2)Gal) was determined by X-ray crystallography to a resolution of 1.6 Å. In this paper, three different structures are presented, which are based on three independent data sets collected from crystals of the same crystal form (Table 1), which were obtained under slightly different crystallization conditions. The three structures were refined to approximately the same resolution (1.6–1.7 Å), but differ with respect to the orientation and occupancy of the inhibitor, providing unique structural insights into the MOA-Z-VAD-fmk complex. In each case, the structure was independently solved by molecular replacement. The asymmetric unit contains a single MOA protomer, and the biological dimer is generated by crystallographic symmetry (Fig 3A). Apart from the presence of the inhibitor, the structures are essentially identical to the MOA structure without Z-VAD-fmk (r. m.s.d. = 0.1 Å for C α coordinates), except for minor differences in the side chain orientations for some surface residues.

The MOA Z-VAD-fmk complex provides the very first structural data on the interaction of MOA with a substrate analogue, and thus represents the first attempt at mapping its active site. The three structures reported in the article differ with respect to the residual electron density at the active site, which allowed fitting the Z-VAD-fmk inhibitor molecule in one or two alternative orientations (Fig 3B–3D). The different orientations of Z-VAD-fmk were defined as ‘direct’ or ‘inverted’, depending on whether the VAD peptide was aligned with (‘direct’, Fig 3B) or against (‘inverted’, Fig 3C) the standard backbone orientation of a PLCP substrate [4]. The preferred orientation of the inhibitor molecule seems to correlate with the dimethyl sulfoxide (DMSO) content of the mother liquor, where a higher amount of DMSO seems to favor the presence of the ‘inverted’ orientation. Here, we refer to the structures containing the inhibitor in ‘direct’, ‘inverted’ or both orientations as ‘ZVAD-direct’, ‘ZVAD-inverted’ and ‘ZVAD-dual’, respectively. All structures show well defined electron density from the fluoromethylketone carbon to the valine residue, while the benzyloxycarbonyl tail (Z) remains substantially less well-defined.

The reaction mechanism of thiol-dependent enzymes with halomethylketone derivatives has not yet been fully elucidated, with two alternative reaction routes proposed [45]. In both cases, the nucleophilic attack of a cysteine residue results in the loss of the halogen atom and the formation of a covalent adduct (a thioether) between the nucleophilic cysteine and the inhibitor. Consistent with known Z-VAD-fmk-bound complexes [30, 31], the inhibitor molecule binds to the active site of MOA forming a thioether with the catalytic cysteine (Cys215). The electron density maps show well-defined, continuous electron density connecting S^Y of Cys215 and the C¹ atom of the fluoromethyl group. The two atoms are placed at a distance of 1.7 Å, in good agreement with a covalent single C-S bond.

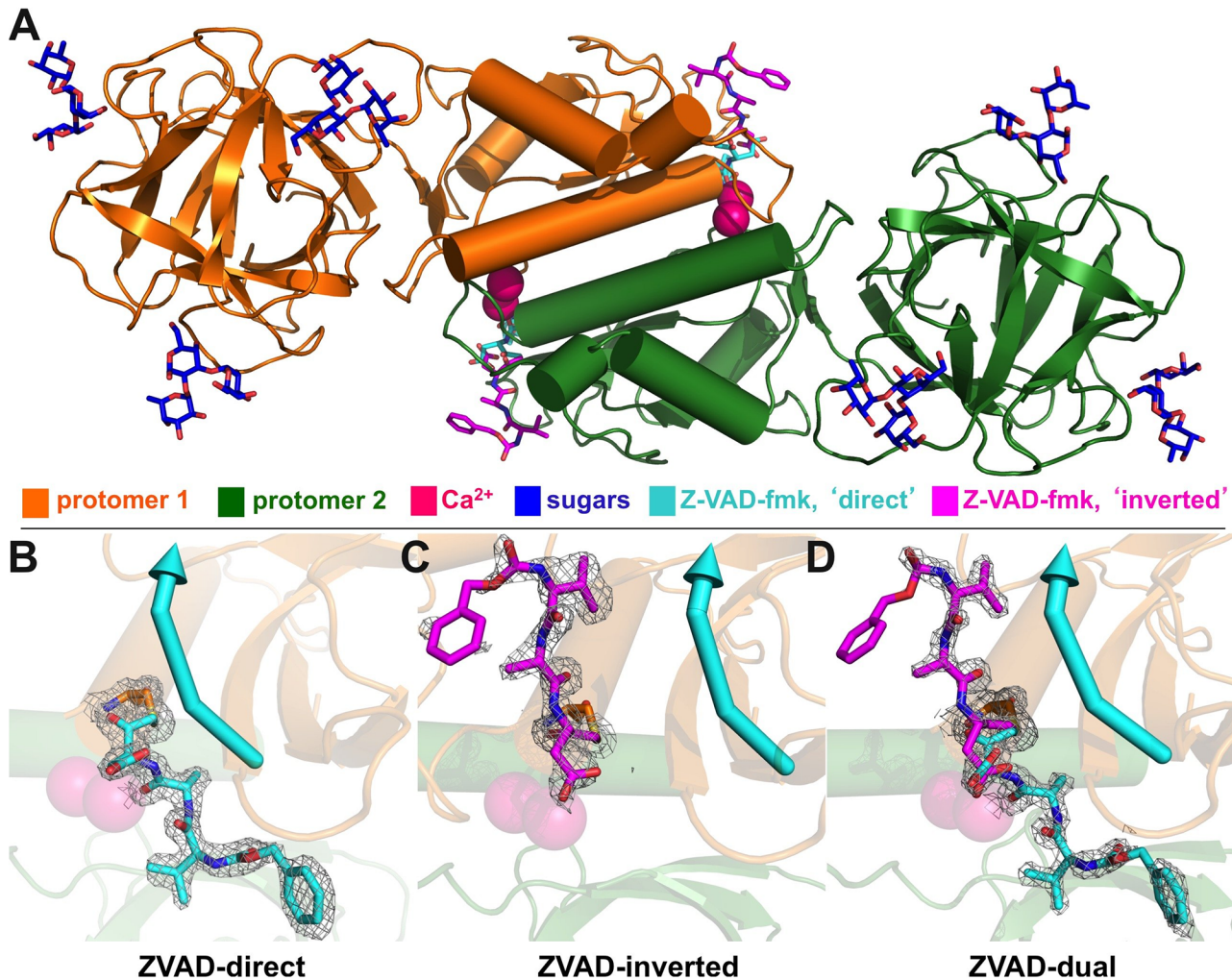


Fig 3. MOA Z-VAD-fmk complex and electron density of the caspase inhibitor. (A) Structure of MOA ('ZVAD-dual') in complex with the Z-VAD-fmk inhibitor (cyan/magenta), three blood group B branched trisaccharide ligands (blue) and two calcium ions (dark magenta). The figure includes the symmetry-related protomer (dark green), shown to match the representation of the functional MOA dimer in Fig 1A. The inhibitor molecule was found in two different orientations (B, 'ZVAD-direct', cyan; C, 'ZVAD-inverted', magenta), interacting with the MOA L- or R-domains, respectively, or in both orientations, with different occupancies (D, 'ZVAD-dual'); the arrow points towards the C-terminus of a natural PLCP substrate. For all three structures, the figure shows the final σ_A -weighted $2mF_o - DF_c$ map for Z-VAD-fmk, contoured at 1σ . For stereo figures and electron density before inclusion of the ligand, see S1 Fig.

doi:10.1371/journal.pone.0149407.g003

Binding geometry of Z-VAD-fmk in the 'direct' orientation

In the 'direct' orientation, the inhibitor molecule extends along the MOA dimerization interface, interacting with both protomers (Fig 3B). The carbonyl group of the aspartate residue lies within hydrogen bonding distance of the backbone NH group of the catalytic cysteine (3.0 Å) and the indole NH of Trp208 (3.0 Å; Fig 4A). The aspartate side chain of the inhibitor points away from the catalytic cleft, and forms a hydrogen bond with the backbone carbonyl oxygen of Ala256 (3.2 Å), which also engages in a hydrogen bonding interaction with the Asp backbone NH group (3.3 Å). The adjacent carbonyl group in the Ala-Asp peptide group interacts strongly with the solvent-exposed Ca^{2+} ion (2.3 Å), replacing a water molecule present in the calcium-bound structure of MOA.

Proceeding further along the chain, the inhibitor molecule interacts with residues provided by both protomers in the MOA dimer. The side chain of the alanine moiety of the VAD

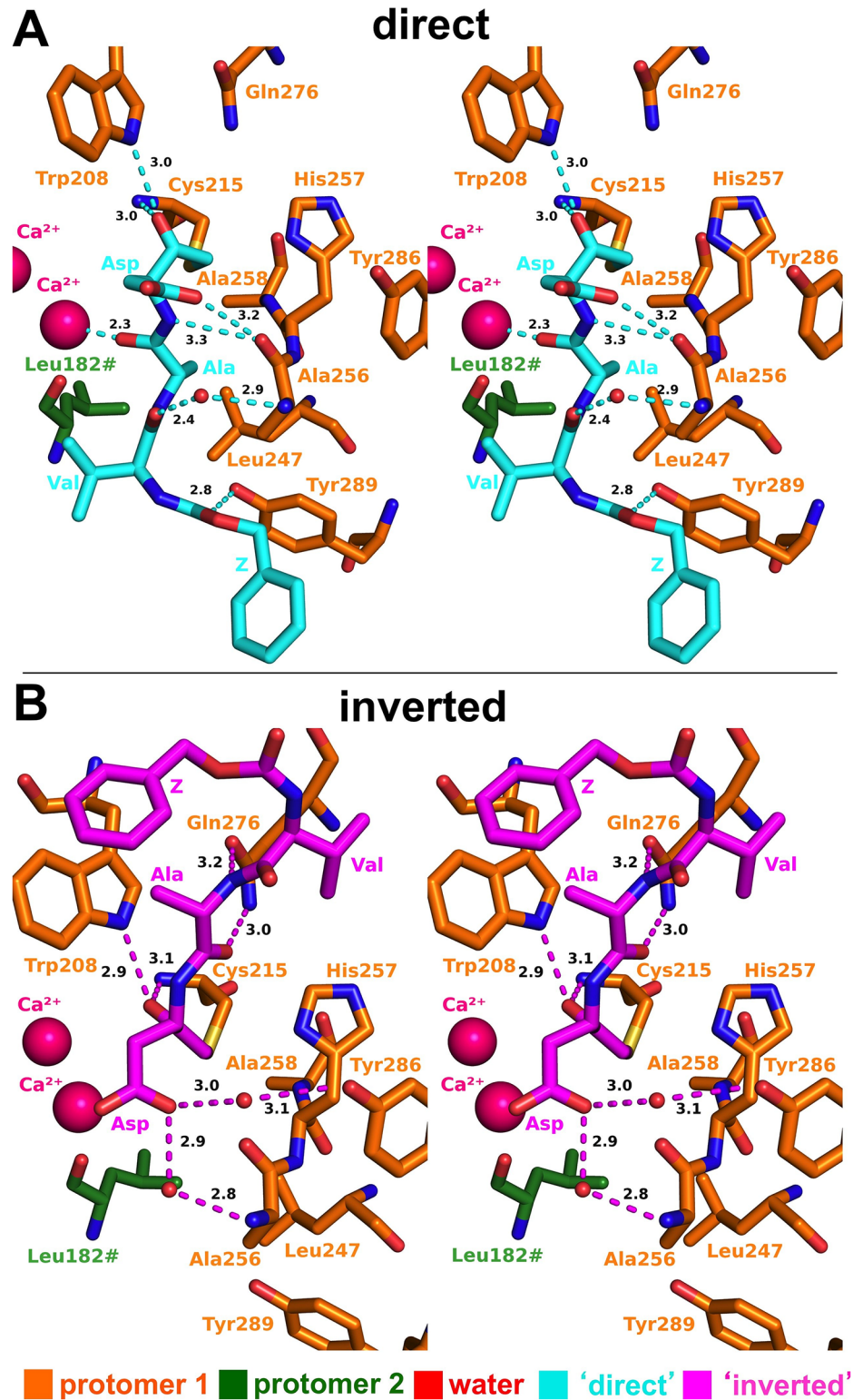


Fig 4. Z-VAD-fmk bound to the MOA active site. Stereographic representation of Z-VAD-fmk in (A) 'direct' (cyan) and (B) 'inverted' (magenta) orientations. Key interactions at the unprimed and primed subsites of the MOA catalytic cleft are depicted.

doi:10.1371/journal.pone.0149407.g004

peptidyl group snugly fits into a hydrophobic cavity lined by residues Leu247, Ala258 and Leu182# (Fig 4A; '#' denotes a residue provided by the symmetry-related protomer). The carbonyl group of the Val-Ala peptide group contacts the backbone NH of Ala256 through a water-mediated interaction (Z-VAD-fmk-Ala3(O)/HOH: 2.4 Å). The valine residue of Z-VAD-fmk is the last well-defined part of the inhibitor. Its side chain points towards a loop in the second protomer (Ile181#-Gly184#), approximately 6 Å away (distance to Ile181#). Beyond that point, the molecule is less ordered in its preferential orientation (S1B Fig), with the carbonyl moiety of the Z group interacting with the hydroxyl group of Tyr289 (2.8 Å; Fig 4A).

Binding geometry of Z-VAD-fmk in the 'inverted' orientation

In the 'inverted orientation', Z-VAD-fmk interacts exclusively with one protomer (Figs 3C and 4B). Close to the newly formed covalent bond, the Z-VAD-fmk keto group interacts with both the backbone NH group of Cys215 (3.1 Å) and the indole NH of Trp208 (2.9 Å; Fig 4B). The side chain of the aspartate moiety points towards the solvent. Its carboxylate group superimposes well with its position in the 'direct' orientation, indirectly interacting with the enzyme through Ala256 (2.9 Å), and indirectly *via* two water molecules. The water molecules bridge the contact with the backbone NH group of Ala256 (Z-VAD-fmk-Asp4-O^{δ1}/HOH: 2.9 Å) and the hydroxyl group of Tyr286 (Z-VAD-fmk-Asp4-O^{δ1}/HOH: 3.0 Å), respectively.

The Ala-Asp peptide group of Z-VAD-fmk is oriented such that the NH group points towards the solvent, while the carbonyl group faces the catalytic cleft. This orientation results in a direct interaction with the side chain amide of Gln276 (Gln276-N^{e2}/Z-VAD-fmk-Ala3-O: 3.2 Å). The alanine side chain is projected towards the left side of the active site (standard view), facing the flat surface of the Trp208 side chain (Fig 4B). The carbonyl group of the Val-Ala peptide group points towards the solvent, while the peptidyl NH is directed towards the enzyme, engaging the oxygen atom of the Gln276 side chain (Gln276-O^{e1}/Z-VAD-fmk-Val2-N: 3.0 Å). The N-terminal end of the inhibitor interacts less strongly with MOA, pointing towards the Phe273-Gly279 loop.

Deciphering the active site of MOA based on the homology with papain

Overall, the inhibitor binding modes observed in the MOA-Z-VAD-fmk complexes are consistent with the PLCP paradigm. Based on the structural alignment with papain, Z-VAD-fmk occupies either the unprimed (S1-S3, 'direct' conformation) or the primed (S1'-S3', 'inverted' conformation) subsites on the catalytic cleft of MOA (Fig 2). As for many PLCP-E-64 complexes [23, 46–49], the 'inverted' Z-VAD-fmk structure provides valuable insight into the fundamental enzyme-substrate interactions, in spite of its reverse backbone orientation. To ease interpretation, we extended the Z-VAD-fmk peptide chain from the 'direct' orientation through the electron density of the inhibitor in the 'inverted' orientation, which resulted in the substrate model shown in Fig 5A. The side-by-side analysis of the catalytic cleft of MOA and papain (Fig 5B, 5C and 5D) provides key pointers to important structural determinants for substrate recognition and catalysis. The combination of the papain-MOA structural alignment with the footprint of the Val-Ala-Asp side chains identifies the catalytic cleft regions corresponding to the S3-S3' subsites of MOA, providing a rationale for the P3-P3' substrate preference observed by Wohlschlager *et al.* [18] (Fig 6).

Non-primed sites (S1-S3)

The non-primed sites (S1-S3) of MOA lie on the dimerization interface and receive contributions from both protomers (Fig 5A). The network of backbone-backbone interactions holding

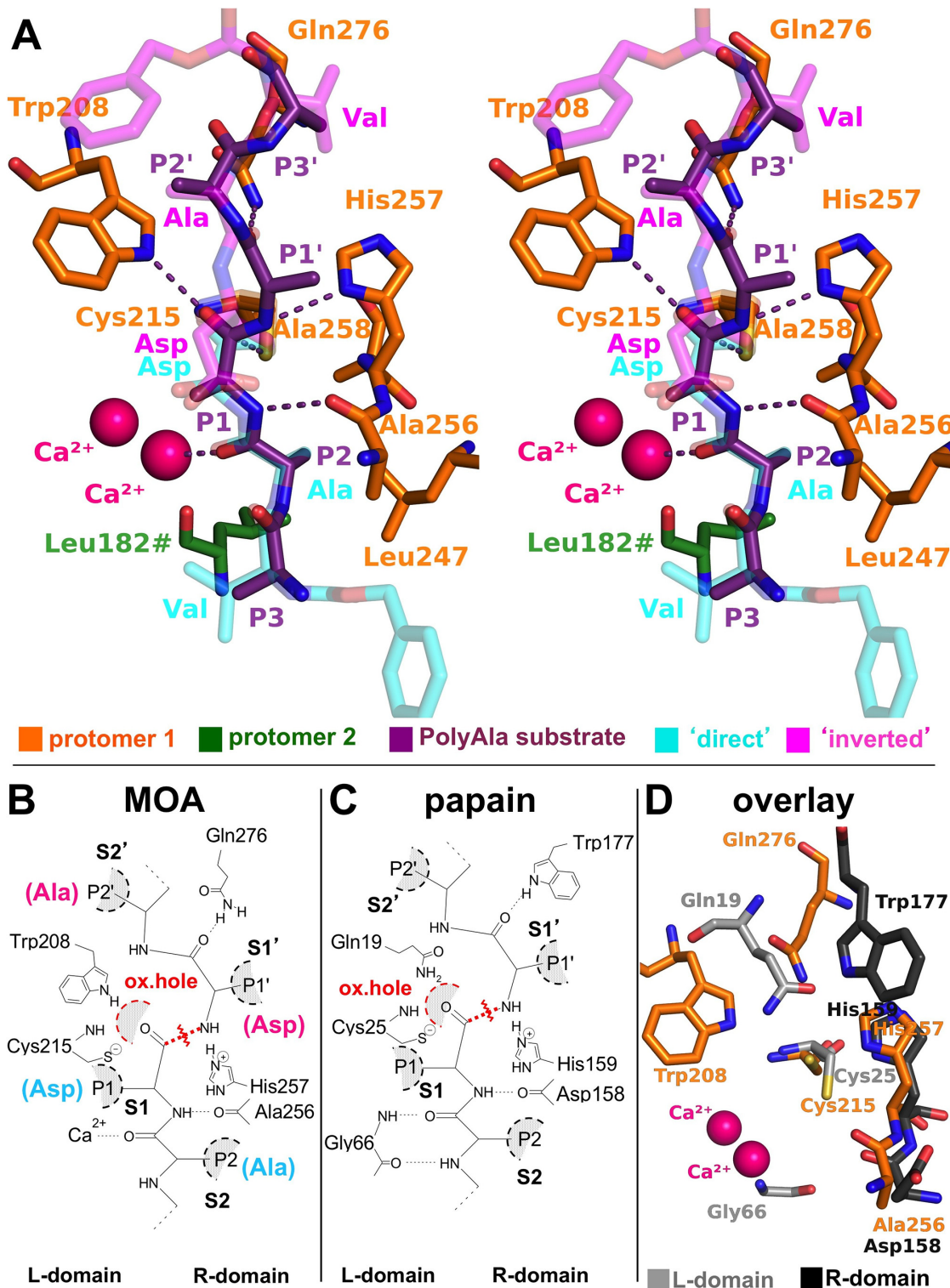


Fig 5. MOA-substrate interactions, derived according to the PLCP paradigm. (A) Stereographic representation of a manually docked polyalanine substrate (purple) to the active site of MOA. The peptide follows the PLCP substrate orientation and takes advantage of the interactions identified by the Z-VAD-fmk-MOA complex. (B,C) Side-by-side schematic representation of the substrate interactions derived for MOA (B) and papain (C); adapted from [5]. The oxyanion hole and the scissile bond are marked in red. (D) Structural superimposition of MOA (PDB ID: 3EF2 [14]) and papain (PDB ID: 1PPN [50]), revealing the Gly66-calcium substitution and the Trp-Gln swap.

doi:10.1371/journal.pone.0149407.g005

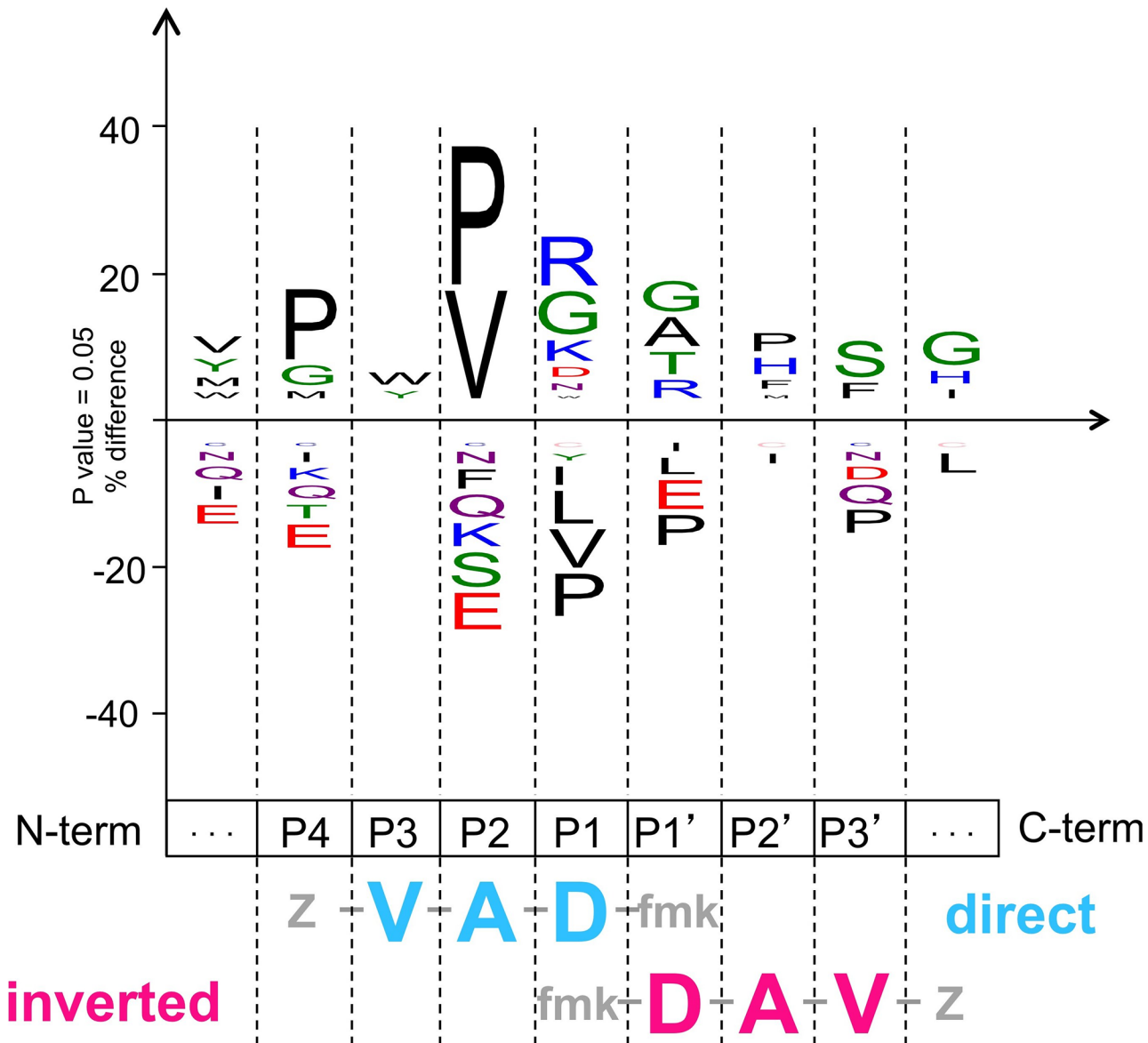


Fig 6. Position-dependent substrate preference of MOA. The diagram was generated *ex novo* through the iceLogo server (<http://iomics.ugent.be/icelogoserver/logo.html>) [51], using a database of peptides derived from the LC-MS analysis of the MOA proteolytic digestion products published by Wohlschlagner *et al.* [18]. The Z-VAD-fmk peptide group in both the direct and inverted orientations is reported underneath the iceLogo diagram for a direct comparison with the binding preferences at each occupied subsite.

doi:10.1371/journal.pone.0149407.g006

the substrate in place forces it into a conformation that has been likened to a strand in a very short β -sheet [52].

S1 subsite. The PLCP S1 subsite lies almost entirely on the L-domain. The most important and better conserved interactions take place between the carbonyl group of the scissile bond on the substrate and the ‘oxyanion hole’ on the enzyme (Fig 5B and 5C). This structural feature is a cavity surrounded by dipoles with the role of stabilizing the transition state during the proteolytic reaction [53, 54]. In PLCPs, the interaction is mediated by the backbone NH of the catalytic cysteine (Cys25 in papain; Fig 5C) and an electron-acceptor group from a nearby residue. The latter is usually provided by the side chain of a glutamine residue (Gln19 in papain),

although some members of the papain superfamily have been shown to carry a different residue (e.g., a tyrosine residue in LapG). A second conserved interaction at the S1 position takes place between the backbone NH from the P1 residue of the substrate and the backbone carbonyl of the residue next to the catalytic His (Asp158 in papain).

In MOA, the latter interaction is conserved, involving the backbone carbonyl of Ala256, while the oxyanion hole exhibits a somewhat different structural framework (Fig 5B). Unlike papain, the NH group of the Trp208 indole ring replaces the glutamine-mediated interaction, whereas the backbone NH group of the catalytic cysteine (Cys215) represents a conserved feature. Gln19 variants of papain retain proteolytic activity [53–55]. This suggests that the oxyanion hole of PLCPs allows for a certain variability, and the conservation of the glutamine residue among family members is not of fundamental importance for the hydrolysis reaction.

Interestingly, the side chain of Asp158, originally thought to be the general acid-base catalyst in the hydrolytic reaction of papain [56] (a role fulfilled by His159), was later determined to play a marginal role in the thiolate-imidazolium pair stabilization [57]. An analysis of the reaction kinetics of different papain variants of Asp158 showed that, while non-essential, substitution of this amino acid with Ala decreased activity to 10% of the wild-type enzyme [58]. In MOA, the equivalent position is occupied by Ala256 (Fig 5B), which might contribute to its relatively low *in vitro* catalytic activity.

The substrate binding conformation in PLCPs forces the side chain of the P1 residue towards the solvent. This structural constraint limits the interaction with the enzyme, which is reflected by the lack of a stringent S1 substrate specificity for PLCPs. The binding surface for the P1 side chain in PLCPs is partly provided by a stretch of residues preceding the catalytic cysteine, and partly by residues from the neighboring Cys63–Gly66 loop. The S1 subsite of MOA is hinted at by the Z-VAD-fmk inhibitor aspartate moiety, which binds to the R-domain through water-mediated interactions. Based on PLCP paradigm, however, the S1 subsite of MOA extends to the L-domain, involving the side chain of Glu210, which would explain the preference for an arginine or a lysine residue at the P1 position (Fig 6).

S2 subsite. The S2 binding site in PLCPs constitutes one of the main selectors for substrate specificity and, as such, is referred to as the ‘*substrate specificity pocket*’ [59, 60] (Figs 2 and 7). In papain, the substrate interacts with the enzyme through the backbone carbonyl and NH groups of a conserved glycine residue (Gly66), with the side chain of the P2 moiety pointing towards a cavity in the active site cleft. The cavity is lined by residues of the R- and the L-domain, the nature and size of which influence the physicochemical properties of the binding pocket. In papain, the walls of the cavity are formed by hydrophobic residues (Pro68, Val133, Ala160) (Fig 7), allowing the processing of substrates with an aliphatic side chain at the P2 position [61, 62]. The residue at the bottom of the cavity (Ser205 in papain) plays a key role in some PLCPs such as cathepsin B, conferring specificity for charged P2 residues, such as arginine [63].

The S2 binding site of MOA introduces two unique features among PLCPs: the presence of the calcium binding site and the dimerization interface (Fig 5A). The S2 specificity pocket of MOA is defined by the P2 alanine side chain of Z-VAD-fmk. Its walls and bottom are lined by the side chains of hydrophobic residues provided by both protomers (Ala258, Leu182# and Leu247) (Fig 7A), defining a more shallow cavity than its papain counterpart (Fig 7B and 7C). The shallow depth and the hydrophobic lining of the S2 pocket provide a structural rationale for the observed substrate preference for a proline or valine residue at the P2 position [18] (Fig 6). A single Ca²⁺ replaces the double backbone-mediated interaction of papain residue Gly66 (Fig 5B). Previous publications linked the calcium-induced opening of its catalytic cleft with the metal-dependence of MOA proteolytic activity [14, 22]. The direct involvement of one of the two metal ions in substrate coordination further reinforces this hypothesis, suggesting a

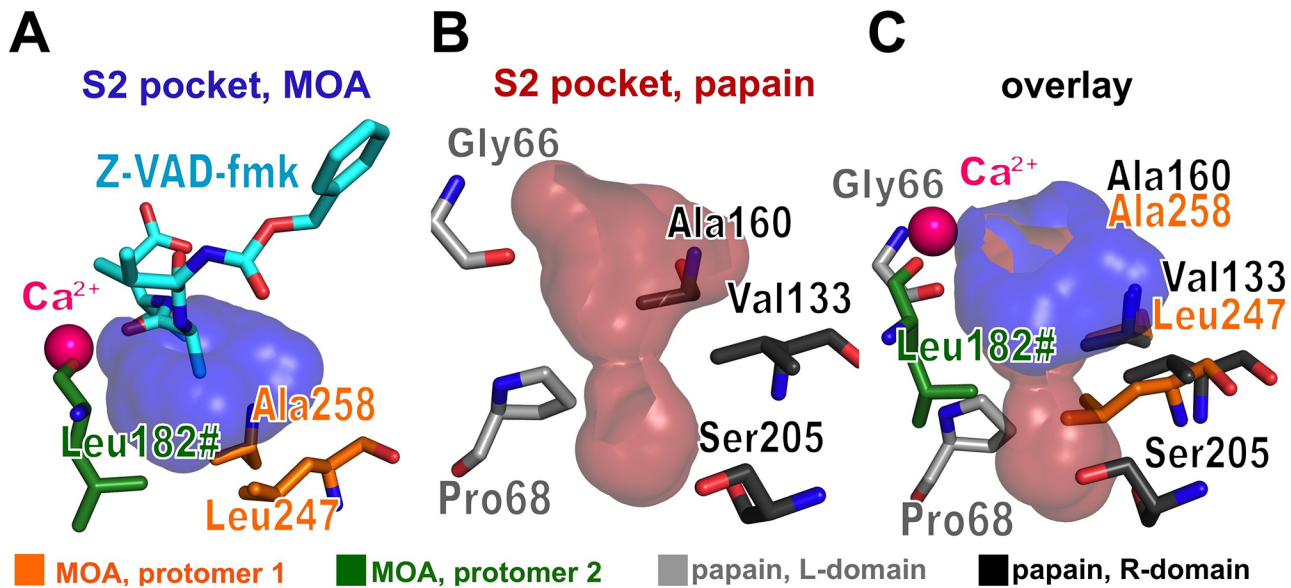


Fig 7. Substrate specificity pockets of MOA and papain. The left and center panels show the solvent-exposed surface at the specificity-determining S2 subsite of papain (A) or MOA (B). A structural alignment of the two proteins (panel C) shows a more shallow S2 binding pocket in MOA compared to papain, explaining the preference of MOA for small P2 residues.

doi:10.1371/journal.pone.0149407.g007

functional involvement in the proteolytic reaction in addition to its structural role. The close distance between the calcium ion and the catalytic center suggests a possible influence of the metal ion on the kinetics of the enzymatic reaction; and in fact the presence (and nature) of the metal ion is essential for the catalytic reaction to occur [22]. The calcium ion is octahedrally coordinated [14] and can be functionally replaced by manganese (II) [22], and possibly other metal ions, with the potential of tuning the catalytic activity of MOA.

S3 subsite. According to the revised definition by Turk *et al.* [4], the S3 binding site primarily involves interactions with the side chains from an R-domain loop (His61-Ala67 in papain). Due to the low sequence conservation of the S3 binding region throughout the PLCP family, this site is usually referred to as ‘binding area’, in contrast to ‘subsite’, which is defined as a spatially well-conserved patch of residues among PLCPs. The corresponding structural feature of MOA is identified by the valine moiety of the Z-VAD-fmk inhibitor. It mainly corresponds to loop Ile181#-Gly184# (part of the L-domain), with a possible contribution from the neighboring Ser160#. The nature of the chemical groups exposed to the S3 subsite, mostly backbone carbonyls, suggests the preference for bulky residues carrying a polar group. A slight preference for tryptophan or tyrosine at the P3 position, as suggested by an analysis of proteolytic products (see iceLogo diagram in Fig 6), supports this notion.

Primed sites (S1'- S3')

In the primed sites, the backbone direction of the VAD tripeptide shows an inverted orientation compared to a PLCP substrate, and analogies are deduced more qualitatively. In the primed subsites of PLCPs, there is only one conserved contact point between the substrate backbone and the enzyme. This interaction involves the P1' carbonyl group from the substrate backbone and the NH group of the indole ring of a conserved tryptophan residue (Trp177 in papain, Fig 5C). In contrast, substrate binding in MOA is mediated by the side chain amide group of a glutamine residue (Glu276). A structure-based sequence alignment of MOA with other papain-like cysteine proteases (not shown) reveals that the conserved glutamine and

tryptophan residues in PLCPs (Gln19 and Trp177 in papain) are spatially and functionally swapped in MOA (Trp208 and Gln276) (Fig 5D). In PLCPs, Trp177 has been suggested to provide shielding of the thiolate-imidazolium pair from the solvent, but at the same time enhancing its nucleophilic character [64]. The much smaller footprint of the glutamine residue in MOA is expected to provide less shielding, consistent with its low enzymatic activity [18].

S1' subsite. The S1' binding site of PLCPs lies entirely on the R-domain. While the S1' site is not a primary specificity selector, it contributes to the enzyme's overall substrate preference [61]. In papain, the S1' site is mostly formed by hydrophobic residues (Ala136, Ala137), with a hydrophilic component provided by two glutamine residues (Gln135, Gln142). Due to the distortion introduced by the inverted backbone orientation, the S1' subsite surface of MOA is only hinted at by the Z-VAD-fmk aspartate side chain. While the terminal carboxylate group of the aspartate residue is engaged in water-mediated interactions with the backbone of Ala256 (Fig 4B), the actual S1' binding surface likely involves a different group of residues (Fig 5A). The S1' subsite provides a shallow groove, lined by the flat surface of the His257 imidazole ring on one side and by the hydroxyl group of an R-domain tyrosine residue (Tyr286) on the other. While the iceLogo diagram suggests glycine or alanine as the most favoured P1' substrate moieties, Tyr286 provides a hydrophilic patch suitable for the interaction with a polar side chain; and indeed, the iceLogo diagram points to a minor preference for threonine or arginine at the P1' position (Fig 6).

S2' subsite. The S2' binding area in PLCPs is identified by a loop capping the oxyanion hole, which directly precedes the catalytic cysteine (residues Gln19 to Ser24 in papain). The alanine moiety of Z-VAD-fmk projects towards MOA residue Trp208 (Fig 5A), which is part of a loop considerably shorter than its papain counterpart. The tryptophan indole ring provides a broad, hydrophobic surface area to interact with the P2' side chain, which could limit the P2' moieties to amino acids with low steric hindrance or alternatively serve as a platform for aromatic stacking interactions. The iceLogo diagram suggests a weak preference for proline, histidine, phenylalanine or methionine as P2' residues (Fig 6), representing both shorter residues similar to the Ala in Z-VAD-fmk, and candidates for stacking.

S3' subsite. Binding sites beyond positions S3' or S2' are not universally defined in PLCPs, as the interaction surface varies for each member, requiring a case-by-case investigation. The valine moiety of the 'inverted' Z-VAD-fmk inhibitor points to an interaction site in the R-domain. While the predicted S3' binding area is lined by the hydrophilic residues Glu274, Gln276 and Asn277, the carboxylate of Glu274 points away from the Val of Z-VAD-fmk, exposing the hydrophobic part of its side chain. Other possible hydrophobic contributions are provided by the neighboring residues Leu281 and Ile284. This site could potentially accommodate a variety of residues with different properties. The iceLogo diagram shown in Fig 6, points to a weak preference for Ser or Phe at the P3' position.

Conclusions

The dimerization domain of MOA represents a novel embodiment of the papain fold with peculiar characteristics, which set it apart from known papain-like proteases. The unique features of MOA include the presence of a binuclear calcium-binding site and a substrate binding cleft contributed by residues from both protomers. The sharing of the binding cleft between the two protomers hints at a functional coupling, which extends beyond their mutual structural support. Another important feature exposed by the MOA Z-VAD-fmk structure is a tryptophan-glutamine structural and functional swap and the direct involvement of one of the metal ions in substrate interaction. While previous data already pointed to the importance of the calcium-induced conformational change in enzyme activation, the direct interaction between one

of the calcium ions and the proteolytic substrate implies additional roles of metal binding in regulating the enzymatic activity. Further structural and biochemical investigations on MOA and homologues are required to explore this possibility and confirm the catalytic role of each active site element, in order to shed further light on the inner workings of this unusual addition to the clan CA cysteine proteases.

Supporting Information

S1 Fig. Stereographic representation of ligand electron density. (A,B) ‘ZVAD-direct’, (C,D) ‘ZVAD-inverted’ and (E,F) ‘ZVAD-dual’ structures. On the left side, panels A, C and E show the final σ_A -weighted $2mFo-DFc$ map (grey, contoured at 1σ) calculated for the coordinates of the Z-VAD-fmk molecule and the catalytic cysteine (Cys215). On the right side, in panels B, D, and F, the σ_A -weighted $mFo-DFc$ difference density map of the same region before the inclusion of the Z-VAD-fmk ligand is shown for comparison (green, contoured at 3σ). Partial occupancy for Asp in the alternative orientation is noticeable in the ‘ZVAD-direct’ and the ‘ZVAD-inverted’ structures when a lower sigma cut-off is applied.

(TIF)

Acknowledgments

We would like to thank the staff at the ESRF and MAX II for assistance and support at synchrotron beamlines ID29 and ID23-2 (ESRF) and I911-3 (MAX II). The work was financed by the University of Oslo and the Norwegian Research Council (grant no. 216625).

Author Contributions

Conceived and designed the experiments: EMG UK GC HCW IJG. Performed the experiments: GC. Analyzed the data: GC AvE. Wrote the paper: GC UK AvE EMG HCW IJG.

References

1. Rawlings ND, Barrett AJ. Evolutionary families of peptidases. *Biochem J.* 1993; 290:205–18. Epub 1993/02/15. PMID: [8439290](#).
2. Rawlings ND, Barrett AJ, Bateman A. *MEROPS*: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 2012; 40(Database issue):D343–D50. Epub 2011/11/17. gkr987 [pii] doi: [10.1093/nar/gkr987](#) PMID: [22086950](#).
3. Berti PJ, Storer AC. Alignment/phylogeny of the papain superfamily of cysteine proteases. *J Mol Biol.* 1995; 246(2):273–83. Epub 1995/02/17. S0022-2836(84)70083-0 [pii] doi: [10.1006/jmbi.1994.0083](#) PMID: [7869379](#).
4. Turk D, Gunčar G, Podobnik M, Turk B. Revised definition of substrate binding sites of papain-like cysteine proteases. *Biol Chem.* 1998; 379(2):137–47. Epub 1998/04/02. PMID: [9524065](#).
5. Turk B, Turk V, Turk D. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem.* 1997; 378(3–4):141–50. Epub 1997/03/01. PMID: [9165064](#).
6. Dubey VK, Pande M, Singh BK, Jagannadham MV. Papain-like proteases: applications of their inhibitors. *Afr J Biotechnol.* 2007; 6(9):1077–86. ISI:000248684900001.
7. Rzychon M, Chmiel D, Stec-Niemczyk J. Modes of inhibition of cysteine proteases. *Acta Biochim Pol.* 2004; 51(4):861–73. Epub 2004/12/31. 045104861. PMID: [15625558](#).
8. O’Farrell PA, Joshua-Tor L. Mutagenesis and crystallographic studies of the catalytic residues of the papain family protease bleomycin hydrolase: new insights into active-site structure. *Biochem J.* 2007; 401(2):421–8. Epub 2006/09/30. BJ20060641 [pii] doi: [10.1042/BJ20060641](#) PMID: [17007609](#).
9. Smith RF, Apple JL, Bottrell DG. The origins of integrated pest management concepts for agricultural crops. In: Apple JL, Smith RF, editors. *Integrated pest management*. New York, NY, U.S.A.: Plenum Press; 1976. p. 1–16.

10. Lecaille F, Kaleta J, Brömme D. Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. *Chem Rev.* 2002; 102(12):4459–88. Epub 2002/12/12. cr0101656 [pii]. PMID: [12475197](#).
11. Shindo T, van der Hoorn RAL. Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. *Mol Plant Pathol.* 2008; 9(1):119–25. doi: [10.1111/J.1364-3703.2007.00439.X](#) ISI:000251813900011. PMID: [18705889](#)
12. Schmitz C, Kinner A, Kölling R. The deubiquitinating enzyme Ubp1 affects sorting of the ATP-binding cassette-transporter Ste6 in the endocytic pathway. *Mol Biol Cell.* 2005; 16(3):1319–29. Epub 2005/01/07. E04-05-0425 [pii] doi: [10.1091/mbc.E04-05-0425](#) PMID: [15635103](#).
13. Futai E, Maeda T, Sorimachi H, Kitamoto K, Ishiura S, Suzuki K. The protease activity of a calpain-like cysteine protease in *Saccharomyces cerevisiae* is required for alkaline adaptation and sporulation. *Mol Gen Genet.* 1999; 260(6):559–68. Epub 1999/02/03. PMID: [9928935](#).
14. Grahn EM, Winter HC, Tateno H, Goldstein IJ, Kregel U. Structural characterization of a lectin from the mushroom *Marasmius oreades* in complex with the blood group B trisaccharide and Calcium. *J Mol Biol.* 2009; 390(3):457–66. Epub 2009/05/12. S0022-2836(09)00541-5 [pii] doi: [10.1016/j.jmb.2009.04.074](#) PMID: [19426740](#).
15. Tsuge H, Nishimura T, Tada Y, Asao T, Turk D, Turk V, et al. Inhibition mechanism of cathepsin L-specific inhibitors based on the crystal structure of papain-CLIK148 complex. *Biochem Biophys Res Commun.* 1999; 266(2):411–6. Epub 1999/12/22. [pii]. PMID: [10600517](#).
16. Heinemann U, Pal GP, Hilgenfeld R, Saenger W. Crystal and molecular structure of the sulfhydryl protease calotropin DI at 3.2 Å resolution. *J Mol Biol.* 1982; 161(4):591–606. Epub 1982/11/15. 0022-2836(82)90410-7 [pii]. PMID: [6759664](#).
17. Kruger RP, Winter HC, Simonson-Leff N, Stuckey JA, Goldstein IJ, Dixon JE. Cloning, expression, and characterization of the Galα1,3Gal high affinity lectin from the mushroom *Marasmius oreades*. *J Biol Chem.* 2002; 277(17):15002–5. Epub 2002/02/12. [pii]. PMID: [11836254](#).
18. Wohlschlager T, Butschi A, Zurfluh K, Vonesch SC, auf dem Keller U, Gehrig P, et al. Nematotoxicity of *Marasmius oreades* agglutinin (MOA) depends on glycolipid binding and cysteine protease activity. *J Biol Chem.* 2011; 286(35):30337–43. Epub 2011/07/16. M111.258202 [pii] doi: [10.1074/jbc.M111.258202](#) PMID: [21757752](#).
19. Bleuler-Martínez S, Butschi A, Garbani M, Wälti MA, Wohlschlager T, Potthoff E, et al. A lectin-mediated resistance of higher fungi against predators and parasites. *Mol Ecol.* 2011; 20(14):3056–70. Epub 2011/04/14. doi: [10.1111/j.1365-294X.2011.05093.x](#) PMID: [21486374](#).
20. Cordara G, Winter HC, Goldstein IJ, Kregel U, Sandvig K. The fungal chimerolectin MOA inhibits protein and DNA synthesis in NIH/3T3 cells and may induce BAX-mediated apoptosis. *Biochem Biophys Res Commun.* 2014; 447(4):586–9. Epub 2014/04/22. S0006-291X(14)00684-6 [pii] doi: [10.1016/j.bbrc.2014.04.043](#) PMID: [24747075](#).
21. Grahn E, Askarieh G, Holmner Å, Tateno H, Winter HC, Goldstein IJ, et al. Crystal structure of the *Marasmius oreades* mushroom lectin in complex with a xenotransplantation epitope. *J Mol Biol.* 2007; 369(3):710–21. Epub 2007/04/20. S0022-2836(07)00329-4 [pii] doi: [10.1016/j.jmb.2007.03.016](#) PMID: [17442345](#).
22. Cordara G, Egge-Jacobsen W, Johansen HT, Winter HC, Goldstein IJ, Sandvig K, et al. *Marasmius oreades* agglutinin (MOA) is a chimerolectin with proteolytic activity. *Biochem Biophys Res Commun.* 2011; 408(3):405–10. Epub 2011/04/26. S0006-291X(11)00605-X [pii] PMID: [21513701](#).
23. Matsumoto K, Mizoue K, Kitamura K, Tse W-C, Huber CP, Ishida T. Structural basis of inhibition of cysteine proteases by E-64 and its derivatives. *Biopolymers.* 1999; 51(1):99–107. Epub 1999/06/25. [pii] doi: [10.1002/\(SICI\)1097-0282\(1999\)51:1<99::AID-BIP11>3.0.CO;2-R](#) PMID: [10380357](#).
24. Howard AD, Kostura MJ, Thornberry N, Ding GJF, Limjuco G, Weidner J, et al. IL-1-converting enzyme requires aspartic acid residues for processing of the IL-1β precursor at two distinct sites and does not cleave 31-kDa IL-1α. *J Immunol.* 1991; 147(9):2964–9. Epub 1991/11/01. PMID: [1919001](#).
25. Thornberry NA, Bull HG, Calaycay JR, Chapman KT, Howard AD, Kostura MJ, et al. A novel heterodimeric cysteine protease is required for interleukin-1β processing in monocytes. *Nature.* 1992; 356(6372):768–74. Epub 1992/04/30. doi: [10.1038/356768a0](#) PMID: [1574116](#).
26. Garcia-Calvo M, Peterson EP, Leiting B, Ruel R, Nicholson DW, Thornberry NA. Inhibition of human caspases by peptide-based and macromolecular inhibitors. *J Biol Chem.* 1998; 273(49):32608–13. Epub 1998/11/26. PMID: [9829999](#).
27. Misaghi S, Pacold ME, Blom D, Ploegh HL, Korbel GA. Using a small molecule inhibitor of peptide:N-glycanase to probe its role in glycoprotein turnover. *Chem Biol.* 2004; 11(12):1677–87. Epub 2004/12/22. S1074-5521(04)00334-5 [pii] doi: [10.1016/j.chembiol.2004.11.010](#) PMID: [15610852](#).
28. Rozman-Pungerčar J, Kopitar-Jerala N, Bogyo M, Turk D, Vasiljeva O, Štefe I, et al. Inhibition of papain-like cysteine proteases and legumain by caspase-specific inhibitors: when reaction mechanism

- is more important than specificity. *Cell Death Differ.* 2003; 10(8):881–8. Epub 2003/07/18. [pii]. PMID: [12867995](#).
29. Fearhead HO, Dinsdale D, Cohen GM. An interleukin-1 β -converting enzyme-like protease is a common mediator of apoptosis in thymocytes. *FEBS Lett.* 1995; 375(3):283–8. Epub 1995/11/20. 0014-5793(95)01228-7 [pii]. PMID: [7498519](#).
 30. Lee J-H, Choi JM, Lee C, Yi KJ, Cho Y. Structure of a peptide-*N*-glycanase-Rad23 complex: insight into the deglycosylation for denatured glycoproteins. *Proc Natl Acad Sci USA.* 2005; 102(26):9144–9. Epub 2005/06/21. 0502082102 [pii] doi: [10.1073/pnas.0502082102](#) PMID: [15964983](#).
 31. Zhao G, Zhou X, Wang L, Li G, Kisker C, Lennarz WJ, et al. Structure of the mouse peptide *N*-glycanase-HR23 complex suggests co-evolution of the endoplasmic reticulum-associated degradation and DNA repair pathways. *J Biol Chem.* 2006; 281(19):13751–61. Epub 2006/02/28. M600137200 [pii] doi: [10.1074/jbc.M600137200](#) PMID: [16500903](#).
 32. Schechter I, Berger A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun.* 1967; 27(2):157–62. Epub 1967/04/20. S0006-291X(67)80055-X [pii]. PMID: [6035483](#).
 33. Brömme D. Papain-like cysteine proteases. *Curr Protoc Protein Sci.* 2001; Chapter 21. Epub 2008/04/23. doi: [10.1002/0471140864.ps2102s21](#) PMID: [18429163](#).
 34. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr.* 2010; 66(Pt 2):125–32. Epub 2010/02/04. S0907444909047337 [pii] doi: [10.1107/S0907444909047337](#) PMID: [20124692](#).
 35. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. *Phaser* crystallographic software. *J Appl Crystallogr.* 2007; 40(4):658–74. Epub 2007/08/01. doi: [10.1107/S0021889807021206](#) PMID: [19461840](#).
 36. Diederichs K, Karplus PA. Improved *R*-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol.* 1997; 4(4):269–75. Epub 1997/04/01. PMID: [9095194](#).
 37. Evans P. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr.* 2006; 62(Pt 1):72–82. Epub 2005/12/22. S0907444905036693 [pii] doi: [10.1107/S0907444905036693](#) PMID: [16369096](#).
 38. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of *Coot*. *Acta Crystallogr D Biol Crystallogr.* 2010; 66(4):486–501. Epub 2010/04/13. S0907444910007493 [pii] doi: [10.1107/S0907444910007493](#) PMID: [20383002](#).
 39. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. *REFMAC5* for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr.* 2011; 67(4):355–67. Epub 2011/04/05. S0907444911001314 [pii] doi: [10.1107/S0907444911001314](#) PMID: [21460454](#).
 40. Chen VB, Arendall WBI, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010; 66(Pt 1):12–21. Epub 2010/01/09. S0907444909042073 [pii] doi: [10.1107/S0907444909042073](#) PMID: [20057044](#).
 41. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010; 66(2):213–21. Epub 2010/02/04. S0907444909052925 [pii] doi: [10.1107/S0907444909052925](#) PMID: [20124702](#).
 42. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr.* 2004; 60(12–1):2256–68. Epub 2004/12/02. S0907444904026460 [pii] doi: [10.1107/S0907444904026460](#) PMID: [15572779](#).
 43. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the *CCP4* suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 2011; 67(4):235–42. Epub 2011/04/05. S0907444910045749 [pii] doi: [10.1107/S0907444910045749](#) PMID: [21460441](#).
 44. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr.* 2012; 68(Pt 4):352–67. Epub 2012/04/17. S0907444912001308 [pii] doi: [10.1107/S0907444912001308](#) PMID: [22505256](#).
 45. Powers JC, Asgian JL, Ekici ÖD, James KE. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev.* 2002; 102(12):4639–750. Epub 2002/12/12. cr010182v [pii]. PMID: [12475205](#).
 46. Varughese KI, Ahmed FR, Carey PR, Hasnain S, Huber CP, Storer AC. Crystal structure of a papain-E-64 complex. *Biochemistry.* 1989; 28(3):1330–2. Epub 1989/02/07. PMID: [2713367](#).
 47. Kerr ID, Lee JH, Pandey KC, Harrison A, Sajid M, Rosenthal PJ, et al. Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. *J Med Chem.* 2009; 52(3):852–7. Epub 2009/01/09. [pii]. PMID: [19128015](#).

48. Hofmann B, Schomburg D, Hecht HJ. Crystal structure of a thiol proteinase from *Staphylococcus aureus* V-8 in the E-64 inhibitor complex. *Acta Crystallogr A Biol Crystallogr*. 1993; 2(49 Supplement): c102.
49. Gomes MTR, Teixeira RD, Lopes MTP, Nagem RAP, Salas CE. X-ray crystal structure of CMS1MS2: a high proteolytic activity cysteine proteinase from *Carica candamarcensis*. *Amino Acids*. 2012; 43(6):2381–91. Epub 2012/05/23. doi: [10.1007/s00726-012-1318-7](https://doi.org/10.1007/s00726-012-1318-7) PMID: [22610687](https://pubmed.ncbi.nlm.nih.gov/22610687/).
50. Pickersgill RW, Harris GW, Garman E. Structure of monoclinic papain at 1.60 Å resolution. *Acta Crystallogr B Struct Sci*. 1992; 48(1):59–67.
51. Colaert N, Helsen K, Martens L, Vandekerckhove J, Gevaert K. Improved visualization of protein consensus sequences by iceLogo. *Nat Methods*. 2009; 6(11):786–7. Epub 2009/10/31. nmeth1109-786 [pii] doi: [10.1038/nmeth1109-786](https://doi.org/10.1038/nmeth1109-786) PMID: [19876014](https://pubmed.ncbi.nlm.nih.gov/19876014/).
52. Madala PK, Tyndall JD, Nall T, Fairlie DP. Proteases universally recognize beta strands in their active sites. *Chem Rev*. 2010; 110(6):PR1–PR31. Epub 2010/04/10. doi: [10.1021/cr900368a](https://doi.org/10.1021/cr900368a) PMID: [20377171](https://pubmed.ncbi.nlm.nih.gov/20377171/).
53. Ménard R, Carrière J, Laflamme P, Plouffe C, Khouri HE, Vernet T, et al. Contribution of the glutamine 19 side chain to transition-state stabilization in the oxyanion hole of papain. *Biochemistry*. 1991; 30(37):8924–8. Epub 1991/09/17. PMID: [1892809](https://pubmed.ncbi.nlm.nih.gov/1892809/).
54. Wolfenden R, Snider MJ. The depth of chemical time and the power of enzymes as catalysts. *Acc Chem Res*. 2001; 34(12):938–45. doi: [10.1021/Ar000058i](https://doi.org/10.1021/Ar000058i) ISI:000172875200002. PMID: [11747411](https://pubmed.ncbi.nlm.nih.gov/11747411/)
55. Ménard R, Plouffe C, Laflamme P, Vernet T, Tessier DC, Thomas DY, et al. Modification of the electrostatic environment is tolerated in the oxyanion hole of the cysteine protease papain. *Biochemistry*. 1995; 34(2):464–71. Epub 1995/01/17. PMID: [7819238](https://pubmed.ncbi.nlm.nih.gov/7819238/).
56. Polgár L, Halász P. Current problems in mechanistic studies of serine and cysteine proteinases. *Biochem J*. 1982; 207(1):1–10. Epub 1982/10/01. PMID: [6758764](https://pubmed.ncbi.nlm.nih.gov/6758764/).
57. Storer AC, Ménard R. Catalytic mechanism in papain family of cysteine peptidases. *Methods Enzymol*. 1994; 244:486–500. Epub 1994/01/01. PMID: [7845227](https://pubmed.ncbi.nlm.nih.gov/7845227/).
58. Taylor MAJ, Baker KC, Connerton IF, Cummings NJ, Harris GW, Henderson IMJ, et al. An unequivocal example of cysteine proteinase activity affected by multiple electrostatic interactions. *Protein Eng*. 1994; 7(10):1267–76. Epub 1994/10/01. PMID: [7855143](https://pubmed.ncbi.nlm.nih.gov/7855143/).
59. Khouri HE, Vernet T, Ménard R, Parlati F, Laflamme P, Tessier DC, et al. Engineering of papain: selective alteration of substrate specificity by site-directed mutagenesis. *Biochemistry*. 1991; 30(37):8929–36. Epub 1991/09/17. PMID: [1892810](https://pubmed.ncbi.nlm.nih.gov/1892810/).
60. Nägler DK, Tam W, Storer AC, Krupa JC, Mort JS, Ménard R. Interdependency of sequence and positional specificities for cysteine proteases of the papain family. *Biochemistry*. 1999; 38(15):4868–74. Epub 1999/04/14. [pii]. PMID: [10200176](https://pubmed.ncbi.nlm.nih.gov/10200176/).
61. Ménard R, Carmona E, Plouffe C, Brömme D, Konishi Y, Lefebvre J, et al. The specificity of the S₁' subsite of cysteine proteases. *FEBS Lett*. 1993; 328(1–2):107–10. Epub 1993/08/09. 0014-5793(93)80975-Z [pii]. PMID: [8344413](https://pubmed.ncbi.nlm.nih.gov/8344413/).
62. Gauthier F, Moreau T, Lalmanach G, Brillard-Bourdet M, Ferrer-Di Martino M, Juliano L. A new, sensitive fluorogenic substrate for papain based on the sequence of the cystatin inhibitory site. *Arch Biochem Biophys*. 1993; 306(2):304–8. Epub 1993/11/01. S000398618371516X [pii]. PMID: [8215429](https://pubmed.ncbi.nlm.nih.gov/8215429/).
63. Brömme D, Bonneau PR, Lachance P, Storer AC. Engineering the S₂ subsite specificity of human cathepsin S to a cathepsin L- and cathepsin B-like specificity. *J Biol Chem*. 1994; 269(48):30238–42. Epub 1994/12/02. PMID: [7982933](https://pubmed.ncbi.nlm.nih.gov/7982933/).
64. Gul S, Hussain S, Thomas MP, Resmini M, Verma CS, Thomas EW, et al. Generation of nucleophilic character in the Cys25/His159 ion pair of papain involves Trp177 but not Asp158. *Biochemistry*. 2008; 47(7):2025–35. Epub 2008/01/30. doi: [10.1021/bi702126p](https://doi.org/10.1021/bi702126p) PMID: [18225918](https://pubmed.ncbi.nlm.nih.gov/18225918/).