Discovery of a non-canonical prototype long-chain monoacylglycerol lipase through a structure-based endogenous reaction intermediate complex

Nikos Pinotsis[1,$,#], Anna Krüger[2,#], Nicolas Tomas[3], Spyros D Chatziefthymiou[1], Claudia Litz[1], Simon Arnold Mortensen[1], Mamadou Daffé[3], Hedia Marrakchi[3], Garabed Antranikian[2] and Matthias Wilmanns[1,4]*

[1] European Molecular Biology Laboratory, Hamburg Unit, Notkestrasse 85, 22607 Hamburg, Germany

[2] Hamburg University of Technology, Kasernenstrasse 12, 21073 Hamburg, Germany

[3] Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, Toulouse, France

[4] University Hamburg Clinical Center Hamburg-Eppendorf, Martinistrasse 52, 20251 Hamburg, Germany

[$] Current address: Department of Chemistry, National and Kapodistrian University of Athens, Zografou, Greece

[#] Equal contribution

Corresponding author: Matthias Wilmanns, Email: matthias.wilmanns@embl-hamburg.de

ABSTRACT

The identification and characterization of enzyme function is largely lacking behind the rapidly increasing availability of large numbers of sequences and associated high-resolution structures. This is often hampered by lack of knowledge on *in vivo* relevant substrates. Here, we present a case study of a high-resolution structure of an unusual orphan lipase in complex with an endogenous C18 monoacylglycerol ester reaction intermediate from the expression host, which is insoluble under aqueous conditions and thus not accessible for studies in solution. The data allowed its functional characterization as a prototypic long-chain monoacylglycerol lipase, which uses a minimal lid domain to position the substrate through a hydrophobic tunnel directly to the enzyme's active site. Knowledge about the molecular details of the substrate binding site allowed us to modulate the enzymatic activity by adjusting protein/substrate interactions, demonstrating the potential of our findings for future biotechnology applications.

INTRODUCTION

The annotation of relevant enzyme function under physiological conditions remains a major challenge, despite the availability of millions of protein sequences and experimentally determined or predicted molecular structures associated with them [1,2]. Enzymes often have different and broader substrate profiles within their natural cellular environment than experimentally accessible under *in vitro* conditions or by using purified protein samples. Hence, albeit powerful tools have been developed to predict enzymatic function, options to prove them *in vivo* remain limited [3-8]. High-resolution structural biology allows the direct visualization of enzyme-ligand complexes at atomic detail and thus presents a highly attractive tool for enzyme function discovery, but it requires the availability of pure compounds for structural analysis. Although protein structures in the presence of large hydrophobic substrates or ligands have been published, often with a requirement of applying target-tailored solubilization protocols [9-12], they still remain exceptional and hence available structure repositories remain mostly populated with protein structures of soluble, small ligands. Alternatively, enzyme structures could also be determined in the presence of endogenous substrates or catalysis intermediates taken up from their cellular environment for direct evidence of enzyme-substrate relationships, but successful instances have remained rare to date [13]. This requires conditions where substrate turnover is inhibited by genetic modification of enzyme targets or experimental conditions that lead to lack of catalytic activity. In this study, we demonstrate the power of an endogenous substrate uptake for structural and subsequent functional characterization of an orphan enzyme with putative lipase/esterase activity, which otherwise would not have been accessible due to lack of solubility.

Lipases are enzymes that reversibly catalyze the hydrolysis and synthesis of a large variety of glycerol esters [14-16]. Knowledge about these enzymes has allowed the development of numerous lipase-catalyzed

51 processes in medical biotechnology, detergent industry, organic synthesis, biodiesel production,
52 agrochemical industry, flavor and aroma industry, and food production [17-19]. In contrast to many other
53 carboxyl ester hydrolases, lipases are active mainly at water-lipid interfaces, a process known as
54 interfacial activation [20,21]. In essence, long-chain glycerol esters are delivered in non-monomeric
55 emulsions rather than water-soluble substrates. Interfacial activation of lipases is thought to require an
56 additional α-helical domain referred to as "lid", alternatively also named "cap", which contains several
57 hydrophilic residues in the vicinity of the active site [22]. In contrast to the catalytic domain (CD), there is
58 only little sequence conservation of the lid. Available structural data on lipases show highly diverse lid
59 topology and structure, suggesting variable mechanisms to promote active site opening for access of acyl
60 substrates and rendering precise prediction of enzymatic substrate specificity challenging [12,23,24]. Lipases
61 from extremophilic microorganisms are of superior interest for applications in industrial processes since
62 demanding conditions such as high temperature or high salt concentration do not impair their activity [25-27].

63 Most of the previous molecular activity studies have been limited to isolated or recombinant lipases.
64 Given that quantitative lipase activity measurements are generally hampered by the limited solubility of
65 long-chain glycerol esters under those experimental conditions it is plausible to assume that the substrate
66 profile differs under cellular conditions.

67 Based on turnover measurements with a broad spectrum of substrates, an enzyme from the thermophilic
68 anaerobic bacterium *Thermoanaerobacter thermohydrosulfuricus (Tth)* was tentatively assigned as a
69 lipase [28]. This strain was originally isolated from Solar Lake located on the Sinai peninsula, which has an
70 extreme marine environment with a temperature range from 16 to 60°C and high level of salinity. These
71 environmental conditions have given rise to complex biochemical phenomena that are linked to cycles of
72 evaporation and infiltration of water external sources. Isolates from Solar Lake have shown remarkable
73 biochemical processes related to degradation of starch, amylose, and pullulan [29,30]. As this enzyme
74 revealed a remarkable level of robustness against a variety of solvents and temperatures, it represents an
75 attractive target for potential biotechnology applications. However, underlying molecular mechanisms
76 explaining the enzyme's activity properties and substrate promiscuity remained unknown.

77 In this work, we determined the structure of the enzyme revealed an unexpected long-chain C18
78 monoacylglycerol (MAG) bound to the active site, which was taken up from its expression host. Its distal
79 part is embedded by a tunnel formed by an unconventional lid domain with minimal α-helix-β-hairpin-α-
80 helix (HBH) topology. In cell extracts that overexpress the enzyme, we found high levels of C16/18 fatty
81 acid (FA) products, in agreement with our structural data, demonstrating that the *Tth* enzyme functions as
82 a long-chain MAG lipase. *Tth* MAG lipase prototypes a new lipase/esterase family with HBH lid domains
83 comprising sequences inserted into the HBH β-hairpin that are highly divergent in length and structure.

Our findings demonstrate a powerful example of using high-resolution structural biology where complex formation with an endogenous reaction intermediate has become achievable to elucidate its function.

RESULTS AND DISCUSSION

Tth MAG lipase reveals an α/β hydrolase fold with a minimal HBH lid domain

To provide insight into the *Tth* MAG lipase active-site topography and shed light on its unique substrate promiscuity, we characterized the enzyme structurally and functionally. Purified *Tth* MAG lipase elutes as a dimer **(Fig. 1a).** Its fold melting temperature measured by nano differential scanning fluorimetry (nanoDSF) is at 77°C, which is within expectation in light of the thermophilic origin of *Tth* **(Fig. 1b, Supplementary Table 1**). Next, we determined its high-resolution crystal structure, both in the presence and absence of the covalent active site serine protease inhibitor phenylmethylsulfonyl fluoride (PMSF) **(Fig. 1c & e & 2**, **Supplementary Table 2**). The overall fold of dimeric *Tth* MAG lipase can also be reliably modeled using the AlphaFold2 and MMseqs2 package **(Supplementary Fig. 1)** [31,32]. However, as there are significant differences in the areas of the bound ligands and the dimeric interface, in the remaining contribution we refer to our experimental structures only.

Both *Tth* MAG lipase structures – the active and phenylmethylsulfonyl (PMS)-inhibited – reveal a two-fold symmetric dimer, in agreement with its dimeric assembly in solution **(Fig. 1a & e)**. The *Tth* MAG lipase dimer is formed by an almost flat interface area which represents about 10 % of the total solvent accessible area **(Supplementary Fig. 2).** The interface is established through an intricate network of charged interactions involving residues from both domains, with the majority of the interactions originating from residues of the CD. It forms a large cleft that connects the two active sites of the two protomers **(Fig. 1e).** Converting the charge of one of the most central interface residues (E72R) renders a monomeric version of the protein, confirming that the interface observed in the structure is relevant **(Fig. 1a)**.

A minimal lid domain comprising the residues 141-183 is inserted between strand β6 and helix αE on top of the CD active-site area (**Fig. 1c-e, Supplementary Fig. 2**). Different from most other lipases, in which the lid is generally formed by a bundle of α-helices [22,24], in *Tth* MAG lipase this domain is compact. It is formed by two small α-helices, followed by a β-hairpin and an additional α-helix, referred to as α-helix/β-hairpin/α-helix (HBH) **(Fig. 1c, Supplementary Fig. 2**). In both *Tth* MAG lipase structures, the orientation of the HBH lid to the CD is basically identical (**Fig. 2d-e**), dismissing any large-scale opening/closing mechanisms previously shown for other lid domain-containing lipases [22,24]. Despite the presence of the lid domain, the MAG lipase active site remains largely accessible **(Fig. 1e)**.

116 *Long chain MAG reaction intermediate is held by a hydrophobic lid tunnel*

117 In the active site of both *Tth* MAG lipase structures, with and without PMS, we observed literally
118 identical density of a glycerol molecule, presenting one of the two final products of esterase/lipase
119 catalysis **(Fig. 2a-b, 3, Supplementary Fig. 3a).** The glycerol molecule tightly interacts via its two C2
120 and C3 hydroxyl groups with the carboxylate group of a non-conserved active-site residue Glu43, specific
121 to *Tth* MAG lipase (**Fig. 2a-b, Supplementary Fig. 5a**). As di- and triacylglycerol (DAG, TAG) esters
122 would require linkages through these two hydroxyl groups, only MAG ester binding via the C1 hydroxyl
123 group of the glycerol backbone is sterically possible. The dihedral conformation of this residue is in an
124 energetically less favored area of the Ramachandran plot in both *Tth* MAG lipase structures
125 (**Supplementary Fig. 6**), suggesting that its role in glycerol binding may supersede any other lower
126 energy conformation.

127 In both protomers of the active *Tth* MAG lipase structure, we found additional extended ligand electron
128 density pointing away from the active site crossing through a 15 Å long tunnel formed by the HBH lid
129 domain (**Supplementary Fig. 3a**). As we did not add any substrate to the crystallization buffer, bound
130 ligands can only originate from the *Escherichia (E.) coli* expression host FA pool. To rule out that the
131 presence of glycerol in the *Tth* MAG lipase structure is due to the addition of glycerol in the cryo-
132 protectant buffer, we determined an additional *Tth* MAG lipase structure without using glycerol in any
133 sample preparation step and cryo-protection prior to X-ray data collection (**Supplementary Table 2**). In
134 this structure, glycerol is found in a literally identical active site position in five out of six protein chains
135 of the respective asymmetric unit, confirming its origin to be independent from a specific crystallization
136 and cryo-protection protocol **(Supplementary Fig. 7**). To independently identify the ligand content of the
137 extended electron density, we analyzed the crystals of the active enzyme by reversed phase liquid
138 chromatography followed by mass spectrometry (LC-MS). A single peak found in the elution diagram
139 corresponded to the glycerol monostearate (2,3-dihydroxypropyloctadecanoate) **(Supplementary Fig.**
140 **3b**), suggesting that under the given sample preparation and crystallization conditions *Tth* MAG lipase
141 products cannot be released.

142 Based on these data, we modeled a monostearate (C18) acyl ligand into the electron density of the active
143 *Tth* MAG lipase structure **(Fig. 2a & Supplementary Fig. 3a)**. Since we observed significant connecting
144 density between the terminal C1 carbon of the C18 monoacyl ligand and both the γ-hydroxyl oxygen of
145 Ser113 and the O1 oxygen of the glycerol moiety, unrestrained refinement produced distances between
146 these atoms in the range of 2.2-2.6 Å. The resulting model, which was confirmed by an omit electron
147 density map, supports the presence of a loosely coordinated tetrahedral MAG reaction intermediate [14,33,34],
148 closely matching the geometry of covalently bound tetrahedral reaction mimics in a previously

149 established MAG lipase [12] (**Fig. 3 & Supplementary Fig. 4**). We also noticed that a conserved solvent
150 molecule required for ester hydrolysis next to the active site triad residue H233 detected in those
151 structures [12] is missing in our *Tth* MAG reaction ligand complex. As *Tth* MAG lipase is exposed to an
152 acidic water/organic solvent mixture in LC-MS experiments after its re-solubilization from crystals
153 favoring esterification [35], our most plausible explanation is a reverse process from the C18 reaction
154 intermediate observed in the crystal structure to glycerol monostearate substrate detected in LC-MS
155 experiments. All subsequent experiments were designed by interpreting our structural and mass
156 spectrometry-based findings as C18 MAG.

157 The distal part of the C18 MAG reaction intermediate involving carbon atoms C11 to C18 is held by a
158 hydrophobic tunnel across the *Tth* MAG lipase HBH lid domain, which is open at both ends **(Fig. 2a & c**
159 **& e)**. Whereas the proximal opening towards the *Tth* MAG lipase active site can be viewed as an
160 extension of the active site area (**Fig. 2a**), the distal lid tunnel exit is exposed and defined by a ring-like
161 structure of side chains from hydrophobic residues (**Fig. 2d-e**). In the presence of the C18 MAG reaction
162 intermediate, there are significant alterations in the packing arrangement of several residues within the lid
163 core when comparing to the PMSF-inhibited structure **(Fig. 2c-e)**. The largest spatial deviations are found
164 in the N-terminal part of the lid domain, which includes the first two lid helices α-LA and α-LB (**Fig. 2d,**
165 **Supplementary Fig. 8**). In the *Tth* MAG lipase ternary complex structure these two helices are separated
166 by a short loop, whereas in the PMSF-inhibited structure both merge into one long helix (residues 142-
167 161) with a sharp kink at residue Gln153 **(Fig. 2d**). As a result, in the PMSF-inhibited structure the side
168 chains of Ser150 and Tyr154 next to the kink block the formation of the lid tunnel (**Fig. 2c-e**). In the
169 presence of the C18 MAG reaction intermediate, the aromatic ring plane of Tyr154 becomes part of the
170 lid tunnel surface. In conclusion, our structural data are strongly suggestive of the hydrophobic lid tunnel
171 to be crucial for positioning long-chain MAG esters to allow efficient turnover by lipase hydrolysis.

172 *Mechanism for variable length MAG ester turnover*

173 To investigate our structure-based findings suggesting the *Tth* enzyme to function as a MAG lipase, we
174 first measured MAD, DAG and TAG ester hydrolysis using an established coupled *in vitro* assay [12]. For
175 C8 MAG, we found a turnover of 16.4 +/- 1.3 U mg^{-1} by the wild-type (wt) *Tth* enzyme. In contrast, only
176 residual turnover was observed for the corresponding C8 DAG and TAG esters **(Fig. 4b, left panel,**
177 **Supplementary Table 3),** demonstrating that the *Tth* enzyme is indeed highly specific for MAGs.
178 Furthermore, we only found residual activity for long chain C18:1 MAG, presumably due to limited
179 solubility.

180 We then screened a number of *Tth* MAG lipase mutants for quantitative analysis using C8 MAG as
181 substrate, (**Fig. 4b, right panel; Supplementary Table 3**). As expected, no activity was found for the

182 S113A active-site mutant, which was subsequently used as negative control. Testing further *Tth* MAG
183 lipase mutants without the lid domain (Δ140-183) and abolishing dimerization (E72R) also resulted in no
184 measurable activity, demonstrating the enzyme to be highly sensitive to structural alterations beyond the
185 catalytic center (**Fig. 2a-b**). We also probed two further residues, we considered to play roles in MAG
186 turnover. The first one was Glu43, which directly interacts with the O1 and O2 positions of the glycerol
187 moiety **(Fig. 2b)** and thus most likely contributes to the high level of MAG specificity as opposed to
188 DAG and TAG substrates. While we found about 1/3 of wt activity for the E43A variant, mutation of the
189 same residue to lysine (E43K) led to complete abolishment of MAG activity. This is supported by our
190 structural data **(Fig. 2b)**, suggesting that steric interference with the intricate interaction network between
191 the glycerol moiety of the MAG ester and residues from the *Tth* enzyme may impair its proper positioning
192 and thus substrate turnover. In contrast, removing the tyrosine group of Tyr154 (Y154A) or adding
193 another bulky side chain (Y154R), designed to test possible effect by residues from the *Tth* lid tunnel
194 **(Fig. 2c-e),** yielded only relative minor activity effects. These findings suggest that there is no significant
195 contribution of the lid to the turnover of C8 MAG, which is supported by our structural observations that
196 only the distal part of the long-chain C18 MAG is bound by the *Tth* lid tunnel **(Fig. 2a).** All *Tth* MAG
197 lipase variants were validated by circular dichroism to be properly folded as well as analyzed for thermal
198 stability and dimeric assembly **(Supplementary Tables 1 and 4, Supplementary Fig. 9 and 10**).

199 For quantitative characterization of *Tth* MAG lipase *in vitro* activity, we subsequently employed a
200 spectrophotometric monoacyl *p*-nitrophenyl acyl ester (*p*NP) assay [28,36]. We used this assay for
201 investigating substrates with variable acyl length ranging from C2 to C18, which is equivalent to the C18
202 MAG reaction intermediate we found in the *Tth* MAG lipase structure (**Fig. 4a**), as a convenient approach
203 for direct detection of turnover and thus allowing reliable quantification. We observed a skewed activity
204 profile for C2-C18 esters, with the highest activities found for C4-C8 esters (**Fig. 4c, Supplementary**
205 **Table 3)**. For C16 and C18 esters, there was only residual activity at a level where quantitative
206 significance estimates were not possible. When directly comparing C8 *p*NP ester hydrolysis with C8
207 MAG turnover (**Fig. 4b, left panel)**, the activity was about 15%, indicating that the presence of an
208 unnatural *p*NP leaving group replacing the glycerol leaving group in MAG substrates could have a
209 diminishing effect on *Tth* enzyme substrate turnover.

210 At this point, we suspected that the established *Tth* activity profiles are largely convolutes of diverting
211 length-dependent substrate binding affinities and catalytic turnover. To address this question, we
212 employed steady-state Michaelis-Menten kinetics for quantitatively determining catalytic efficiency
213 k_{cat}/K_M, again using the *p*NP ester turnover assay due to its superior statistical reliability **(Fig. 4d-e;**
214 **Supplementary Table 6**). We found stronger binding affinities with increasing *p*NP ester length in wt

215 *Tth* MAG lipase, ranging from residual C2 ester binding (> 2 mM) to substantially higher binding
216 affinities of around 0.1 mM for the longest measurable *p*NP esters (C12, C14) **(Fig. 4d, left panel).** As
217 expected, the catalytic rate constants (k_{cat}) for C2-C14 *p*NP esters show a similar profile as for overall
218 activity turnover measurements **(Fig. 4d, central panel; Supplementary Table 6**). As a result, the
219 catalytic efficiencies for C4-C12 *p*NP esters are all in the same range of 12-21 s^{-1} mM^{-1}, with insignificant
220 differences of less than a factor of 2 **(Fig. 4c, right panel; Supplementary Table 6)**. The turnover
221 efficiency was considerably less for the shortest substrate used in these experiments (C2) and medium- to
222 long-chain *p*NP esters (> C12), indicative that the catalytic efficiency of medium- to long-chain MAG
223 esters is hampered due to insufficient solubility under the given experimental conditions (**Supplementary**
224 **Table 5**).

225 Finally, we investigated the release of FAs by different *Tth* MAG lipase variants expressed in cells (*E.*
226 *coli* through the analysis of their respective fatty acid methyl esters (FAMEs). When analyzing the FA
227 composition using gas chromatography–mass spectrometry (GC-MS), wt and S113A mutant *Tth* MAG
228 lipase as well as void vector displayed the same composition with mostly hexadecanoic (palmitic) acid
229 (C16:0), cis-Δ11-octadecenoic (cis-vaccenic) acid (C18:1) and smaller contributions of ethylenic
230 (C16:1)/cyclopropanic C17 (9-10 methylene hexadecanoic) and C19 (11-12 methylene octadecanoic)
231 acids **(Supplementary Fig. 11).** Further, we used high performance thin layer chromatography **(Fig. 4f,**
232 **left panel; Supplementary Fig. 11)** to assess the FAMEs released in MAG lipase variants. When
233 comparing the data from wt *Tth* MAG lipase with the S113A negative control variant, wt *Tth* MAG lipase
234 displayed about two-fold higher FA amounts, demonstrating the enzyme's ability to turn over the
235 respective acyl esters under cellular conditions **(Fig. 4f, right panel; Supplementary Table 7).**
236 Interestingly, the *Tth* MAG lipase E43A variant showed a significant two-fold increase of FA turnover
237 with respect to wt *Tth* MAG lipase, while the same *Tth* variant in the *in vitro* MAG turnover assay
238 showed reduced activity when compared with the wt enzyme. In agreement with our *in vitro* data as well,
239 expression of the (Δ140-183) *Tth* MAG lipase variant without the lid domain led to only residual level of
240 FA turnover, similar to the S113A control mutant. Interestingly, the *Tth* MAG lipase Y154R variant also
241 showed a tendency for increased FA turnover.

242 *Tth MAG lipase is a minimal prototype for HBH lid containing lipases*

243 Moving beyond the specific functional properties of *Tth* MAG lipase, we wondered about the enzyme's
244 relations to other previously characterized lipases and esterases. A search against the UniProtKB50
245 protein sequence data base revealed the top scoring 1000 sequence families to be almost exclusively from
246 bacterial and archaeal phyla **(Supplementary Fig. 12a, Supplementary Data 1)**. Amongst these
247 sequences were only two with known 3D structures: cinnamoyl esterase from *Lactobacillus johnsonii*

248 (UNIPROT code D3YEX6 [https://www.uniprot.org/uniprotkb/D3YEX6/entry], PDB code 3S2Z
249 [https://doi.org/10.2210/pdb3S2Z/pdb], plus additional codes of the same sequence), and feruloyl esterase
250 from *Butyrivibrio proteoclasticus* (UNIPROT code D2YW37 [https://www.uniprot.org/uniprotkb/
251 D2YW37/entry], PDB code 2WTM [https://doi.org/10.2210/pdb2WTM/pdb], plus additional codes of the
252 same sequence). Cinnamoyl and feruloyl esters are found as CoA ester intermediates in phenylpropanoids
253 metabolic pathways [37]. Even though these enzymes share a high degree of similarity with *Tth* MAG lipase
254 minor active site sequence differences can explain the inability of *Tth* MAG lipase to hydrolyze
255 phenylpropanoid ester substrates.

256 The relations of cinnamoyl esterase from *Lactobacillus johnsonii* (UNIPROT code D3YEX6
257 [https://www.uniprot.org/uniprotkb/D3YEX6/entry], PDB code 3S2Z
258 [https://doi.org/10.2210/pdb3S2Z/pdb], plus additional codes of the same sequence) and feruloyl esterase
259 from *Butyrivibrio proteoclasticus* (UNIPROT code D2YW37 [https://www.uniprot.org/uniprotkb/
260 D2YW37/entry], PDB code 2WTM [https://doi.org/10.2210/pdb2WTM/pdb], plus additional codes of the
261 same sequence) to the sequence and structure of *Tth* MAG lipase are at a comparable level, as illustrated
262 by rmsd values of 1.2–1.3 Å and structure-based sequence identities of 30-32 % (**Fig. 5, Supplementary**
263 **Fig. 5, Supplementary Table 8**). The lid domain in these three structures comprises a virtually identical
264 two-stranded antiparallel β-hairpin (residues 163-165 and 170-172 in the *Tth* MAG lipase), which is
265 connected by a short hairpin motif of five or less residues (**Fig. 1c, Fig. 5a**). This hairpin shields a highly
266 conserved histidine-glycine-phenylalanine (HGF) motif within the catalytic domain loop connecting β-
267 strand 3 and α-helix B (residues 35-37 in the *Tth* MAG lipase structure) and the small α-helix B'
268 (residues 77-79 in the *Tth* MAG lipase structure) next to the active site (**Fig. 5b, Supplementary Fig. 5**).
269 Helix B' comprises a phenylalanine-serine-glutamate/aspartate (FSE/D) motif that is conserved in the
270 sequences of the related esterases from *L. johnsonii, B. proteoclasticus*, and *T. thermohydrosulfuricus*
271 (*Tth*) but not in other related esterase/lipases **(Supplementary Fig. 5).** Phe37, Phe77 and Phe/Tyr175
272 from the C-terminal lid domain helix α-LC form a highly conserved aromatic residue cluster in these
273 structures and play a crucial role in shaping the substrate-binding sites of the respective enzymes (**Fig. 2a,**
274 **Fig. 6, Supplementary Fig. 5**).

275 Based on these relations it is surprising that *Tth* MAG lipase did not show measurable activity against
276 phenylpropanoid FA ester substrates of the esterases from *L. johnsonii* and *B. proteoclasticus* (**Fig. 5a**)
277 [40,41]. Inspection of their structures revealed a plausible explanation of the distinct substrate specificities:
278 First, in these esterases, there is an additional second small β-hairpin interaction following the first lid
279 helix α-LA, which is not found in any other HBH esterase/lipase lid domains including the structure of
280 *Tth* MAG lipase **(Fig. 5a, Fig. 6).** This hairpin hovers over the central part of the active site of these two

281 esterases to shield bound substrates and thereby further restricts the size of the substrate binding pocket.
282 In the *Tth* MAG lipase this hairpin structure is replaced by a second lid helix α-LB, which is more distal
283 to the enzyme's active site **(Fig. 6a).** Secondly, there are two residues in the *Tth* MAG lipase lid structure,
284 Leu145 from the N-terminal lid helix α-LA and Phe175 from the C-terminal lid helix α-LC that are
285 replaced by residues with functional polar groups, aspartate (Asp138, only in the sequence of the *L.*
286 *johnsonii* esterase sequence) and tyrosine (Tyr169 in the *L. johnsonii* and *B. proteoclasticus* esterase
287 sequences) respectively **(Fig. 6).** Structures of the esterase from *L. johnsonii* in the presence of several
288 ferulate derivatives (3PFB [https://doi.org/10.2210/pdb3PFB /pdb], 3PFC
289 [https://doi.org/10.2210/pdb3PFC/pdb], 3S2Z [https://doi.org/10.2210/pdb3S2Z/pdb]) demonstrate the
290 distinct roles of these residues in substrate binding, and removal of the polar side chain of Asp138
291 strongly impairs substrate turnover [41]. In conclusion, the absence of these functional residue groups in the
292 *Tth* MAG lipase well explains the lack of any measurable activity towards hydroxycinnamic esters.

293 Next, we searched for more distantly related esterase/lipase sequences by exploiting available structural
294 data. The structure-based identities of all 16 selected target sequences (except those from *L. johnsonii* and
295 *B. proteoclasticus* with closer relations) are in the range of 15-23 % and the root mean square deviation
296 values of the superimposed structures are between 1.9 and 2.2 Å, suggesting their relations to be distant
297 with uncertain phylogenetic relations **(Supplementary Fig. 12b, Supplementary Table 8)**. Invariant
298 residue positions are confined to the active site including the catalytic triad and four glycine positions
299 (Gly71, Gly111, Gly115, Gly116), which also contribute to the active site oxyanion hole, another
300 hallmark of α/β hydrolases **(Supplementary Fig. 5)** [24,38].

301 Remarkably, we found an accumulation of esterase/lipase sequences with non-canonical HBH lid
302 domains **(Fig. 5, Supplementary Table 8, Supplementary Fig. 5).** This motif is not present in most
303 previously characterized MAG lipases that comprise lid domains with one or more helices without a β-
304 sheet inserted [22], we define here as canonical. These include human MAG lipase, which has been used as
305 model in various studies [15,39] (**Fig. 5a**). Comparison of these HBH lid structures showed that the terminal
306 helices and adjacent loops function as flexible hinges, even in structures of the same sequence with
307 different ligands **(Fig. 5b)** [22,24]. On the other hand, the hydrogen bond pattern found in the conserved lid
308 β-hairpin provides an additional element of local fold rigidity, which is not found in other more generic
309 lid domains that are formed from α-helices only. While the small antiparallel β-hairpin motif is conserved
310 in HBH lid-containing esterases/lipases, there is substantial divergence with respect to the length of
311 inserted sequences **(Fig. 5)**. Different from the short hairpins found in the *Tth* MAG lipase and the most
312 closely related esterases from *L. johnsonii* and *B. proteoclasticus,* in other more remotely related
313 sequences there are various levels of extensions, allowing to cover extensive surface areas of the CD

active site. The most extreme case is found in the hydrolase-1 domain-containing protein from *B. bassiana,* which contains an insert of 43 residues with a 12-residue helix and an extensive loop structure, allowing to shield the complete active site of this enzyme (PDB code 7D79 [https://doi.org/10.2210/pdb7D79/pdb]) **(Fig. 5a, Supplementary Table 9**).

We finally investigated to what extent esterase/lipases with a HBH lid domain may resemble properties we observed in the structure of *Tth* MAG lipase. Remarkably, none of the structures with a short HBH lid domain **(Fig. 5a),** except for *Tth* MAG lipase, reveals any tunnel-like structure across the lid domain. This is in accordance with data indicating most efficient substrate turnover for esters of small organic compounds with additional polar groups, as well demonstrated for hydroxycinnamic acid family esterases [40,41]. In contrast, HBH lid tunnels are observed in at least two structures with long HBH lid domains the *Paleococcus ferrophilus* MAG lipase (6QE2 [https://doi.org/10.2210/pdb6QE2/pdb]) and the *Bacillus* sp. H257 lipase (4KE8 [https://doi.org/10.2210/pdb4KE8/pdb]/4KE9 [https://doi.org/10.2210/pdb4KE9/pdb]), which both have been annotated as long-chain MAG lipases **(Supplementary Table 8)**. In one of these structures of the MAG lipase from *Bacillus* sp. (4KE9 [https://doi.org/10.2210/pdb4KE9/pdb]), the distal end of a 16-carbon MAG analogue binds through a short tunnel of its HBH lid domain [12], with some resemblance of the HBH lid path found in the structure of *Tth* MAG lipase. As key residues of the *Bacillus* sp. enzyme involved in regulation of MAG lipase turnover (Ile145, Glu156) are positioned on segments of the long HBH lid loop insert, which is missing in *Tth* MAG lipase **(Fig. 5b),** the precise mechanism of MAG lipase turnover regulation appears to be unrelated. A more detailed structural comparison is not possible as there is only very little conservation of residues contributing to long-chain MAG binding.

Taking the data together, HBH lid domains seem to play a crucial role in esterase/lipase function. Supported by our *Tth* MAG lipase structural and functional data, they promote binding of especially long chain MAG substrates. MAG lipase binding and turnover is promoted by a toolbox of specific lid domain features: largely varying levels of extensive substrate-shielding areas as demonstrated by vastly diverting HBH β-hairpin sequence inserts, and the formation of lid domain tunnels to hold especially the distal part of long-chain MAG esters. Ultimately, our structural data on *Tth* MAG lipase demonstrate that the presence of a minimal HBH lid domain is sufficient to allow turnover of long-chain MAG esters. Thus, the structure of *Tth* MAG lipase presents a minimal prototype for a growing number of esterases and lipases with HBH lid domains of different length. Our data illustrate that a profound understanding of MAG lipase catalysis goes well beyond classical enzymology with well-defined active site / substrate relations. Knowledge on enzyme-specific contributions by lid structure and dynamics next and topping the active site are essential for our understanding of substrate specificity and possible future applications

347 on lipase design with further improved turnover properties. Our structure data also reveal the importance
348 on the way how the glycerol backbone of lipase acyl substrates is bound by specific active site residues,
349 ultimately defining its activity profile as MAG lipase, which is not permissive to act as a DAG or TAG
350 lipase. HBH lid lipases with the ability for tunnel formation, as defined in this contribution, may offer
351 distinct opportunities for additional evolvement to recognize promiscuous lipidic substrates of variable
352 length, which may even be extracted directly from lipid membranes [21,39].

353 In conclusion, this study presents a rare case where a high-resolution enzyme structure in the presence of
354 an endogenous long-chain lipid reaction intermediate led to its functional annotation as a MAG lipase.
355 We attribute the opportunity of capturing such complex to the extremophilic properties of its natural host
356 *Tth*, which are sufficiently different from the growth conditions of the target enzyme's heterologous
357 expression host (*E. coli*). Due to the high level of lid diversity and dynamics, homology modeling tools
358 successfully advocated for various other discoveries of enzymatic function [1,2] would have failed, as the
359 closest related enzymes with known structures to *Tth* MAG lipase have specific phenylpropanoid FA
360 esterase activities. Ultimately, this study highlights the strength of high-resolution structural biology to
361 identify and characterize unexplored enzymatic functions under physiological conditions that allowed
362 effective inhibition of substrate turnover. It may inspire future investigations of enzyme function by
363 systematically searching for suitable conditions, under which endogenous substrates could be directly
364 captured.

365 **METHODS**

366 *Protein expression and purification*

367 For X-ray crystallography studies, wt *Tth* MAG lipase was expressed and purified according to a
368 previously established protocol [28]. In this contribution, the protein was expressed using a pETBlue-1
369 (Novagen) vector in *E. coli* (DE3) strain. The purification was performed by liquid chromatography using
370 a Sepharose column followed by Size Exclusion Chromatography (SEC) step. For SEC profile analysis, a
371 Superdex 75 increase 3.2/300 column (Cytiva) was used. The protein was eluted in a 50 mM Tris-HCl pH
372 8.0 buffer containing 170 mM NaCl and the pure fractions were pooled together and concentrated up to
373 20 mg ml^{-1}. For production of the seleno-L-methionine (SeMet)-incorporated *Tth* MAG lipase, the
374 corresponding expression vector was used to transform the methionine auxotroph *E. coli* strain
375 B834(DE3). A preculture grown in LB broth medium was used for the expression of the protein in 2 L
376 M9 minimal medium including trace elements. After a 3-hour starvation step a L-aminoacid mixture was
377 added (40 mg ml^{-1} each except methionine) and SeMet (60 mg ml^{-1}). The culture was induced with 1 mM
378 isopropyl-β-D-thiogalactopyranoside (IPTG) when entering the logarithmic growth phase and was further
379 grown overnight. The SeMet protein was purified using the same protocol as the native protein.

380 To generate different protein variants for *in vitro* assays we sub-cloned the gene encoding *Tth* MAG
381 lipase into a pETM-14 vector (EMBL). For all single residue *Tth* variants used in this study, we
382 performed PCR-based mutagenesis on wt *Tth* MAG lipase. For the *Tth* MAG lipase (Δ140-183) construct
383 two primers were designed, the first one in forward orientation for the C-terminal segment of the catalytic
384 domain expressing residues 183-259 and the second one in reverse orientation for the N-terminal segment
385 of the catalytic domain expressing residues 1-140 (**Supplementary Table 9**). Both primers generated a
386 unique KpnI restriction site. PCR amplification followed by KpnI restriction and ligation yielded the *Tth*
387 MAG lipase (Δ140-183) construct.

388 The genes encoding wt *Tth*, E43A, E43K, E72R, Y154A, and Y154R were cloned into the pETM-14
389 expression vector, and for expression *E. coli* BL21-CodonPlus(DE3)-RIL cells (Agilent Technologies)
390 were used. Cells were grown at 37 °C in lysogeny broth (LB-Lennox) supplemented with kanamycin and
391 chloramphenicol. Cells were induced by addition of 0.5 mM IPTG at an OD_{600} of 0.6-0.8 and harvested
392 after continuous shaking overnight at 18 °C. For purification, the cells were resuspended in lysis buffer
393 (25 mM Tris-HCl (pH 8.0), 300 mM NaCl, 20 mM imidazole, and DNaseI), sonicated, and centrifuged at
394 40.000 g for 30 minutes at 4 °C. The supernatant was filtered and loaded onto Ni-NTA beads, which had
395 been equilibrated in wash buffer (25 mM Tris-HCl (pH 8.0), 300 mM NaCl, 20 mM imidazole). After
396 loading, beads were washed with wash buffer and eluted in elution buffer (25 mM Tris-HCl (pH 8.0), 300
397 mM NaCl, 500 mM imidazole). Each *Tth* variant was concentrated using an Amicon® Ultra concentrator
398 with a 10 kDa cutoff (membrane - regenerated cellulose) and further purified by size exclusion
399 chromatography, using a Superdex 75 10/300 column (GE Healthcare), equilibrated in 25 mM Tris-HCl
400 (pH 8.0), 50 mM NaCl.

401 *Tth MAG lipase reductive methylation*

402 *Tth* MAG lipase, purified protein was concentrated to 5 mg ml^{-1} and for each 1 ml of enzyme solution, 20
403 μl 1 M dimethylaminoborane (DMAB) was added followed by 40 μl 1 M formaldehyde, as previously
404 described [42]. DMAB and formaldehyde were added once more after 2 hours of incubation. After another 2
405 hours, another 10 μl 1 M DMAB was added and incubated for 18 hours. During the entire incubation time
406 the sample was kept in the dark and at 4 °C. To terminate the reaction, DMAB and formaldehyde were
407 removed via a three-step centrifugation buffer exchange to a 50 mM Tris-HCl pH 8.0 and 100 mM NaCl
408 buffer using a filter with a 10 kDa molecular-weight cutoff. Methylated and non-methylated *Tth* MAG
409 lipase were compared by SDS-PAGE (**Supplementary Fig. 13**). The level of methylation of *Tth* MAG
410 lipase was quantified by quadrupole time of flight (Q-TOF) mass spectrometry **(Supplementary Table**
411 **11).**

412 For these experiments, protein samples were acidified using 1% trifluoroacetic acid prior to injection onto
413 the Acquity UPLC System (Waters Corporation). Approximately 2 µg of each sample were loaded onto a
414 protein separation column (Acquity UPLC Protein BEH C4 column, 2.1 mm x 150 mm, 1.7 µm). The
415 outlet of the analytical column was coupled directly to a Q-TOF Premier mass spectrometer
416 (Waters/Micromass) using the standard Electrospray Ionisation source in positive ion mode. Solvent A
417 was water, 0.1% formic acid, and solvent B was acetonitrile, 0.1% formic acid. For Q-TOF experiments,
418 a spray voltage of 3.5 kV was applied with a cone voltage of 35 V and extraction cone at 5 V. The
419 desolvation temperature was set at 350 °C, with source temperature 120 °C. Desolvation gas was nitrogen
420 at a flow rate of 500 L/min. Collision energy was set at 5 eV with argon in the collision cell at a pressure
421 of 4.5·10-5 mbar. Data were acquired in continuum mode over the mass range 500–3500 m/z with a scan
422 time of 0.5 s and an interscan delay of 0.1 s. Data were externally calibrated against an intact protein
423 reference standard (Thermo Fisher), acquired immediately prior to sample data acquisition. Spectra from
424 the chromatogram protein peak were summed and intact mass was calculated using the MaxEnt1
425 maximum entropy algorithm (Waters/Micromass) to give the zero charge deconvoluted molecular weight.

426 Catalytic activity against C8 *p*-nitrophenyl MAG ester using the assay as described in section "*In vitro*
427 assay to determine lipolytic enzyme activity" was close to non-methylated *Tth* MAG lipase (2,3 U mg^{-1},
428 *cf.* **Fig. 4b, left panel, and Supplementary Table 3**).

429 *Tth MAG lipase crystallization*

430 Crystallization experiments using commercial screens for the sitting drop vapor diffusion technique were
431 performed at the EMBL-Hamburg crystallization facility at 20 ˚C [43]. Protein solution at 20 mg ml^{-1} was
432 mixed with the precipitant solution to a final volume of 0.4 nl (0.2 + 0.2 nl). Single rod-shaped crystals of
433 about 100 µm in length grew from precipitant solution containing 5% PEG-6000 in 0.1M Tris-HCl (pH
434 8.0) buffer.

435 We also used a methylated version of the protein at 20 mg ml^{-1} for *Tth* MAG lipase crystallization. Well
436 diffracting crystals with a maximum size of 50 x 50 x 100 µm appeared after 7-10 days using precipitant
437 solution of 2.4 M ammonium sulfate and 0.1 M Tris (pH 7.5) buffer. For the PMS-bound structure, the
438 protein was incubated for 2 hours with 10 mM PMSF followed by the same crystallization protocol.
439 Identical crystallization conditions were used for methylated SeMet-substituted protein.

440 *X-ray data collection, structure solution, and validation*

441 Prior to X-data collection, the methylated *Tth* MAG lipase crystals were immersed in mother liquor
442 containing 30 % (v/v) glycerol and flash cooled in liquid nitrogen. All data sets were collected at 100 K.
443 An anomalous dispersion experiment of a peak (λ=0.97873 Å) and an inflection (λ=0.97903 Å) data-set

using a SeMet-containing methylated *Tth* MAG lipase crystal was carried out on the tunable beamline BM14 (EMBL/ESRF, Grenoble, France). X-ray data of PMS-inhibited *Tth* MAG lipase were also collected on the same beamline (λ=0.95372 Å). The SeMet crystal belongs to the space group $P4_32_12$ while the PMS-bound to the space group $P4_322$, both estimated to contain a protein dimer per asymmetric unit.

All data sets were integrated, scaled and merged using the HKL suite [44]. Selenium atoms were located using the SHELXD program and initial density modification was performed by SHELXE [45]. The correct heavy atom substructure was then used together with the anomalous data from HKL as input for the PHENIX AutoSol wizard, which recalculated the experimental phases and carried out density modification [46]. The density modified phases where then merged with the high resolution data-set from the same SeMet crystal using the program CAD and were then used as an input for model building by ARP/wARP [47,48]. The produced model was further improved by manual corrections and modifications in COOT [49]. ARP/wARP was used again for adding ordered solvent molecules. Additional ligands were added manually after inspecting the electron density maps in COOT. Molrep was used to determine the remaining structures (PMS-bound and native)[50]. Refinement using molecular dynamics and TLS was performed by PHENIX [51]. The data collection and refinement statistics are reported in **Supplementary Table 2**.

Native *Tth* MAG lipase crystals were immersed in crystallization solution containing in addition 15 % MPD as cryoprotectant and then immediately mounted at the beamline's goniometer cryo-cooled at 100 K. X-ray data to 2.42 Å resolution were collected at a wavelength λ=0.9763 Å at the EMBL/DESY beamline BW7A at the DORIS III synchrotron storage ring integrated and scaled as above. We determined initial phases by molecular replacement using the Se-Met substituted and lysine methylated *Tth* MAG lipase coordinates (7Q4J). Structure refinement was carried out as described above with the addition of non-crystallographic symmetry restraints.

The quality of all models was assessed by Ramachandran plots with 97.5 %, 97.2 %, 97.6% of the residues in the most favored positions for the active Se-Met, PMS-inhibited and native structures while no residues were present in disallowed positions (**Supplementary Fig. 6**).

A structure-based sequence alignment was computed by PDBeFold [52]**.** Figures of the 3D structural model were prepared using PYMOL v.2.5.

AlphaFold2 (ColabFold version1.5.2) was used for modeling the *Tth* MAG lipase structure from the sequence, by forcing a dimeric arrangement. The resulting models superimposed onto the *Tth* MAG lipase crystal structures with a root mean square deviation of about 0.45 Å.

Analysis of crystal content

476 About 30 crystals of Se-Met *Tth* MAG lipase were carefully washed with water and dissolved in 200 µl 1
477 M NaCl. A chloroform/methanol mixture of 250/500 µl was then added to the solution. After 10 minutes
478 of intensive vortexing, 250 µl chloroform were added and the solution mixed for 1 minute. Then, 250 µl
479 water were added and the solution mixed again for 1 minute. The mixture was centrifuged for 10 minutes
480 at maximum speed and the upper water phase was carefully discarded. The organic phase was evaporated
481 under vacuum centrifugation at room temperature and resuspended in 20 µl hexafluoroisopropanol. LC-
482 MS measurements injecting 20 µl of the sample in a C18 column and applying a 50-95% acetonitrile
483 gradient in water, revealed a single peak eluted from the column, which in the MS spectrum gave a peak
484 at 359 in positive ionization mode (358+1).

485 *Monoacylglycerol (MAG) ester assay*

486 Enzyme activity towards MAG esters was measured as follows [53]. The protein concentration of the
487 purified enzyme variants was adjusted with sample buffer comprising 50 mM Tris (pH 8.0) and 100 mM
488 NaCl to a suitable concentration range of 0.05 to 0.005 mg ml^{-1} to perform measurements in the linear
489 range of the assay. 1-Octanoyl-*rac*-glycerol and 1-oleoyl-*rac*-glycerol were used as short- and long-chain
490 MAG substrates, respectively. 1,2-Dioctanoyl-*sn*-glycerol and tricaprylin were used as DAG and TAG
491 substrates. All substrates were purchased from Sigma-Aldrich if not stated otherwise. Prior to use, the
492 substrate stocks were diluted with assay buffer comprising 50 mM Tris/HCl (pH 8.0) to a final
493 concentration of 1 mM and supplemented with defatted bovine serum albumin at a molar ratio of 1:1. All
494 values were determined in triplicates and corrected by blank subtraction. The activity assay was
495 performed at 40 °C for 10 minutes followed by addition of 100 μl chloroform, centrifugation (10,000xg
496 for 10 minutes) and determination of the released glycerol in 10 μl of the upper phase with a commercial
497 Glycerol Assay Kit according to the manufacturer's instructions (Sigma-Aldrich).

498 *Mono p-nitrophenyl acyl ester (pNP) assay*

499 Cleavage of *p*-nitrophenyl acyl (*p*NP) esters was determined as in previous studies [28]. The protein
500 concentration of the purified enzyme variants was adjusted with sample buffer (25 mM Tris, pH 8.0 and
501 50 mM NaCl) to a suitable concentration ranging from 0.02 to 0.05 mg ml^{-1} to perform measurements in
502 the linear range of the assay. The substrates were stored as 10 mM stock solutions in 100 % ethanol
503 protected from light at -20 °C. Prior to use, the substrate stocks were diluted with assay buffer (50 mM
504 Tris-HCl, pH 8.0) additionally containing 1 mg ml^{-1} gum arabicum (Sigma Aldrich). All values were
505 determined in quadruplicates and corrected for autohydrolysis using the experimentally determined pH-
506 dependent extinction coefficient of *p*-nitrophenol in microtiter plates (ε = 9.3024 mL μmol^{-1}). The activity
507 assay was performed at 40 °C in a semi-automated high-throughput setup in micro-scale using clear, flat-
508 bottom 96-well plates (Corning). Applying a Freedom EVO® screening robot equipped with a Safire2

509 plate reader (Tecan Group Ltd.), *p*NP substrates were used at a final concentration of 0.9 mM, and
510 continuous measurements were performed at $\lambda = 410$ nm for 10 minutes in a total assay volume of 200 μl
511 per well. Blanks with sample buffer, but without enzyme, were simultaneously measured for each
512 substrate.

513 *Tth* MAG lipase Michaelis-Menten kinetics were performed by quadruple measurements of enzyme
514 activity with *p*NP esters at different substrate concentrations (0.005 mM, 0.0075 mM, 0.01 mM, 0.025
515 mM, 0.05 mM, 0.075 mM, 0.1 mM, 0.25 mM, 0.5 mM, 0.75 mM and 1 mM). The respective activity
516 values were fitted to the Michaelis-Menten equation applying a non-linear regression tool in Prism
517 version 9.1.2. Obtained K_M and v_{max} values of each Michaelis-Menten fit were then used to calculate
518 averages and standard deviations of these basic kinetic parameters. Finally, k_{cat} and K_M yielded the
519 catalytic efficiency **(Supplementary Table 6)**.

520 *Cellular Tth MAG lipase activity by cell extract analysis*

521 Wt *Tth* and Y154A were cloned into the pETBlue-1 vector, transformed into *E. coli* Tuner™ (DE3) pLacI
522 cells and grown at 37 °C in lysogeny broth (LB-Lennox) supplemented with ampicillin and
523 chloramphenicol. Cells were induced by addition of 1 mM IPTG at an OD_{600} of 0.6-0.8 and harvested
524 after continuous shaking at 37 °C for 1 hour. Cell pellets were washed with PBS prior to lipid extraction.

525 For lipid isolation and analysis, total lipid extracts (TLE) were prepared from wet *E. coli* cell pellets by
526 three successive extractions using distinct chloroform/methanol mixtures (1:2, 1:1, and 2:1 v/v). The
527 fractions were pooled, washed with water, and dried under air stream. For chromatography analyses, FAs
528 were methylated using TMS-diazomethane (Sigma), purified on a florisil batch column and analyzed by
529 thin layer chromatography (TLC) (Merck silica Gel 60) with dichloromethane as developing solvent.

530 For lipid semi-quantitative analysis, equivalent masses of total lipid fraction from each strain were spotted
531 onto a high-performance TLC (HPTLC) silica gel 60 plate (Merck) with a Camag ATS4 apparatus. The
532 plates were developed in dichloromethane using a Camag ADC2 device and stained by immersion using a
533 Camag CID3 apparatus with 10 % primuline in acetone/water (84:16). Lipids were revealed and
534 quantified by determining their relative percentage using the Chemidoc Imaging System and ImageLab
535 software (Biorad).

536 To determine FA composition by GC-MS, the FAMEs were purified by preparative thin-layer
537 chromatography (TLC) using a silica gel coated plate G60 (Merck) with dichloromethaneide as
538 developing solvent. Visualization was performed by spraying the plate with rhodamine B, and FAMEs
539 were scraped off the plate and extracted twice with diethyl ether. The purified FAMEs were then analyzed
540 by GC-MS (Thermo Scientific gas chromatograph, fitted with a 15 m x 0.25 mm TG1MS fused silica

541 capillary column and connected to an ISQ single quadrupole mass spectrometer), and identified according
542 to their retention time and their fragmentation pattern [54].

543 *Nano differential scanning fluorimetry*

544 Temperature-induced aggregation curves were measured using Prometheus NT.48 (NanoTemper
545 Technologies GmbH) equipped with a scattering module. Proteins were prepared in 25 mM Tris-HCl (pH
546 8.0), 50 mM NaCl and measured in three technical replicates at a concentration of 1.5 mg ml^{-1}. Excitation
547 power was pre-adjusted to get fluorescence readings above 2000 RFU for F330 and F350, and samples
548 were heated from 15 °C to 95 °C with a slope of 1 °C per minute. The processed data were exported from
549 the manufacturer's software (PR.ThermControl) and analyzed with MoltenProt to obtain aggregation
550 temperatures [55].

551 *Circular dichroism*

552 Far-UV CD spectrums for the *Tth* samples were recorded on a Chirascan Circular Dichroism
553 Spectrometer upgraded with an Active Nitrogen Management System (Applied Photophysics).
554 Measurements were performed at a controlled temperature of 20 °C in a 1 mm QS cuvette, between 280-
555 190 nm using a bandwidth of 1 nm. The *Tth* samples were measured at a concentration of 0.25 mg ml^{-1} in
556 20 mM Tris-HCl (pH 8.0) in three technical replicates. Signal from buffer was subtracted for all measured
557 samples. The estimation of the secondary structure contents was calculated using CONTIN [56].

558 *Statistics and Reproducibility*

559 All enzymatic experiments were carried out at least in triplicate as follows: MAG ester assay: n=3 or 4;
560 *p*NP assay: n=3 or 4; cellular *Tth* MAG lipase activity by cell extract analysis: n=3, 4, 6 or 7; nano
561 differential scanning fluorimetry: n=3; circular dichroism: n=3. Statistical significance of the cell extract
562 analysis data **(Fig. 4f)** was evaluated using two-tailed Student's t-test and p values of less than 0.05 were
563 deemed significant. Statistical analysis was performed using GraphPad Prism 5.04.

564 **DATA AVAILABILITY**

565 The coordinates of the X-ray structures determined in this contribution have been deposited in the Protein
566 Data Bank under accession codes 7Q4J [https://doi.org/10.2210/pdb7Q4J/pdb], 7Q4H
567 [https://doi.org/10.2210/pdb7Q4H/pdb], 8B9S [https://doi.org/10.2210/pdb8B9S/pdb]. All other data
568 generated in this study are provided in the Supplementary Information and a Source Data file.

569 The coordinates of the following X-ray structures used for analyzing the results in this work are also
570 available from the Protein Data Bank with accession codes: 3S2Z [https://doi.org/10.2210/pdb3S2Z/pdb],
571 2WTM [https://doi.org/10.2210/pdb2WTM/pdb], 7D79 [https://doi.org/10.2210/pdb7D79/pdb], 3HJU

572 [https://doi.org/10.2210/pdb3HJU/pdb], 3LLC [https://doi.org/10.2210/pdb3LLC/pdb], 4KE8
573 [https://doi.org/10.2210/pdb4KE8/pdb], 4KE9 [https://doi.org/10.2210/pdb4KE9/pdb], 3PFB
574 [https://doi.org/10.2210/pdb3PFB/pdb], 3PFC [https://doi.org/10.2210/pdb3PFC/pdb], 5CML
575 [https://doi.org/10.2210/pdb5CML/pdb], and 6QE2 [https://doi.org/10.2210/pdb6QE2/pdb]).

576 **REFERENCES**

577 1. Jacobson, M.P., Kalyanaraman, C., Zhao, S. & Tian, B. Leveraging structure for enzyme function
578 prediction: methods, opportunities, and challenges. *Trends Biochem Sci* **39**, 363-71 (2014).

579 2. Mills, C.L., Beuning, P.J. & Ondrechen, M.J. Biochemical functional predictions for protein
580 structures of unknown or uncertain function. *Comput Struct Biotechnol J* **13**, 182-91 (2015).

581 3. Bastard, K. et al. Revealing the hidden functional diversity of an enzyme family. *Nat Chem Biol*
582 **10**, 42-9 (2014).

583 4. Calhoun, S. et al. Prediction of enzymatic pathways by integrative pathway mapping. *Elife* **7**(2018).

584 5. Zhao, S. et al. Discovery of new enzymes and metabolic pathways by using structure and genome
585 context. *Nature* **502**, 698-702 (2013).

586 6. Carter, M.S. et al. Functional assignment of multiple catabolic pathways for D-apiose. *Nat Chem*
587 *Biol* **14**, 696-705 (2018).

588 7. Sévin, D.C., Fuhrer, T., Zamboni, N. & Sauer, U. Nontargeted in vitro metabolomics for high-
589 throughput identification of novel enzymes in Escherichia coli. *Nat Methods* **14**, 187-194 (2017).

590 8. Zhao, S. et al. Prediction and characterization of enzymatic activities guided by sequence similarity
591 and genome neighborhood networks. *Elife* **3**(2014).

592 9. Lee, C.W. et al. Structural basis for the ligand-binding specificity of fatty acid-binding proteins
593 (pFABP4 and pFABP5) in gentoo penguin. *Biochem Biophys Res Commun* **465**, 12-8 (2015).

594 10. Lücke, C., Qiao, Y., van Moerkerk, H.T., Veerkamp, J.H. & Hamilton, J.A. Fatty-acid-binding
595 protein from the flight muscle of Locusta migratoria: evolutionary variations in fatty acid binding.
596 *Biochemistry* **45**, 6296-305 (2006).

597 11. Vasiliauskaité-Brooks, I. et al. Structural insights into adiponectin receptors suggest ceramidase
598 activity. *Nature* **544**, 120-123 (2017).

599 12. Rengachari, S. et al. Conformational plasticity and ligand binding of bacterial monoacylglycerol
600 lipase. *J Biol Chem* **288**, 31093-104 (2013).

601 13. Kim, J. et al. Structure-guided discovery of the metabolite carboxy-SAM that modulates tRNA
602 function. *Nature* **498**, 123-6 (2013).

603 14. Jaeger, K.E., Dijkstra, B.W. & Reetz, M.T. Bacterial biocatalysts: molecular biology, three-
604 dimensional structures, and biotechnological applications of lipases. *Annu Rev Microbiol* **53**, 315-
605 51 (1999).

606 15. Labar, G., Wouters, J. & Lambert, D.M. A review on the monoacylglycerol lipase: at the interface
607 between fat and endocannabinoid signalling. *Curr Med Chem* **17**, 2588-607 (2010).

608 16. Salhi, A., Carriere, F., Grundy, M.M.L. & Aloulou, A. Enzymes Involved in Lipid Digestion. in
609 *Bioaccessibility and Digestibility of Lipids from Food* (eds. Grundy, M.M.L. & Wilde, P.J.) 3-28
610 (Springer International Publishing, Cham, 2021).

611 17. Jaeger, K.E. & Eggert, T. Lipases for biotechnology. *Curr Opin Biotechnol* **13**, 390-7 (2002).

612 18. Borrelli, G.M. & Trono, D. Recombinant Lipases and Phospholipases and Their Use as Biocatalysts
613 for Industrial Applications. *Int J Mol Sci* **16**, 20774-840 (2015).

614 19. Anobom, C.D. et al. From structure to catalysis: recent developments in the biotechnological
615 applications of lipases. *Biomed Res Int* **2014**, 684506 (2014).

616 20. Reis, P., Watzke, H., Leser, M., Holmberg, K. & Miller, R. Interfacial mechanism of lipolysis as
617 self-regulated process. *Biophys Chem* **147**, 93-103 (2010).

618 21. Aloulou, A. et al. Exploring the specific features of interfacial enzymology based on lipase studies.
619 *Biochim Biophys Acta* **1761**, 995-1013 (2006).

620 22. Khan, F.I. et al. The Lid Domain in Lipases: Structural and Functional Determinant of Enzymatic
621 Properties. *Front Bioeng Biotechnol* **5**, 16 (2017).

622 23. Aschauer, P. et al. Crystal structure of the Saccharomyces cerevisiae monoglyceride lipase Yju3p.
623 *Biochim Biophys Acta* **1861**, 462-70 (2016).

624 24. Bauer, T.L., Buchholz, P.C.F. & Pleiss, J. The modular structure of α/β-hydrolases. *Febs j* **287**,
625 1035-1053 (2020).

626 25. López-López, O., Cerdán, M.E. & González Siso, M.I. New extremophilic lipases and esterases
627 from metagenomics. *Curr Protein Pept Sci* **15**, 445-55 (2014).

628 26. Salameh, M.A. & Wiegel, J. Purification and characterization of two highly thermophilic alkaline
629 lipases from Thermosyntropha lipolytica. *Appl Environ Microbiol* **73**, 7725-31 (2007).

630 27. Littlechild, J.A. Improving the 'tool box' for robust industrial enzymes. *J Ind Microbiol Biotechnol*
631 **44**, 711-720 (2017).

632 28. Royter, M. et al. Thermostable lipases from the extreme thermophilic anaerobic bacteria
633 Thermoanaerobacter thermohydrosulfuricus SOL1 and Caldanaerobacter subterraneus subsp.
634 tengcongensis. *Extremophiles* **13**, 769-83 (2009).

635 29. Klingeberg, M., Hippe, H. & Antranikian, G. Production of novel pullulanases at high
636 concentrations by two newly isolated thermophilic clostridia. *FEMS Microbiol Lett* **57**, 145-52
637 (1990).

638 30. Jørgensen, B.B. & Cohen, Y. Solar Lake (Sinai). 5. The sulfur cycle of the bcnthic cyanobacterial
639 mats 1. *Limnology and oceanography* **22**, 657-666 (1977).

640 31. Mirdita, M., Steinegger, M. & Söding, J. MMseqs2 desktop and local web server app for fast,
641 interactive sequence searches. *Bioinformatics* **35**, 2856-2858 (2019).

642 32. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589
643 (2021).

644 33. Holmquist, M. Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr*
645 *Protein Pept Sci* **1**, 209-35 (2000).

646 34. Reis, P., Holmberg, K., Watzke, H., Leser, M.E. & Miller, R. Lipases at interfaces: a review. *Adv*
647 *Colloid Interface Sci* **147-148**, 237-50 (2009).

648 35. Fischer, F., Mutschler, J. & Zufferey, D. Enzyme catalysis with small ionic liquid quantities. *J Ind*
649 *Microbiol Biotechnol* **38**, 477-87 (2011).

650 36. Winkler, U.K. & Stuckmann, M. Glycogen, hyaluronate, and some other polysaccharides greatly
651 enhance the formation of exolipase by Serratia marcescens. *J Bacteriol* **138**, 663-70 (1979).

652 37. Barros, J. et al. Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nat Plants* **2**,
653 16050 (2016).

654 38. Nardini, M. & Dijkstra, B.W. Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr*
655 *Opin Struct Biol* **9**, 732-7 (1999).

656 39. Scalvini, L., Piomelli, D. & Mor, M. Monoglyceride lipase: Structure and inhibitors. *Chem Phys*
657 *Lipids* **197**, 13-24 (2016).

658 40. Goldstone, D.C. et al. Structural and functional characterization of a promiscuous feruloyl esterase
659 (Est1E) from the rumen bacterium Butyrivibrio proteoclasticus. *Proteins* **78**, 1457-69 (2010).

660 41. Lai, K.K. et al. An inserted α/β subdomain shapes the catalytic pocket of Lactobacillus johnsonii
661 cinnamoyl esterase. *PLoS One* **6**, e23269 (2011).

662 42. Rypniewski, W.R., Holden, H.M. & Rayment, I. Structural consequences of reductive methylation
663 of lysine residues in hen egg white lysozyme: an X-ray analysis at 1.8-A resolution. *Biochemistry*
664 **32**, 9851-8 (1993).

665 43. Mueller-Dieckmann, J. The open-access high-throughput crystallization facility at EMBL
666 Hamburg. *Acta Cryst*. **D62**, 1446-1452 (2006).

667 44. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode.
668 *Macromolecular Crystallography, Pt A* **276**, 307-326 (1997).

669 45. Sheldrick, G.M. A short history of SHELX. *Acta Crystallogr A* **64**, 112-22 (2008).

670 46. Terwilliger, T.C. et al. Iterative model building, structure refinement and density modification with
671 the PHENIX AutoBuild wizard. *Acta Crystallographica Section D* **64**, 61-69 (2008).

672 47. Langer, G., Cohen, S.X., Lamzin, V.S. & Perrakis, A. Automated macromolecular model building
673 for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**, 1171-9 (2008).

674 48. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol*
675 *Crystallogr* **67**, 235-42 (2011).

676 49. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D*
677 *Biol Crystallogr* **60**, 2126-32 (2004).

678 50. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr D Biol*
679 *Crystallogr* **66**, 22-5 (2010).

680 51. Afonine, P.V. et al. Towards automated crystallographic structure refinement with phenix.refine.
681 *Acta Crystallogr D Biol Crystallogr* **68**, 352-67 (2012).

682 52. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein
683 structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256-68 (2004).

684 53. Heier, C. et al. Identification of Yju3p as functional orthologue of mammalian monoglyceride
685 lipase in the yeast Saccharomycescerevisiae. *Biochim Biophys Acta* **1801**, 1063-71 (2010).

686 54. Lanéelle, M.A., Spina, L., Nigou, J., Lemassu, A. & Daffé, M. Lipid and Lipoarabinomannan
687 Isolation and Characterization. *Methods Mol Biol* **2314**, 109-150 (2021).

688 55. Kotov, V. et al. In-depth interrogation of protein thermal unfolding data with MoltenProt. *Protein*
689 *Sci* **30**, 201-217 (2021).

690 56. Provencher, S.W. & Glöckner, J. Estimation of globular protein secondary structure from circular
691 dichroism. *Biochemistry* **20**, 33-7 (1981).

692

693 **ACKNOWLEDGEMENTS**

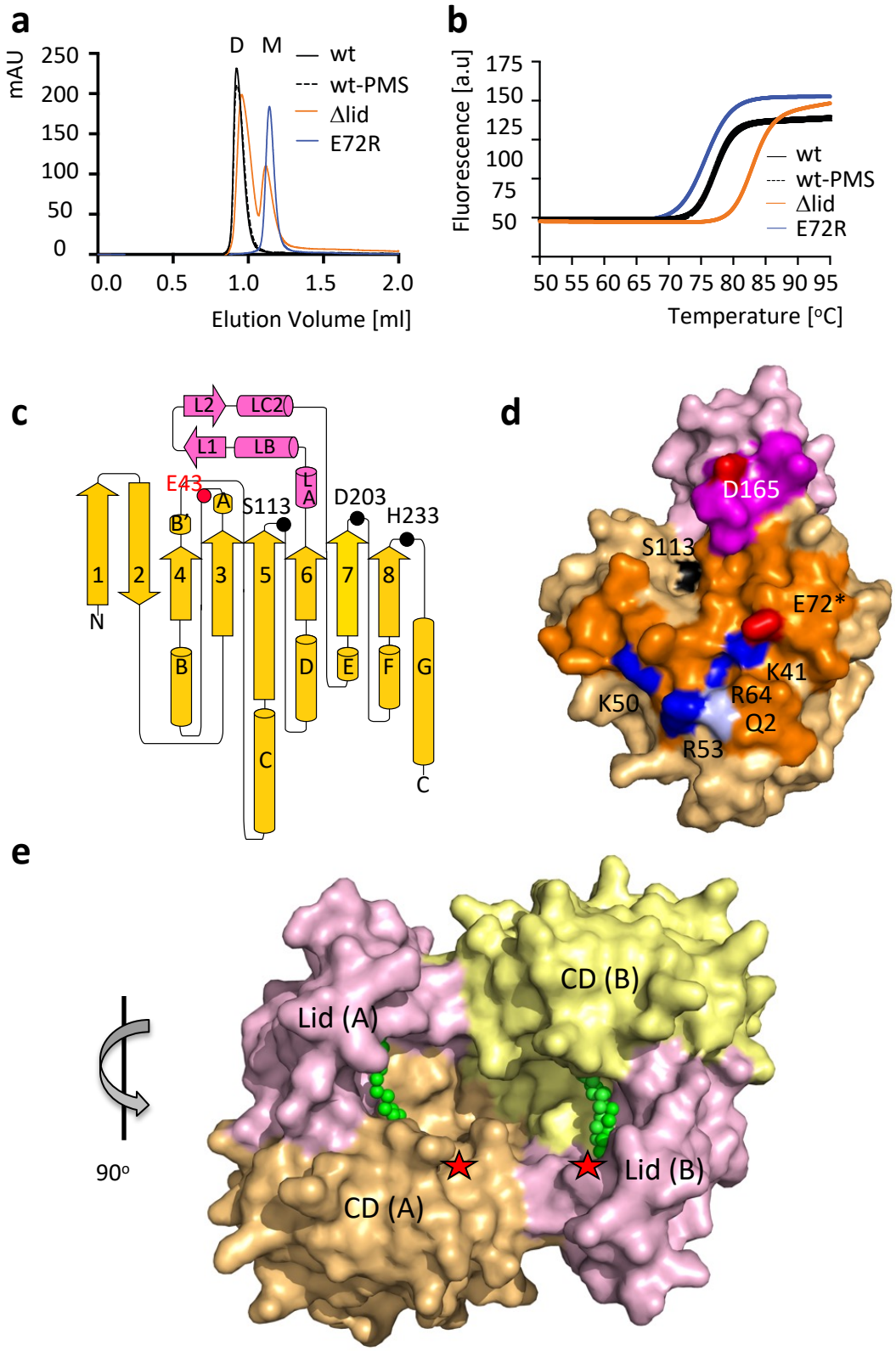
We acknowledge the technical support by the Sample Preparation and Characterization facility of the EMBL Hamburg Unit, the EMBL Proteomics Core Facility and the Proteomics Core Unit of the University of Münster, Germany. We also acknowledge the contributions of Marina Royter during the early phase of the project, as well as the technical contributions of Anke Peters and Barbara Klippel, all from the Hamburg University of Technology, as well as Lucie Spina and Marie-Antoinette Laneelle from the IPBS, University of Toulouse, CNRS, UPS, for contribution in methodology. We are grateful to Dr Martin Walsh at the European Synchrotron Radiation Facility (ESRF), Grenoble, France, for providing assistance in using the beamline BM14 for our X-ray data collection experiments. We thank Sihyun Sung, Grzegorz Chojnowski and Barbara Ramsak from the EMBL Hamburg Unit for technical and scientific support, and Chris Meier from the University of Hamburg for advice during the manuscript revision process.

AUTHOR CONTRIBUTIONS

NP, AK, HM and MW designed the project. NP, AK, NT, SDC and CL performed the experiments. NP, AK, SDC, SAM, HM and MW analyzed the data. NP, AK, SAM, HM and MW wrote the manuscript. MD, AG and MW supported the project.

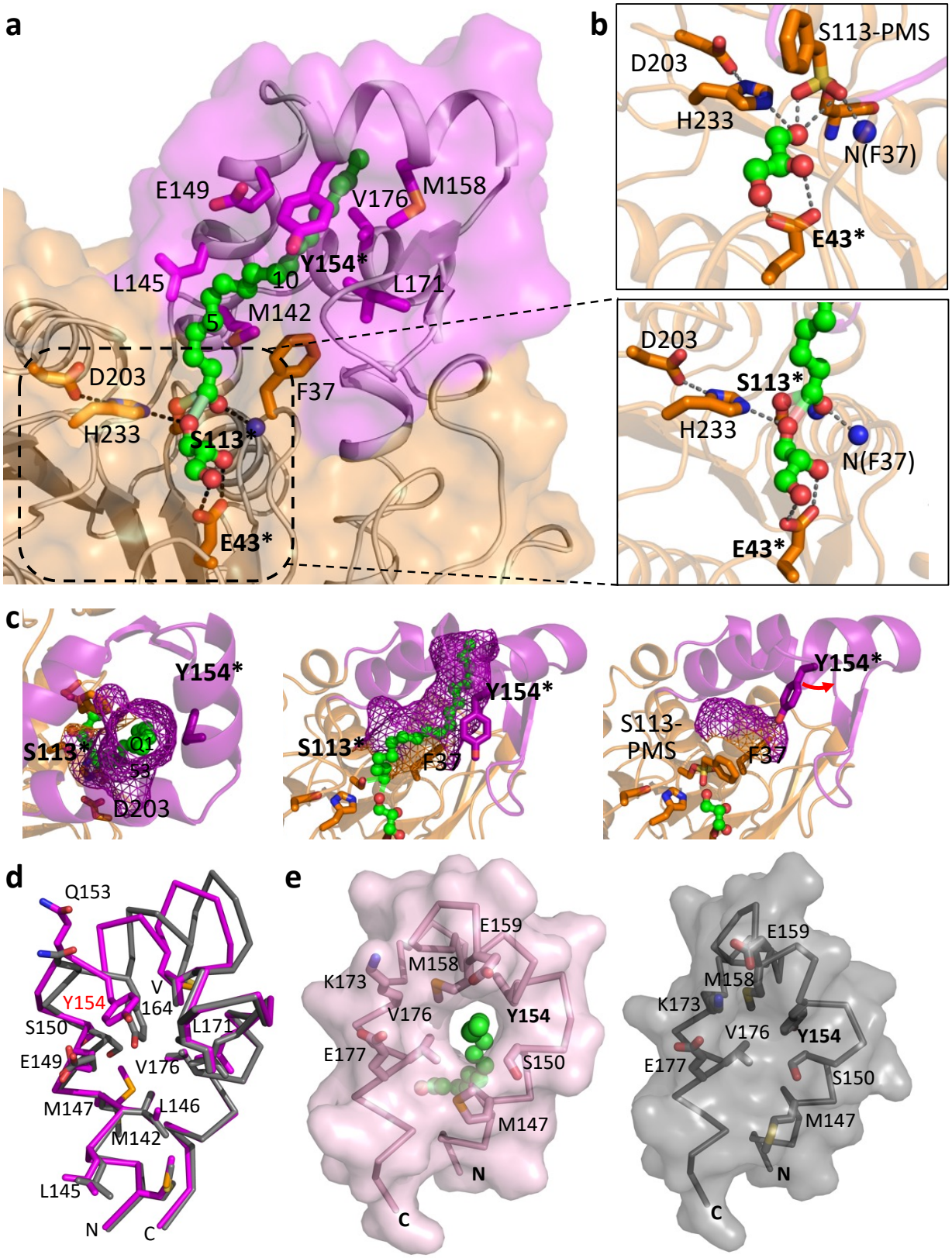
COMPETING INTERESTS

The authors declare no competing interests.


a
mAU
250
200
150
100
50
0
D
M
wt
wt-PMS
Δlid
E72R
0.0
0.5
1.0
1.5
2.0
Elution Volume [ml]
b
Fluorescence [a.u]
175
150
125
100
75
50
wt
wt-PMS
Δlid
E72R
50 55 60 65 70 75 80 85 90 95
Temperature [°C]
c
L2
LC2
L1
LB
E43
A
B'
S113
LA
D203
H233
1
2
4
3
5
6
7
8
N
B
D
E
F
G
C
C
d
D165
S113
E72*
K41
K50
R64
Q2
R53
e
90°
Lid (A)
CD (B)
CD (A)
Lid (B)

711

712 **Fig. 1. Biophysical and structural properties of *Tth* MAG lipase.** a, Size exclusion chromatograms
713 indicating dimeric (D) and monomeric (M) fractions and b, nano differential scanning fluorimetry profiles
714 of selected *Tth* MAG lipase variants. For further analysis, see **Supplementary Table 1**. c, topology
715 diagram of the molecular structure of the *Tth* MAG lipase protomer. CD, yellow; lid domain, pink.
716 Secondary structural elements are in proportion to their length and are labeled. The positions of key active
717 site residues **(*cf.* Supplementary Fig. 5)** are highlighted. The active site glycerol interacting residue
718 Glu43 is shown in red. While in the text three-letter codes were used for residue specifications, for the
719 reason of clarity one-letter codes were used in all Figures. d, surface presentation of *Tth* MAG lipase
720 protomer, with the dimeric interface surface highlighted in stronger colors. Polar residues that contribute
721 to specific interactions to the dimeric interface are shown in red (glutamate, aspartate), blue (arginine,
722 lysine), and pale blue (glutamine), and are labeled. Glu72, which was mutated into an arginine for probing
723 the dimeric interface, is highlighted by an asterisk. The surface of the active site residue Ser113 is shown
724 in black, as point of reference for locating the active site. e, *Tth* MAG lipase dimer in surface
725 presentation. The catalytic domains (CD) and lid domains of each protomer are labeled and colored
726 separately. The two *Tth* MAG lipase protomers form a large joint substrate binding site, as indicated by
727 the bound C18 MAG ligands **(Fig. 2a)** and two asterisks in red, indicating the two active sites. In this
728 presentation, the proximal part of the C18 MAG ligands is visible **(*cf.* Fig. 2)**. The orientation of *Tth*
729 MAG lipase protomer A is rotated by about 90 degrees around a vertical axis with respect to the *Tth*
730 MAG lipase monomer shown in panel d, as indicated. Source data are provided as a Source Data file.


a
E149
V176
M158
Y154*
L145
10
L171
M142
5
D203
F37
H233
S113*
E43*
b
S113-PMS
D203
H233
N(F37)
E43*
D203
S113*
H233
N(F37)
E43*
c
Y154*
S113*
Q1
53
D203
Y154*
S113*
F37
Y154*
S113-
PMS
F37
d
Q153
Y154
V
164
S150
L171
E149
V176
M147
L146
M142
L145
N
C
e
E159
M158
K173
V176
Y154
E177
S150
M147
N
C
E159
M158
K173
V176
Y154
E177
S150
M147
N
C

731

732 **Fig. 2: *Tth* MAG lipase active site and lid domain tunnel for long chain MAG esters.** a, surface
733 presentation of the *Tth* MAG lipase long-chain MAG binding site, in part generated by a tunnel across the
734 lid domain. Bound C18 MAG intermediate is shown in green. Atom-specific colors are used for oxygen
735 (red) and nitrogen (blue). Every fifth methylene group of C18 MAG is numbered. Those parts of the C18
736 MAG that are sequestered by the lid domain tunnel (C11-C18) appear to be darker due to the semi-
737 transparent *Tth* MAG lipase surface. Side chains of key residues contributing to ligand binding and the
738 *Tth* active site are shown and labeled. Residues used for probing function are labeled with an asterisk. The
739 central part of the active site is zoomed in a box to the right, indicating specific polar interactions of the
740 glycerol product with *Tth* MAG lipase Glu43. b, (upper inlet), active site of the PMSF-inhibited *Tth* MAG
741 lipase structure for comparison. c, mesh surface representation of the C18 MAG lid domain tunnel from
742 the top of the lid domain, indicating that it is open at the most distal end (left panel), side view of the lid
743 tunnel (central panel) similar to the orientation in a. Right panel: for comparison, part of the surface
744 generated by residues Phe37 from the CD and Tyr154 from the lid domain is shown for the distinct
745 conformation of the PMSF-inhibited *Tth* MAG lipase structure, due a conformational change of Tyr154,
746 which is labeled in red. This change leads to blockage of the lid tunnel found in the *Tth* MAG lipase
747 structure without PMSF. d, superposition of *Tth* MAG lipase lid domain structures, without (violet) and
748 with PMSF (grey). e, surface presentation of *Tth* MAG lipase lid domain structures, without (violet, left
749 panel) and with PMSF (grey, right panel). Residues involved in the formation of the proximal lid
750 substrate tunnel opening and polar residues forming an outer half-ring wrapping the substrate tunnel exit
751 are labeled. The proximal end of the C18 MAG reaction intermediate bound to the *Tth* MAG lipase
752 structure without PMSF is shown in green. Comparison of the two conformations shows that the lid
753 tunnel is only opened in the *Tth* MAG lipase structure without PMSF, when C18 MAG ester ligand is
754 present. For reasons of optimizing illustration of the lid tunnel the orientation is different from the one in
755 panel b.

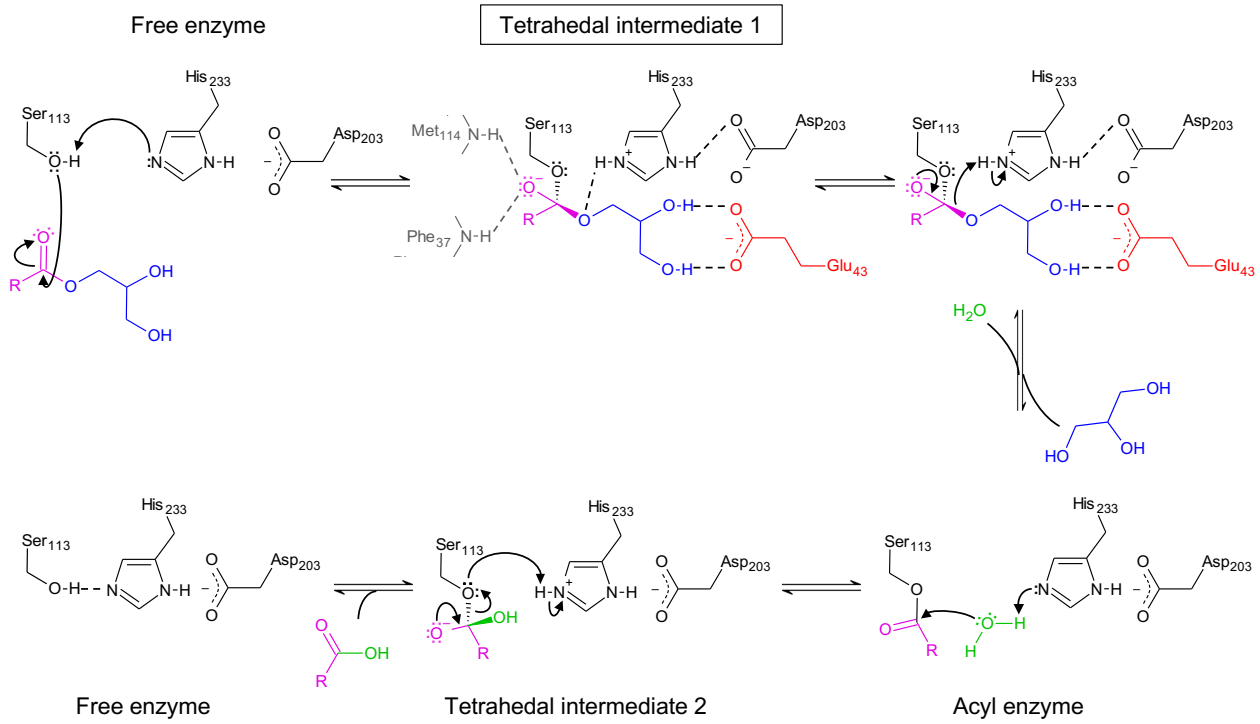


Fig. 3: *Tth* MAG lipase reaction scheme. Reaction intermediates are indicated. Color codes: acyl group, pink; glycerol group, blue; reactive water, green; side chain of Glu43, red. The ligand density found C18 MAG ester reaction intermediate found in the *Tth* MAG lipase structure active site **(Supplementary Fig. 3)** is closest to loosely coordinated tetrahedral transition intermediate 1 (boxed).

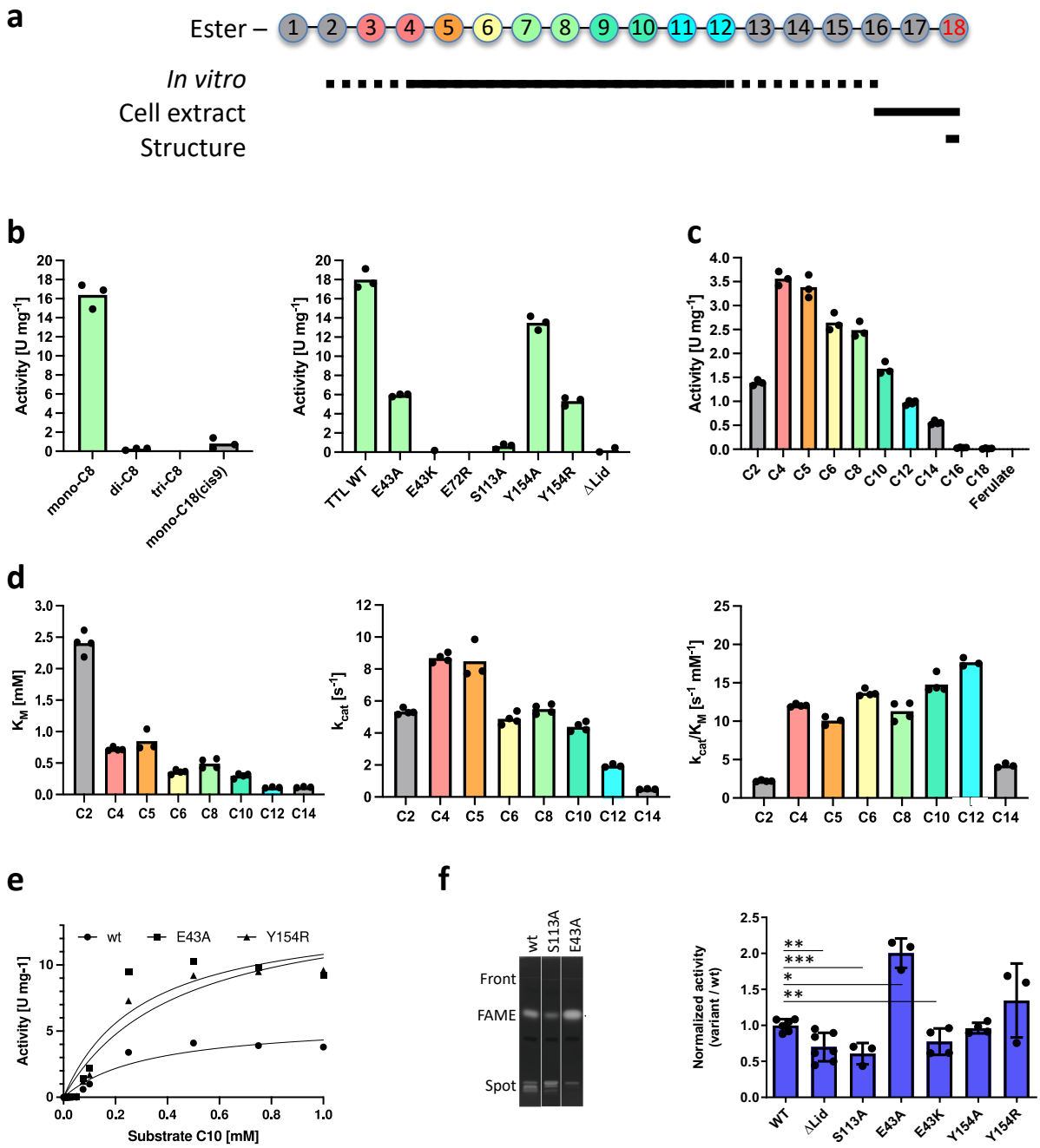


Fig. 4: *Tth* MAG lipase activities for MAG esters. a, *Tth* MAG lipase activities for MAG esters of variable length. Coloring scheme of esters characterized *in vitro* assays (**b-d**). MAG esters found in cell extracts (**f**) and by structural data (**Fig. 2, Supplementary Fig. 3**) are also indicated, demonstrating complementarity of different experimental approaches. b, left panel: *in vitro* turnover of C8 MAG, DAG and TAG esters and C18:1 MAG. b, right panel: C8 MAG turnover of different *Tth* enzyme variants. c, *Tth* MAG lipase activity for several *p*NP acyl esters of C2-C18 FAs and ferulate, indicating a broad range of possible substrates. d, steady-state Michaelis-Menten analysis for MAG esters with quantifiable

activity profiles. Left panel, apparent substrate binding affinities for C2-C14 MAG esters measured by Michaelis constants K_M, indicating increase of binding with increasing MAG ester length; central panel, catalytic rates (k_{cat}) for C2-C14 MAG esters, showing an overall profile comparable to that measured for overall substrate turnover (central panel, *cf*. panel b, left); right panel, catalytic efficiencies k_{cat}/K_M for the same substrate spectrum, showing only minor variations for most of the MAG esters used (C4-C12). e, left panel: experimental steady-state kinetic data of selected *Tth* variants (wt, Y154A, E43A) to estimate Michaelis-Menten values for C6 MAG esters. f, Fatty Acid Methyl Esters (FAMEs) analysis by high-performance thin-layer chromatography (HPTLC) of *E. coli* total lipid extracts (TLE) expressing *Tth* MAG lipase. Left panel, 1, *Tth* MAG lipase (wt); 2, *Tth* MAG lipase (S113A); 3, *Tth* MAG lipase (E43A). Right panel: FAMEs quantification for selected *Tth* MAG lipase mutants, normalized to *Tth* MAG lipase (wt). Data statistics (b-e): Data are presented as mean values and individual data points, $n = 3$-4 (technical replicates). Data statistics (f): Data are presented as means ± SD, and individual data points. Error bars represent the standard deviation from biologically independent replicates (n=3 for variants S113A, E43A, Y154R; n=4 for variants E43K, Y154A; n=6 for WT; and n=7 for ΔLid). Unpaired t-test was used to evaluate statistical significance applying p values for a two-tailed confidence interval as follows: $p < 0.001$ (***), $p < 0.01$ (**), $p < 0.05$ (*). Definition p values**:** p=0.0055 ΔLid, p=<0.0001 S113A, p=0.0288 E43A, p=0.0013 E43K, p=0.4985 Y154A, p=0.1285 Y154R. Source data and uncropped image are provided as a Source Data file.

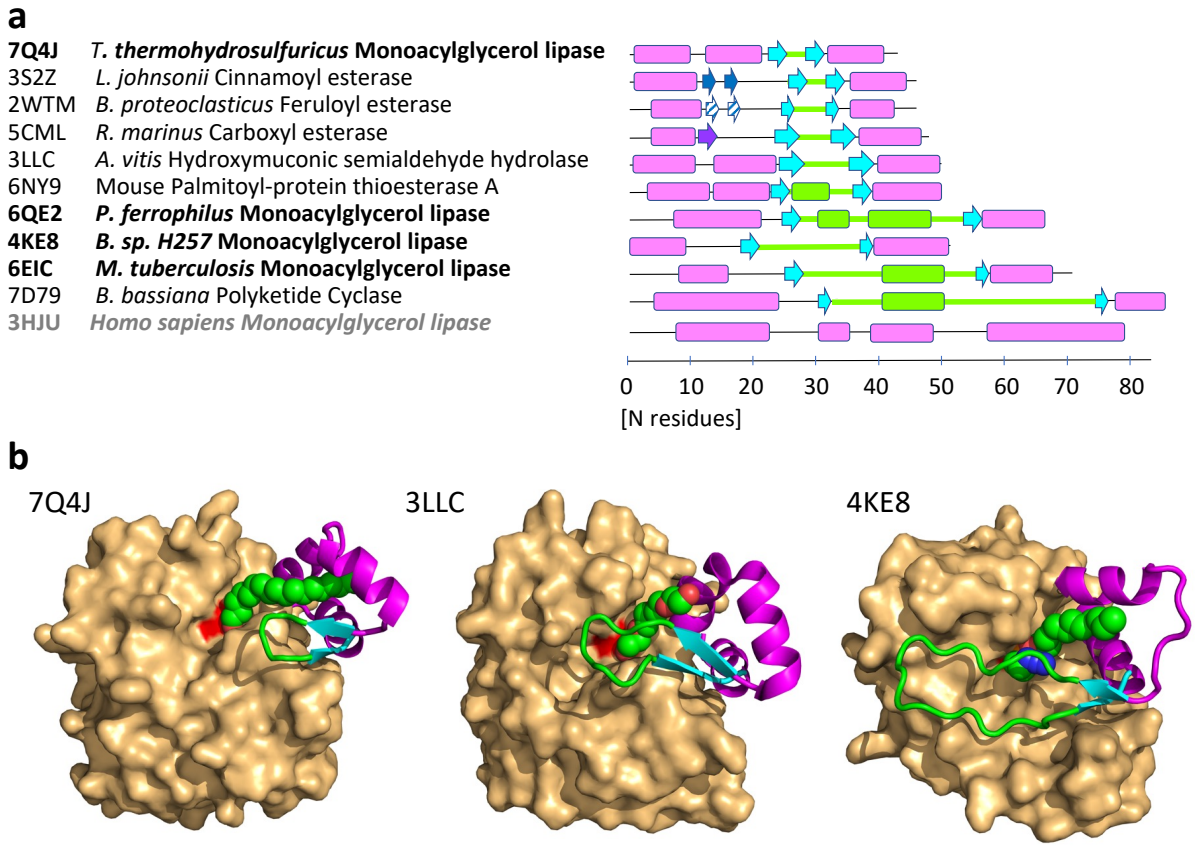


Fig. 5: Sequence, structural and functional relations of esterase/lipases with divergent HBH lid domains. a, Scheme of HBH lid domain sequences related to *Tth* MAG lipase (**Fig. 1c**) with secondary structural elements indicated proportional to sequence length. The two-stranded HBH lid β-hairpins are in cyan, sequence and secondary structural elements inserted between strands βL1 and βL2 are in green. The additional two-stranded β-hairpin structure hovering over the active site in the esterase from *L. johnsonii* (PDB code: 3S2Z [https://doi.org/10.2210/pdb3S2Z/pdb]) is shown in dark blue. The related β-hairpin-like structure in the esterase from *B. proteoclasticus* (PDB code: 2WTM [https://doi.org/10.2210/pdb2WTM/pdb]) is shown in the same color with a diagonal stripe pattern. The lid structure of the Osmotically inducible protein C from *R. marinus* (PDB code: 5CML [https://doi.org/10.2210/pdb5CML/pdb]) comprises an additional third β-strand colored in dark violet, which interacts with the first β-strand of the HBH β-hairpin. Human MAG lipase (3HJU [https://doi.org/10.2210/pdb3HJU/pdb]), which does not have a HBH lid motif and presents a canonical MAG lipase, is shown for comparison as well. Further details are summarized in **Supplementary Table 8.** b, HBH lid structures of three representative esterase/lipases with their PDB codes indicated (**Supplementary Fig. 12b**), demonstrating common elements (N- and C-terminal α-helices, two-stranded β-hairpin) and diversity of the inserted sequences in terms of length and structure. The CDs of these

804 structures are shown by surface presentation in light orange, and the HBH domains are illustrated by
805 cartoon presentation. The surface covered by the catalytic triad serine (Ser113 in *Tth* MAG lipase) is
806 shown in red, as a point of reference to locate the active site. Active site ligands are shown in sphere
807 presentation in atom-specific colors (carbon, green; oxygen, red; nitrogen, blue). Only in *Tth* MAG lipase,
808 the C18 ligand crosses the lid domain via a hydrophobic tunnel. The ligands in 3LLC
809 [https://doi.org/10.2210/pdb3LLC/pdb] (tetraethylene glycol) and 4KE8
810 [https://doi.org/10.2210/pdb4KE8/pdb] (tetradecyl hydrogen (R)-(3-azidopropyl)phosphonate, chain A)
811 are in similar positions next to the active site but they are too short to cross the respective lid domains
812 **(Supplementary Fig. 4a).**

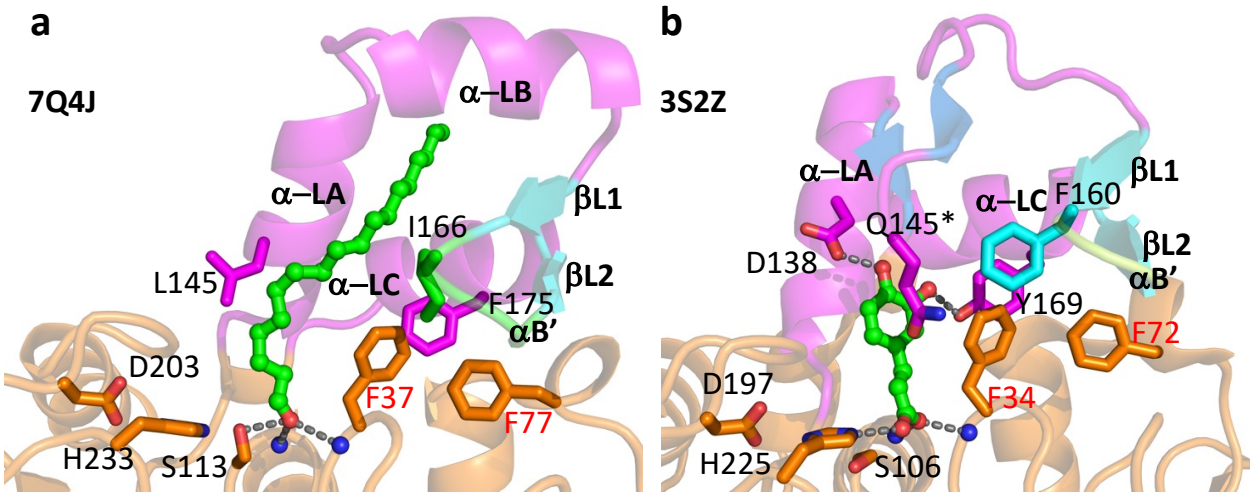


Fig. 6: Structural comparison of the ligand bound HBH lid domains and active site area, indicating different substrate specificities. a, *Tth* MAG lipase in the presence of C18 MAG (left); b, esterase from *L. johnsonii* (PDB code: 3S2Z) in the presence of ferulate (right). Two residues critical for specific hydrogen bond interactions with ferulate in structure of the esterase from *L. johnsonii* (Asp138, Tyr169) are substituted by hydrophobic residues in *Tth* MAG lipase (equivalent residues Leu145, Phe175). Other key residues with critical roles for substrate binding are also shown. In the structure of the esterase from *L. johnsonii*, an additional N-terminal lid β-sheet is formed (colored in blue) that hovers over the CD active site. At the end of the short turn connecting the two β-strands, there is a highly exposed residue, which is Gln145 in this structure. Residues that are part of the conserved HGF motif (residues 35-37 in *Tth* MAG lipase, residues 32-34 in the *L. johnsonii* esterase) and less conserved FSE/D motif (residues 77-79 in *Tth* MAG lipase, residues 70-72 in in the *L. johnsonii* esterase) are labeled in red (**Supplementary Fig. 2).**