Structural basis for the inhibitor and substrate specificity of the unique Fph serine hydrolases of *Staphylococcus aureus*

Matthias Fellner[1]*, Christian S. Lentz[2,3], Sam A. Jamieson[1], Jodi L. Brewster[1], Linhai Chen[2,4], Matthew Bogyo[2], Peter D. Mace[1]

[1]Biochemistry Department, School of Biomedical Sciences, University of Otago, Dunedin 9054, New Zealand

[2]Pathology, Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305, United States

[3]Centre for New Antibacterial Strategies (CANS) and Research Group for Host-Microbe Interactions, Department of Medical Biology, UiT – The Arctic University of Norway, Tromsø N-9037, Norway

[4]National Center for Drug Screening, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

*To whom correspondence should be addressed

Abstract

Staphylococcus aureus is a prevalent bacterial pathogen in both community and hospital settings, and its treatment is made particularly difficult by resilience within biofilms. Within this niche, serine hydrolase enzymes play a key role in generating and maintaining the biofilm matrix. Activity-based profiling has previously identified a family of serine hydrolases, designated fluorophosphonate-binding hydrolases (Fphs), some of which contribute to the virulence of *S. aureus in vivo*. These ten Fph proteins have limited annotation, and have few, if any, characterized bacterial or mammalian homologs. This suggests unique hydrolase functions even within bacterial species. Here we report structures of one of the most abundant Fph family members, FphF. Our structures capture FphF alone, covalently bound to a substrate analog, and bound to small molecule inhibitors that occupy the hydrophobic substrate-binding pocket. In line with these findings, we show that FphF has promiscuous esterase activity towards hydrophobic lipid substrates. We present docking studies that characterize interactions of inhibitors and substrates within the active site environment, which can be extended to other Fph family members. Comparison of FphF to other esterases and the wider Fph protein family suggest that FphF forms a new esterase subfamily. Our data suggest that other Fph enzymes, including the virulence factor FphB, are likely to have more restricted substrate profiles than FphF. This work demonstrates a clear molecular rationale for the specificity of fluorophosphonate probes that target FphF and provides a structural template for the design of enhanced probes and inhibitors of the Fph family of serine hydrolases.

Keywords: *Staphylococcus aureus*, activity-based probe, serine hydrolases, acyl, inhibitor

Introduction

Staphylococcus aureus populates mucosal tissues or skin of about 30% of the world's population. It is a common cause of a variety of diseases ranging from local skin or soft tissue infections to invasive infections such as bacteremia, pneumonia or endocarditis [1-2]. Increased occurrence of community-acquired antibiotic-resistant *S. aureus*, such as Methicillin-resistant *S. aureus* (MRSA) [3], is a looming health threat requiring urgent development of new diagnostic and therapy options [4]. *S. aureus* often occurs in human tissue as biofilms, a biomolecular matrix that is largely impermeable to the immune system and many traditional antibiotics [5-6]. Bacteria within the biofilm have reduced metabolic activity, which exacerbates antibiotic tolerance and can promote chronic infection [7-8]. In order to

successfully target bacteria in this restricted-growth state, it is imperative to understand the underlying bacterial physiology and druggable enzymes that contribute to biofilm formation and maintenance.

Serine hydrolyses are one of the largest enzyme families in nature, are highly druggable, and execute a wide range of biological functions. Serine hydrolases degrade their substrates by the hydrolytic cleavage of ester, thioester or amide bonds, allowing them to act as proteases, peptidases, lipases, esterases and amidases. However, the role of serine hydrolases in bacterial homeostasis, survival at the host-pathogen interface or in biofilm-associated growth is not well explored. In order to identify new serine hydrolase targets to counter virulence and infectivity, we have recently performed a cell-based chemical proteomics study in *S. aureus*, employing activity-based probes (ABPs) [9]. ABPs are functionalized enzyme inhibitors that can be used for selective labelling of active enzymes to characterize their physiological roles *in vitro* and *in vivo*. Target-specific probes can be used to create selective imaging agents, provide leads for small molecule inhibitors and enable direct analysis of drug efficacy or specificity.

Using fluorophosphonate-probes we identified 12 serine hydrolyses targets that are enzymatically active under biofilm forming conditions. These enzymes include lipase1 and 2 (SAL1, SAL2) as well as 10 uncharacterized hydrolases that we termed fluorophosphonate-binding hydrolases (Fph) A–J [9]. FphA–FphJ have limited annotation, and lack characterized bacterial or mammalian homologs. A detailed investigation of one particular hydrolase, FphB, has since revealed that it facilitates infection of specific tissues in a mouse model of systemic *S. aureus* infection [9]. As part of the initial studies, we identified selective inhibitors and fluorescent activity-based probes for several Fph proteins [9-10]. However, further optimization of probes for specific Fph proteins to determine their biological roles is currently limited by a general lack of structural knowledge for this potentially unique family of hydrolases.

In this study, we present the structure-function characterization of an abundantly expressed Fph member during biofilm-associated growth conditions—FphF [9]. Around two thirds of *Staphylococcus* species contain an FphF homolog (at least 50% sequence identity), with some species containing more than one [11]. FphF is a 29 kDa hydrolase encoded by a gene previously annotated as *estA*. In the UniProt database, this protein (Q2FUY3) is putatively annotated as a tributyrin esterase and has a GO-annotation as an S-formylglutathione hydrolase based on similarity with the human orthologue. Our biochemical characterization and a structure in complex with a substrate suggest that FphF is likely not a specific tributyrin esterase but rather is a promiscuous esterase enzyme. Crystal structures of FphF in complex with triazole urea-based inhibitors KT129 and KT130 show inhibitors act by covalent modification of the active site serine, with the inhibitor acyl-group and aliphatic tail of substrate analogs competing for the same binding pocket. These findings illustrate the structural basis for the substrate and inhibitor selectivity of FphF, which is reinforced by docking and modelling studies comparing FphF to other esterases and the remaining Fph proteins. This study enables future precise inhibitor and probe development for the new Fph serine hydrolase family.

Results

Overall structure of FphF

To begin our efforts to understand FphF substrate and inhibitor specificity, we recombinantly expressed the full-length enzyme in *E. coli*, and established crystallization conditions. FphF crystallized in several different conditions (Table S1) in three distinct crystal forms. The overall oligomeric structure was nearly identical across the three crystal forms showing a tetramer (Figure 1A and 1B) either within the asymmetric unit or across symmetry mates. In solution, FphF appears to be a dimer of ~58 kDa based on gel-filtration experiments (Figure S1). There are two significant dimer interfaces in both tetramers. One predominantly formed by an antiparallel association of the first β-

strand β1, while the second one is formed around interactions of the second helix α2. FphF oligomer analysis in PISA [12] indicates a similar area for both interfaces but the β-strand dimer interface has a higher complexation significance score, suggesting it might be the relevant dimer contact in solution. There was no significant difference between the FphF monomers (Cα RMSD of 0.25-0.35 Å between the four chains of the apo structure PDB ID 6VH9 at 1.71 Å resolution). The biological role of the oligomer remains uncertain, potentially contributing to overall stability.

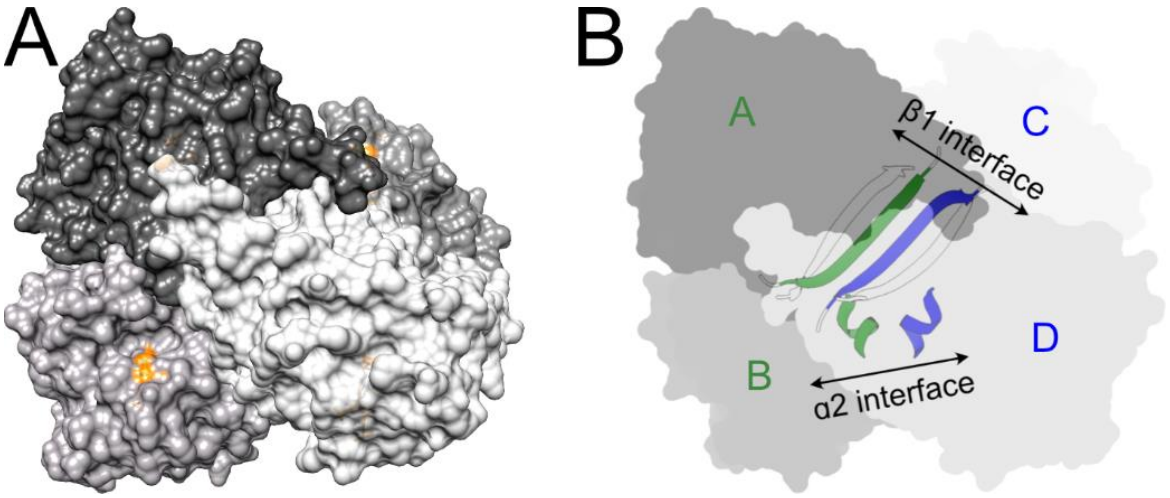


Figure 1: Overall structure of FphF. A) FphF tetramer (PDB ID 6VH9). The surfaces of the four chains are colored in different shades of grey with active site triads in orange. B) The two different dimer interfaces within the tetramer, β1-β1 and α2-α2, are indicated by blue and green coloring.

Active site of FphF

FphF is a member of the α/β hydrolase superfamily [13], characterized by a core of 8 β strands connected by several α helices (Figure 2A and S2). The FphF serine hydrolase catalytic triad consists of Ser121, His234 and Asp205 (Figure 2A and 2B). Ser121 is located on the connecting loop between β5 and α6 within the conserved GXSXG motif [14], present in all Fph proteins. His234 sits on a loop between β8 and the conserved C-terminal helix α11 and Asp205 on a loop between β7 and α10. Across all apo-FphF crystal forms and all crystallization conditions, Ser121 is found proximal to additional electron density consistent with three additional atoms. We modelled this density as a sodium ion, which interacts with the oxyanion hole residues [15] Leu48 and Met122 chelated by two molecules of water (for details on this density assignment see methods section, Figure S3).

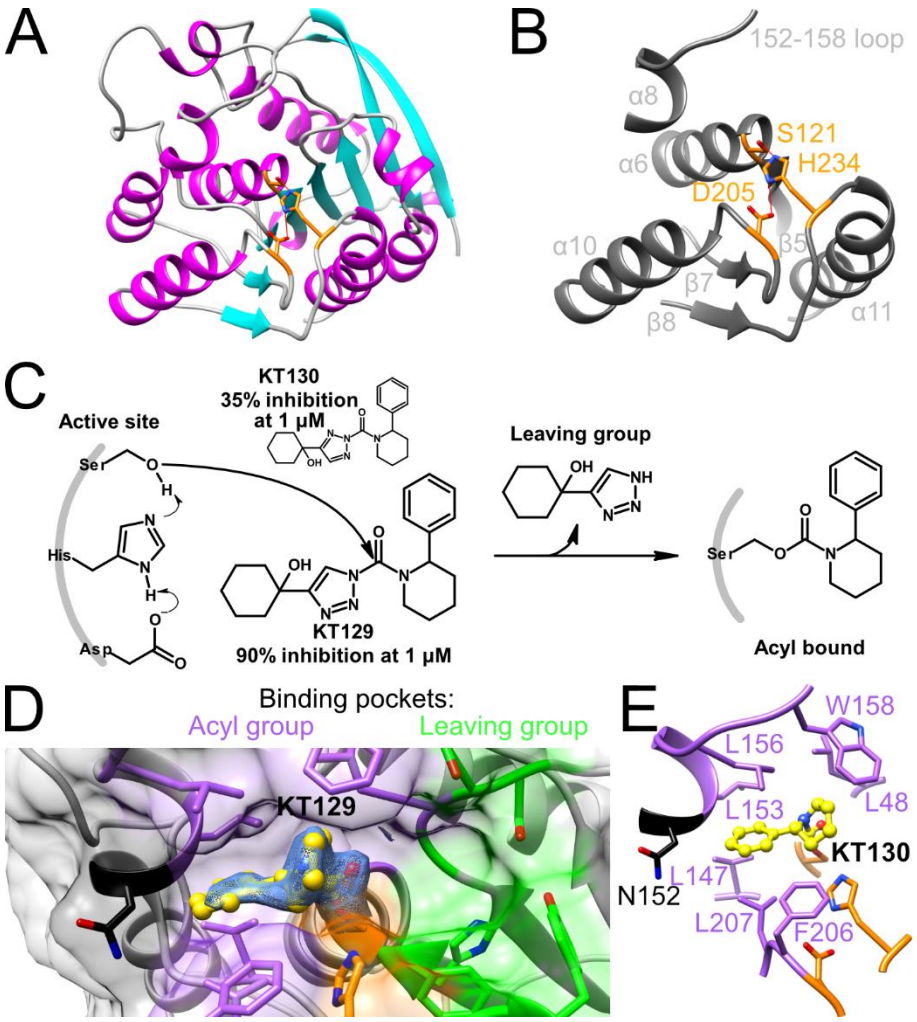


Figure 2: Covalent inhibitor bound to FphF crystal structures. A) Ribbon representation of the FphF monomer showing the location of the active site triad. β-strands are colored in cyan, α-helices in magenta and the active site Ser-His-Asp triad in orange. B) Closeup of the active site triad. C) The mode of inhibition of KT129 and KT130, with the % of inhibition at 1 μM concentration [10]. D) FphF KT129 inhibitor bound crystal structure (PDB ID 6VHD) is shown with the inhibitor carbon atoms in yellow, the active site triad in orange, the acyl binding pocket in purple, the Asn125 gate in black and the leaving group binding pocket in green. The *2FO−FC* map for KT129 is shown as blue mesh at 1 σ. E) KT130 inhibitor bound crystal structure (6VHE).

Covalent inhibitor bound structure of FphF

We previously identified several 1,2,3-triazole urea-based inhibitors in a dose-response screening assay of S. aureus cell lysate from biofilm-promoting growth conditions which preferentially targeted FphF over other Fph proteins [10]. Two isomers, KT129 and KT130, which differed in the constitution of the triazole urea linkage had distinguishable altered potencies, with KT129 inhibiting FphF at 1 μM by 90% and KT130 at 1 μM by 35% (Figure 2C). To characterize their mode of action we obtained structures of them in complex with FphF and later used docking to explain their potency differences. We determined co-crystal structures of FphF with KT129 (PDB ID 6VHD at 1.98 Å resolution) and KT130 (6VHE at 1.94 Å). In both structures, all protomers of the tetramer contained electron density (Figure 2D) consistent with the inhibitor within a hydrophobic pocket alongside the active site. As expected based on the enzymatic mechanism (Figure 2C), the triazole urea linkage that distinguishes KT129 and KT130 was decomposed, leaving the identical 2-phenylpiperidine-1-carbonyl moiety

covalently attached to the terminal side chain oxygen atom of Ser121 revealing the acyl binding pocket of FphF (Figure 2D and E).

Several hydrophobic residues surround the phenylpiperidine: Leu48, Val147, Leu153, Leu156, Trp158, Phe206 and Leu207. Asn152 gates the entrance to the pocket (Figure 2E), however, its sidechain is turned towards the surface with its uncharged backbone interacting with the phenylpiperidine. Residues that form the binding pocket are generally similar to their position in the apo-protein, with only minor displacement of the sidechains of Trp158 and Phe206 suggesting a predominantly direct, non-induced fit of the inhibitor (Figure S4). In both structures, electron density evidence of a possible minor second inhibitor conformation is observed (for details see supporting information, Figure S5). This minor conformation rearranges the 152–158 active site loop and does not match the pocket of the apo-protein structure, so could represent active-site plasticity that is potentially relevant to inhibitor design.

One outstanding question that can be inferred from the structures is the binding-site of the triazole leaving group of KT129 and KT130. The leaving group is not present in the crystal structures but appears well suited to occupy a polar pocket adjacent to the active site Ser121, opposite the covalently attached acyl-group (Figure 2D; green residues). In all structures, water molecules form a hydrogen bonding network within this pocket, which includes Ser49, Ser50 and His120. Nearby sidechains of Tyr236 and Trp239, adjacent to the active site H234 on the C-terminal helix α11, could also play important roles. While our structures do not yet enable structure-based optimization of leaving group character, the observation of this pocket can provide qualitative information about how leaving group chemistry contributes to the specificity of Fph-targeting probes.

FphF substrate profile and substrate bound structure

To establish how the active site architecture from the inhibitor bound structures relates to enzymatic function we next established the substrate specificity profile of FphF. In order to first establish if recombinant FphF is enzymatically active, we used a fluorescent fluorophosphonate-probe (FP-TMR) that covalently binds to the active site serine of α/β hydrolases. We previously used this probe to label FphF along with the other newly identified Fph enzymes in live *S. aureus* [9]. The probe labelled a single protein of the expected molecular weight of the recombinant protein including the N-terminal His_6-tag (Figure 3A). In addition, we previously identified the sulfonyl fluoride JCP678 as a selective inhibitor of FphF in *S. aureus*. Pre-treatment of the purified protein with this covalent inhibitor, blocked labelling by the FP-TMR probe (Figure 3A). These results suggest that recombinant FphF protein is enzymatically active and has labeling/inhibition profiles that resemble those observed previously for native protein in intact cells [9]. Next, we tested the substrate preference of FphF using a panel of commercially available fluorogenic substrates (Figure 3B). We found that the protein cleaved lipid ester substrates, but was unable to process phosphate, phosphonate or glycosidic substrates. FphF showed a promiscuous specificity profile, cleaving hydrophobic saturated lipid substrates with acyl chain lengths ranging from C2 to C10 with the highest activity for C7 (Figure 3B).

To better understand the reason for the broad substrate selectivity of FphF, we determined the crystal structure of FphF in complex with the preferred C7 model substrate 4-Methylumbelliferyl (4-MU) heptanoate (Figure 3C). In our structure (PDB ID 6WCX at 2.89 Å resolution), clear electron density is visible corresponding to the heptyl acyl moiety of the substrate covalently linked to Ser121 in all FphF chains. The side chain orientations in the acyl binding pocket also matched the inhibitor bound state, with the exception of Phe206, which can adopt several orientations between different protomers of the tetramer (Figure 3C). This residue may be flexible and play a role in opening up the pocket to facilitate a binding event. In order to gain a broader insight into the structure-function relationship of

FphF and the other Fph serine hydrolases, we performed docking studies based on our structural data defining the active site, the position of the substrate acyl and the predicted leaving group location.

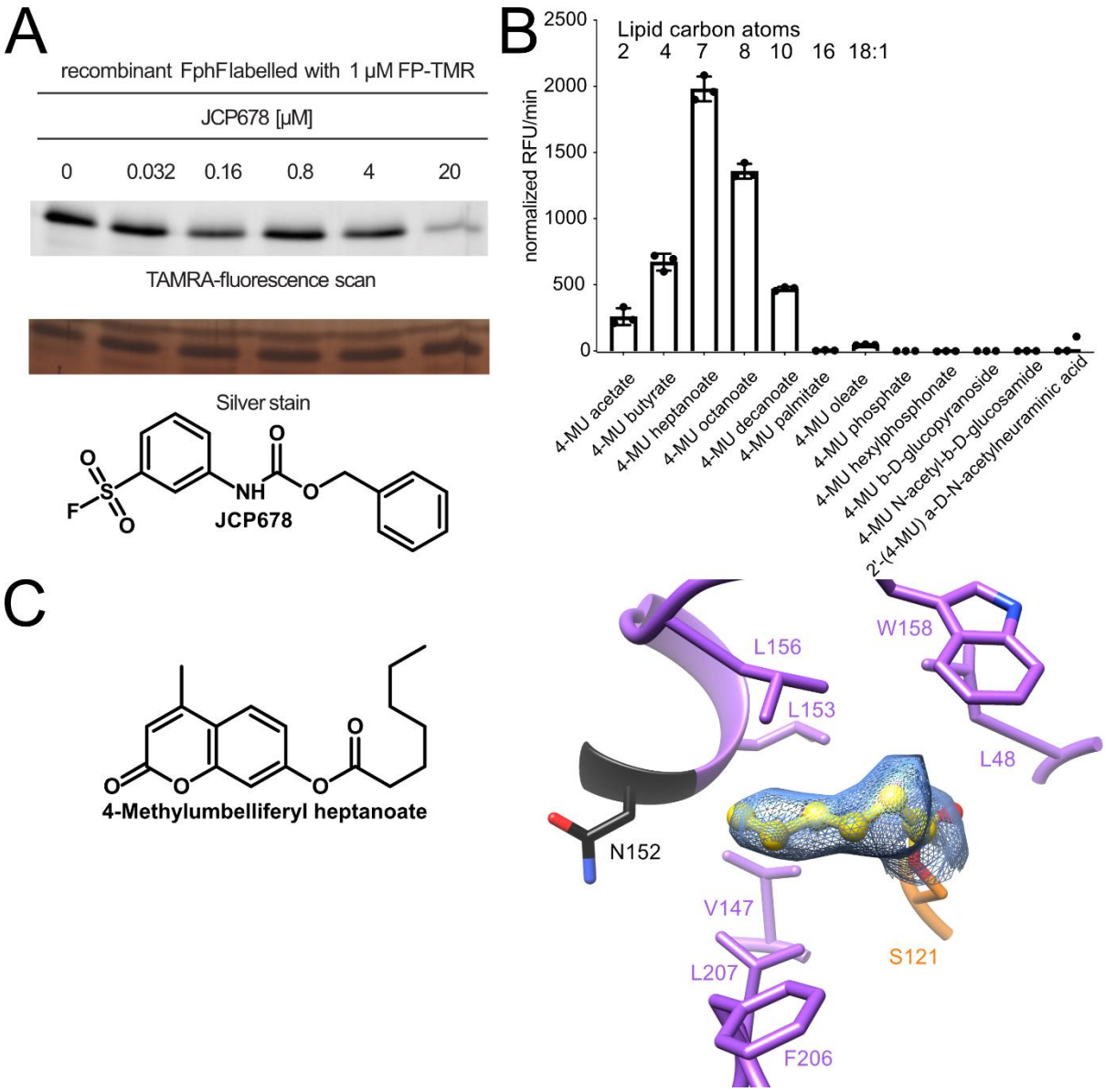


Figure 3: **Inhibition, activity and substrate bound structure of recombinant FphF.** A) Purified recombinant FphF was pretreated with different concentrations of the selective inhibitor JCP678 before labelling with the fluorescent ABP FP-TMR (for the complete gel including molecular weight markers see Figure S6). B) Assessment of the substrate specificity profile using a library of 4-MU based fluorogenic substrates. The graph shows the turnover rates for each substrate as relative fluorescence units (RFU)/min and depicts the mean +/- standard deviation of 3 independent reactions. In addition, the experiment was repeated twice with similar results. C) Covalent substrate (4-MU heptanoate) bound FphF crystal structure (PDB ID 6WCX). The *2FO−FC* map for the heptyl acyl intermediate is shown as blue mesh at 1 σ.

Probing the FphF hydrophobic and leaving group binding pockets with docking studies

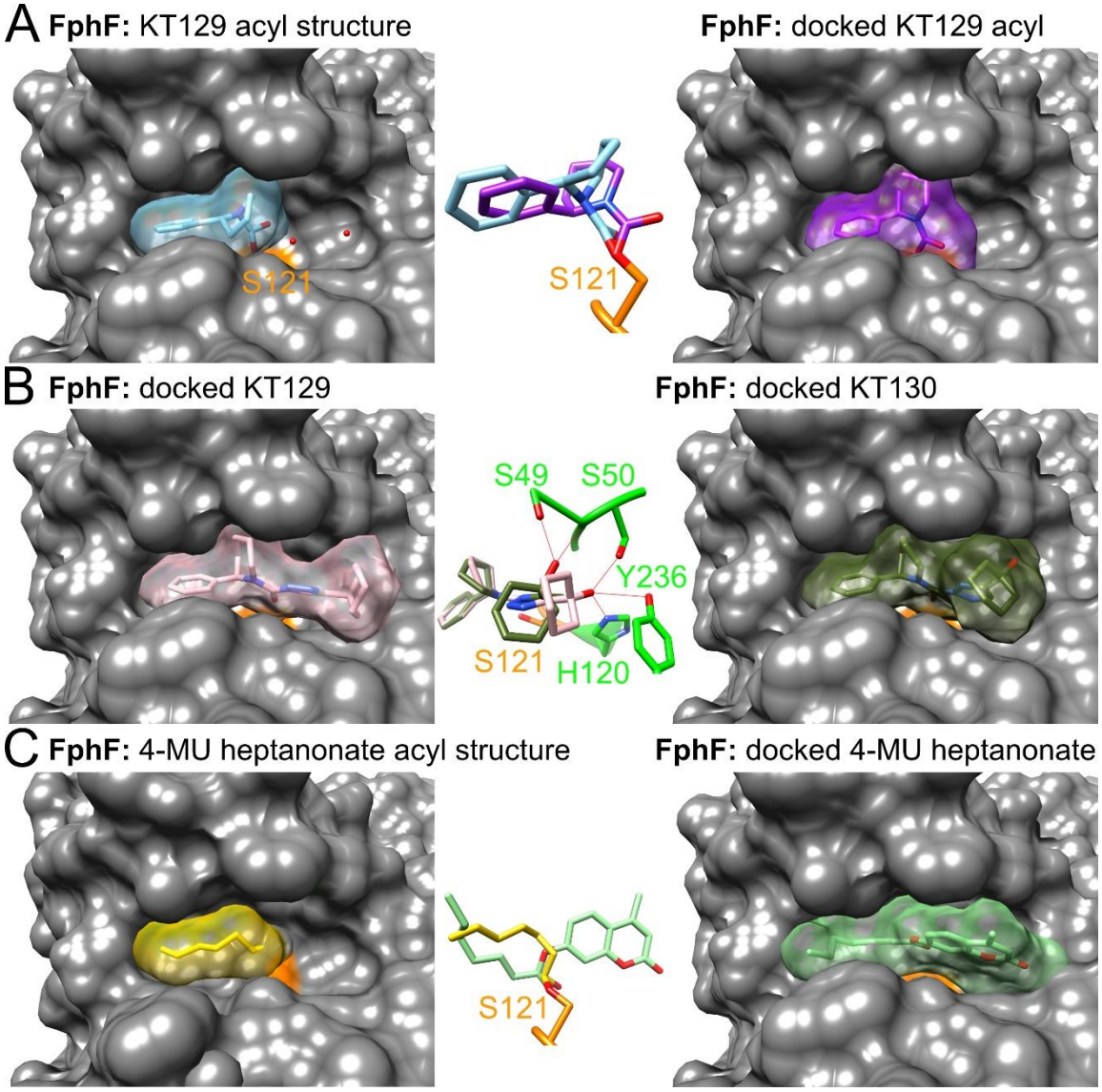


Figure 4: FphF docking studies. A) FphF-KT129 structure (PDB ID 6VHD) (left) and docking comparison (right). The protein surface is shown in grey, with the active site Ser in orange and water molecules in the leaving group pocket illustrated as red spheres. The ligand and ligand surface are shown in light blue in the crystal structure or purple in the docked model. An overlay of the two is also shown (center). B) Comparison of full-length KT129 (pink) and KT130 (dark green) docked into 6VHD-S121G. The middle insert shows an overlay of the two with differing hydrogen bonds to the ligand hydroxyl group in green. C) Heptyl acyl structure and docking comparison. Heptyl acyl ligand (yellow) structure (6WCX) (left) and full length 4-MU heptanoate (green) docked into 6VHD-S121G (right). The middle insert shows overlay of the positions of the two.

We first tested computational docking using the GOLD suite [16] for the acyl group of KT129 using the protein structure after removing the ligand. Specifying only the covalent linkage between the acyl group and Ser121, the docking gave a similar structural model for the ligand position as was observed in the crystal structure (Figure 4A). Docking of the full length KT129 inhibitor into a S121G model of FphF (Figure 4B) showed a similar fit for the acyl group, with the remaining part of the ligand occupying the predicted leaving group binding pocket. The triazole ring is orientated such that the nitrogen atoms are surface exposed, with the carbon atoms on the opposite side of the triazole ring facing the protein. and the lone hydroxyl group is stabilized by hydrogen bonds with Ser50, His120

and Tyr236. Docking of KT130 resulted in unique orientations for the hydroxyl group induced by the triazole ring (Figure 4B). As KT130 contains nitrogen atoms on both sides of the triazole ring, one nitrogen atom must face the protein, resulting in a tilt of the remaining ligand in this binding pocket. This results in the hydroxyl group forming hydrogen bonds with Ser49 and its adjacent backbone. 1,4-isomers of inhibitors such as KT129 were more potent against FphF in general when compared to the 2,4-isomers found in KT130 [10], suggesting that a nitrogen facing the protein is unfavorable for binding in the leaving group pocket.

Comparison of the FphF-substrate structure with full length 4-MU heptanoate docked into a S121G model of FphF (Figure 4C) both show the heptanoate group occupying the acyl pocket. The pocket is fully occupied by the C7 chain with several side chains clearly defining the space. The leaving group binding pocket is again identified for the 4-MU moiety. In line with our substrate-specificity profile there is some room for longer chains and indeed C8 chain docking fits similarly well, though the terminal atom is more surface exposed (Figure S7). Docking an even longer C10 substrate into FphF results in a poorer fit of either end of the ligand as there is not enough room to accompany them. In docking of shorter chain lengths (C2 and C4) the ligand is never correctly positioned over the active site Ser121 residue. Together, these data support a promiscuous substrate profile of FphF for hydrophobic lipid chains, with a preference for a C7 chain length.

Relating acyl-binding pocket character to substrate preference

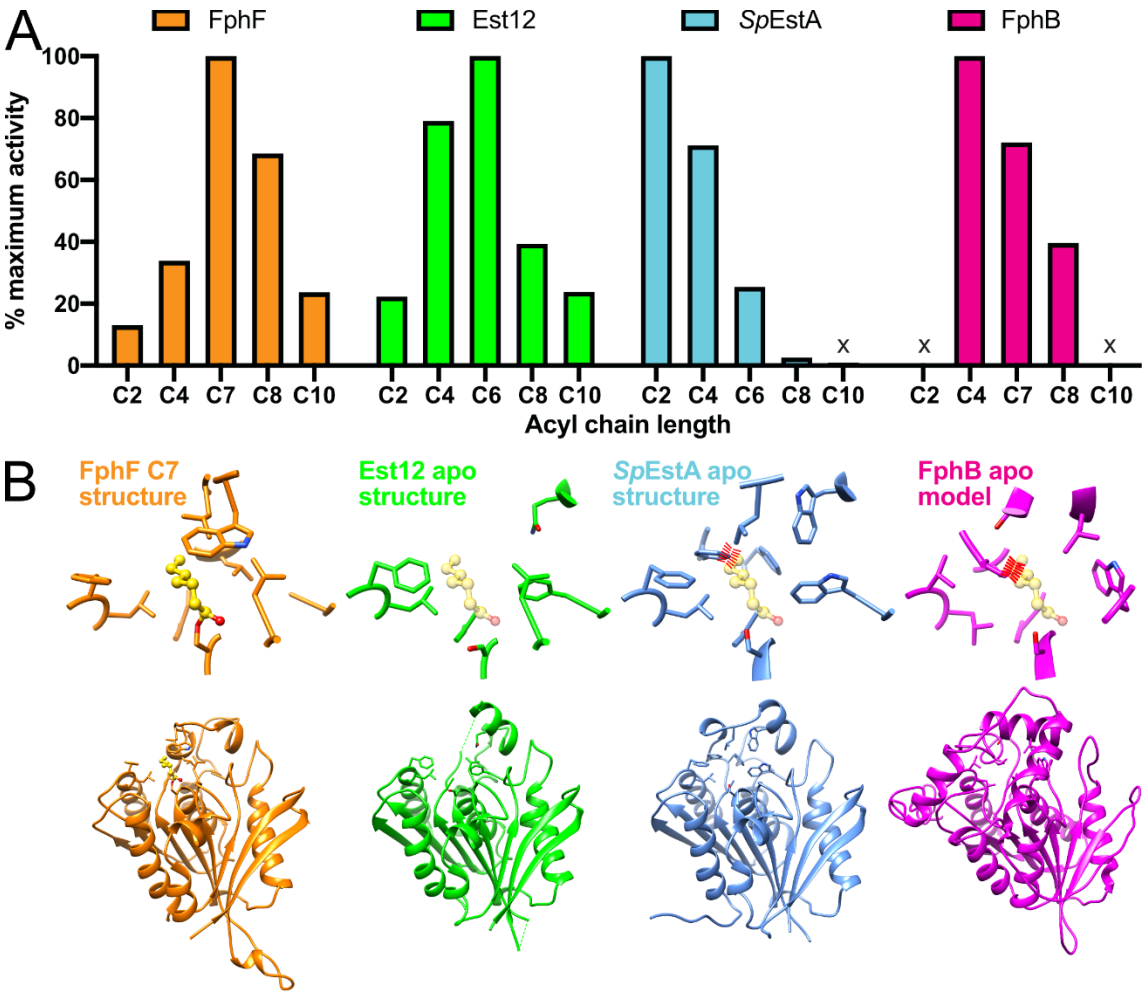


Figure 5: Functional and structural comparison of FphF to characterized esterases. A) The relative enzymatic activity of FphF, FphB, *Sp*EstA and Est12. The activity data of FphF and those

reported previously for Est12 [17], *Sp*EstA [18] and FphB [9] were normalized by setting activity data measured for the preferred substrate to 100% and the data of the other substrates as the percentage of maximum activity. X denotes inactivity of the enzyme against the indicated substrate. B) Active site and overall structure ribbon comparison (based on Cα alignment) of FphF heptyl acyl (C7) structure (PDB ID 6WCX), Est12 apo structure (4RGY [17]), *Sp*EstA apo structure (2UZ0[19]) and FphB predicted model. The close up show the heptyl acyl in yellow within FphF and the active sites of related esterases. Red marks indicate clashes with the heptyl acyl in *Sp*EstA and FphB.

To understand if the structure-function profile of FphF matches other characterized esterases we performed structural and literature searches. Comparing FphF to known structures, Dali [20] searches (Table S2) revealed one significant identity hit (39%) in Est12 [17]. Est12 was discovered from a marine sediment metagenomic fosmid library with an unknown species origin. It was proposed that Est12 represents a new family or subfamily of bacterial lipolytic enzymes, which is distinct from the EstA family. Interestingly, Est12 also exhibits an ability to cleave a relatively broad range of substrates of different acyl chain length, in line with structural similarity to FphF (Figure 5). The closest related structures after Est12 (Table S2) were all esterases with a preference for C2 substrates belonging to the EstA/S-formylglutathione hydrolase family (sequence alignment Figure S8), with the top hit being tributyrin esterase (*Sp*EstA) from *Streptococcus pneumoniae* (PDB ID 2UZ0 [19]). EstA and S-formylglutathione hydrolases [21-29], including the *Homo sapiens* esterase D [23], show a preference for a C2 acetate chain length and all share a trend of decreasing activity with longer chains (Figure 5A). Comparison of the FphF heptyl acyl structure with *Sp*EstA (Figure 5B) indicates that a phenylalanine which is absent in FphF and Est12 will directly clash with long carbon chain substrates. As this phenylalanine is conserved across EstA/S-formylglutathione hydrolases with C2 substrate preference, it is a feature clearly distinguishing this family from FphF.

Next, we compared FphF with FphB, the only other Fph protein for which we have previously established a substrate profile [9] (Figure 5A). FphB and FphF share 21% amino acid sequence identity (sequence alignment Figure S9) and are predicted to have the same overall protein fold. The demonstrated carboxylic acid esterase activity of FphF is similar to that of FphB [9]. However, FphB has a narrower substrate specificity such that it is unable to cleave C2 and C10 lipids and prefers C4 substrates [9]. We used FphF as a template to model FphB using I-Tasser [30], which combines the FphF input structure with up to nine additional structures from the protein data bank [31] (Figure 5B and Table S3). Fph proteins all belong to the α/β hydrolase superfamily [13] and all of the predicted Fph protein models show this fold containing the core β strands and several of the conserved helices, including FphB. Superposition of the FphF structure with the FphB model demonstrated similar positions of the catalytic triad residues, giving confidence to the prediction. While the FphB acyl binding pocket is also hydrophobic, an overlay of our FphF heptyl acyl structure suggests that the ligand would clash with an isoleucine in FphB. Neither the C7 acyl group, nor the acyl group of KT129/KT130 would fit well into the predicted FphB model, confirming experimental observations that KT129/KT130 and inhibitors with a similar sized acyl group are very poor inhibitors of FphB [10]. However, FphB is still able to process C7 and C8 substrates, suggesting that the predicted model might show a closed state of the active site. But overall these findings further support an active site in FphB that has a narrower specificity compared to the promiscuous FphF.

Among the discussed esterases, our substrate and inhibitor bound structures are the only ligand bound structures available and enable the first comparisons of an acyl bound moiety. These comparisons show a clear relationship between substrate specificity and the architecture of a hydrophobic acyl binding pocket. Hydrolases such as FphF and Est12 exhibit a broad substrate specificity based on their large acyl binding pocket, whereas those with a tighter terminal pocket (such as FphB and *Sp*EstA) display more restricted substrate profiles.

Comparison of FphF and FphB with other Fph proteins

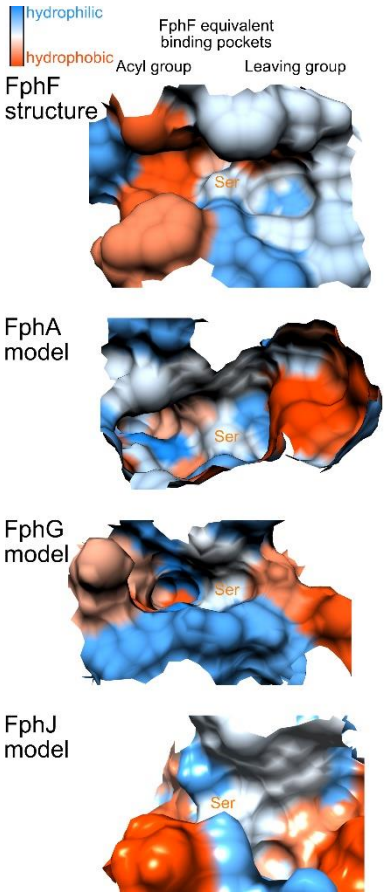


Figure 6: Fph protein active sites. FphF active site compared to examples of predicted active sites of other Fph proteins.

Similar to FphB we used the FphF crystal structure, in combination with available data bank entries, to predict the structure of all Fph (A-J) proteins (Figure S10 and Table S3). While our models predict that the general protein fold is conserved across all the Fph proteins, alignment in Clustal Omega [32] indicated that Fph proteins have an overall low sequence relationship, with most having under 20% identity. Overall, the C-terminus is structurally conserved and always contains a terminal α-helix following the active site histidine loop. The N-terminus shows some variability. Compared to FphF, some of the Fph enzymes may have additional (FphA) or fewer (FphC) core β strands, which could influence their oligomerization. FphB and FphD stand out with predicted additional N-terminal extensions that contain two to three α-helices specific to these proteins. In the case of FphB these helices are predicted to close the active site, acting as a lid, a characteristic of certain lipases enabling interfacial activation [33]. A similar helical lid is predicted to close the active sites of FphC and FphE, here the lid stems from an extension of a loop equal to the FphF 152–158 residues area. Apart from these distinctions, other differences stem from variations in length and fold of loops and α-helices between the core β strands, some having a direct effect on the active site environments and all dictating the range of overall protein size between the Fph proteins (FphA 52kD – FphJ 22kD).

When comparing the predicted active sites of all the Fph proteins in relationship to FphF, it appears that in contrast to FphB and FphF, most Fph proteins may not prefer hydrophobic acyl groups. The acyl binding pocket of FphA, FphC, FphD, FphE, FphG and FphI have at least one and often multiple hydrophilic side chains pointing towards the acyl binding pocket. The remaining FphH and FphJ (Figure 6) do not have a defined acyl binding pocket but a surface exposed area. FphA (Figure 6) introduces a range of hydrophobic and hydrophilic residues in a narrow but surface exposed pocket. The fold and hydrophobicity of this pocket suggest that FphA may have a highly selective substrate binding mechanism. This could possibly explain why our triazole inhibitor screening only identified modest inhibitors for FphA [10]. FphG (Figure 6) has the greatest diversity of residues predicted to be at the acyl pocket and also forms a narrow pocket which appears to be designed for a specific substrate. Overall, FphB and FphF appear to have the most similar hydrophobic acyl binding pockets with all the other Fph proteins showing greater variety. The leaving group binding pocket is predicted to be just as diverse. The overarching conclusion from these modelling studies is that there is little evidence for redundancy with respect to the Fph proteins. The low sequence identity is reflected in their predicted active sites, suggesting a wide range of substrate specificities, and thus potentially diverse biological functions.

Discussion

Given that serine hydrolases can be effectively targeted by chemical probes, they are promising targets for diagnosis, therapeutics and monitoring of treatment of *S. aureus* infections. However, it remains difficult to specifically target a particular bacterial enzyme over related bacterial family members or host enzymes. A clear understanding of the structure-function relationships is necessary to overcome this problem. The structural characterization of *S. aureus* FphF presented here uncovers the mechanisms of both inhibitor and substrate binding. Structural analysis suggests that the related virulence factor FphB (from *S. aureus*), and the virulence factor *Sp*EstA (from *S. pneumoniae*) [19, 34], both appear to be related and act on hydrophobic lipid substrates. However, several key differences suggest that FphF is a promiscuous esterase, while *Sp*EstA is mainly a deacetylase, and FphB may be specific for a yet to be discovered substrate.

Although our findings do not rule out that FphF can deacetylate native substrates, our data suggest that alternative, preferred physiological substrates of FphF may exist. Because the function of FphF is not likely to be similar to *Sp*EstA, we propose to change the gene annotation of FphF to *fphF* from *estA*, including highly identical homologs in *Staphylococci*. Est12 is the only characterized putative homolog of FphF, however as its organism origin is unknown, it does not give any further insights into a potential biological function. FphF belongs to the recently defined Est12 esterase subfamily [17], which is distinct from the eight original families I-VIII of bacterial lipolytic enzymes [35], and unlike the EstA subfamily.

While we showed that FphF can process a variety of lipid substrates, our structural analysis of the large, surface exposed acyl pocket suggests that FphF may be able to process a wide range of hydrophobic acyl esters, potentially including branched lipids and chains containing carbon rings. Our analysis shows that substrate specificity is determined by residues on both sides of the scissile bond, in the hydrophobic and leaving group binding pockets. The FphF structures and docking studies presented here indicate how a covalent ligand targeting the active site serine interacts with the acyl and leaving group binding pocket. The specificity introduced by residues within these pockets dictates binding and the structural observations fit with features of previously identified inhibitors [9-10]. The triazole based KT129 could be converted into an FphF specific probe by expanding on the observed binding mode. Such a probe could be used as a chemical tool for the single-cell phenotypic characterization of enzyme activity levels in *S. aureus*. Both rings of the phenylpiperidine can be modified by attaching a label on the surface exposed side. Further optimization may also consider

reducing off-target toxicity by increasing specificity over human serine hydrolases, as several have been successfully targeted with triazole inhibitors [36-37]. A starting point could be a parallel investigation with the closest human homolog esterase D [23] which could be expanded to human proteome selectivity analysis, able to screen multiple potential off-target serine hydrolases [38].

In addition, esterases are also industrially relevant [35, 39] and as the new Est12 esterase subfamily offers new esterase candidates our characterization of FphF gives the opportunity to explore FphF for industrial purposes including directed evolution approaches [40].

Conclusion

FphF is active during biofilm-associated growth conditions [9] and activities of esterases like EstA and S-formylglutathione hydrolase have been directly connected to biofilms [27, 41] Although further studies are necessary to elucidate the precise biological role of FphF, using the knowledge from our FphF structures, combined with available structure prediction and docking tools, we can begin to define similarities and differences between FphF and the other Fph proteins. This analysis suggests that the Fph enzymes have specific roles in *S. aureus* facilitated by their distinct active site environments. Expanding the structure-function relationship, optimally using additional native substrates, will be crucial to understand the specific roles of each Fph protein. The results shown here enable the specific probe targeting of FphF and combined with its high level of expression during biofilm-associated growth conditions future studies will be able to elucidate the role of FphF in *S. aureus* homeostasis and virulence *in vivo*.

Methods

FphF cloning, expression and purification for activity measurements

The full length *fph*F (currently annotated as *estA*, gene loci SAOUHSC_02962, NWMN_2528, SAUSA300_2564) was amplified from the *S. aureus* ATCC35556 genome using primers 5' ATGAGGATCCGCTTATATTTCATTAAACTATCA 3' and 5' GAAACTCGAGTTAATCATTCACCATCCATGTT 3' that introduced XhoI and BamHI restriction sites, respectively. The PCR product was gel-purified and extracted before double-digestion with XhoI and BamHI-HF and dephosphorylation with Antarctic Phosphatase. The resulting gene fragment was ligated into XhoI- and BamHI-HF-double digested pET28a using T4 DNA ligase (NEB). The ligation mixture was transformed into NEB 5-α-competent *Escherichia coli*.

For expression of recombinant full length-FphF the pET28a-SAOUHSC02962 plasmid was transformed into BL21 (DE3) competent *E. coli* and grown on LB-Kanamycin (35 μg/mL) selection media. An overnight culture of the transformed bacteria in selection medium was diluted 1:250 into 500 mL LB-Kanamycin in a 2 L flask and grown at 37°C, 220rpm until the culture reached an OD_{600nm} of 0.5. Recombinant protein expression was induced by addition of 10 μM Isopropyl β- d-1-thiogalactopyranoside. Cultures were continued to shake at 27°C for 4 h, before cells were harvested by centrifugation (8000 g, 10 min, 4°C). The cell pellet was re-suspended in a small volume of LB medium, transferred to two 50 mL polypropylene tubes and centrifuged again (4000g, 10 min, 4°C). The supernatant was discarded and cell pellets were frozen at -80°C. Each pellet was re-suspended in 8 mL Lysis Buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole) and lysed by sonication. The lysate was cleared by centrifugation at 4,350 g, 30 min, 4°C and the supernatant was added to 1 mL Ni-NTA-Agarose. The sample was incubated and mixed by rotation at 4°C for 1-2 h. The resin was washed 3x with wash buffer (Lysis buffer with 20 mM imidazole) before His_6-tagged protein was

eluted with elution buffer (Lysis buffer with 250 mM imidazole) in 8 fractions of 1 mL. Eluates were pooled and concentrated 10-fold in 10,000 MWCO spin columns. 10 mL of 50 mM Tris-HCl, pH8.0, 300 mM NaCl, 20% glycerol were added and the samples were concentrated 10-fold. Samples were combined and protein concentration determined by OD_{280nm} measurements using the sequence-specific calculated extinction coefficient (E1% = 14.24).

ABP-labeling of recombinant FphF

Recombinant FphF was diluted into PBS/0.01% SDS (50 nM) and was pre-incubated with JCP678 or DMSO at 37°C for 30 min before FP-TMR was added (1 μM final concentration) for fluorescent labelling of active protein at 37°C for an additional 30 min. After addition of 4x SDS-PAGE sample buffer, the samples were boiled at 95 C for 10 min, cooled down and 35 ng of protein were analyzed by SDS-PAGE. The gel was scanned for TMR fluorescence on a Typhoon 9410 variable mode imager (λ_{ex}= 380 nm, 580 BP filter). Subsequently, the gel was subjected to silver staining and photographed over a transilluminator.

FphF enzymatic activity assays

The hydrolytic activity of purified recombinant FphF protein was tested using a series of 4-Methylumbelliferyl(4-MU)-based fluorogenic substrates as described previously [9]. In brief, 0.3 μL of fluorogenic substrates (10 mM in DMSO) was added to the wells of an opaque flat-bottom 384 well plate. 30 μL of a 10 nM solution of FphF in PBS/0.01%TritonX-100 was added and fluorescence (λ_{ex} = 365 nm and λ_{em} = 455 nm) was read at 37°C at 1 min intervals on a Cytation 3 imaging reader (BioTek, Winooski, VT, USA) for 60 min. In the linear phase of the reaction (10 – 40 min) turnover rates were calculated using Gen5 software (BioTek) as RFU/min and were normalized by subtracting background hydrolysis rates measured for each substrate in reaction buffer in the absence of protein.

FphF cloning, expression, and purification for crystallization

FphF DNA construct was designed based on UniProt sequence Q2FUY3 omitting the starting methionine with overhangs for ligation-independent cloning [42]. The construct was synthesized by Integrated DNA Technology (IDT) and cloned into modified pET28a-LIC vectors incorporating an N-terminal His_6-tag and a 3C protease cleavage site.

E. coli BL21(DE3) cells in 1 L cultures (1X Luria-Bertani with 50 μg/mL kanamycin) at 37 °C and 200 rpm shaking were grown until OD_{600} reached 0.6. Cultures were induced with 0.2 mM Isopropyl β- d-1-thiogalactopyranoside and grown overnight at 18 °C and 200 rpm shaking. Bacterial pellets were harvested via centrifugation, suspended in 50 mM Tris pH 8.0, 300 mM NaCl and stored at -20 °C.

For purification thawed suspended pellets were incubated for 30 min on ice with ~20 μg/mL lysozyme and ~4 μg/mL DNase. Cells were lysed via sonication (Sonifier Heat Systems Ultrasonics). FphF protein was initially purified by Ni^{2+} affinity chromatography (HIS-Select resin, Sigma-Aldrich) using an elution buffer containing 50 mM Tris pH 8.0, 300 mM NaCl, 300 mM Imidazole, 10% (v/v) glycerol and 10% (w/v) sucrose. Elution fractions were incubated with 3C protease and 2 mM DTT overnight at 4 °C. FphF was further purified by size-exclusion chromatography using anion exchange (RESOURCE Q) and/or Superdex 75 or 200 Increase columns (GE Life Sciences). Anion exchange (10-20 mM HEPES pH 7.5, gradient from 10 to 1000 mM NaCl) separated two FphF species, potentially representing two different dimer forms. Both anion exchange peaks looked identical on SDS-PAGE analysis and both yielded crystals. The first eluted major peak (~80% of protein) was used for most experiments and resulted in the presented datasets. Anion exchange was followed by size-exclusion chromatography (10-20 mM HEPES pH 7.5, 10-50 mM NaCl), which resulted in a single dimer peak. Purified protein was either used directly for crystal drops or snap frozen in liquid nitrogen.

Inhibitor synthesis

KT129 and KT130 were synthesized in-house as previously described [10], and their purity was 98% and 96% by HPLC analysis.

FphF crystallization

FphF was broad screened for crystallization resulting in multiple hits with details given in supporting Table S1. After optimization, the following three datasets were obtained, fully refined and deposited.

For the FphF apo-form crystal structure 0.2 µL of ~8.5 mg/mL FphF (20 mM HEPES pH 7.5, 10 mM NaCl) were mixed with 0.1 µL FphF crystal seeds (in 54.4% Tacsimate pH 7.0, 0.1 M Bis-Tris propane pH 6.5, 8% Polypropylene glycol P 400) and 0.2 µL of reservoir solution. The sitting drop reservoir contained 50 µL of 2.8 M sodium acetate. Crystals were soaked for ~20 seconds in 75% reservoir solution and 25% ethylene glycol prior to freezing in liquid nitrogen.

For the FphF KT129 bound crystal structure 0.2 µL of ~7.5 mg/mL FphF+KT129 (0.12 mM KT129, 12% DMSO, 18 mM HEPES pH 7.5, 8 mM NaCl) were mixed with 0.2 µL of reservoir solution. The sitting drop reservoir contained 50 µL of 0.2 M sodium citrate, 0.1 M Bis-tris propane pH 6.5, 20% w/v PEG 3350. Crystals were soaked for ~20 seconds in 75% reservoir solution and 25% glycerol prior to freezing in liquid nitrogen.

For the FphF KT130 bound crystal structure 0.25 µL of ~6.6 mg/mL FphF+KT130 (0.2 mM KT130, 20% DMSO, 8 mM HEPES pH 7.5, 40 mM NaCl) were mixed with 0.2 µL of reservoir solution. The sitting drop reservoir contained 50 µL of 0.8 M sodium formate, 10% w/v PEG 8000, 10% w/v PEG 1000 and 0.1 M HEPES pH 7.0. Crystals were soaked for ~20 seconds in 75% reservoir solution and 25% glycerol prior to freezing in liquid nitrogen.

For the FphF heptyl acyl bound crystal structure 0.4 µL of ~8.0 mg/ml FphF (10 mM HEPES pH 7.5, 10 mM NaCl) were mixed with 0.07 µL ligand solution (~0.5 mM 4-MU heptanoate in 100% DMSO) and 0.4 µL of reservoir solution. Sitting drop reservoir contained 50 µL of 0.8 M Sodium formate, 10 % w/v PEG 8000, 10 % w/v PEG 1000 and 0.1 M Tris pH 7.5. Crystals were soaked for ~20 seconds in 75% reservoir solution and 25% glycerol prior to freezing in liquid nitrogen.

FphF data collection and processing

X-ray diffraction data were collected at the Australian synchrotron MX1 [43] and MX2 [44] beamlines. Datasets were processed with XDS [45], merging and scaling were performed using AIMLESS [46]. Phases were solved with Phenix Phaser molecular replacement [47] using a model created with BALBES [48]. Model building and refinement were conducted in COOT [49] and Phenix [50]. Ligand creation and restraint generation utilized Jligand [51] and eLBOW [52]. Statistics for the datasets are listed in Table 1 and 2. Structure figures, analysis and alignments were created with ChemSketch [53], UCSF Chimera [54] and Procheck [55].

Table 1: FphF data collection and processing

Values for the outer shell are given in parentheses.

	apo	KT129 bound	KT130 bound	heptyl acyl bound
PDB ID	6VH9	6VHD	6VHE	6WCX
Diffraction source	Australian synchrotron MX2	Australian synchrotron MX2	Australian synchrotron MX1	Australian synchrotron MX1
Wavelength (Å)	0.954	0.954	0.954	0.954
Detector	DECTRIS EIGER X 16M	DECTRIS EIGER X 16M	DECTRIS EIGER X 9M	DECTRIS EIGER X 9M
Space group	P 6_1 2 2	P 6_1 2 2	P 6_1 2 2	P 6_1 2 2
a, b, c (Å)	87.1, 87.2, 453.6	87.2, 87.2, 455.2	87.1, 87.1, 454.9	87.0 87.0 454.7
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution range (Å)	49.16 – 1.71 (1.74 – 1.71)	49.28 – 1.98 (2.02 – 1.98)	49.24 – 1.94 (1.98 – 1.94)	49.19 – 2.89 (3.07 – 2.89)
Total No. of reflections	3,064,967 (144,842)	1,980,193 (107,296)	3,049,489 (145,969)	217,394 (26,252)
No. of unique reflections	112,281 (5,332)	73,040 (4,307)	77,616 (4,466)	23,922 (3,641)
Completeness (%)	99.9 (97.8)	99.9 (98.5)	100.0 (99.5)	99.4 (97.4)
Redundancy	27.3 (27.2)	27.1 (24.9)	39.3 (32.7)	9.1 (7.2)
$\langle I/\sigma(I)\rangle$	16.5 (1.5)	16.7 (1.5)	23.1 (2.2)	7.7 (1.6)
$CC_{1/2}$	0.999 (0.581)	0.999 (0.711)	1.000 (0.804)	0.988 (0.524)
$R_{merge.}$	0.128 (3.155)	0.151 (2.676)	0.152 (1.957)	0.223 (1.137)
$R_{p.i.m.}$	0.025 (0.605)	0.029 (0.535)	0.024 (0.342)	0.069 (0.380)

Table 2: FphF structure solution and refinement

Values for the outer shell are given in parentheses.

	apo	KT129 bound	KT130 bound	heptyl acyl bound
PDB ID	6VH9	6VHD	6VHE	6WCX
Resolution range (Å)	43.57 – 1.71 (1.73 – 1.71)	49.28 – 1.98 (2.00 – 1.98)	49.24 – 1.94 (1.96 – 1.94)	49.19 – 2.89 (2.96 – 2.89)
Final R_{cryst}	0.177 (0.316)	0.176 (0.301)	0.168 (0.272)	0.225 (0.305)
Final R_{free}	0.201 (0.342)	0.213 (0.333)	0.209 (0.332)	0.261 (0.272)
Protein residues	1020	1020	1020	1020
Ligands	4 (Na)	4 (KT129)	4 (KT130)	4 (heptyl acyl)
Water	457	258	499	17
R.m.s. deviations				
Bonds (Å)	0.009	0.010	0.007	0.007
Angles (°)	1.190	1.399	0.948	0.947
Average *B* factors (Å^2)	39.5	46.9	38.0	43.9
Ligands	42.4	56.0	46.0	42.4
Water	45.0	48.3	40.9	36.0
Ramachandran plot				
Most favored (%)	96.2	96.3	96.7	96.2
Outlier (%)	0	0	0	0

Structure prediction and docking

Structure prediction was done by the I-Tasser [30] webserver using chain A of the FphF apo structure (PDB ID 6VHD). The only other input was a sequence obtained from UniProt: FphA (UniProt ID Q2FVG3, predicted molecular weight 52.0 kDa); FphB (Q2FV90, 36.8); FphC (Q2FYZ3, 35.3); FphD (Q2G2D6, 33.2); FphE (Q2FV39, 31.0); FphF (Q2FUY3, 29.1); FphG (Q2G2V6, 28.4); FphH (Q2G025, 28.1); FphI (Q2G0V7, 27.4); FphJ (Q2FVA9, 21.8). The resulting models were manually inspected and compared to FphF. The best model was picked by examining the prediction of the core secondary structure elements, the location of the active site triad, as well as statistics provided by I-Tasser (Table S3). Ligand creation, conversion and manipulation used the tools PRODRG [56], Open Babel [57] and COOT [49]. Molecular docking experiments were performed with GOLD [16]. Default

parameters were used unless stated otherwise, with hydrogens added to the protein structure using the gold_serine_protease_VS template. The binding site was centered at the active site serine residue. For uncleaved ligands, the active site serine was mutated to a glycine. For docking triazole inhibitors ligand flexibility options were utilized. For covalent docking, the GOLD instructions were followed, adding a "link oxygen atom" to the ligands and defining the link between the active site serine and the link oxygen atom.

Supporting information: Contains additional figures and tables, including detailed information for specific points raised in the main text. This information is available free of charge on the ACS Publications website

PDB ID Codes: The datasets generated and analyzed during the current study are available in the worldwide Protein Data Bank under PDB ID **6VH9** (FphF apo), PDB ID **6VHD** (FphF KT129 bound), PDB ID **6VHE** (FphF KT130 bound) and PDB ID **6WCX** (FphF heptyl acyl bound). Authors will release the atomic coordinates and experimental data upon article publication.

Corresponding Author Information: To whom correspondence should be addressed: Matthias Fellner: Department of Biochemistry, School of Biomedical Sciences, University of Otago, PO Box 56, Dunedin 9054, New Zealand; matthias.fellner@otago.ac.nz; T: +64 22 4975643

Author Contributions: M.F. and P.M. conceptualization; M.F., C.L. and S.J. protein purification; C.L. activity and inhibition assays; M.F. and S.J. crystallization; M.F. and J.B. data collection and processing; M.F. data deposition, curation and analysis; M.F. model prediction; M.F. docking studies; L.C. chemistry analysis; M.F. manuscript writing; M.F., C.L., J.B., L.C., M.B. and P.M. manuscript review and editing; M.F., C.L., L.C. and P.M. visualization; M.F., C.L., M.B. and P.M. funding acquisition; M.F., C.L., M.B. and P.M. project administration.

Abbreviations:

Fph: fluorophosphonate-binding hydrolases

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

Acknowledgments: This work was supported by a University of Otago Health Sciences Postdoctoral Fellowship No. HSCDPD1703 (to M.F.); a postdoctoral research fellowship of the German Research Foundation DFG (to C.L.); a China Scholarship Council Fellowship No. 201604910274 (to L.C.); a Youth Innovation Promotion Association of the Chinese Academy of Sciences grant No. 2017329 (to L.C.).

References

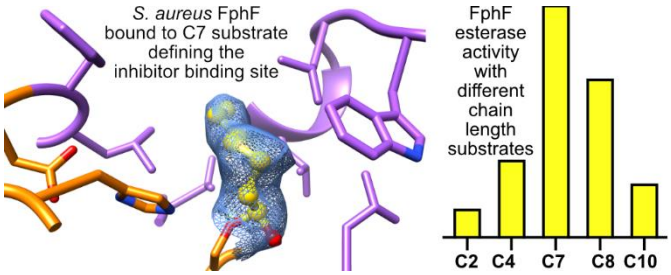
1. Wertheim, H. F.; Melles, D. C.; Vos, M. C.; van Leeuwen, W.; van Belkum, A.; Verbrugh, H. A.; Nouwen, J. L., The role of nasal carriage in *Staphylococcus aureus* infections. *Lancet Infect Dis* **2005,** *5*, 751-762. DOI: 10.1016/S1473-3099(05)70295-4.
2. Tong, S. Y.; Davis, J. S.; Eichenberger, E.; Holland, T. L.; Fowler, V. G., Jr., *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* **2015,** *28*, 603-661. DOI: 10.1128/CMR.00134-14.
3. Turnidge, J. D.; Kotsanas, D.; Munckhof, W.; Roberts, S.; Bennett, C. M.; Nimmo, G. R.; Coombs, G. W.; Murray, R. J.; Howden, B.; Johnson, P. D. R.; Dowling, K.; Cooperative, A. N. Z., *Staphylococcus aureus* bacteraemia: a major cause of mortality in Australia and New Zealand. *Med J Australia* **2009,** *191*, 368-373.
4. Tong, S. Y. C.; Davis, J. S.; Eichenberger, E.; Holland, T. L.; Fowler, V. G., *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* **2015,** *28*, 603-661. DOI: 10.1128/Cmr.00134-14.
5. Laupland, K. B.; Lyytikainen, O.; Sogaard, M.; Kennedy, K. J.; Knudsen, J. D.; Ostergaard, C.; Galbraith, J. C.; Valiquette, L.; Jacobsson, G.; Collignon, P.; Schonheyder, H. C.; International Bacteremia Surveillance, C., The changing epidemiology of *Staphylococcus aureus* bloodstream infection: a multinational population-based surveillance study. *Clin Microbiol Infect* **2013,** *19*, 465-471. DOI: 10.1111/j.1469-0691.2012.03903.x.
6. de la Fuente-Nunez, C.; Reffuveille, F.; Fernandez, L.; Hancock, R. E. W., Bacterial biofilm development as a multicellular adaptation: antibiotic resistance and new therapeutic strategies. *Curr Opin Microbiol* **2013,** *16*, 580-589. DOI: 10.1016/j.mib.2013.06.013.
7. Waters, E. M.; Rowe, S. E.; O'Gara, J. P.; Conlon, B. P., Convergence of *Staphylococcus aureus* Persister and Biofilm Research: Can Biofilms Be Defined as Communities of Adherent Persister Cells? *Plos Pathog* **2016,** *12*, e1006012. DOI: 10.1371/journal.ppat.1006012.
8. Lebeaux, D.; Ghigo, J. M.; Beloin, C., Biofilm-related infections: bridging the gap between clinical management and fundamental aspects of recalcitrance toward antibiotics. *Microbiol Mol Biol Rev* **2014,** *78*, 510-543. DOI: 10.1128/MMBR.00013-14.
9. Lentz, C. S.; Sheldon, J. R.; Crawford, L. A.; Cooper, R.; Garland, M.; Amieva, M. R.; Weerapana, E.; Skaar, E. P.; Bogyo, M., Identification of a *S. aureus* virulence factor by activity-based protein profiling (ABPP). *Nat Chem Biol* **2018,** *14*, 609-617. DOI: 10.1038/s41589-018-0060-1.
10. Chen, L.; Keller, L. J.; Cordasco, E.; Bogyo, M.; Lentz, C. S., Fluorescent triazole urea activity-based probes for the single-cell phenotypic characterization of *Staphylococcus aureus*. *Angew Chem Int Ed Engl* **2019,** *58*, 5643-5647. DOI: 10.1002/anie.201900511.
11. Keller, L. J.; Lentz, C. S.; Chen, Y. E.; Metivier, R. J.; Weerapana, E.; Fischbach, M. A.; Bogyo, M., Characterization of Serine Hydrolases Across Clinical Isolates of Commensal Skin Bacteria Staphylococcus epidermidis Using Activity-Based Protein Profiling. *ACS Infect Dis* **2020,** *6*, 930-938. DOI: 10.1021/acsinfecdis.0c00095.
12. Krissinel, E.; Henrick, K., Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **2007,** *372*, 774-797. DOI: 10.1016/j.jmb.2007.05.022.
13. Ollis, D. L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S. M.; Harel, M.; Remington, S. J.; Silman, I.; Schrag, J.; Sussman, J. L.; Verschueren, K. H. G.; Goldman, A., The alpha/beta hydrolase fold. *Protein Eng* **1992,** *5*, 197-211. DOI: 10.1093/protein/5.3.197.
14. Schrag, J. D.; Cygler, M., Lipases and alpha/beta hydrolase fold. *Methods Enzymol* **1997,** *284*, 85-107. DOI: 10.1016/s0076-6879(97)84006-2.
15. Nardini, M.; Lang, D. A.; Liebeton, K.; Jaeger, K. E.; Dijkstra, B. W., Crystal structure of pseudomonas aeruginosa lipase in the open conformation. The prototype for family I.1 of bacterial lipases. *J Biol Chem* **2000,** *275*, 31219-31225. DOI: 10.1074/jbc.M003903200.
16. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R., Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* **1997,** *267*, 727-748. DOI: 10.1006/jmbi.1996.0897.
17. Hu, Y.; Liu, Y.; Li, J.; Feng, Y.; Lu, N.; Zhu, B.; Xue, S., Structural and functional analysis of a low-temperature-active alkaline esterase from South China Sea marine sediment microbial metagenomic library. *J Ind Microbiol Biotechnol* **2015,** *42*, 1449-1461. DOI: 10.1007/s10295-015-1653-2.

18. Kahya, H. F.; Andrew, P. W.; Yesilkaya, H., Deacetylation of sialic acid by esterases potentiates pneumococcal neuraminidase activity for mucin utilization, colonization and virulence. *Plos Pathog* **2017,** *13*, e1006263. DOI: 10.1371/journal.ppat.1006263.
19. Kim, M. H.; Kang, B. S.; Kim, S.; Kim, K. J.; Lee, C. H.; Oh, B. C.; Park, S. C.; Oh, T. K., The crystal structure of the estA protein, a virulence factor from *Streptococcus pneumoniae*. *Proteins* **2008,** *70*, 578-583. DOI: 10.1002/prot.21680.
20. Holm, L., Benchmarking fold detection by DaliLite v.5. *Bioinformatics* **2019,** *35*, 5326-5327. DOI: 10.1093/bioinformatics/btz536.
21. Legler, P. M.; Kumaran, D.; Swaminathan, S.; Studier, F. W.; Millard, C. B., Structural characterization and reversal of the natural organophosphate resistance of a D-type esterase, Saccharomyces cerevisiae S-formylglutathione hydrolase. *Biochemistry* **2008,** *47*, 9592-9601. DOI: 10.1021/bi8010016.
22. van Straaten, K. E.; Gonzalez, C. F.; Valladares, R. B.; Xu, X.; Savchenko, A. V.; Sanders, D. A., The structure of a putative S-formylglutathione hydrolase from Agrobacterium tumefaciens. *Protein Sci* **2009,** *18*, 2196-2202. DOI: 10.1002/pro.216.
23. Wu, D.; Li, Y.; Song, G.; Zhang, D.; Shaw, N.; Liu, Z. J., Crystal structure of human esterase D: a potential genetic marker of retinoblastoma. *FASEB J* **2009,** *23*, 1441-1446. DOI: 10.1096/fj.08-125286.
24. Alterio, V.; Aurilia, V.; Romanelli, A.; Parracino, A.; Saviano, M.; D'Auria, S.; De Simone, G., Crystal structure of an S-formylglutathione hydrolase from Pseudoalteromonas haloplanktis TAC125. *Biopolymers* **2010,** *93*, 669-677. DOI: 10.1002/bip.21420.
25. Legler, P. M.; Leary, D. H.; Hervey, W. J. t.; Millard, C. B., A role for His-160 in peroxide inhibition of S. cerevisiae S-formylglutathione hydrolase: evidence for an oxidation sensitive motif. *Arch Biochem Biophys* **2012,** *528*, 7-20. DOI: 10.1016/j.abb.2012.08.001.
26. Lemak, S.; Tchigvintsev, A.; Petit, P.; Flick, R.; Singer, A. U.; Brown, G.; Evdokimova, E.; Egorova, O.; Gonzalez, C. F.; Chernikova, T. N.; Yakimov, M. M.; Kube, M.; Reinhardt, R.; Golyshin, P. N.; Savchenko, A.; Yakunin, A. F., Structure and activity of the cold-active and anion-activated carboxyl esterase OLEI01171 from the oil-degrading marine bacterium Oleispira antarctica. *Biochem J* **2012,** *445*, 193-203. DOI: 10.1042/BJ20112113.
27. Chen, N. H.; Counago, R. M.; Djoko, K. Y.; Jennings, M. P.; Apicella, M. A.; Kobe, B.; McEwan, A. G., A glutathione-dependent detoxification system is required for formaldehyde resistance and optimal survival of Neisseria meningitidis in biofilms. *Antioxid Redox Signal* **2013,** *18*, 743-755. DOI: 10.1089/ars.2012.4749.
28. Wefers, D.; Cavalcante, J. J. V.; Schendel, R. R.; Deveryshetty, J.; Wang, K.; Wawrzak, Z.; Mackie, R. I.; Koropatkin, N. M.; Cann, I., Biochemical and Structural Analyses of Two Cryptic Esterases in Bacteroides intestinalis and their Synergistic Activities with Cognate Xylanases. *J Mol Biol* **2017,** *429*, 2509-2527. DOI: 10.1016/j.jmb.2017.06.017.
29. Lee, C. W.; Yoo, W.; Park, S. H.; Le, L.; Jeong, C. S.; Ryu, B. H.; Shin, S. C.; Kim, H. W.; Park, H.; Kim, K. K.; Kim, T. D.; Lee, J. H., Structural and functional characterization of a novel cold-active S-formylglutathione hydrolase (SfSFGH) homolog from *Shewanella frigidimarina*, a psychrophilic bacterium. *Microb Cell Fact* **2019,** *18*, 140. DOI: 10.1186/s12934-019-1190-1.
30. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y., The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **2015,** *12*, 7-8. DOI: 10.1038/nmeth.3213.
31. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J. L., The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res* **2007,** *35*, D301-303. DOI: 10.1093/nar/gkl971.
32. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W. Z.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **2011,** *7*. DOI: 10.1038/msb.2011.75.
33. Jaeger, K. E.; Dijkstra, B. W.; Reetz, M. T., Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. *Annu Rev Microbiol* **1999,** *53*, 315-351. DOI: 10.1146/annurev.micro.53.1.315.
34. Kang, E. H.; Gebru, E.; Kim, M. H.; Cheng, H.; Park, S. C., EstA protein, a novel virulence factor of *Streptococcus pneumoniae*, induces nitric oxide and pro-inflammatory cytokine production

in RAW 264.7 macrophages through NF-kappaB/MAPK. *Microb Pathog* **2009,** *47*, 196-201. DOI: 10.1016/j.micpath.2009.07.002.
35. Arpigny, J. L.; Jaeger, K. E., Bacterial lipolytic enzymes: classification and properties. *Biochem J* **1999,** *343 Pt 1*, 177-183.
36. Janssen, A. P. A.; van Hengst, J. M. A.; Béquignon, O. J. M.; Deng, H.; van Westen, G. J. P.; van der Stelt, M., Structure Kinetics Relationships and Molecular Dynamics Show Crucial Role for Heterocycle Leaving Group in Irreversible Diacylglycerol Lipase Inhibitors. *Journal of Medicinal Chemistry* **2019,** *62*, 7910-7922. DOI: 10.1021/acs.jmedchem.9b00686.
37. Adibekian, A.; Martin, B. R.; Wang, C.; Hsu, K.-L.; Bachovchin, D. A.; Niessen, S.; Hoover, H.; Cravatt, B. F., Click-generated triazole ureas as ultrapotent in vivo–active serine hydrolase inhibitors. *Nature Chemical Biology* **2011,** *7*, 469-478. DOI: 10.1038/nchembio.579.
38. Ahn, K.; Boehm, M.; Brown, M. F.; Calloway, J.; Che, Y.; Chen, J.; Fennell, K. F.; Geoghegan, K. F.; Gilbert, A. M.; Gutierrez, J. A.; Kalgutkar, A. S.; Lanba, A.; Limberakis, C.; Magee, T. V.; O'Doherty, I.; Oliver, R.; Pabst, B.; Pandit, J.; Parris, K.; Pfefferkorn, J. A.; Rolph, T. P.; Patel, R.; Schuff, B.; Shanmugasundaram, V.; Starr, J. T.; Varghese, A. H.; Vera, N. B.; Vernochet, C.; Yan, J., Discovery of a Selective Covalent Inhibitor of Lysophospholipase-like 1 (LYPLAL1) as a Tool to Evaluate the Role of this Serine Hydrolase in Metabolism. *ACS Chem Biol* **2016,** *11*, 2529-2540. DOI: 10.1021/acschembio.6b00266.
39. Liu, X.; Kokare, C., Microbial enzymes of use in industry. In *Biotechnology of microbial enzymes*, Elsevier: 2017; pp 267-298.
40. Jiang, H.; Zhang, S.; Gao, H.; Hu, N., Characterization of a cold-active esterase from Serratia sp. and improvement of thermostability by directed evolution. *BMC biotechnology* **2016,** *16*, 7.
41. Wilhelm, S.; Gdynia, A.; Tielen, P.; Rosenau, F.; Jaeger, K. E., The autotransporter esterase EstA of Pseudomonas aeruginosa is required for rhamnolipid production, cell motility, and biofilm formation. *J Bacteriol* **2007,** *189*, 6695-6703. DOI: 10.1128/JB.00023-07.
42. Luna-Vargas, M. P.; Christodoulou, E.; Alfieri, A.; van Dijk, W. J.; Stadnik, M.; Hibbert, R. G.; Sahtoe, D. D.; Clerici, M.; Marco, V. D.; Littler, D.; Celie, P. H.; Sixma, T. K.; Perrakis, A., Enabling high-throughput ligation-independent cloning and protein expression for the family of ubiquitin specific proteases. *J Struct Biol* **2011,** *175*, 113-119. DOI: 10.1016/j.jsb.2011.03.017.
43. Cowieson, N. P.; Aragao, D.; Clift, M.; Ericsson, D. J.; Gee, C.; Harrop, S. J.; Mudie, N.; Panjikar, S.; Price, J. R.; Riboldi-Tunnicliffe, A.; Williamson, R.; Caradoc-Davies, T., MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. *J Synchrotron Radiat* **2015,** *22*, 187-190. DOI: 10.1107/S1600577514021717.
44. Aragao, D.; Aishima, J.; Cherukuvada, H.; Clarken, R.; Clift, M.; Cowieson, N. P.; Ericsson, D. J.; Gee, C. L.; Macedo, S.; Mudie, N.; Panjikar, S.; Price, J. R.; Riboldi-Tunnicliffe, A.; Rostan, R.; Williamson, R.; Caradoc-Davies, T. T., MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J Synchrotron Radiat* **2018,** *25*, 885-891. DOI: 10.1107/S1600577518003120.
45. Kabsch, W., Xds. *Acta Crystallogr D Biol Crystallogr* **2010,** *66*, 125-132. DOI: 10.1107/S0907444909047337.
46. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011,** *67*, 235-242. DOI: 10.1107/S0907444910045749.
47. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *J Appl Crystallogr* **2007,** *40*, 658-674. DOI: 10.1107/S0021889807021206.
48. Long, F.; Vagin, A. A.; Young, P.; Murshudov, G. N., BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **2008,** *64*, 125-132. DOI: 10.1107/S0907444907050172.
49. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** *66*, 486-501. DOI: 10.1107/S0907444910007493.
50. Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D., Towards

automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **2012,** *68*, 352-367. DOI: 10.1107/S0907444912001308.
51. Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N., JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr D Biol Crystallogr* **2012,** *68*, 431-440. DOI: 10.1107/S090744491200251X.
52. Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D., electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* **2009,** *65*, 1074-1080. DOI: 10.1107/S0907444909029436.
53. ACD/ChemSketch, version 2018.1.1. *Advanced Chemistry Development, Inc., Toronto, ON, Canada* **2018,** *www.acdlabs.com*.
54. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **2004,** *25*, 1605-1612. DOI: 10.1002/jcc.20084.
55. Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M., Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *Journal of Applied Crystallography* **1993,** *26*, 283-291. DOI: Doi 10.1107/S0021889892009944.
56. Schuttelkopf, A. W.; van Aalten, D. M., PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **2004,** *60*, 1355-1363. DOI: 10.1107/S0907444904011679.
57. O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R., Open Babel: An open chemical toolbox. *J Cheminform* **2011,** *3*, 33. DOI: 10.1186/1758-2946-3-33.

For Table of Contents Only



Supporting information

Structural basis for the inhibitor and substrate specificity of the unique Fph serine hydrolases of Staphylococcus *aureus*

Authors

Matthias Fellner[1]*, Christian S. Lentz[2,3], Sam A. Jamieson[1], Jodi L. Brewster[1], Linhai Chen[2,4], Matthew Bogyo[2], Peter D. Mace[1]

[1]Biochemistry Department, School of Biomedical Sciences, University of Otago, Dunedin 9054, New Zealand

[2]Pathology, Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305, United States

[3]Centre for New Antibacterial Strategies (CANS) and Research Group for Host-Microbe Interactions, Department of Medical Biology, UiT – The Arctic University of Norway, Tromsø N-9037, Norway

[4]National Center for Drug Screening, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

*To whom correspondence should be addressed: Matthias Fellner: Department of Biochemistry, School of Biomedical Sciences, University of Otago, PO Box 56, Dunedin 9054, New Zealand; matthias.fellner@otago.ac.nz; T: +64 22 4975643

Table of contents:

Table S1: Crystallization conditions of FphF.

Highest diffracting crystal in Å, with datasets presented in this publication originating from crystal form 1. Crystal form 2 showed significant anisotropy and translational noncrystallographic symmetry with a significant Patterson peak at 0 0 0.5.

(Å)	Compound 1	Compound 2	Compound 3
Crystal form 1*			
1.7	2.8 M Sodium acetate		
1.9	0.2 M Trisodium citrate	0.1 M Bis-tris propane pH 6.5	20% w/v PEG 3350
1.9	0.8 M Sodium formate	0.1 M Tris pH 7.5 or HEPES pH 7.0	10% w/v PEG 8000 10% w/v PEG 1000
2.2	0.7-0.8 M Sodium formate	0.1 M Tris pH 7.5 or HEPES pH 7.0	25% w/v PEG MME 2000
2.2	1.5 M Li_2SO_4	0.1 M Tris pH 8.0 or 8.5	
~4	2.0 M $(NH_4)_2SO_4$	0.1 M Sodium cacodylate pH 6.5	0.2 M NaCl
Crystal form 2**			
2.1	40-75% Tacsimate	0-0.1 M Bis-tris propane pH 6.5-8.0	0-8% w/v Polypropylene glycol
2.6	0.8 M Sodium formate	0-0.1 M Tris pH 7.5	25% w/v PEG MME 2000
2.7	0.8-1 M Sodium formate	0-0.1 M Sodium cacodylate pH 6.8 or HEPES pH 7.0	7-10% w/v PEG 8000 7-10% w/v PEG 1000
2.9	0.2 M Trisodium citrate	0.1 M Bis-tris propane pH 6.5	20% w/v PEG 3350
3.0	1.4 M Disodium malonate	0.1 M Bis-tris propane pH 7.0	
~8	0.2 M Ammonium acetate	0.1 M HEPES pH 7.5	25% w/v PEG 3350
Crystal form 3***			
3.1	2.7-2.8 M Sodium acetate	0-0.1 M Sodium cacodylate pH 6.8 M or Tris pH 7.5	
3.2	0.8 M Sodium formate	0-0.1 M Tris pH 7.5 or HEPES pH 7.0	10% w/v PEG 8000 10% w/v PEG 1000
3.3	0.7-0.9 M Sodium formate	0.1 M Sodium cacodylate pH 6.4 or Tris pH 7.5 or 8.5	25% w/v PEG MME 2000
3.6	60% Tacsimate	0-0.1 M Bis-tris propane pH 7.0	0-0.1 M $(NH_4)_2SO_4$ or Disodium malonate
3.8	1.4-2.4 M Disodium malonate	0-0.1 M Bis-tris propane pH 7.0	
3.8	1-1.4 M Trisodium citrate	0.1 M Bis-tris propane pH 7.0 or Sodium cacodylate pH 6.5 or HEPES pH 7.0	
~4	0.2 M Trisodium citrate	0.1 M Bis-tris propane pH 6.5	20% w/v PEG 3350
~5	0.8 M Sodium formate	0.1 M Tris pH 7.5	25% w/v PEG MME 2000
~7	0.2 M Ammonium acetate	0.1 M HEPES pH 7.5	25% w/v PEG 3350
~9	1.8 M Li_2SO_4	0.1 M Tris pH 8.8	

*Crystal form 1: $P6_1$ 2 2; a, b, c (Å); α, β, γ (°) = ~87, 87, 454; 90, 90, 120; 4 chains

**Crystal form 2: $P6_2$ 2 2; a, b, c (Å); α, β, γ (°) = ~88, 88, 222; 90, 90, 120; 2 chains

***Crystal form 3: $P6_4$ 2 2; a, b, c (Å); α, β, γ (°) = ~88, 88, 110; 90, 90, 120; 1 chain

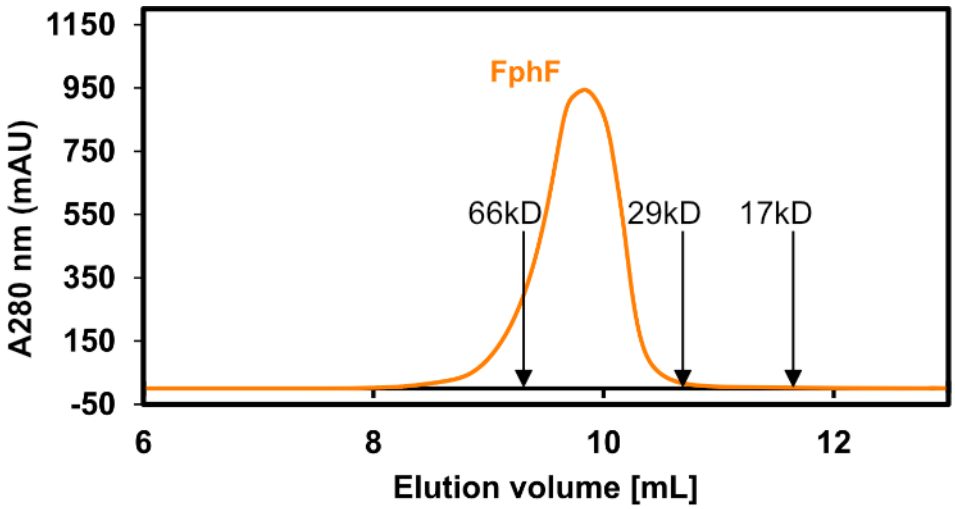


Figure S1: FphF gel-filtration chromatogram. A280 Superdex 75 chromatogram comparison of 58 kD FphF dimer (29 kD monomer) compared to molecular weight standard proteins bovine serum albumin 66 kD; carbonic anhydrase 29 kD and diubiquitin 17 kD.

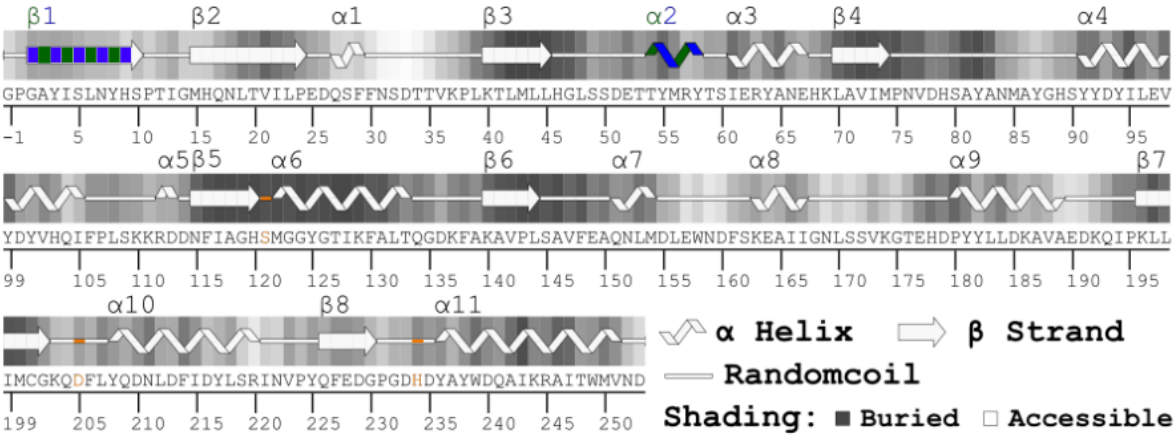


Figure S2: Sequence and corresponding secondary structure elements of FphF. The active site triad is shown in orange. The two different dimer interfaces within the tetramer indicated by the blue and green coloring of β1-β1 and α2-α2.

Sodium binding in the FphF apo-protein crystal structure

All FphF apo-protein crystal structures contained a three-atom modification to Ser121. The electron density peak for the central atom ~2.4 Å from the OG of serine was always significantly larger, than what could be accounted for by a water molecule. No anomalous dispersion peak was observed in electron density maps at this site, ruling out a heavy metal. Purified FphF and FphF crystals were subjected to mass spectrometry analysis, but no covalent post-translational modification could be detected. Sodium chloride was the only compound present in across all crystallization conditions. The density was therefore modelled as a sodium ion with two water molecules (Figure S3). Two other α/β hydrolases with similar active site serine bound sodium ions from *Coxiella burnetii* (PDB ID 3TRD [1]) and *Pseudomonas fluorescens* (3T4U unpublished) were identified using DALI [2] and rcsb.org [3] searches.

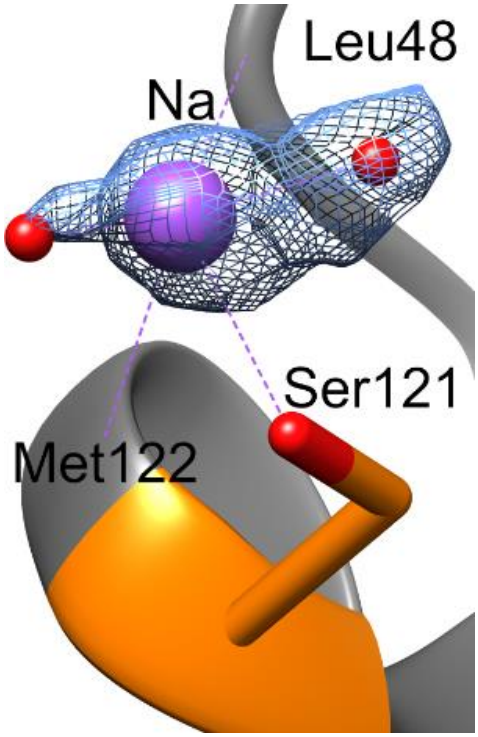


Figure S3: Sodium binding in the FphF apo-protein. The *2FO−FC* maps for sodium atom (purple sphere) and chelating water molecules (red sphere) are shown, electron density is contoured at 1 σ (blue mesh) (PDB ID 6VH9). Sodium coordination to Ser121, the backbone of Met122 and the backbone of Leu48 is illustrated as purple dashes.

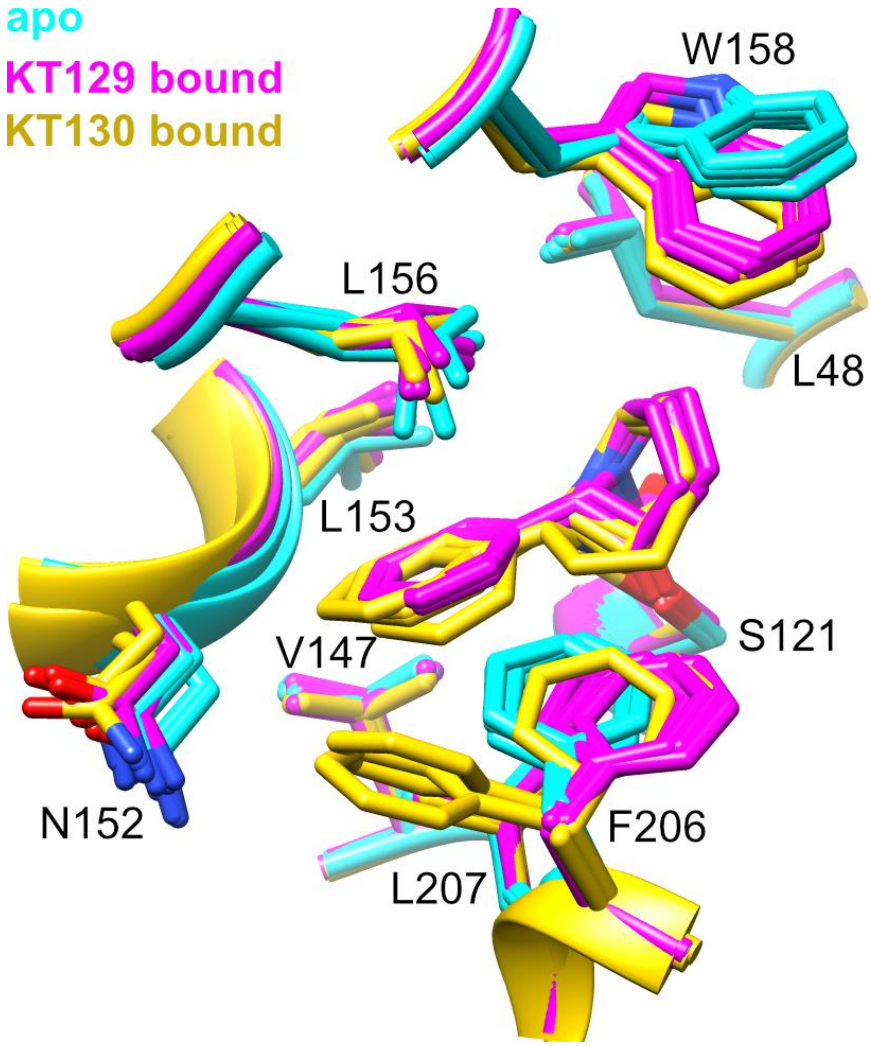


Figure S4: Comparison of unbound and bound active sites. Active site comparison based on Cα alignment of four chains of the FphF tetramer. Apo in cyan (PDB ID 6VH9), KT129 inhibitor bound in magenta (6VHD) and KT130 inhibitor bound in yellow (6VHE).

Alternate ligand conformations in FphF crystal structures

KT129 (PDB ID 6VHD) and KT130 (6VHE) crystal structures contain evidence of a second minor conformation at some active sites. However, in the majority of cases evidence to build a model for the minor conformation was not strong enough. In the KT130 structure chain C showed the most convincing remaining density and a second alternate minor conformation for the ligand was built at the same active site. The minor conformation induces several protein changes, especially the backbone and side chains of the 152-158 loop are influenced. The conformations refined to 55% major and 45% minor occupancy (Figure S4). The heptyl acyl group (6WCX) also hinted at a possible second conformation in some chains but the refinement of alternate conformations of the ligand were not satisfying.

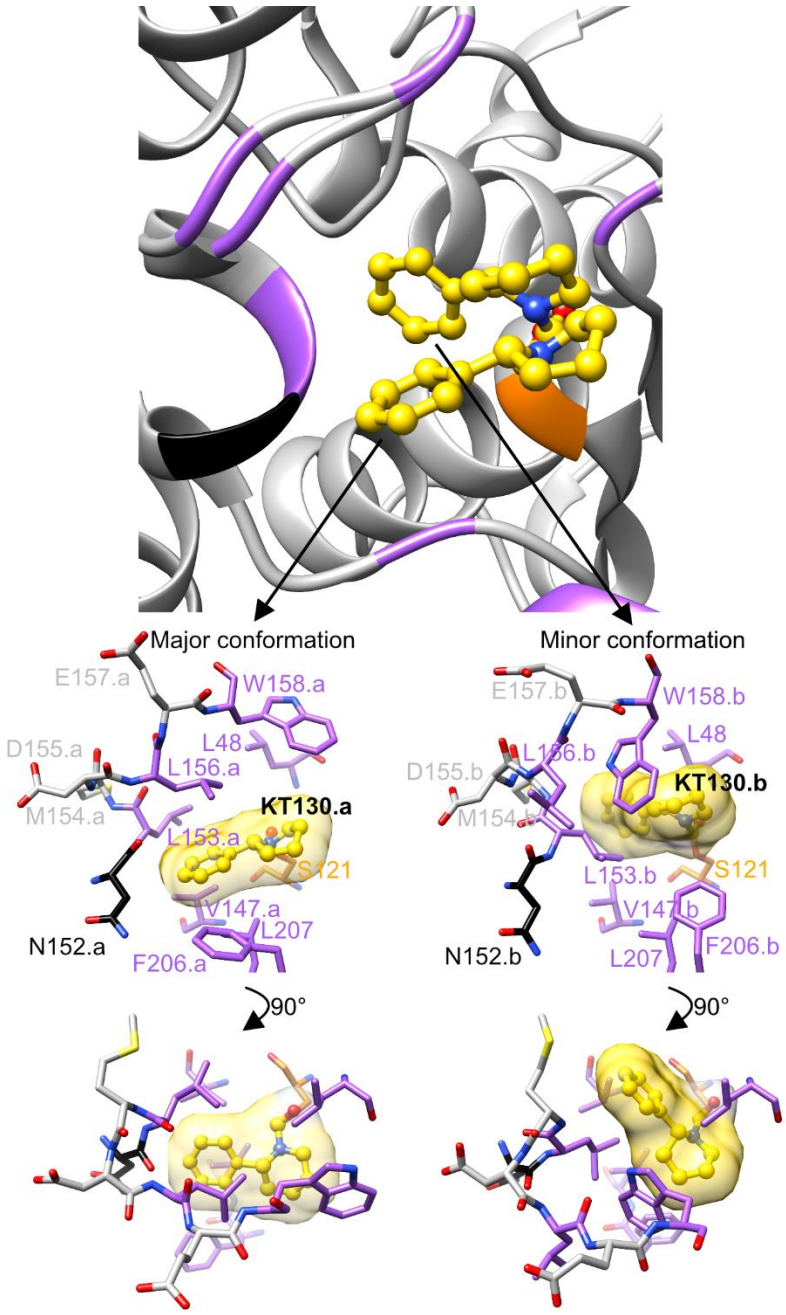


Figure S5: Comparison of major and minor conformation of KT130. KT130 crystal structure chain C (PDB ID 6VHE) with two alternate conformations. Separated view of the major and minor conformation with transparent ligand surface.

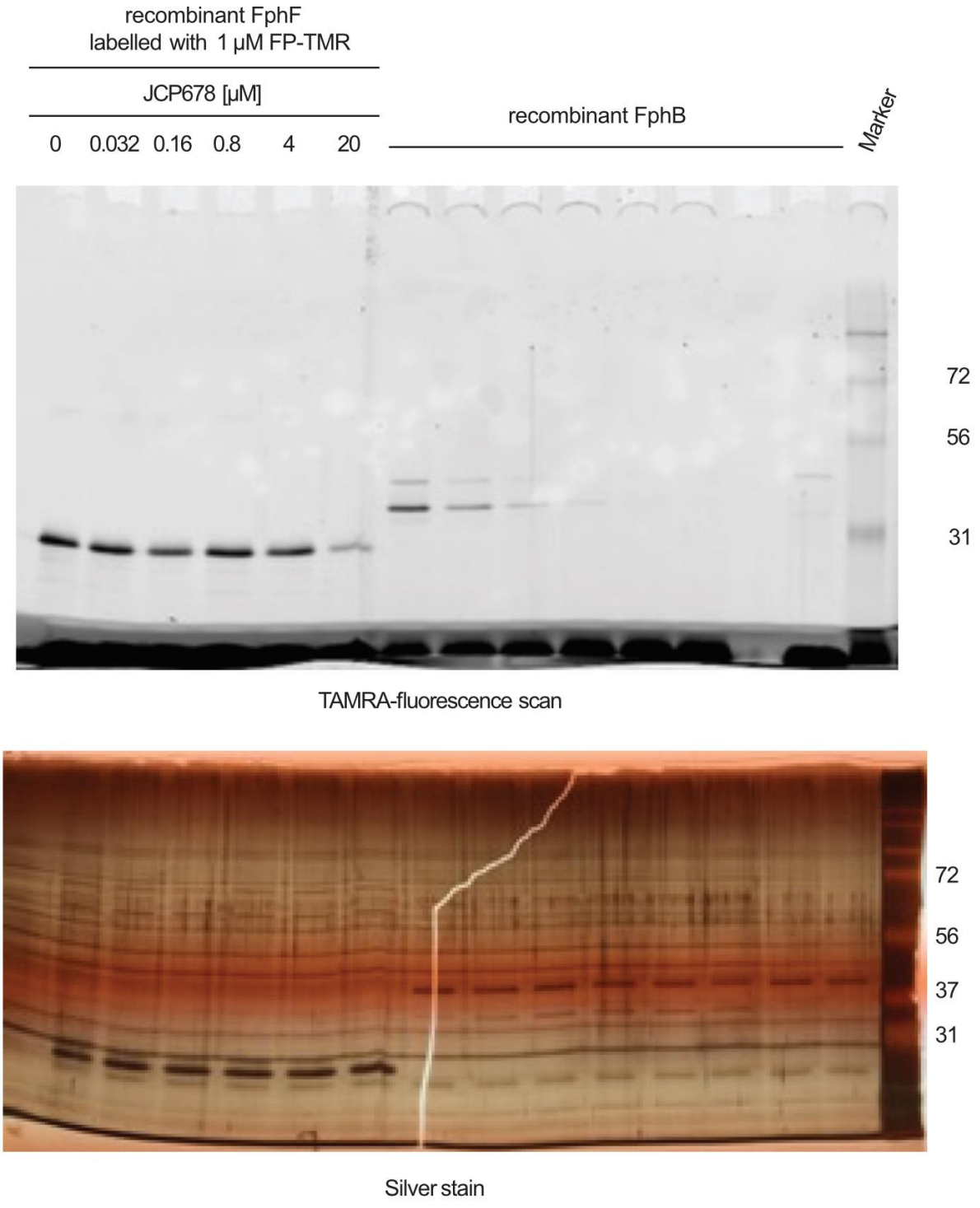


Figure S6: FphF inhibition and labelling - complete gel. Complete gel corresponding to figure 3A with molecular weight marker proteins kD on the right.

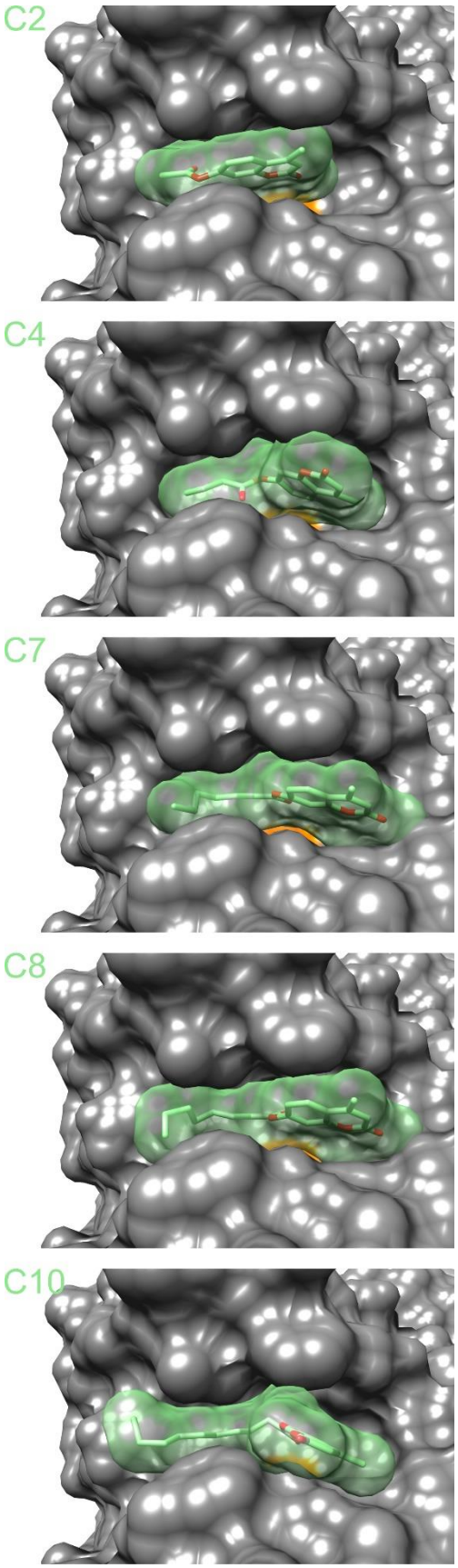


Figure S7: FphF substrate docking comparison. Full length 4-Methylumbelliferyl (4-MU) acetate (C2), 4-MU butyrate (C4), 4-MU heptanoate (C7), 4-MU octanoate (C8) and 4-MU decanoate (C10) docked into 6VHD-S121G.

Table S2: Esterases similar to FphF.

The Dali server was used for protein structure database searching, the table is ranked by the Dali Z-score [2].

#	Name	Organism	FphF % Identity	PDB ID	Z-score	rmsd	lali	Reference
Dali hits:								
1	Est12	*uncultured bacterium FLS12*	39	4RGY	35.7	1.8	236	[4]
2	*Sp*EstA[a]	*Streptococcus pneumoniae*	27	2UZ0	31.8	2.0	238	[5]
3	*Spy*EstA	*Streptococcus pyogenes*	30	4ROT	30.9	2.1	237	unpublished
4	*Ph*Est[b]	*Pseudoalteromonas haloplanktis*	21	3LS2	29.8	2.0	235	[6]
5	*Ol*EI01171[a]	*Oleispira antarctica*	23	3S8Y, 3i6y	29.4	2.0	235	[7]
6	*Nm*EstD[b]	*Neisseria meningitidis*	24	4B6G	29.2	2.0	234	[8]
7	*Sf*SFGH[a]	*Shewanella frigidimarina*	23	6JZL	29.1	2.1	235	[9]
8	*Bi*Fae1A[b]	*Bacteroides intestinalis*	22	5VOL	29.1	2.0	230	[10]
9	*Hs*ESD[b]	*Homo sapiens*	22	3FCX	28.8	2.1	232	[11]
10	*Atu*SFGH[a]	*Agrobacterium tumefaciens*	21	3E4D	28.8	2.1	232	[12]
11	*Sc*SFGH[a]	*Saccharomyces cerevisiae*	19	1PV1, 3C6B, 4FLM, 4FOL	27.1	2.7	239	[13, 14]

a…Experimental determined preference of C2 substrates over longer substrates.

b…Experimental determined activity towards C2 substrate without investigations of longer substrates.

lali…number of aligned residue pairs

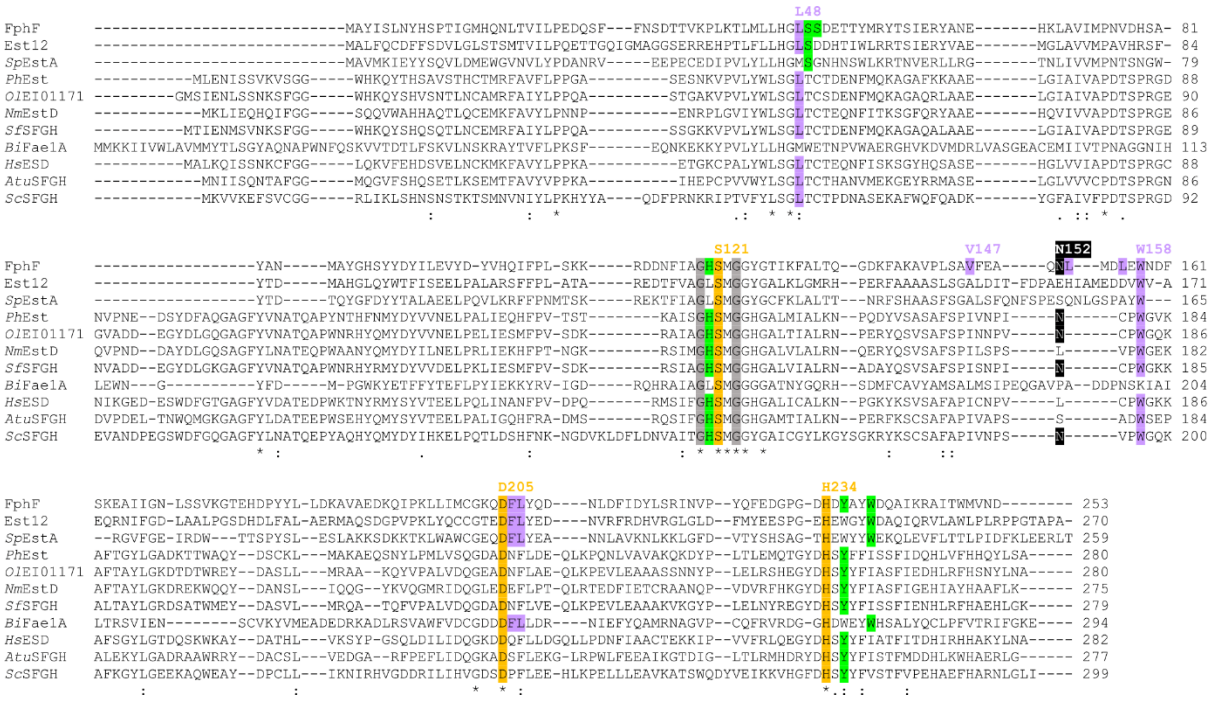


Figure S8: Sequence alignment of FphF with other esterases. Alignment of FphF (UniProt ID Q2FUY3), Est12 (B8Y562), *Sp*EstA (A0A0H2UNZ8), *Ph*Est (Q3IL66), *Ol*EI01171 (D0VWZ4), *Nm*EstD (Q9JZ43), *Sf*SFGH (Q07XK4), *Bi*Fae1A (B3CET1), *Hs*ESD (P10768), *Atu*SFGH (A9CJ11) and *Sc*SFGH (P40363). Selected residues of FphF labelled are above and the level of conservation indicated below. The conserved active site triad is highlighted in orange and G-X-S-X-G motif in grey. The conserved acyl binding pocket residues (relative to FphF) are highlighted in purple, with the N152 gate in black and the leaving group binding pocket residues in green.

Figure S9: Sequence alignment of FphF and FphB. Alignment of *S. aureus* FphF (UniProt ID Q2FUY3) and *S. aureus* FphB (Q2FV90). Selected residues of FphF are labelled above and the level of conservation indicated below. The conserved active site triad is highlighted in orange (FphF S-H-D, FphB S-H-E) and the conserved G-X-S-X-G motif in grey. The conserved acyl binding pocket residues (relative to FphF) are highlighted in purple, with the N152 gate in black and the leaving group binding pocket residues in green.

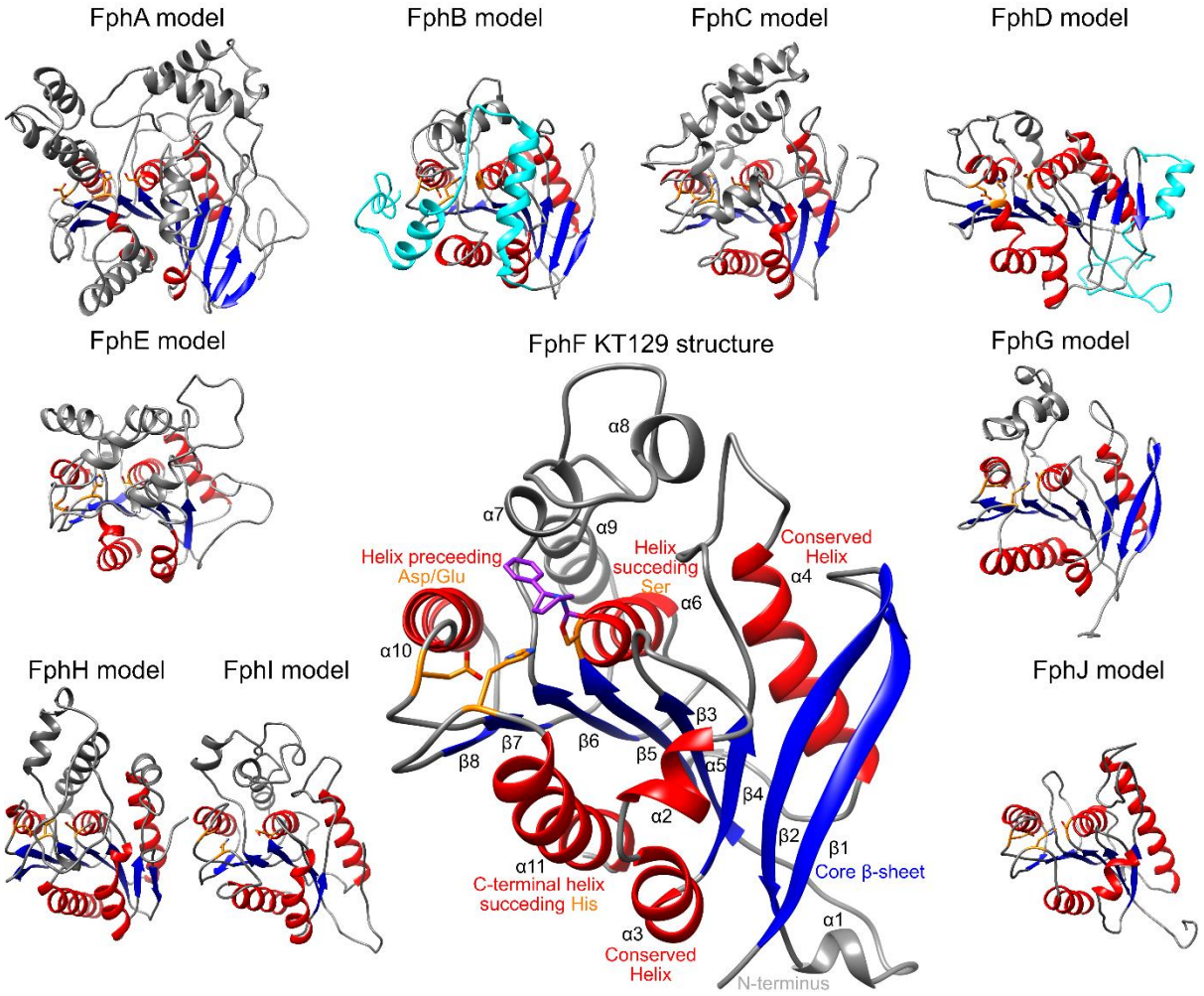


Figure S10: Predicted structures of *S. aureus* Fph proteins. Structure of FphF with KT129 bound in comparison to the homology models of the other Fph proteins, generated by I-Tasser [15]. Figures are based on full length Cα alignments. The β-sheet core is shown in blue, conserved α-helices in red, active site Ser-His-Asp/Glu triad in orange. The N-terminal extension present in FphB and FphD is shown in cyan. Secondary structure labelling is indicated for FphF and the bound substrate KT129 is colored purple.

Table S3: I-Tasser final predicted structures C-score and 1-10 threading templates statistics.

	PDB hit	Iden1	Iden2	Cov	Norm. Z-score
FphA C-score = 0.33					
1	user (FphF)	0.14	0.1	0.44	10
2	1qe3A	0.32	0.33	0.95	3.53
3	1c7jA	0.32	0.33	0.98	5.52
4	1c7jA	0.32	0.33	0.98	4.5
5	2pm8	0.25	0.28	0.98	2.61
6	2pm8	0.26	0.28	0.98	1.99
7	1qe3A	0.32	0.33	0.95	4.52
8	1qo9	0.23	0.27	0.98	2.91
9	6qaaA	0.27	0.29	0.99	8.42
10	1qe3A	0.33	0.33	0.95	4.06
FphB C-score = 0.47					
1	user (FphF)	0.17	0.17	0.69	10
2	2wirA	0.18	0.25	0.9	4.18
3	2wirA	0.18	0.25	0.92	2.74
4	2wirA	0.18	0.25	0.91	3.93
5	2wirA	0.18	0.25	0.91	2.86
6	3zwqA	0.18	0.25	0.93	3.24
7	5l2pA	0.14	0.18	0.9	2.81
8	2c7b	0.18	0.2	0.86	1.12
9	2yh2A	0.18	0.24	0.91	3.16
10	4krxA	0.13	0.24	0.9	2.82
FphC C-score = -0.67					
1	user (FphF)	0.16	0.19	0.74	10
2	5o7gA	0.27	0.29	0.99	2.65
3	5o7gA	0.26	0.29	0.99	3.55
4	5o7gA	0.26	0.29	0.99	2.7
5	5o7g	0.26	0.29	0.98	0.8
6	5o7g	0.26	0.29	0.99	0.7
7	5o7gA	0.26	0.29	0.99	2.88
8	5o7g	0.27	0.29	0.98	0.91
9	5oluA	0.27	0.29	0.99	4.81
10	5o7gA	0.26	0.29	0.99	2.42
FphD C-score = -0.35					
1	user (FphF)	0.15	0.17	0.73	10
2	1pv1A	0.15	0.18	0.85	1.31
3	3fleA	0.61	0.51	0.8	1.88
4	3fleA	0.64	0.51	0.8	2.82
5	3fleA	0.63	0.53	0.83	2.06
6	3bf7	0.18	0.15	0.73	0.73
7	3fle	0.61	0.51	0.8	0.9
8	3fleA	0.61	0.51	0.8	2.48
9	3fle	0.61	0.51	0.8	0.99
10	3fleA	0.63	0.53	0.83	4.64

FphE C-score = -0.79					
1	user (FphF)	0.1	0.19	0.73	10
2	5egnA	0.16	0.21	0.93	2.5
3	4q3lA	0.15	0.2	0.95	3.21
4	4q3lA	0.16	0.19	0.95	2.34
5	3wzl	0.2	0.21	0.91	0.9
6	2qmq	0.12	0.2	0.93	0.64
7	4q3lA	0.16	0.19	0.95	2.39
8	4mea	0.19	0.25	0.94	0.94
9	2xuaA	0.15	0.16	0.92	4.81
10	5egnA	0.15	0.21	0.93	2.05
FphG C-score = -0.44					
1	user (FphF)	0.11	0.21	0.89	10
2	4hxeB	0.15	0.25	1	2.06
3	5l8sA	0.18	0.19	1	4.23
4	4hxeB	0.15	0.25	0.99	2.25
5	4hxeB	0.15	0.25	1	2.74
6	4hxeB	0.15	0.25	1	2.1
7	4hxgA	0.14	0.25	0.97	3.33
8	4hxeB	0.15	0.25	0.99	4.06
9	4hxeB	0.15	0.25	0.99	3.91
10	4hxeB	0.15	0.25	1	3.51
FphH C-score = 0.22					
1	user (FphF)	0.17	0.21	0.82	10
2	5xksA	0.24	0.26	1	2.64
3	4diuA	0.57	0.56	0.99	3.3
4	1tqhA	0.57	0.56	0.98	2.56
5	3dkr	0.27	0.28	0.96	0.8
6	3dkr	0.25	0.28	0.96	0.59
7	5xksA	0.24	0.26	1	2.73
8	3dkr	0.27	0.28	0.96	0.91
9	1tqhA	0.58	0.56	0.98	4.17
10	5xksA	0.24	0.26	1	2.29
FphI C-score = 0.36					
1	user (FphF)	0.15	0.16	0.86	10
2	5xksA	0.27	0.28	0.99	2.46
3	4diuA	0.29	0.3	0.99	3.28
4	1tqhA	0.3	0.3	0.98	2.47
5	3dkr	0.25	0.26	0.96	0.81
6	3dkr	0.25	0.26	0.97	0.59
7	5xksA	0.26	0.28	0.98	2.25
8	3dkr	0.26	0.26	0.96	0.9
9	5xksA	0.31	0.28	0.89	3.65
10	3rm3A	0.24	0.26	0.99	2.04
FphJ C-score = 0.48					
1	user (FphF)	0.14	0.23	0.9	10
2	3u0vA	0.17	0.22	0.99	2.33
3	3u0v	0.17	0.22	0.98	0.84

4	3u0v	0.16	0.22	0.99	0.64
5	3u0vA	0.17	0.22	0.98	2.72
6	3fcxA	0.14	0.21	0.99	1.49
7	3u0vA	0.16	0.22	0.98	3.44
8	3u0vA	0.17	0.22	0.99	2.29
9	3u0vA	0.18	0.22	0.98	2.43
10	5kreA	0.16	0.22	0.99	3.23

Final predicted I-Tasser structures C-score: C-score is a confidence score for estimating the quality of predicted models by I-TASSER. It is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. C-score is typically in the range of [-5,2], where a C-score of higher value signifies a model with a high confidence and vice-versa.

1-10 I-Tasser threading templates statistics: Rank of templates represents the top ten threading templates used by I-TASSER.

Ident1 is the percentage sequence identity of the templates in the threading aligned region with the query sequence

Ident2 is the percentage sequence identity of the whole template chains with query sequence.

Cov represents the coverage of the threading alignment and is equal to the number of aligned residues divided by the length of query protein.

Norm. Z-score is the normalized Z-score of the threading alignments. Alignment with a Normalized Z-score >1 mean a good alignment and vice versa.

Supporting information references

[1] Franklin, M. C., Cheung, J., Rudolph, M. J., Burshteyn, F., Cassidy, M., Gary, E., Hillerich, B., Yao, Z. K., Carlier, P. R., Totrov, M., and Love, J. D. (2015) Structural genomics for drug design against the pathogen *Coxiella burnetii*, *Proteins 83*, 2124-2136.

[2] Holm, L. (2019) Benchmarking fold detection by DaliLite v.5, *Bioinformatics 35*, 5326-5327.

[3] Berman, H., Henrick, K., Nakamura, H., and Markley, J. L. (2007) The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data, *Nucleic Acids Res 35*, D301-303.

[4] Hu, Y., Liu, Y., Li, J., Feng, Y., Lu, N., Zhu, B., and Xue, S. (2015) Structural and functional analysis of a low-temperature-active alkaline esterase from South China Sea marine sediment microbial metagenomic library, *J Ind Microbiol Biotechnol 42*, 1449-1461.

[5] Kim, M. H., Kang, B. S., Kim, S., Kim, K. J., Lee, C. H., Oh, B. C., Park, S. C., and Oh, T. K. (2008) The crystal structure of the estA protein, a virulence factor from *Streptococcus pneumoniae*, *Proteins 70*, 578-583.

[6] Alterio, V., Aurilia, V., Romanelli, A., Parracino, A., Saviano, M., D'Auria, S., and De Simone, G. (2010) Crystal structure of an S-formylglutathione hydrolase from Pseudoalteromonas haloplanktis TAC125, *Biopolymers 93*, 669-677.

[7] Lemak, S., Tchigvintsev, A., Petit, P., Flick, R., Singer, A. U., Brown, G., Evdokimova, E., Egorova, O., Gonzalez, C. F., Chernikova, T. N., Yakimov, M. M., Kube, M., Reinhardt, R., Golyshin, P. N., Savchenko, A., and Yakunin, A. F. (2012) Structure and activity of the cold-active and anion-activated carboxyl esterase OLEI01171 from the oil-degrading marine bacterium Oleispira antarctica, *Biochem J 445*, 193-203.

[8] Chen, N. H., Counago, R. M., Djoko, K. Y., Jennings, M. P., Apicella, M. A., Kobe, B., and McEwan, A. G. (2013) A glutathione-dependent detoxification system is required for formaldehyde resistance and optimal survival of Neisseria meningitidis in biofilms, *Antioxid Redox Signal 18*, 743-755.

[9] Lee, C. W., Yoo, W., Park, S. H., Le, L., Jeong, C. S., Ryu, B. H., Shin, S. C., Kim, H. W., Park, H., Kim, K. K., Kim, T. D., and Lee, J. H. (2019) Structural and functional characterization of a novel cold-active S-formylglutathione hydrolase (SfSFGH) homolog from *Shewanella frigidimarina*, a psychrophilic bacterium, *Microb Cell Fact 18*, 140.

[10] Wefers, D., Cavalcante, J. J. V., Schendel, R. R., Deveryshetty, J., Wang, K., Wawrzak, Z., Mackie, R. I., Koropatkin, N. M., and Cann, I. (2017) Biochemical and Structural Analyses of Two Cryptic Esterases in Bacteroides intestinalis and their Synergistic Activities with Cognate Xylanases, *J Mol Biol 429*, 2509-2527.

[11] Wu, D., Li, Y., Song, G., Zhang, D., Shaw, N., and Liu, Z. J. (2009) Crystal structure of human esterase D: a potential genetic marker of retinoblastoma, *FASEB J 23*, 1441-1446.

[12] van Straaten, K. E., Gonzalez, C. F., Valladares, R. B., Xu, X., Savchenko, A. V., and Sanders, D. A. (2009) The structure of a putative S-formylglutathione hydrolase from Agrobacterium tumefaciens, *Protein Sci 18*, 2196-2202.

[13] Legler, P. M., Leary, D. H., Hervey, W. J. t., and Millard, C. B. (2012) A role for His-160 in peroxide inhibition of S. cerevisiae S-formylglutathione hydrolase: evidence for an oxidation sensitive motif, *Arch Biochem Biophys 528*, 7-20.

[14] Legler, P. M., Kumaran, D., Swaminathan, S., Studier, F. W., and Millard, C. B. (2008) Structural characterization and reversal of the natural organophosphate resistance of a D-type esterase, Saccharomyces cerevisiae S-formylglutathione hydrolase, *Biochemistry 47*, 9592-9601.

[15] Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction, *Nat Methods 12*, 7-8.