

Structural basis for substrate specificity of the peroxisomal acyl-CoA hydrolase MpaH' involved in mycophenolic acid biosynthesis

Cai You¹, Fengwei Li¹, Xingwang Zhang¹, Li Ma¹, Yu-Zhong Zhang¹, Wei Zhang¹ and Shengying Li^{1,2} 

¹ State Key Laboratory of Microbial Technology, Shandong University, Qingdao, China

² Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, China

Keywords

biosynthesis; crystal structure; mycophenolic acid; peroxisome; α/β hydrolase

Correspondence

S. Li, State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong 266237, China
 Tel: +86-532-58632466
 E-mail: lishengying@sdu.edu.cn

(Received 28 January 2021, revised 30 March 2021, accepted 9 April 2021)

doi:10.1111/febs.15874

Mycophenolic acid (MPA) is a fungal natural product and first-line immunosuppressive drug for organ transplantations and autoimmune diseases. In the compartmentalized biosynthesis of MPA, the acyl-coenzyme A (CoA) hydrolase MpaH' located in peroxisomes catalyzes the highly specific hydrolysis of MPA-CoA to produce the final product MPA. The strict substrate specificity of MpaH' not only averts undesired hydrolysis of various cellular acyl-CoAs, but also prevents MPA-CoA from further peroxisomal β -oxidation catabolism. To elucidate the structural basis for this important property, in this study, we solve the crystal structures of the substrate-free form of MpaH' and the MpaH'^{S139A} mutant in complex with the product MPA. The MpaH' structure reveals a canonical α/β -hydrolase fold with an unusually large cap domain and a rare location of the acidic residue D163 of catalytic triad after strand β 6. MpaH' also forms an atypical dimer with the unique C-terminal helices α 13 and α 14 arming the cap domain of the other protomer and indirectly participating in the substrate binding. With these characteristics, we propose that MpaH' and its homologs form a new subfamily of α/β hydrolase fold protein. The crystal structure of MpaH'^{S139A}/MPA complex and the modeled structure of MpaH'/MPA-CoA, together with the structure-guided mutagenesis analysis and isothermal titration calorimetry (ITC) measurements, provide important mechanistic insights into the high substrate specificity of MpaH'.

Introduction

Mycophenolic acid (MPA) was discovered in 1893 from the culture of a *Penicillium brevicompactum* strain. This fungi-specific antibiotic was subsequently re-isolated from many other species of *Penicillium* [1,2]. Among the wide spectrum of its biological activities, the highly selective inhibitory activity of MPA toward human inosine-5'-monophosphate dehydrogenase has led this old compound to be developed into a first-line immunosuppressive drug. Clinically, MPA

has been broadly used to suppress the immunologic rejections during organ transplantations and also to treat various autoimmune diseases [2,3]. Of note, the annual sales of MPA-derived drugs are still over \$ 1 billion after the patent cliff [4,5].

Recently, we elucidated the full biosynthetic pathway of MPA in *P. brevicompactum* NRRL 864, which featuring the compartmentalized cooperation between biosynthetic and β -oxidation catabolic machineries and

Abbreviations

ASA, accessible surface area; CoA, coenzyme A; HPLC, high-performance liquid chromatography; ITC, isothermal titration calorimetry; MMF, mycophenolate mofetil; MPA, mycophenolic acid; RMSD, root-mean-square deviation; SAD, single-wavelength anomalous dispersion; TEV, tobacco etch virus.

the organelle-associated catalytic mechanisms [6]. In this pathway, the acyl-coenzyme A (CoA) hydrolase MpaH' that is subcellularly located in peroxisomes plays a key role in determining the specific production of MPA. In peroxisomes, MpaH' efficiently and highly selectively catalyzes the hydrolysis of MPA-CoA, thereby preventing nonspecific hydrolysis of other acyl-CoAs and the excessive chain-shortening degradation of MPA-CoA by the peroxisomal β -oxidation process (Fig. 1). Thus, MpaH' exerts a critical valve-like function in the MPA biosynthetic pathway.

Sequence analysis indicates that MpaH' belongs to α/β hydrolase superfamily, which represents one of the largest families of structurally related proteins [7]. The family members catalyze a vast variety of reactions, such as hydrolysis, redox, acyl-transfer, isomerization, and lyase-type reactions, using a conserved protein fold and catalytic core [8]. The canonical α/β hydrolase fold comprises a mostly parallel eight-stranded β -sheet in the core with α helices surrounded on both sides (Fig. 2A). The substrate binding pocket usually consists of several inserts forming a 'cap/lid' domain on top of the active site, with the 'lid' mostly referred to small mobile elements, which can undergo a

conformational transition, while the 'cap' defined as a large immobile module [9]. This 'cap/lid' domain often gives rise to the structural complexity and substrate diversity within the superfamily [10,11]. Another shared property of canonical α/β hydrolases is the catalytic triad consisting of a nucleophile, an acidic residue, and a histidine, with the nucleophile normally located in a tight turn after $\beta 5$, thus being called the 'nucleophile elbow' [7].

We previously identified the putative catalytic triad of MpaH' to be S139-D163-H365 by multiple protein sequence alignment. However, considering that MpaH' exhibits low sequence identities to the structure-known α/β hydrolase family members, it is difficult to accurately locate the catalytic core and also to further reveal the catalytic mechanism solely based on sequence analysis. Thus, herein, we report the crystal structures of the substrate-free form of MpaH' and its inactive mutant MpaH'^{S139A} in complex with product MPA. Together with structural modeling and mutagenesis analysis, we elucidated the catalytic mechanism and the structural basis for the high substrate specificity of MpaH'. We also identified the unusual dimeric formation and other unique features of MpaH', which

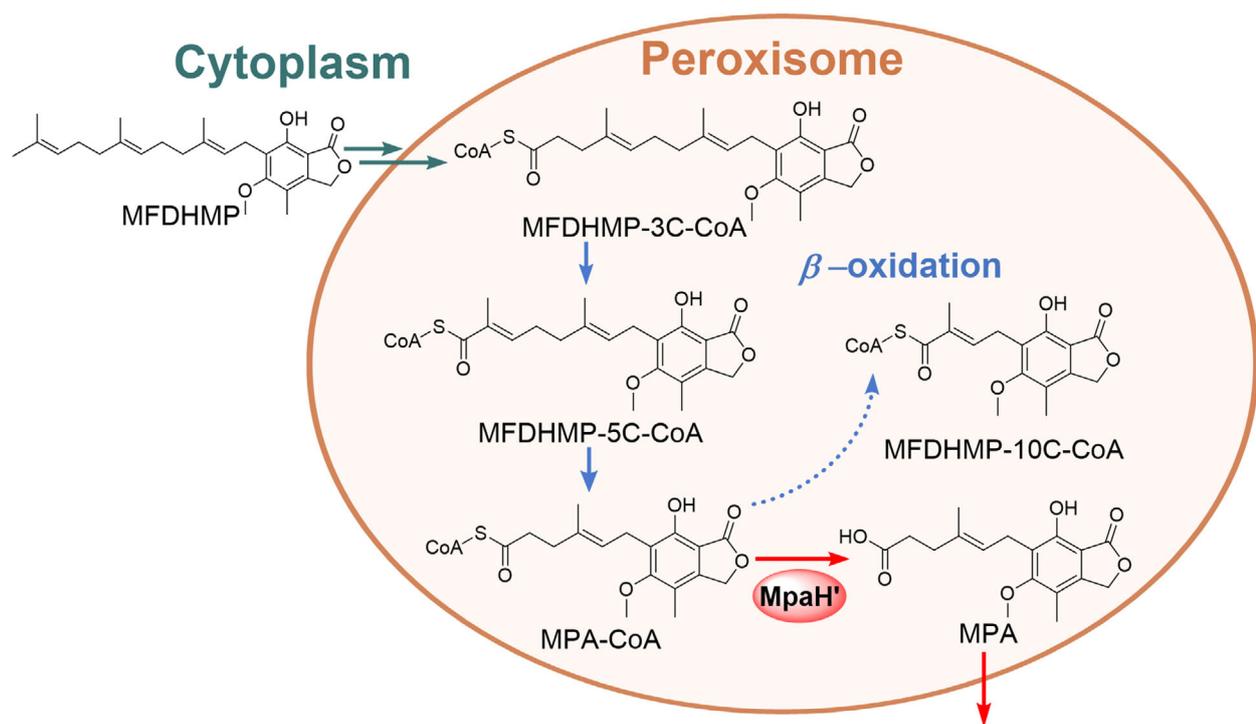


Fig. 1. Part of the compartmentalized MPA biosynthesis mediated by the β -oxidation machinery and the acyl-CoA hydrolase MpaH' in fungal peroxisomes. The solid arrows indicate the major MPA biosynthetic steps. The dashed arrow denotes the shunt pathway when MpaH' is absent. MFDHMP: 5-O-methyl-4-farnesyl-3,5-dihydroxy-6-methylphthalide. MFDHMP-3C, MFDHMP-5C, and MFDHMP-10C represent the reduction of 3, 5, and 10 carbon atoms from MFDHMP, respectively.

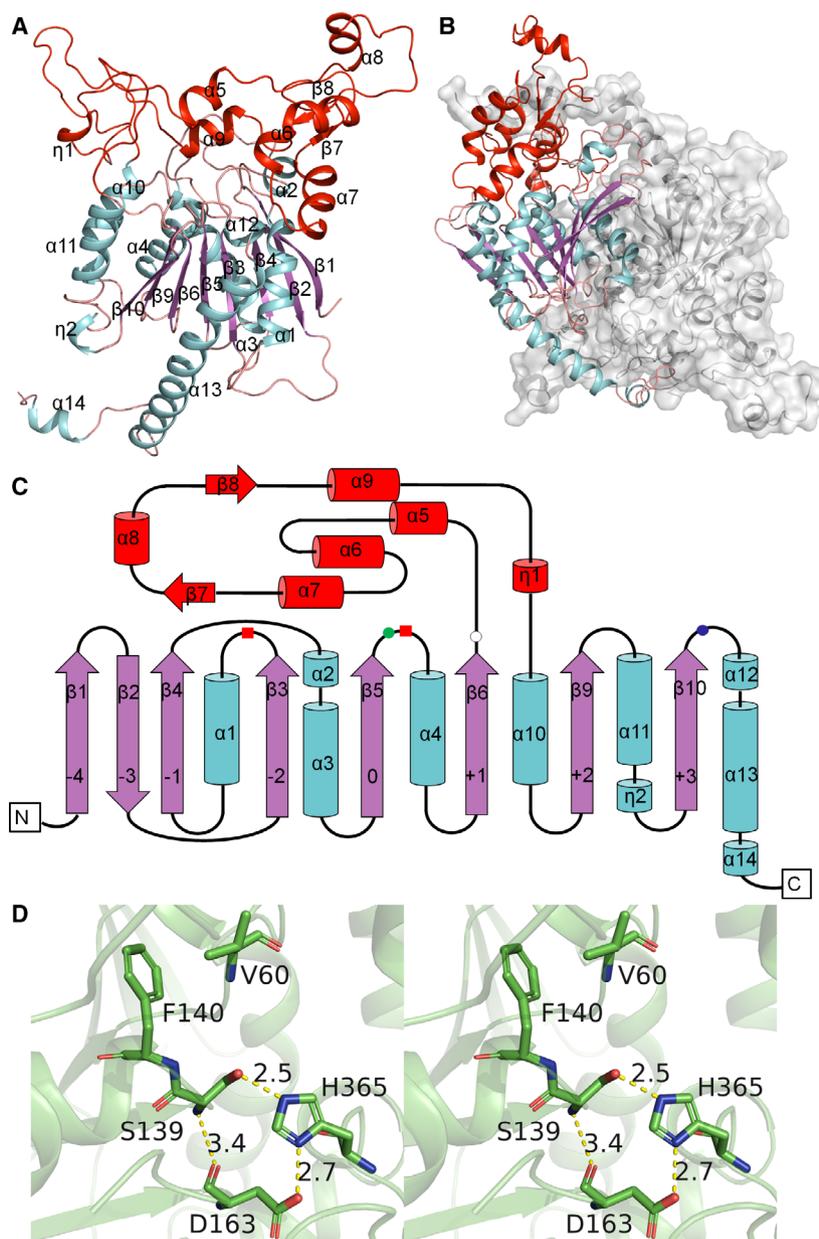


Fig. 2. Structural analysis of MpaH'. (A) Cartoon representation of an MpaH' protomer with the secondary structural elements indicated. The unique large cap domain is colored red, and the core domain is colored by secondary structures with α helices in cyan, β -sheets in magenta, and loops in salmon. (B) The dimeric state of MpaH'. One protomer is colored as presented in (A), while the other protomer is colored gray in both cartoon and surface representation. (C) Topology representation of the MpaH' structure. The cap domain is highlighted in red, and the core domain is colored by secondary structures with α -helices in cyan and β -strands in magenta. α -Helices are shown as barrels and β -strands as arrows with secondary structures can be numbered β_0 , and the *N*- and *C*-terminal β -strands are correspondingly assigned as β_{-1} , β_{-2} , β_{-3} , β_{-4} , β_{+1} , β_{+2} , and β_{+3} according to the architecture analysis and indicated at the bottom of the arrows. The catalytic triad S139-D163-H365 is shown as green, white, and blue circles, respectively. The oxyanion hole residues are marked as red rectangles. (D) Stereoview of the catalytic triad of MpaH'. Catalytic residues are displayed as sticks. Hydrogen bonds are displayed as yellow dotted lines with distances indicated in angstrom.

make this protein and its homologs form a new sub-family of α/β -hydrolase fold proteins.

Results and Discussion

Structure of substrate-free MpaH'

We first solved the crystal structure of substrate-free MpaH' at 1.99 Å resolution in space group $P2_1$. Due to the unavailability of the structure of homologous proteins (the most similar protein with a structure, PDB ID code: 2Y6V, is only 13% identical to MpaH'), the structure of MpaH' was solved by the

selenomethionine single-wavelength anomalous dispersion (SAD) method, with four molecules in one asymmetric unit forming two dimers (PDB ID code: 7DBI). Although our construct comprises the full-length sequence of MpaH' for crystallization, we were only able to model residues 3–422 for chain A and chain C, 4–420 for chain B, and 3–420 for chain D. The remaining *N*- and *C*-terminal residues lacked defined electron density, supposedly due to their flexibility. The majority of modeled residues exhibited excellent electron density and geometry. The best modeled molecule chain B was used for the subsequent analysis.

The overall structure of MpaH' exhibits a typical α/β -hydrolase fold, comprised of a core domain and a cap domain (Fig. 2A,B). The core domain of MpaH' mainly comprises a half-barrel consisting of eight β strands in the order of $\beta 1$ - $\beta 2$ - $\beta 4$ - $\beta 3$ - $\beta 5$ - $\beta 6$ - $\beta 9$ - $\beta 10$, with strand $\beta 1$ almost perpendicular to $\beta 10$. The half-barrel is flanked by α helices on both sides, with helices $\alpha 1$, $\alpha 2$, $\alpha 12$, and $\alpha 13$ packing onto the defined convex face of the half-barrel, and helices $\alpha 3$, $\alpha 4$, $\alpha 10$, and $\alpha 11$ packing onto the concave face. The cap domain is inserted between strand $\beta 6$ and helix $\alpha 10$ (residues L168–R301), mainly consisting of five helices ($\alpha 5$ - $\alpha 9$), two antiparallel β strands ($\beta 7$ and $\beta 8$), and several long loops (Fig. 2C). This domain, covering on top of the core domain (active site), usually participates in substrate binding and varies a lot in different α/β hydrolases. It is noticed that half of the long helix $\alpha 13$ extends away from the core region and connects to the last helix $\alpha 14$ via a long loop, which extends further to arm the cap domain of the other protomer (Fig. 2B). Residues F389–R401, which continue helix $\alpha 13$, make this last helix in the core domain longer than those of other known α/β hydrolases. These features make MpaH' quite different from other α/β hydrolases.

The active site of MpaH'

The active site of a canonical α/β hydrolase usually contains a catalytic triad. Structural analysis (Fig. 2D) and sequence alignment (Fig. 3) identified S139-D163-H365 as the catalytic triad of MpaH', corresponding to the nucleophile, acidic amino acid, and histidine residue. The nucleophile S139 is located at a tight turn after strand $\beta 5$, namely the nucleophile elbow consisting of $^{137}\text{GHSFG}^{141}$, which is commonly conserved in α/β hydrolases with a consensus sequence of G-X-Nuc-X-G [7,12]. The acidic residue D163 is located at the C terminus of strand $\beta 6$ although its counterpart in other α/β hydrolases is typically located after strand $\beta 7$ (Fig. S1) [9,10]. In common with other α/β hydrolases, the catalytic H365 is located at the interface of the core and the cap domain, specifically, in the loop between $\beta 10$ and $\alpha 12$ of MpaH' (Fig. 2C). H365 is

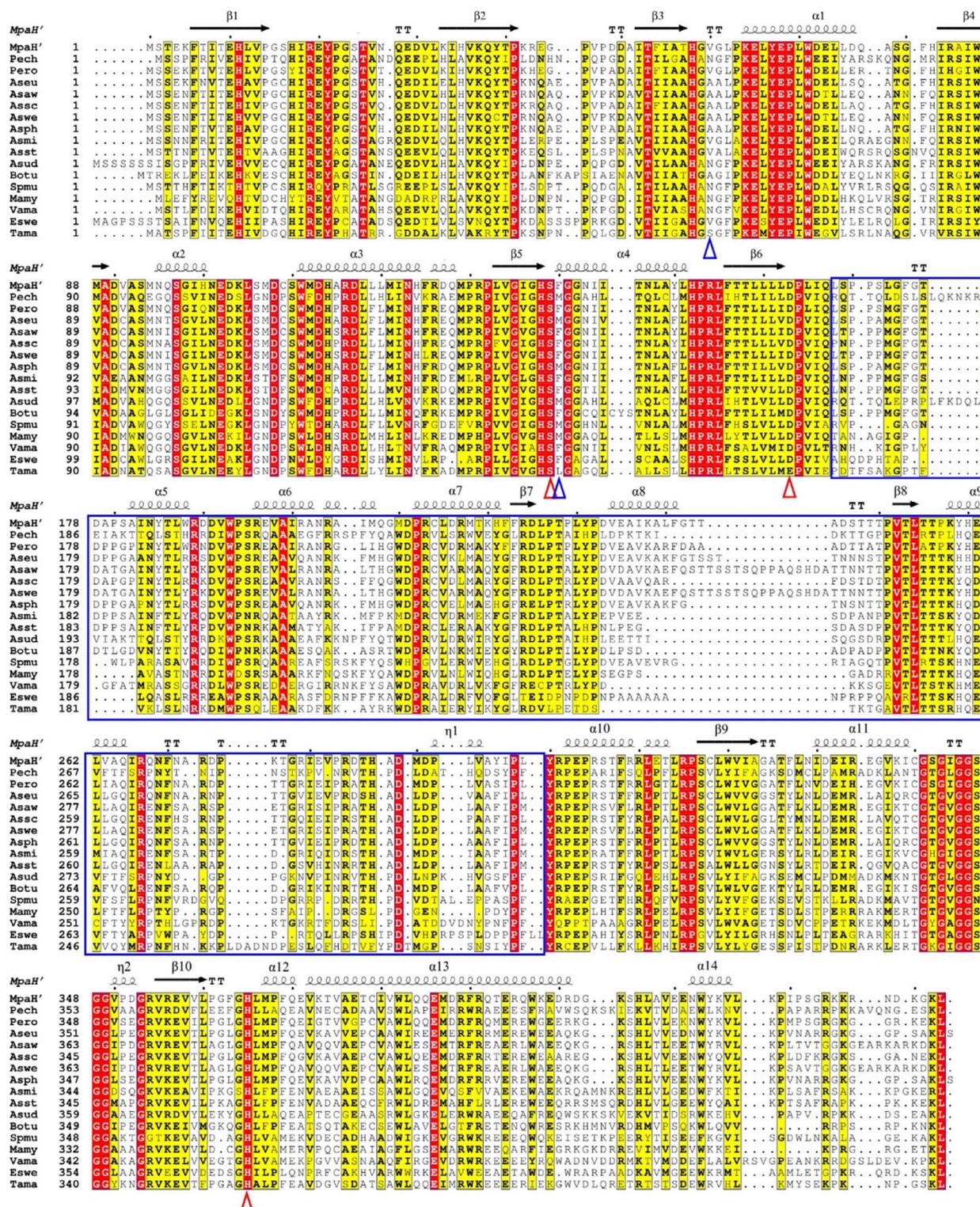
hydrogen bonded to D163 with a distance of 2.7 Å between D163 O $\delta 2$ and H365 N $\delta 1$ atoms to correctly position H365 for catalysis. With the responsibility for activation of the nucleophile, the N $\epsilon 2$ atom of H365 forms a hydrogen bond with O γ atom of S139 with a distance of 2.5 Å (Fig. 2D). The essential roles of the catalytic triad residues were determined by mutating each individual residue to alanine, and the effects of mutations on substrate conversion ratio were evaluated by high-performance liquid chromatography (HPLC) analysis. At a catalyst concentration of 20 nM, substitution of the nucleophilic S139 or histidine H365 to alanine completely abolished the hydrolytic activity of MpaH', while the acidic residue mutant D163A lost 49.5% of the catalytic activity when compared with the wild-type enzyme (Fig. S2).

The structure of mutant MpaH'^{S139A} in complex with MPA

Initially, we attempted to solve the structure of the active site mutant S139A of MpaH' (MpaH'^{S139A}) in complex with the native substrate MPA-CoA in order to elucidate the structural basis for the substrate specificity of MpaH'. However, both co-crystallization and substrate soaking using MPA-CoA turned to form the product–enzyme complexes. To reason this unexpected result, we conducted HPLC analysis of the reaction mixture of MpaH'^{S139A} and MPA-CoA under the crystallization conditions, revealing that the residual activity of this mutant was responsible for hydrolyzing MPA-CoA into MPA and CoASH at the high concentrations of enzyme and substrate (Fig. S3).

The crystal of MpaH'^{S139A}/MPA complex grew in the same crystal form (PDB ID code: 7DBL with a resolution of 1.84 Å) as the substrate-free MpaH'. Again, four protomers form two dimers in each crystallographic symmetric unit, with parts of N- and C-terminal residues lacking electron density (Fig. 4A). Chain B was used for the subsequent analysis. Structural overlay of the MpaH'^{S139A}/MPA complex and the substrate-free MpaH' revealed an RMSD value of 0.11 Å for the aligned C α atoms, indicative of no large

Fig. 3. Sequence alignment of MpaH' and its homologous proteins from other fungi. Pech, *Penicillium chrysogenum* (XP_002564597); Pero, *P. roqueforti* (W6QL41); Aseu, *Aspergillus eucaalypticola* (XP_025387189); Asaw, *A. awamori* (GCB20447); Assc, *A. sclerotii carbonarius* (PY108358); Aswe, *A. welwitschiae* (XP_026629444); Asph, *A. phoenicis* (RDK42238); Asmi, *A. minisclerotigenes* (KAB8270075); Asst, *A. steynii* (XP_024710389); Asud, *A. udagawae* (GAO83358); Botu, *Botrytis tulipae* (TGO14402); Spmu, *Sphaceloma murrayae* (PNS19944); Mamy, *Madurella mycetomatis* (KXX74861); Vama, *Valsa mali* (KUI69808); Eswe, *Escovopsis weberi* (KOS19841); and Tama, *Talaromyces marneffeii* (XP_002152328). Red and yellow backgrounds indicate sequence identity and similarity, respectively. The catalytic triad residues are marked with red triangles, and the residues constituting the oxyanion hole are indicated by blue triangles. The region of the cap domain is boxed in blue rectangles. Protein sequence alignments were performed using MUSCLE [31] integrated within MEGA 5.1 [32].



conformational changes (Fig. 4B). A large cavity is located at the interface between the core domain and the cap domain of MpaH'. Expectedly, the product

MPA is well accommodated in this cavity defined by clear electron density (Fig. S4). This binding pocket is predominantly hydrophobic attributed to residues

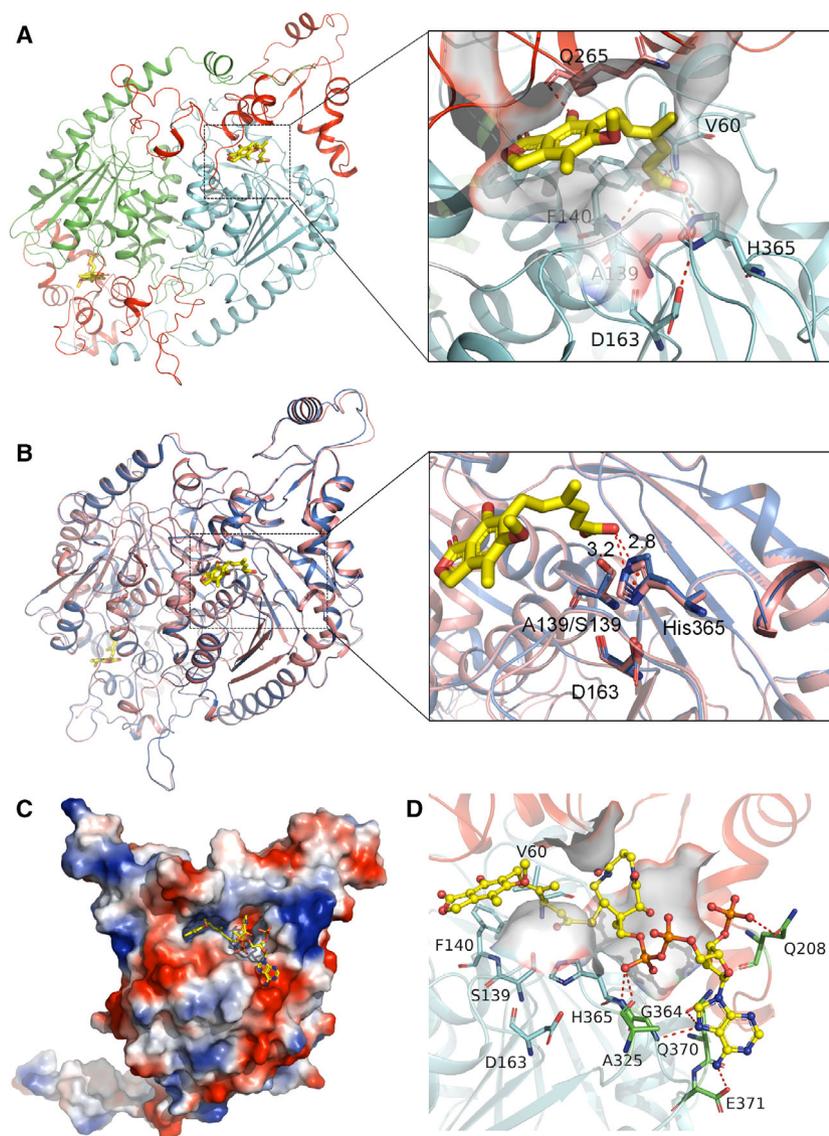


Fig. 4. Analysis of the MpaH'^{S139A}/MPA complex structure and the MpaH'/MPA-CoA complex model. (A) The overall structure of MpaH'^{S139A}/MPA complex as a homodimer. The two cap domains are colored red in both protomers, while the core domains are colored cyan and green. The two molecules of product MPA are presented as yellow sticks. The enlarged view of the substrate binding pocket is shown in the right panel. The catalytic residues and oxyanion residues are shown as sticks. The hydrophobic binding pocket formed by residues A139, L165, Q167, F327, L328, P171, A182, and I183 is shown as surface. (B) Superimposition of the MpaH'^{S139A}/MPA complex and the substrate-free MpaH' structure, which are colored marine and salmon, respectively. MPA is shown in yellow stick. The right panel is a close-up of the active sites in the two aligned structures. Hydrogen bonds are represented as red dotted lines with distances indicated in angstrom. The pseudoatoms of the imidazole ring of residue H365 in both structures are shown as red spheres. (C) Overall structure of the MpaH'/MPA-CoA complex model with MpaH' in electrostatic representation and MPA-CoA depicted as yellow ball and stick. (D) A close-up of MPA-CoA binding pocket. MpaH' is displayed as cyan and red cartoons for the core domain and cap domain, respectively. The catalytic and oxyanion hole residues are shown as cyan sticks. Residues interacting with the phosphoadenosine group of CoA moiety are shown as green sticks with hydrogen bonds indicated as red dot lines. Residues L62, I206, A182, F327, and L366 that hydrophobically interact with the CoA moiety are represented as 40% transparent surface.

A139, L165, Q167, F327, and L328 from the core domain and residues P171, A182, and I183 from the cap domain (Fig. 4A).

The bicyclic moiety of MPA accepts one hydrogen bond from the backbone oxygen of Q265 in the cap domain. The O6 atom is hydrogen bonded to Nε2 atom

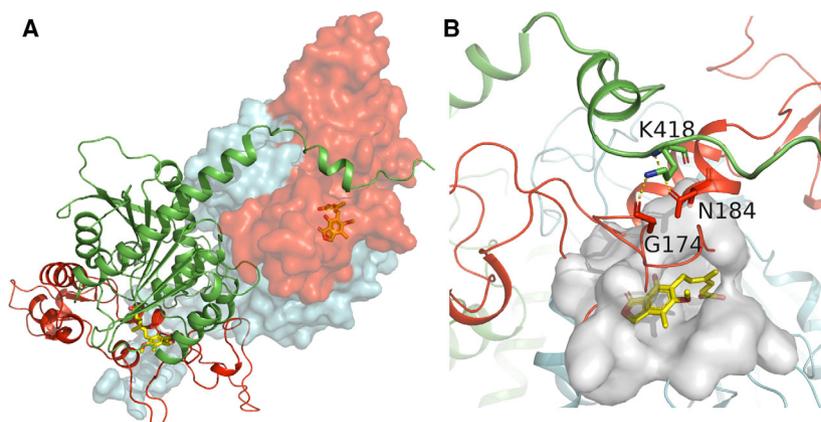


Fig. 5. Unusual dimerization of MpaH'. (A) Overall structure of MpaH' dimer from the top view. The core domains in the two molecules of the dimer are colored green and cyan, respectively, and both cap domains are colored red. One protomer is presented as cartoon, while the other as surface. The product MPA is presented as yellow stick. (B) A close-up of the interactions between the extended C-terminal part from one protomer and the substrate binding cavity of the other. The extended C-terminal part is highlighted as green cartoon. The substrate binding cavity is shown in gray surface. Hydrogen bonds between C-terminal residue K418 and the substrate binding pocket residues G174 and N184 are indicated as yellow dotted lines.

of H365 with a distance of 2.8 Å. It is worth noting that the imidazole ring of H365 is positioned toward the hydroxyl group of MPA to form a strong hydrogen bond, with a shift of 0.5 Å compared with the substrate-free MpaH' structure (Fig. 4A,B). Another two hydrogen bonds are formed between the carboxycarbonyl oxygen of MPA and the backbone nitrogen of V60 and F140; these two residues are defined as oxyanion hole residues of MpaH' (Figs 3 and 4A). The oxyanion hole is another notable feature of the α/β hydrolase machinery, functioning to polarize the carbonyl group in substrate to stabilize the catalytic intermediate (Fig. S5A). In most cases, it is formed by the main chain nitrogen atoms of two residues, one located at strand β_4 and the other adjacent to the nucleophile. The oxyanion hole residues vary a lot among the aligned enzymes (Fig. 3). However, their roles would not be affected since the hydrogen bonds are donated by the main chain nitrogen atoms.

In previous reports, mutation of the serine nucleophile to alanine completely abolished the catalytic activity [13–15]. However, the MpaH'^{S139A} mutant was found to retain a weak thioesterase activity in this study (Fig. S3). For canonical esterase catalysis, mechanistically, the catalytic histidine first deprotonates the nucleophile; then, the nucleophilic attack on the ester carbonyl carbon and restoration of the carbon–oxygen double bond yield the acyl-enzyme intermediate; finally, a water molecule binds to the intermediate and is further deprotonated by histidine to hydrolytically release the product and reset the resting enzyme state [8] (Fig. S5A). Interestingly, it was also suggested that

water can directly attack the substrate without forming an acyl-enzyme intermediate [16]. This alternative mechanism may reason the weak activity of MpaH'^{S139A} (Fig. S5B). Additionally, MPA-CoA can slowly spontaneously hydrolyze in aqueous solution (Fig. S6). Furthermore, the natural Ala/His/Asp triad without a nucleophile was previously identified in some α/β -hydrolase fold proteins, with histidine-mediated substrate deprotonation in the first crucial step [11,17]. Taken together, with the conformation stabilized by the pocket residues and carbonyl group polarized by the oxyanion hole, MPA-CoA could be alternatively hydrolyzed without a nucleophile, as H365 may directly deprotonate water to attack the thioester bond in the substrate (Fig. S5B). Supporting this, a water molecule was found in the active site of the MpaH'^{S139A}/MPA complex structure, which may serve as a nucleophile (Fig. S4).

Modeling of MpaH' in complex with MPA-CoA

Despite our greatest efforts, MpaH'^{S139A} and the substrate failed to co-crystallize. Thus, we used AUTODOCK VINA 1.1.2 to dock MPA-CoA into the structure of MpaH' in order to gain an insight into the binding mode of the CoA moiety, using the coordinate of product as reference. The modeling of the MpaH'/MPA-CoA complex revealed a quite open substrate binding pocket, and the CoA moiety is almost completely accessible to solvent (Fig. 4C). Residues L62, I206, A182, F327, and L366 contribute side chains to form an additional hydrophobic cavity that neighbors the MPA cavity to enclose the hydrophobic portion of CoA, mainly the

cysteine and pantothenate moieties, while the backbone oxygen atom from A325 and the backbone amide nitrogen atom from H365 form two weak hydrogen bonds to harbor the diphosphate group (Fig. 4D). The phosphoadenosine group located at the entrance of the active site is mainly positioned by polar residues with strong hydrogen bonds formed by residues Q208, G364, Q370, and E371 (Fig. 4D). Based on the complex structure, we reason that the substrate may easily slide down the tunnel to the active site with a large open cavity and the MPA moiety likely plays an essential role in accurate substrate binding and orientation by extensively interacting with the MPA binding pocket. The cavity specialized to accommodate MPA appears to be improper to accept other acyl-CoA esters to a right position for efficient hydrolysis. Supporting this, in our previous report, the k_{cat} values of MpaH' for MPA-CoA were two orders of magnitude higher than the corresponding values for ten other CoA esters [6]. Moreover, as we predicted, MpaH' significantly hydrolyzed mycophenolate mofetil (MMF, i.e., 2-morpholinoethyl ester of MPA) [18], a clinically used MPA derivative, into MPA and 2-morpholinoethanol by acting as an esterase (Fig. S7A). Comparatively, by monitoring the consumption of substrates, the apparent reaction velocity (V_{app}) at 200 μM of MPA-CoA and MMF was determined to be $92.60 \pm 4.26 \mu\text{M}/\text{min}$ and $10.16 \pm 1.07 \mu\text{M}/\text{min}$, respectively. The rate constants (k) of MPA-CoA and MMF were deduced to be $0.46 \pm 0.02 \text{ min}^{-1}$ and $0.05 \pm 0.01 \text{ min}^{-1}$, respectively. It is evident that MpaH' hydrolyzes its natural substrate MPA-CoA much faster than MMF (Fig. S7B).

To further evaluate the contributions of the MPA and CoA moieties of MPA-CoA to MpaH' substrate binding, we determined the dissociation constants (K_D) of MPA and CoASH using isothermal titration calorimetry (ITC). Of note, the binding affinity of MPA-CoA to MpaH' could not be determined by this method because the hydrolysis of MPA-CoA would occur during titrations. As shown in Fig. S8, MPA binds to MpaH' with a K_D of $103 \pm 10.5 \mu\text{M}$ ($\Delta H = -7.94 \pm 0.479 \text{ kcal}\cdot\text{mol}^{-1}$, $T\Delta S = -2.50 \text{ kcal}\cdot\text{mol}^{-1}$), while no appreciable binding was observed when 4 mM CoASH was titrated into a solution containing 20 μM MpaH'. These results well support the conclusion that the MPA portion of MPA-CoA plays a central role in productive substrate binding by acting as an anchor head.

Unusual dimerization of MpaH'

PISA analysis identified a large interface between the two protomers of MpaH', with the accessible surface area (ASA) buried in the interface amounting to

4003 \AA^2 , corresponding to approximately one-fifth of the total ASA (20 161 \AA^2), indicating a stable dimeric state of MpaH' (Fig. 2B). Some 765 \AA^2 of the ASA (3.8% of the total ASA) is buried between the two dimers attributing to the close crystal packing. The dimerization of MpaH' involves extensive hydrogen bonds, salt bridges, and hydrophobic interactions, mainly contributed by the residues from the long loop between $\beta 1$ and $\beta 2$, the loops prior to $\alpha 3$, $\alpha 5$, and $\alpha 9$, the loops after $\alpha 9$ and $\alpha 11$, and the extended helices $\alpha 13$ and $\alpha 14$.

The function of the extended C-terminal part, including part of helix $\alpha 13$, the last helix $\alpha 14$, and a long loop connecting $\alpha 13$ and $\alpha 14$, is unclear as it neither constitutes the core domain nor interacts with the substrate (Fig. 5A). Of note, this feature is uncommon among α/β hydrolases. Thus, we constructed a truncated version MpaH' $^{\Delta 401-433}$ with the protrudent C-terminal part removed. The gel filtration experiment still indicated a dimeric form of MpaH' $^{\Delta 401-433}$, suggesting a nonessential role of this extended part in dimerization (Fig. S9). However, this part indeed makes significant contacts with the cap domain, including extensive hydrophobic interactions and hydrogen bonds (Fig. 5). For example, residue K418 in the C-terminal loop forms two hydrogen bonds with G174 and N184 from the loop after strand $\beta 6$ and helix $\alpha 5$, which contributes to the substrate binding pocket (Fig. 5B). Therefore, we assume that the extended C-terminal part may indirectly participate in the substrate binding of the adjacent molecule by stabilizing the cap domain and coordinating the substrate binding pocket residues. Supporting this, the MPA-CoA hydrolytic activity of MpaH' $^{\Delta 401-433}$ decreased to 75.6% of the wild-type enzyme (Fig. S10).

The unique features of MpaH' and its homologs

A BLAST search revealed a large array of MpaH' homologs from different fungal species, with the protein sequence identity higher than 60%. These proteins were previously annotated as α/β hydrolase, peroxisomal membrane protein Lpx1, toxin biosynthesis protein-like protein, or proteins with undefined functions. Multiple protein sequence alignment (Fig. 3) indicates that these homologous proteins likely adopt structures similar to MpaH', especially a similar architecture of the core domain. The major differences exist in the cap domain, especially in the region of helix $\alpha 8$, which forms a bulge in MpaH' and protruded from the main body with unknown function. Variations of the cap domains are supposed to correlate with the substrate specificity [12,19]. The catalytic triad is strictly

conserved in the homologs with the same location of the acidic residue at the C terminus of strand $\beta 6$.

Submitting MpaH' structure to Dali server retrieved a number of structural homologs including peroxisomal membrane protein Lpx1, methylesterase, and lipase. The best homolog turned out to be Lpx1 from *Saccharomyces cerevisiae* (PDB ID code: 2Y6V) with a root-mean-square deviation (RMSD) of 3.3 Å for the aligned C α atoms [20]. Superimposition of the structures of MpaH' and Lpx1 revealed a similar overall fold of the core domain and some obvious differences in the cap domain. Besides, the last helix $\alpha 13$ in MpaH' core domain extends much longer than that in Lpx1 (Fig. S11A). Moreover, a deviation of the hydroxyl side chain of the nucleophile S145 in Lpx1 can be observed when compared with S139 in MpaH', which makes it out of the hydrogen bond range of the catalytic His323 (Fig. S11B).

The ESTHER database is a widely used database dedicated to proteins of the α/β hydrolase fold, which now contains >30 000 manually curated proteins [21]. A BLAST search within the ESTHER database using the MpaH' sequence classified it to block X, and this α/β hydrolase specifically falls into a poorly characterized family, the 6_AlphaBeta_hydrolase family, with most family members functionally unassigned.

A recent report divided α/β hydrolases into 12 different architectures from the perspective of structural modules [9]. The topological structure of MpaH' indicates that it should be classified as the single cap family with the cap domain inserted into the defined β -strands of β_{+1} and β_{+2} (the β -strand preceding nucleophile elbow can be defined as the central β -sheet, β_0 ; the β -strands in the direction of the N and C termini are assigned negative and positive numbers accordingly; Figs 2C and S1). However, the cap domain of MpaH' is extraordinarily large in size with a length of 134 amino acids, which is almost double the size of the statistical single cap (75 ± 18 amino acids). Particularly, MpaH' functions as an unusual dimer as described above, which is a unique feature of MpaH' and its homologous structure Lpx1. Another remarkable feature is the unusual location of the acidic residue of catalytic triad in MpaH' and the well-aligned homologous proteins (Figs 2C and 3). With these unique features, we annotated MpaH' and its homologs as a new subfamily of the single cap α/β hydrolase family.

Conclusion

In this study, we determined the crystal structures of the important hydrolase MpaH' in the

compartmentalized biosynthesis of MPA, including its substrate-free form and its mutant MpaH'^{S139A} in complex with the product MPA. MpaH' functions as an unusual dimer with the protruding C-terminal part indirectly participating in the substrate binding, which is supposed to be a novel folding of α/β hydrolases. On the basis of further structural and sequence analysis, we propose MpaH' and its homologs forming a new subfamily of α/β hydrolases. We also provide important mechanistic insights into the high substrate specificity of MpaH' by analysis of the structure of MpaH'^{S139A}/MPA complex and the modeled structure of MpaH'/MPA-CoA, together with mutagenesis analysis and ITC measurements. In summary, our results broaden the understanding of α/β hydrolase superfamily, and as the first structure reported in MPA gene cluster, this will provide theoretical and experimental basis for the future design and construction of artificial biosynthetic pathways using the key catalytic elements in MPA biosynthesis.

Materials and methods

Gene cloning

The DNA sequence-encoding MpaH' was amplified by standard PCR method using *P. brevicompactum* NRRL 864 cDNA as template. The PCR fragment was inserted into the expression vector pETM11 (EMBL) using *Nco*I and *Xho*I restriction sites, and the mutants were constructed using the QuikChange site-directed mutagenesis method (Stratagene). All the recombinant proteins were expressed to contain an N-terminal His₆ tag and a tobacco etch virus (TEV) protease cleavage site.

Protein expression and purification

The sequence-verified constructs were transformed into *Escherichia coli* BL21 (DE3) for recombinant overexpression. The bacteria were cultured in Luria-Bertani (LB) medium at 37 °C until the value of OD₆₀₀ reached about 0.6, and then, 0.2 mM IPTG was added to induce protein expression, with further culturing at 16 °C overnight. Cells were harvested and disrupted by sonication in lysis buffer containing 50 mM Tris/HCl, 300 mM NaCl, pH 8.0. The lysates were clarified by centrifugation (15 000 g, 60 min) at 4 °C, and the supernatant was applied to a nickel-chelating Sepharose affinity chromatography column (GE Healthcare, Marlborough, MA, USA). The N-terminal His₆ tag was removed by TEV protease, and further purification was performed by size-exclusion chromatography using a HiLoad 16/60 Superdex 200 Column (GE Healthcare) in a buffer containing 50 mM Tris/HCl, 300 mM NaCl, pH 7.5. The purified proteins were concentrated to about 5–10 mg/mL

as determined by absorbance at 280 nm and stored in aliquots at -80°C after flash freezing in liquid nitrogen. The Se-Met-derived MpaH' (Se-Met-MpaH') was produced using the metabolic inhibition method as described previously [22]. The purification method for Se-Met-derived MpaH' and the four mutants (MpaH'^{D163A}, MpaH'^{H365A} and MpaH'^{S139A}, and MpaH'^{Δ401-433}) was the same as that for the wild-type MpaH'. For oligomeric state analysis, the purified proteins were subjected to a Superdex 200 Increase (HiLoad 10/300) column. The results indicated that all the mutants were still in dimeric forms as the wild-type enzyme (Fig. S9).

Crystallization and structural determination

The crystallization experiments were performed with sitting drop vapor diffusion method at 20°C . The drops contained equal volume of protein sample and reservoir solution. After optimization, the best crystals of MpaH' and Se-Met-MpaH' were obtained in reservoir solution containing 8% PEG8000 (w/v) and 200 mM calcium acetate in 100 mM imidazole, pH 7.5. The best crystals of MpaH'^{S139A} were grown in reservoir solution containing 8% PEG8000 (w/v) and 200 mM calcium acetate in 100 mM imidazole, pH 7.0. Crystals of MpaH'^{S139A}/MPA complexes were obtained by soaking MpaH'^{S139A} crystals in solution containing 8% PEG8000 (w/v), 200 mM calcium acetate, and 2 mM MPA-CoA in 100 mM imidazole, pH 7.0, for about 15 min. The crystallization solution added with 15% glycerol served as the cryoprotectant for all the crystals. Suitable crystals were quickly soaked in the cryoprotectant solution and frozen directly in liquid nitrogen.

The diffraction data were collected on the BL19U1 beamlines at Shanghai Synchrotron Radiation Facility (SSRF) [23]. The data were processed by XDS [24]. The structure of MpaH' was solved by the SAD method using Se-Met-derived crystals. Phasing and initial model building were accomplished with Crank 2.0 in CCP4i2 software package [25,26]. The structure was refined with Refmac 5 [27,28]; alternatively, model building on Coot was performed [29]. Structures of MpaH'^{S139A}/MPA complexes were solved by molecular replacement with the wild-type MpaH' structure as the search model using experimental Phaser [30]. All the crystallographic data are summarized in Table S1.

Sequence alignment

Protein sequence alignments were performed using MUSCLE [31] integrated within MEGA 5.1 [32]. Superimposition of secondary structures of MpaH' (PDB ID code: 7DBI) and rendering was performed on ESPript server [33]. The protein interface of the potential oligomers was analyzed using PISA program [34]. The protein–ligand interactions were analyzed with LigPlus [35]. The structural

superimpositions were performed using Align. All of the structural images were generated using PyMOL (Schrödinger).

Hydrolytic assays of MpaH'

The hydrolytic activities of the wild-type MpaH', the point mutants MpaH'^{D163A}, MpaH'^{H365A}, and MpaH'^{S139A}, and the truncated version MpaH'^{Δ401-433} were determined by following the previously established protocol with some modifications. Briefly, the typical assay containing 20 nM enzyme and 0.2 mM substrate MPA-CoA in 80 μL reaction buffer (50 mM Tris, 300 mM NaCl, pH 7.5) was performed at 30°C for 20 min and quenched with 20 μL methanol. The organic extracts were used for HPLC analysis after centrifugation (15 000 *g*, 10 min) at 4°C .

The mutant MpaH'^{S139A} was previously reported to completely lose catalytic activity. However, in the crystallization experiment, the addition of MPA-CoA substrate into the MpaH'^{S139A} solution produced the hydrolytic product MPA. The activity assay of enzyme under crystallization condition was performed as follows: 75 μM MpaH'^{S139A} was mixed with 1 mM MPA-CoA in 20 μL reaction buffer (4% PEG8000 (w/v) and 100 mM calcium acetate, 50 mM imidazole, 25 mM Tris, 150 mM NaCl, pH 7.5) at 20°C for 30 min. The assay was also conducted under low concentrations in parallel: 15 nM enzyme and 1 mM MPA-CoA in 100 μL reaction buffer containing 50 mM Tris, 300 mM NaCl, pH 7.5, were incubated at 20°C for 30 min. The organic extracts were subject to HPLC analysis after centrifugation (15 000 *g*, 10 min) at 4°C .

For MPA-CoA stability test, 0.2 mM MPA-CoA in reaction buffer (50 mM Tris, 300 mM NaCl, pH 7.5) was placed at 25°C for 1 and 24 h, and then subjected to HPLC analysis. For the hydrolytic activity of MpaH' against MMF, the assay was performed under the conditions as follows: 4 μM MpaH' was mixed with 1.6 mM MMF in 50 μL reaction buffer at 25°C for 20 min, and the reaction was quenched with equal volume of methanol. The organic extract was subject to HPLC analysis after centrifugation (15 000 *g*, 10 min) at 4°C .

The substrate consumption assays were performed at 30°C , which containing 20 nM purified MpaH' and 0.2 mM substrate (MPA-CoA or MMF) in 80 μL reaction buffer (50 mM Tris, 300 mM NaCl, pH 7.5). The reactions were quenched at 1, 2, 3, 5, 10, 15, 20, and 25 min by the addition of 20 μL methanol. The organic extracts were used for HPLC analysis after high-speed centrifugation (15 000 *g*, 10 min) at 4°C .

The preparation of the substrate MPA-CoA is the same as we reported before [6]. Briefly, 160 mg MPA was added to 3 mL acetonitrile and followed by the addition of 120 mg carbonyldiimidazole with stirring. The mixture was stirred for 30 min at room temperature under N_2 and

followed by dropwise addition of CoA solution (320 mg in 1 mL H₂O) at pH 8.0 with saturated NaHCO₃ solution. The mixture was further stirred for 6 h at room temperature, and concentrated and purified by preparative HPLC.

Molecular docking of MPA-CoA to MpaH'

To investigate the MpaH'/MPA-CoA complex, molecular docking of the substrate to chain B was performed using AutoDock Vina 1.1.2 and MGLTools 1.5.6 [36,37]. The simulation box was fixed around the location of MPA in the MpaH'^{S139A}/MPA complex structure, with box size 30 Å in all three dimensions, and exhaustiveness and num_modes setting to 32 and 50, respectively. Collected conformations were clustered according to the RMSD against the conformation with the lowest binding energy [38]. The 50 generated conformations were further evaluated based on docking energy and the position of the MPA moiety of MPA-CoA. The selected model was used for structural analysis.

Isothermal titration calorimetry

ITC measurements were carried out using a MicroCal PEAQ-ITC (Malvern Panalytical) at 25 °C. For MPA titrations, 40 or 20 μM MpaH' was exchanged to an ITC buffer containing 50 mM HEPES, pH 7.5, 300 mM NaCl, 10% DMSO using PD-10 desalting columns (GE Healthcare). MPA was first dissolved in DMSO and then diluted into the ITC buffer to a final concentration of 4 mM. For CoASH titrations, 20 μM MpaH' was exchanged to ITC buffer without DMSO (50 mM HEPES, pH 7.5, 300 mM NaCl), and CoASH was dissolved in the same buffer. All the samples were centrifuged at 12 000 *g* for 10 min prior to experiments. For titrations, after an initial delay of 60 s, the first injection of 0.4 μL was made and followed by 12 further injections of 3.0 μL with a spacing time of 150 s to allow for equilibration. The cell was stirred at 750 rpm, and the reference power is 10 μcal·s⁻¹. For the blank titrations, substrates were injected into the corresponding ITC buffer. All experiments were performed for at least two times. Data were processed using MicroCal PEAQ-ITC analysis software and fitted into a one-site model.

Acknowledgements

This work was supported by the National Key Research and Development Program of China 2019YFA0905704 (to WZ), 2019YFA0706900 and 2019YFA0905100 (to SL), China Postdoctoral Science Foundation 2020M672049 (to CY), National Natural Science Foundation of China 32025001 and 31872729 (to SL), 32071266 (to LM), 31800664 (to FL), and 82022066 (to WZ), the Natural Science Foundation of Shandong

Province, China (ZR2019ZD20, to SL), the Laboratory for Marine Drugs and Bioproducts of Pilot National Laboratory for Marine Science and Technology (Qingdao) (LMDBKF-2019-01, to SL), and the State Key Laboratory of Bio-organic and Natural Products Chemistry (SKLBNPC18242, to WZ). We thank Prof. Bian Wu from the Institute of Microbiology, Chinese Academy of Sciences for the assistance in AutoDock analysis. We thank the beamline BL19U1 of the Shanghai Synchrotron Radiation Facility (SSRF, Shanghai, China) for X-ray diffraction data collection. We also thank Jingyao Qu, Jing Zhu, and Zhifeng Li from the State Key Laboratory of Microbial Technology of Shandong University for help and guidance in PEAQ-ITC.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

CY, WZ, and SL conceived this study. CY and FL performed molecular cloning, protein purification, crystallographic experiments, and structural analysis. XZ synthesized MPA-CoA. CY and LM performed mutagenesis analysis. CY, FL, WZ, YZ, and SL analyzed the data. CY and SL wrote the manuscript.

Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/febs.15874>.

Data availability statement

The atomic coordinates and structural factors have been deposited in the Protein Data Bank (<https://www.rcsb.org>) with the accession codes: 7DBI (substrate-free MpaH') and 7DBL (MpaH'^{S139A}/MPA complex).

References

- 1 Bentley R (2000) Mycophenolic Acid: a one hundred year odyssey from antibiotic to immunosuppressant. *Chem Rev* **100**, 3801–3826.
- 2 Geris R & Simpson TJ (2009) Meroterpenoids produced by fungi. *Nat Prod Rep* **26**, 1063–1094.
- 3 Zhang W, Cao S, Qiu L, Qi F, Li Z, Yang Y, Huang S, Bai F, Liu C, Wan X *et al.* (2015) Functional characterization of MpaG', the O-methyltransferase involved in the biosynthesis of mycophenolic acid. *ChemBioChem* **16**, 565–569.

- 4 Marzano AV, Dassoni F & Caputo R (2006) Treatment of refractory blistering autoimmune diseases with mycophenolic acid. *J Dermatol Treat* **17**, 370–376.
- 5 de Winter BCM & van Gelder T (2008) Therapeutic drug monitoring for mycophenolic acid in patients with autoimmune diseases. *Nephrol Dial Transplant* **23**, 3386–3388.
- 6 Zhang W, Du L, Qu Z, Zhang X, Li F, Li Z, Qi F, Wang X, Jiang Y, Men P *et al.* (2019) Compartmentalized biosynthesis of mycophenolic acid. *Proc Natl Acad Sci USA* **116**, 13305–13310.
- 7 Carr PD & Ollis DL (2009) Alpha/beta hydrolase fold: an update. *Protein Peptide Lett* **16**, 1137–1148.
- 8 Rauwerdink A & Kazlauskas RJ (2015) How the same core catalytic machinery catalyzes 17 different reactions: the serine-histidine-aspartate catalytic triad of α/β -hydrolase fold enzymes. *ACS Catal* **5**, 6153–6176.
- 9 Bauer TL, Buchholz PCF & Pleiss J (2020) The modular structure of α/β -hydrolases. *FEBS J* **287**, 1035–1053.
- 10 Heikinheimo P, Goldman A, Jeffries C & Ollis DL (1999) Of barn owls and bankers: a lush variety of α/β hydrolases. *Structure* **7**, R141–R146.
- 11 Wullich SC, Kobus S, Wienhold M, Hennecke U, Smits SHJ & Fetzner S (2019) Structural basis for recognition and ring-cleavage of the *Pseudomonas* quinolone signal (PQS) by AqdC, a mycobacterial dioxygenase of the α/β -hydrolase fold family. *J Struct Biol* **207**, 287–294.
- 12 Holmquist M (2000) Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr Protein Pept Sci* **1**, 209–235.
- 13 Xing Y, Li Z, Chen Y, Stock JB, Jeffrey PD & Shi Y (2008) Structural mechanism of demethylation and inactivation of protein phosphatase 2A. *Cell* **133**, 154–163.
- 14 McCulloch KM, Mukherjee T, Begley TP & Ealick SE (2010) Structure determination and characterization of the vitamin B6 degradative enzyme (E)-2-(acetamidomethylene)succinate hydrolase. *Biochemistry* **49**, 1226–1235.
- 15 Xu D, Gao Y, Sun B, Ran T, Zeng L, He J, He J & Wang W (2020) Structure and catalytic mechanism of a pyrethroid carboxylesterase PytH from *Sphingobium faniae* JZ-2. *Appl Environ Microbiol* **86**, e02971–e3019.
- 16 Schleberger C, Sachelaru P, Brandsch R & Schulz GE (2007) Structure and action of a C-C bond cleaving alpha/beta-hydrolase involved in nicotine degradation. *J Mol Biol* **367**, 409–418.
- 17 Wullich SC, Arranz San Martín A & Fetzner A (2020) An α/β -hydrolase fold subfamily comprising *pseudomonas* quinolone signal-cleaving dioxygenases. *Appl Environ Microbiol* **86**, e00279–20.
- 18 Regueira TB, Kildegaard KR, Hansen BG, Mortensen UH, Hertweck C & Nielsen J (2011) Molecular basis for mycophenolic acid biosynthesis in *Penicillium brevicompactum*. *Appl Environ Microbiol* **77**, 3035–3043.
- 19 Bains J, Kaufman L, Farnell B & Boulanger MJ (2011) A product analog bound form of 3-oxoadipate-enol-lactonase (PcaD) reveals a multifunctional role for the divergent cap domain. *J Mol Biol* **406**, 649–658.
- 20 Thoms S, Hofhuis J, Thöing C, Gärtner J & Niemann HH (2011) The unusual extended C-terminal helix of the peroxisomal α/β -hydrolase Lpx1 is involved in dimer contacts but dispensable for dimerization. *J Struct Biol* **175**, 362–371.
- 21 Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P & Chatonnet A (2013) ESTHER, the database of the α/β -hydrolase fold superfamily of proteins: tools to explore diversity of functions. *Nucleic Acids Res* **41**, 423–429.
- 22 Doublé S (2007) Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods Mol Biol* **363**, 91–108.
- 23 Zhang W-Z, Tang J-C, Wang S-S, Wang Z-J, Qin W-M & He J-H (2019) The protein complex crystallography beamline (BL19U1) at the Shanghai synchrotron radiation facility. *Nucl Sci Tech* **30**, 170
- 24 Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D* **66**, 133–144.
- 25 Pannu NS, Waterreus WJ, Skubák P, Sikharulidze I, Abrahams JP & de Graaff RAG (2011) Recent advances in the CRANK software suite for experimental phasing. *Acta Crystallogr D* **67**, 331–337.
- 26 Potterton L, Agirre J, Ballard C, Cowtan K, Dodson E, Evans PR, Jenkins HT, Keegan R, Krissinel E, Stevenson K *et al.* (2018) CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D* **74**, 68–84.
- 27 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53**, 240–255.
- 28 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* **67**, 355–367.
- 29 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* **66**, 486–501.
- 30 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
- 31 Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.
- 32 Tamura K, Peterson D, Peterson N, Stecher G, Nei M & Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.
- 33 Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, 320–324.

- 34 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
- 35 Laskowski RA & Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778–2786.
- 36 Trott O & Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455–461.
- 37 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS & Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* **30**, 2785–2791.
- 38 Shao J, Tanner SW, Thompson N & Cheatham TE (2007) Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *J Chem Theory Comput* **3**, 2312–2334.
- Supporting information**
- Additional supporting information may be found online in the Supporting Information section at the end of the article.
- Fig. S1.** Topology diagram of the typical single cap α/β -hydrolase family.
- Fig. S2.** Mutagenesis analysis of the catalytic triad of MpaH'.
- Fig. S3.** HPLC analysis (254 nm) of the activities of the wild-type MpaH' and the MpaH'^{S139A} mutant under different conditions.
- Fig. S4.** Structure of MPA bound to MpaH'S139A.
- Fig. S5.** Proposed catalytic mechanisms of MpaH' to hydrolyze MPA-CoA.
- Fig. S6.** HPLC analysis (254 nm) of the stability of MPA-CoA.
- Fig. S7.** The hydrolytic activity of MpaH' on the unnatural substrate MMF.
- Fig. S8.** ITC measurements of the binding affinities between MpaH' and different compounds.
- Fig. S9.** Superdex 200 Increase gel filtration chromatography of the wild-type and mutant MpaH' enzymes.
- Fig. S10.** The hydrolytic activities of the wild-type MpaH' and the truncated form MpaH' ^{Δ 401–433}.
- Fig. S11.** Comparison of MpaH' with the most similar structure Lpx1.
- Table S1.** Diffraction data collection and refinement statistics.