

The structural basis of fatty acid elongation by the ELOVL elongases

Elisabeth Carpenter (✉ liz.carpenter@sgc.ox.ac.uk)

University of Oxford <https://orcid.org/0000-0001-9138-2937>

Laiyin Nie

SGC

Ashley Pike

Structural Genomics Consortium <https://orcid.org/0000-0001-9661-2607>

Tomas Pascoa

University of Oxford

Gian Filippo Ruda

University of Oxford

Amy Chu

University of Oxford

Victoria Cole

University of Oxford

Simon Bushell

University of Oxford

Andrew Quigley

University of Oxford

David Speedman

University of Oxford

Tiago Moreira

University of Oxford

Shubhashish Mukhopadhyay

University of Oxford

Nicola Burgess-Brown

Structural Genomics Consortium, University of Oxford

James Love

New York Structural Biology Center

Paul Brennan

University of Oxford

Keywords: Very long chain fatty acids, 3-keto acyl-CoA synthases, fatty acid elongation

Posted Date: November 4th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-96918/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Structural & Molecular Biology on June 10th, 2021. See the published version at <https://doi.org/10.1038/s41594-021-00605-6>.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

The structural basis of fatty acid elongation by the ELOVL elongases

Authors

Laiyin Nie^{1,5,12}, Ashley C. W. Pike^{1,12}, Tomas C. Pascoa^{1,12}, Simon R. Bushell^{1,6}, Andrew Quigley², Gian Filippo Ruda^{1,7}, Amy Chu^{1,8}, Victoria Cole^{1,9}, David Speedman¹, Tiago Moreira¹, Leela Shrestha¹, Shubhashish M.M. Mukhopadhyay¹, Nicola A. Burgess-Brown¹, James D. Love^{3,10}, Paul E. Brennan^{1,4}, Elisabeth P. Carpenter^{1,11*}.

¹ Structural Genomics Consortium, Centre For Medicines Discovery, University of Oxford, Oxford, OX3 7DQ, UK.

² Membrane Protein Laboratory, Research Complex at Harwell, Harwell, UK.

³ Albert Einstein College of Medicine, Department Of Biochemistry, 1300 Morris Park Avenue, Bronx, NY 10461-1602, USA.

⁴ Alzheimer's Research UK Oxford Drug Discovery Institute, Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, UK.

⁵ Present address: CRELUX GmbH, am Klopferspitz 19A, 82152 Planegg-Martinsried, Germany.

⁶ Present address: Orbit Discovery, Schrödinger Building, Oxford OX4 4GE, UK.

⁷ Present address: Evotec Ltd, 114 Innovation Dr, Milton Park, Milton, Abingdon OX14 4RZ, UK.

⁸ Present address: Department of Biochemistry, Oxford University, Oxford OX1 3QU, UK.

⁹ Present address: Genomics Centre, South African Medical Research Council, Francie van Zijl Drive, 7505, Cape Town, South Africa.

¹⁰ Present address: Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark.

¹¹ Present address: Vertex Pharmaceuticals Ltd, Milton Park, Jubilee Drive, Milton Park, Abingdon OX14 4RZ, UK.

27 ¹²These authors contributed equally: Laiyin Nie, Ashley C. W. Pike, Tomas C. Pascoa.

28 *e-mail: liz.carpenter@cmd.ox.ac.uk

29

30 **Abstract**

31 Very long chain fatty acids (VLCFAs) are essential building blocks for synthesis of the
32 ceramides and sphingolipids required for nerve, skin and retina function and 3-keto acyl-CoA
33 synthases (ELOVL elongases) perform the first step in the FA elongation cycle. Although
34 ELOVLs are implicated in common diseases including insulin resistance, hepatic steatosis and
35 Parkinson's, their underlying molecular mechanisms are unknown. Here we report the structure
36 of the human ELOVL7 elongase, which includes an inverted transmembrane barrel structure
37 surrounding a 35 Å long tunnel containing a covalently-attached product analogue. The
38 structure reveals the substrate binding sites in the tunnel and an active site deep in the membrane
39 including the canonical ELOVL HxxHH sequence. This indicates a ping-pong mechanism for
40 catalysis, involving unexpected covalent histidine adducts. The unusual substrate-binding
41 arrangement and chemistry suggest mechanisms for selective ELOVL inhibition, relevant for
42 diseases where VLCFAs accumulate such as X-linked adrenoleukodystrophy.

43

44 **Main**

45 The seven human 3-keto acyl-CoA synthases (elongation of very long chain fatty acids
46 proteins: ELOVL1-7 elongases) catalyse the first, rate-limiting step in the cycle that adds two
47 carbon units to the acyl chains of fatty acids (FAs) with 12 or more carbons per chain (Fig. 1a).
48 These long and very long chain FAs (LCFAs: 12C:20C and VLCFAs: >20C) ^{1,2} are the
49 precursors for synthesis of ceramides, sphingolipids and sphingolipid signalling molecules ³.
50 VLCFAs are essential for the myelin sheaths of nerves ^{4,5}, the skin permeability barrier ^{6,7},
51 retina ⁸ and liver function ^{4,9}. Mutations in ELOVL elongases cause severe genetic diseases

52 including Stargardt syndrome ¹⁰ and spinocerebellar ataxia ^{11,12}. Mouse knockouts suggest
53 ELOVL involvement in hepatic steatosis ¹³ and insulin resistance ¹⁴, and in particular ELOVL7
54 is implicated in cancer ¹⁵⁻¹⁷, early-onset Parkinson's disease ¹⁸ and necroptosis ¹⁹. However,
55 very little is known about the molecular mechanisms underlying this key step in fatty acid and
56 lipid synthesis by the ELOVLs.

57 In order to understand the molecular basis of acyl chain elongation by the ELOVL elongases,
58 we purified and solved the crystal structure of human ELOVL7 in complex with a copurified,
59 covalently bound product analogue and performed intact protein mass spectrometry
60 experiments to probe its catalytic mechanism. The results obtained enabled us to propose a
61 ping-pong type mechanism, achieved through the formation of an unusual acyl-imidazole
62 intermediate upon reaction with the first acyl-CoA substrate.

63

64 **Results**

65 **Structure of human ELOVL7**

66 We have solved the structure of human ELOVL7 by X-ray crystallography to 2.6 Å
67 resolution (see Methods, Table 1 and Extended Data Fig. 1a), revealing an inverted dimer with
68 a small and unconserved interaction surface (870 Å²) in the crystals which was also observed
69 in solution (Extended Data Fig. 1b-f).

70 Overall ELOVL7 has seven transmembrane (TM) helices (TM1-TM7) with TM2-7
71 forming a six TM inverted barrel surrounding a narrow tunnel (Fig. 1b-f), formed by two units
72 of three helices (TM2-4 and TM5-7), each forming an antiparallel three helix arrangement. The
73 two units are assembled as an inverted repeat around the central tunnel (Fig. 1f) with TM1 lying
74 against TM3/4, outside the barrel. This fold is unlike the GPCR 7TM fold and we did not find
75 any six or seven TM protein structures with a similar fold (DALI²⁰) (Extended Data Fig. 2).
76 The 35 Å long central tunnel has a narrow (8-10 Å wide) opening on the cytoplasmic face of

77 the protein and is sealed at its ER lumen end by the short buried loop between TM4-5, which
78 connects the two halves of the barrel (Fig. 1b-e). This loop is separated from the ER lumen by
79 the two disulphide-linked ER loops between TMs2-3 and TMs6-7.

80

81 **ELOVL7 copurifies with a bound acyl-CoA**

82 Surprisingly, the electron density maps clearly showed a long hooked density from the
83 closed ER end of the tunnel to the cytoplasmic surface of the protein. This electron density is
84 consistent with there being a bound acyl-CoA within the active site tunnel spanning its entire
85 length (Fig. 2a,b and Extended Data Fig. 3). Denaturing intact mass spectrometry analysis of
86 protein samples taken during purification showed two species: an apo form with the mass of
87 the purified protein construct (34132.78 Da) and a form with a mass adduct of 1073.66 Da (Fig.
88 2c and Extended Data Fig. 1g,h). This mass adduct was observed both with protein obtained
89 from over-expression in *Spodoptera frugiperda* (*Sf9*) insect cells or mammalian (Expi293F)
90 cells. The form with the mass adduct was more stable during purification, so that following size
91 exclusion chromatography (SEC), it represented >90% of the purified protein (Fig. 2c). Based
92 on the electron density and the observed mass adduct, we have modelled a covalently bound 3-
93 keto eicosanoyl-CoA (Mw 1074.02 Da), an analogue of the reaction product which co-purified
94 with ELOVL7. The electron density clearly shows that the adduct is covalently attached to the
95 imidazole rings of both H150 and H181 (Fig. 2d and Extended Data Fig. 3); the C5 carbon of
96 the elongated 3-keto FA is attached to the N ϵ of H150 and the C2 carbon atom is attached to
97 the N ϵ of H181. When we introduced a H150A mutation we were not able to detect any adducts
98 (Fig. 3a,b). In contrast, with a H181A mutation we detected both unmodified protein and
99 adducts of 1032.51 and 1059.29Da., consistent with covalent attachment of the acyl-CoAs
100 C18:0 -CoA or C20:1 -CoA, respectively (Fig. 3c).

101 ELOVLs catalyse the first step of the LCFA and VLCFA elongation cycle through a
102 condensation reaction between an acyl-CoA and malonyl-CoA to yield a 3-keto acyl-CoA²¹.
103 This product then undergoes a series of reduction and dehydration reactions catalysed by three
104 other enzymes to yield an acyl-CoA with two additional carbons in the acyl chain (Fig. 1a). We
105 have confirmed that detergent-purified ELOVL7 converts the substrates stearoyl-CoA (C18:0)
106 and malonyl-CoA to give the product 3-keto-eicosanoyl (C20)-CoA, by mass spectrometry
107 (Extended Data Fig. 4a,b), indicating that although ~90% of the purified enzyme is covalently
108 modified by the 3-keto acyl-CoA and therefore inhibited, the remaining unmodified protein is
109 active.

110

111 **The substrate binding site lies within an extended tunnel**

112 The covalently bound acyl-CoA adduct seen in the crystal structure conveniently marks
113 out the binding sites for both the acyl and CoA components of the substrates and products (Fig.
114 4), and reveals the architecture of the active site (Fig. 5). The acyl chain binding site is at the
115 upper, occluded end of the tunnel and is terminated by very short TM4-5 loop (Fig. 4a,b). It is
116 lined predominantly by hydrophobic sidechains (W158, G161, G170, F238 and I242), and is
117 curved, allowing both unsaturated and saturated acyl chains to bind (Fig. 4b,c). The acyl chain
118 of the 3-keto acyl-CoA seen in the structure has 20 ordered carbons and it does not occupy the
119 entire pocket, with space for 2 additional carbons, suggesting that 3-keto acyl-CoA products
120 with up to 22 carbons could be accommodated. ELOVL7 has a preference for C18-CoA
121 substrates; it can accept fatty acyl-CoA substrates with carbon chains up to C20 in length^{3,21},
122 consistent with the observed acyl chain length in the structure. The seven ELOVLs have
123 different preferences for acyl chain length and number of double bonds³, and sequence
124 alignments of the human ELOVL elongases suggest that the residues lining the acyl chain

125 binding site are not well conserved (Fig. 4c; Extended Data Fig. 5b), reflecting their abilities to
126 accommodate different acyl chains.

127 The CoA binding site lies at the open, cytoplasmic end in the tunnel, with the CoA
128 almost completely buried within the core of the enzyme (95% of the 1400 Å² ligand surface
129 area is buried). The 3'-phospho-ADP/pantothenate portion of CoA adopts a 'U' shaped
130 conformation wrapped around the sidechain of Y188 (Fig. 4d). The adenine ring is inserted in
131 a narrow cleft between TM2, 3 and 5. N72 on TM2 forms hydrogen bonds to both N5 and N6
132 on the adenine base and D130 on TM3b provides a third hydrogen bond to the base (Fig. 4d).
133 The 3' phosphate projects towards the solvent and forms a salt bridge with R137. The
134 diphosphate forms salt bridges with K204 and R266 and a hydrogen bond with the sidechain of
135 Y187. The pantothenate (Pan) moiety follows a helical path and is hydrogen bonded to the
136 sidechains of T208, Q211 and H147. The cysteamine is elongated and sits in a hydrophobic
137 pocket surrounded by M184 and F253. Together these extensive interactions ensure a highly
138 specific binding site for the CoA portion of the substrates.

139

140 **Fatty acid extension proceeds via a ping-pong type mechanism**

141 Acyl transferase reactions generally involve either a ping-pong or a ternary complex
142 mechanism. In the first, the two substrates bind one after the other and a covalent adduct is
143 involved, whereas in the second both substrates bind at the same time and the acyl group is
144 transferred directly from one substrate to the other. The narrow substrate binding tunnel does
145 not provide sufficient space for both substrates to bind at the same time (Fig. 2b), excluding a
146 ternary complex mechanism. The narrow tunnel is however consistent with a ping-pong type
147 mechanism, where the first product must leave before the second substrate binds (Fig. 5a).
148 Therefore we propose that in the first step the acyl-CoA binds, with the FA tail occluded at the
149 closed ER end of the tunnel and the CoA moiety near the entrance. Then the CoA is cleaved
150 and can dissociate, leaving behind the buried acyl chain. In the second step the malonyl-CoA

151 binds at the more accessible CoA binding site, with the malonyl unit in a side pocket in the
152 central active site. Decarboxylation and carbon-carbon bond formation will then yield the 3-
153 keto acyl-CoA product. The entry and exit route for acyl-CoA and 3-keto acyl-CoA is likely to
154 involve lateral movement of the hydrophobic acyl chain directly into the membrane, between
155 two TM helices, rather than having the acyl chain move along the whole of the tunnel, parts of
156 which are hydrophilic. The structure suggests that this entry/exit portal lies between TM4 and
157 TM7, between the two 3TM halves of the barrel (Fig. 1f).

158 The catalytic site of ELOVL elongases lies at the centre of the membrane and the
159 residues lining this section of the tunnel are highly conserved across all seven elongases (Fig.
160 4c). A site directed mutagenesis study in the ELOVL homologue EloA from the social amoeba
161 *Dictyostelium discoideum* identified a series of residues that were essential for activity
162 (equivalent to K124, D130, H150, H151, N177, H181 and Q211 in ELOVL7)²² and all of these
163 residues lie within the active site. The active site is lined with histidine residues (Fig. 5b,c),
164 including the canonical ELOVL elongase HxxHH motif²¹. Although this type of histidine motif
165 is sometimes associated with metal ion binding, for example in fatty acyl desaturases²³, the
166 electron density maps show no indication of bound metal ions and the geometry is not suitable
167 for metal ion binding, indicating that metal-assisted catalysis is unlikely. Although cysteines
168 are often involved in acyl transferase reactions, there are no cysteine residues in or near the
169 active site that could be involved in catalysis. The structure shows the CoA thiol sits just below
170 a triad of histidines – H150 and H151 on TM4 from the conserved HxxHH motif and H181 on
171 TM5, all conserved residues required for ELOVL activity²² (Fig. 5b). In the crystal structure
172 H150 and H181 are covalently attached to the copurified product analogue, and mutation of
173 these residues affects the *in vivo* modification of heterologously expressed ELOVL7. The
174 H150A mutation gave only unmodified protein, suggesting that H150 could be the first residue
175 to be modified, whereas the H181A mutation resulted in modification by substrate analogues,
176 suggesting that H181 could be involved in the subsequent catalytic steps (Fig. 3b,c).

177

178 **ELOVL7 catalysis proceeds via an acyl-imidazole intermediate involving**

179 **H150**

180 Using the ELOVL7 structure with the bound product analogue we modeled the positions
181 of each substrate and product during the reaction, indicating how the reaction could proceed
182 (Fig. 6a-c). Once the first substrate, the fatty acyl-CoA has bound, then H150 is ideally placed
183 to act as a nucleophile to attack the carbonyl group of the fatty acyl-CoA. The nucleophilicity
184 of H150 is enhanced by a bonding network involving H151 and D130 that ensures that the N ϵ
185 atom is unprotonated (Fig. 6a). The sidechains of N177 and H181, together with a water
186 molecule (which hydrogen bonds to N177 and Q214), are suitably positioned to stabilise the
187 oxyanion prior to release of CoA. This transient tetrahedral intermediate would then collapse
188 back to the carbonyl, ejecting the CoA, which could acquire a hydrogen atom from H181,
189 providing protonation of the CoA thiol and completing the transacylation step of the reaction
190 (Fig. 6b). The formation of a covalent acyl-enzyme intermediate at the end of the first step was
191 confirmed by denaturing intact mass spectrometry analysis of the enzyme after incubation with
192 an acyl-CoA substrate (C18:0-CoA). We observed a 266.51 Da adduct (Fig. 6d and Extended
193 Data Fig. 6a,b), consistent with the transfer of a C18:0-acyl group to the protein. However,
194 incubation with both substrates, C18:0-CoA and malonyl-CoA, greatly reduced accumulation
195 of the adduct intermediate, consistent with the reaction having gone to completion (Fig. 6d and
196 Extended Data Fig. 6c). Similarly, incubation with the unsaturated C18:3(n3)-CoA substrate
197 alone resulted in the formation of a 260.74 Da adduct, whereas incubation with C18:3(n3)-CoA
198 and malonyl-CoA did not give an adduct, consistent with covalent addition of the unsaturated
199 acyl chain of the substrate as the first step, followed by completion of the reaction with the
200 second substrate (Extended Data Fig. 6f,g). Addition of EDTA or EGTA did not interfere with

201 covalent modification of the protein by either substrate, supporting the hypothesis that metal
202 ions are not required for catalysis (Extended Data Fig. 6d-g).

203 Many fatty acid synthases and other acyltransferases require the formation of acyl-
204 enzyme intermediates with either oxy- or thio-ester covalent linkages²⁴, usually involving
205 cysteine or serine²⁵ residues. However, there are no cysteines or serines in or near the ELOVL
206 active site. The structure indicated that H150 was a candidate for this role and a H150A
207 mutation in ELOVL7 gave protein that was not modified upon incubation with C18:3(n3)-CoA
208 (Extended Data Fig. 7a,b) (we were unable to use C18:0-CoA as substrate due to the presence
209 of background peaks in the MS traces). In contrast, incubation of the H181A protein with
210 C18:3(n3)-CoA gave a covalent adduct of 261.97 Da (Extended Data Fig. 7c,d). These results
211 are consistent with the H150 sidechain being the site where the fatty acyl unit is covalently
212 attached. Covalent acyl adducts to histidines are uncommon, but they are not entirely without
213 precedent: HlyC, the enzyme that activates the *Escherichia coli* toxin prohemolysin, catalyses
214 acyl transfer via a histidine^{24,26} and the enzymes of the complement system, C3 and C4, utilise
215 a histidine for internal acyl transfer^{27,28}. In addition artificial enzyme systems have been created
216 in which a nucleophilic histidine is acylated to form a covalent acyl-imidazole intermediate,
217 confirming that such additions are possible²⁹⁻³¹.

218 In the second ‘condensation’ step, malonyl-CoA is combined with the bound acyl chain
219 through a Claisen condensation. This reaction requires malonyl-CoA binding, decarboxylation
220 to form an enolate, followed by reaction with the acyl-enzyme intermediate to form the 3-keto
221 acyl-CoA product that is elongated by 2 carbon units. Initially malonyl-CoA would bind to the
222 CoA binding site, and modelling suggests that the malonyl unit would be positioned so that the
223 carboxylate would lie in a pocket at the side of the tunnel between TM5 and TM6 (Fig. 5c).
224 This shallow pocket is lined by a semicircle of amide groups from N177, Q211 and Q214,
225 which would hydrogen bond to the carboxylate and, together with H181, could catalyse
226 decarboxylation (Fig. 6c and Extended Data Fig. 8). Modelling suggests that His181 lies

227 adjacent to the thioester oxo group of malonyl-CoA, where it could promote decarboxylation
228 by stabilizing enolate formation. Mutation of these residues in EloA gave either inactive
229 enzyme (H181, N177 and Q211 (ELOVL7 numbering)), or enzyme with 10% of WT activity
230 (Q214). The resulting nucleophilic enolate carbanion would then be well-placed to react with
231 the acyl-CoA adduct on H150, to form the 3-keto acyl-CoA product (Extended Data Fig. 8)
232 which would then be released from the enzyme.

233 Interestingly, the crosslinked version of the product analogue seen in the crystal
234 structure is covalently linked to the C5 of the 3-keto acyl product analogue inhibitor, not directly
235 the C3-carbonyl. We speculate that this crosslink could have been the result of ELOVL7's
236 conjugate addition to an unsaturated acyl-CoA substrate or breakdown product, with an
237 adjacent double bond in the acyl chain (e.g. 2,3-*trans*-enoyl-CoA), thus allowing the formation
238 of a displaced acyl-imidazole covalent bond. Acyl-CoAs with a double bond in this location are
239 formed both during the FA elongation cycle and during breakdown of VLCFAs in the
240 peroxisomes.

241

242 **Discussion**

243 The structure of ELOVL7 reveals a narrow extended tunnel which is involved in
244 binding and extension of acyl-CoAs. Previous studies have highlighted a series of histidine
245 residues that are essential for ELOVL enzyme activity and these are clustered halfway along
246 the membrane-embedded tunnel, at the midpoint of the bilayer, where they are poised for
247 catalysis. Through a combination of mutagenesis and intact protein MS analysis, our results
248 show that the elongation reaction proceeds through a stable acyl-imidazole intermediate formed
249 between the substrate acyl-CoA and the second His in the ELOVL HxxHH motif. We propose
250 that a second group of sidechains previously demonstrated to be required for elongase activity
251 (N177, H181, Q211, and Q214)²² lie adjacent to the modified histidine, where they could play
252 a role in the decarboxylation of malonyl-CoA, the second substrate. The resulting reactive

253 enolate nucleophile could then react with the acyl-enzyme intermediate giving rise to the 3-keto
254 acyl-CoA product.

255 Failure of the VLCFA elongation process through ELOVL mutations leads to loss of
256 ceramides and sphingolipids required for myelin, skin barrier and retinal function. Mutations in
257 the ELOVL4 gene cause Stargardt disease-3¹⁰, ichthyosis, intellectual disability, spastic
258 quadriplegia³², and spinocerebellar ataxia 34 (SCA34)^{11,33}. ELOVL5 mutations could cause
259 spinocerebellar ataxia 38 (SCA38)¹². *Elovl3* knockout mice suffer severely from hair loss and
260 have an imbalance in the lipid species of the sebum^{34,35}; and mouse knockout studies of
261 ELOVL5 / ELOVL6 suggest associations with hepatic steatosis¹³ and obesity-induced insulin
262 resistance¹⁴. ELOVL7, the most recently discovered ELOVL elongase, is associated with
263 prostate^{15,16} and gynaecological¹⁷ cancer and early onset Parkinson's disease¹⁸. ELOVL7
264 knockdown reduces cell death and membrane permeabilization during necroptosis, a form of
265 programmed cell death¹⁹. The unusual architecture, active site and chemistry revealed by our
266 ELOVL elongase structure provides exciting new avenues for design of modulators of
267 ELOVLs. These may be of value for patients with X-linked adrenoleukodystrophy (X-ALD),
268 as they are unable to break down VLCFAs, leading to accumulation in cells and further
269 elongation by ELOVL1³⁶. X-ALD affects the adrenal cortex and nervous system, leading to
270 adrenocortical insufficiency and myelopathy, with progressive demyelination of nerves in
271 severe cases. The formulation known as Lorenzo's oil has been shown to alleviate symptoms
272 of X-ALD and is reported to inhibit ELOVL1³⁷. The design of specific ELOVL elongase
273 inhibitors may offer a route to reduction of VLCFAs, thus providing a novel approach to therapy
274 for this devastating disease.

275

276

277

278

279

280 **Methods**

281 **Cloning**

282 The *Homo sapiens ELOVL7* gene, encoding the ELOVL7 protein (Uniprot ID:
283 A1L3X0) was cloned into the baculovirus transfer vector pFB-CT10HF-LIC (available from
284 The Addgene Nonprofit Plasmid Repository) for expression in *Sf9* cells (Thermo-Fisher
285 Scientific, 11496015) using the primers shown in Extended Data Table 1. A C-terminal TEV-
286 cleavable His10-FLAG tag was fused to the protein for purification. For mammalian
287 expression, the same construct was cloned into the pHTBV1.1-LIC baculovirus transfer vector
288 (The BacMam vector backbone (pHTBV1.1), kindly provided by Professor Frederick Boyce,
289 Massachusetts General Hospital, Cambridge, MA and adapted for ligation independent cloning
290 in-house for expression in Expi293F cells (Thermo-Fisher Scientific) which similarly confers
291 a TEV cleavable C-terminal His10-FLAG tag.

292 The mutagenesis was performed using Q5[®] Site-Directed Mutagenesis Kit (NEB)
293 according to manufacturer's instructions using primers listed in Extended Data Table 1.

294 **Protein Production**

295 For both insect and mammalian expression, baculoviral DNA was produced by
296 transposition of DH10Bac with either the ELOVL7-pFB-CT10HF-LIC or ELOVL7-
297 pHTBV1.1-LIC transfer vectors, which were then used to transfect *Sf9* cells to produce
298 baculovirus for transduction using the transfection reagent Insect GeneJuice[®]
299 (MerckMillipore). The virus amplification was performed by infecting mid-log *Sf9* cells (2×10^6
300 cells ml⁻¹) grown in *Sf*-900IITM media with 2% fetal bovine serum. For large scale protein
301 production, *Sf9* cells at the density of 2×10^6 cell ml⁻¹ were infected with 5 ml of P2 (second

302 passage) recombinant baculovirus in *Sf*-900™ II in 2 L roller bottles (Biofil) and incubated for
303 72 h at 27°C. Cells were harvested by centrifugation at 900 x g for 15 mins, washed with
304 phosphate buffered saline (PBS), and pelleted again prior to flash freezing in liquid N₂, then
305 stored at -80 °C for further use.

306 For mammalian (Expi293F) expression, 1 L of Expi293F cell cultures (2 x 10⁶ cells ml⁻¹)
307 in Freestyle 293™ Expression Medium (Thermo-Fisher) were transduced with 30 ml of P3
308 baculovirus (third passage) in the presence of 5 mM sodium butyrate in a 2 L roller bottle
309 (Biofil). Cells were grown in a humidity controlled orbital shaker for 48 hours at 37 °C with
310 8% CO₂ before being harvested using the same process as for insect cells.

311 **Protein Purification**

312 All the following steps were performed at 4°C unless otherwise indicated. Cell pellets
313 were resuspended in lysis buffer (50 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 5% v/v
314 glycerol, 1 mM TCEP-NaOH, Roche protease inhibitor cocktail EDTA-free) at the ratio of 50
315 ml / L equiv. original cell culture. The resuspension was then passed twice through an
316 EmulsiFlex-C5 homogenizer (Avestin) at 10000 psi. Membrane proteins were extracted from
317 the cell lysate with 1% w/v octyl glucose neopentyl glycol (OGNG; Generon, Cat. No. NG311)
318 / 0.1% cholesteryl hemisuccinate tris salt (CHS; Sigma-Aldrich, Cat. No. C6512) and rotated
319 for 2 h. Cell debris was removed by centrifugation at 35,000 x g for 1 h. The supernatant was
320 supplemented with 5 mM imidazole pH 8.0 before incubation with Co²⁺ charged TALON resin
321 (Clontech) for 1 h on a rotator (1 ml resin slurry per L original culture volume). The Talon resin
322 was collected by centrifugation at 700 x g for 5 mins and washed with 30 column volumes of
323 washing buffer (50 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 1 mM TCEP-NaOH, 0.12%
324 w/v OGNG / 0.012% w/v CHS and 20 mM imidazole pH 8.0) before the target protein was
325 eluted with elution buffer (washing buffer supplemented with 250 mM imidazole pH 8.0). The

326 eluted protein was then exchanged into lysis buffer supplemented with 0.15% w/v OGNG /
327 0.015% w/v CHS by passing over a pre-equilibrated Sephadex PD-10 desalting column (GE
328 Life Sciences). TEV protease was added to the desalted protein at a weight ratio of 1:10 for
329 overnight tag cleavage. The His-tagged TEV protease was removed by a second Talon resin
330 binding step for 1 h, and the flow through was collected and concentrated to <1ml using a 100
331 kDa molecular weight cutoff (MWCO) Vivaspin 20 centrifugal concentrator (GE Life
332 Sciences) and further purified by size exclusion chromatography on a Superdex 200 10/300
333 Increase GL column (GE Healthcare) in size exclusion chromatography (SEC) buffer (20 mM
334 HEPES-NaOH, pH 7.5, 200 mM NaCl, 1 mM TCEP-NaOH, 0.08% w/v OGNG/ 0.008% w/v
335 CHS). The H150A and H181A mutated protein were expressed and purified in the same way
336 as the wild type protein. In addition we attempted to produce a His151Ala version of the protein,
337 but this variant proved to be too unstable to purify.

338 For structural studies, the flow through from the reverse Talon step was incubated with
339 50 mM iodoacetamide (IAM) (Merck Millipore) for 20 mins at room temperature. IAM was
340 removed by passing the reaction mixture down a PD-10 desalting column prior to concentration.
341 After SEC, fractions containing the highest concentration of ELOVL7 were pooled and
342 concentrated to 12-25 mg/ml using a 100 kDa MWCO Vivaspin 20 centrifugal concentrator.

343 **Crystallization**

344 Initial protein crystals were grown at 4°C in condition E10 of the MemGold2-ECO
345 Screen (Molecular Dimensions; 0.05 M Na-acetate pH 4.5, 0.23 M NaCl, 33 % v/v polyethylene
346 glycol (PEG) 400) in 3-well sitting-drop crystallisation plates (SwissCi) with 150 nl drops and
347 2:1 and 1:1 protein to reservoir ratios. Crystals appeared after 4-7 days and grew to full size
348 within 3-4 weeks. Two microlitre hanging drops were set up in 24-well XRL plates (Molecular
349 Dimensions) at protein to reservoir ratios of 2.5:1, 2:1 and 1.5:1. The best crystals grew in 0.1

350 M Na-acetate pH 4.5, 0.23 M NaCl, 34-38% v/v PEG400 using the IAM-modified protein at a
351 concentration of 5-8 mg/ml and were harvested after 12-14 days. Crystals could also be grown
352 for the unmodified native protein but were limited in size and poorly reproducible despite using
353 seeding. Prior to vitrification, crystals were sequentially transferred to mother liquor solutions
354 with an increasing amount of PEG400 to a final concentration of PEG400 of 46% v/v over 10-
355 15 mins. For heavy atom derivatisation, crystals were looped into drops containing reservoir
356 solution supplemented with 10 mM mercury chloride and soaked for 10 mins. Hg-soaked
357 crystals were then treated to the same PEG400 escalation strategy using Hg-free solutions
358 before being vitrified in liquid nitrogen.

359 **Data Collection**

360 All crystals were screened and X-ray diffraction data were collected on the I24
361 microfocus beamline at the Diamond Light Source (Didcot, UK) from vitrified crystals at 100
362 K. Multiple data sweeps were collected from crystals using 0.2° oscillation and a beam-size of
363 20 µm x 20 µm. Crystals were translated by at least 25 µm between collection areas. Native
364 data were collected at a wavelength of 0.9686 Å. The high resolution native dataset was
365 assembled from four overlapping 100° wedges of data collected from separate volumes of a
366 single crystal. 960° of anomalous data were also collected from a mercury-derivatised crystal
367 at a wavelength close to the Hg L_{III} edge ($\lambda = 0.992$ Å) in a single pass.

368 Unmerged scaled data from the *xia2.dials* automated processing pipeline³⁸ was
369 truncated and merged using AIMLESS (CCP4^{39,40}). ELOVL7 crystals are monoclinic and
370 contain two copies of the enzyme in each asymmetric unit. All diffraction data were highly
371 anisotropic and limited to between 3.4 - 4.5 Å resolution in the worse direction and 2.05 – 3 Å
372 in the best direction (Table 1). Data truncated in AIMLESS with isotropic resolution limits was
373 used for phasing and initial cycles of model building. The high resolution native dataset

374 (nominal overall resolution 2.6 Å) was further processed with STARANISO (v1.0.4⁴¹) to the
375 maximum resolution of 2.05 Å and used for refinement (see Table 1 for details).

376 **Phasing**

377 Phasing was carried out in PHENIX⁴² using SIRAS with the Hg-peak data and a 3Å
378 isomorphous lower resolution native dataset. Two Hg²⁺ sites were located with *phenix.hyss*
379 using data to 4.5 Å. The resulting 3 Å phased electron density map had clear protein density
380 allowing the identification of the NCS relationship between the two ELOVL7 molecules in the
381 asymmetric unit. After two-fold averaging using RESOLVE⁴³, the resultant map was of
382 sufficient quality to manually model all the TM helices. Initial phases were further improved
383 by cross-crystal averaging with a non-isomorphous, less anisotropic and slightly higher
384 resolution dataset (unit cell dimensions 64.14 Å x 71.93 Å x 111.74 Å $\beta = 106.7^\circ$) using
385 DMMULTI⁴⁴. The resultant map was of excellent quality and the majority of the structure could
386 be built automatically with BUCCANEER⁴⁵.

387 **Model Building and Refinement**

388 The BUCCANEER-built model was manually rebuilt in COOT⁴⁶ and refined using
389 BUSTER (version 2.10.3, 29-NOV-2019⁴⁷). The model was refined with LSSR restraints⁴⁸ and
390 a single TLS group was refined for each protein chain. Ligand dictionaries were generated using
391 GRADE (Global Phasing Ltd). All data to 2.05 Å from the STARANISO processing was used
392 in refinement despite the low completeness and truncation to lower resolutions resulted in a
393 refined model with worse geometry. The final model comprises residues 16-269 (chain B, 14-
394 269), a covalently bound 3-keto-CoA acyl lipid, four OGNG detergent molecules and 112
395 solvent molecules. None of the resolved cysteines appear to be modified with IAM. Data
396 collection and refinement statistics are reported in Table 1. Molecular models for the complexes
397 with substrate, covalently bound acyl-imidazole intermediate, malonyl-CoA and 3-keto-

398 eicosanoyl product were built in COOT based on the crystal structure using simple manual
399 placement and correction of any steric clashes.

400 **Denaturing Intact Mass Spectrometry**

401 The denaturing intact mass spectrometry measurements were performed using an
402 Agilent 1290 Infinity LC System in-line with an Agilent 6530 Accurate-Mass Q-TOF LC/MS
403 (Agilent Technologies Inc.). The solvent system was consisted of 0.1% Optima™ LC/MS grade
404 formic acid (Fisher Chemical) in HPLC electrochemical grade water (Fisher Chemical) (solvent
405 A) and 0.1% formic acid in Optima™ LC/MS grade methanol (Fisher Chemical) (solvent B).
406 Typically, 1-2 µg of protein sample was diluted to 60 µl with 30% methanol in 0.1% formic
407 acid. 60 µl of sample was injected onto a ZORBAX StableBond 300 C3 column (Agilent
408 Technologies Inc.) by an auto sampler. The flowrate of the LC system was set to 0.5 ml/min.
409 30% of solvent B was applied in the beginning and the sample elution was initiated by a linear
410 gradient from 30% to 95% of solvent B over 7 min. 95% B was then applied for 2 min, followed
411 by 2 min equilibration with 30% B. The mass spectrometer was in positive ion, 2 GHz detector
412 mode and spectra were recorded with capillary, fragmentor and collision cell voltages of 4000
413 V, 250 V and 0 V, respectively. The drying gas was supplied at 350°C with flow rate of 12
414 l/min and nebulizer at 60 psi. The data was acquired from 100-3200 m/z. Data analysis was
415 performed using MassHunter Qualitative Analysis Version B.07.00 (Agilent) software.

416 Assignment of the deconvoluted mass peak to the apo form of the protein was achieved
417 by identifying a -89.62 Da mass shift relative to the theoretical mass of the purified protein.
418 This corresponds to the loss of the initiator methionine, followed by acetylation of the new N-
419 terminus (theoretical mass shift of -89.16 Da). Further characterisation of adduct formation and
420 IAM modification was carried out by analysing the deconvoluted mass shift from the apo
421 protein peak to those corresponding to covalently modified forms of the protein. This enabled

422 the determination of a mass shift of +1073.66 Da for the copurified adduct, and the observation
423 that 1, 2 or 3 sites can be simultaneously modified by IAM (predicted +57.07 Da per site
424 modified).

425 In order to trap the covalent acyl-enzyme intermediate, the purified, tagged, wild-type
426 ELOVL7 protein at 1.5 mg/mL (obtained after the desalting step that followed IMAC elution)
427 was incubated with 100 μ M C18:0-CoA (Avanti Polar Lipids, Cat. No. 870718) or C18:3(n3)-
428 CoA (Avanti Polar Lipids, Cat. No. 870732) for 2 hours at 37 °C, in the presence and absence
429 of 1 mM EDTA, 1 mM EGTA, or 100 μ M Malonyl CoA (Sigma-Aldrich, Cat. No. M4263).
430 The reaction was terminated by dilution into 30% methanol in 0.1% formic acid, as described
431 above. Covalent acyl-enzyme intermediate formation was identified by monitoring the presence
432 of a mass shift upon incubation with the substrate, corresponding to the addition of the substrate
433 acyl chain through attachment of the histidine imidazole to the thioester carbonyl, resulting in
434 thioester cleavage and loss of CoA (predicted +266.47 Da upon reaction with C18:0-CoA;
435 +260.42 Da upon reaction with C18:3(n3)-CoA). The site of covalent modification was probed
436 by carrying out this experiment with the His150Ala and His181Ala mutants, which allowed
437 identification of His150 as the nucleophile involved in covalent intermediate formation.

438 **Product formation detected by mass spectrometry**

439 To follow the enzymatic reaction, a small molecule mass spectrometry experiment was
440 performed using LC-MS with a Waters system equipped with a Waters 2545 binary gradient
441 module, a Waters SQ Detector 2, Waters 2489 UV/visible detector, and a Waters 2424 ELS
442 detector. Masslynx 4.0 software by Waters (Beverly, MA) was applied for data processing.
443 Analytical separation of malonyl-CoA, C18:0-CoA and the 3-keto-eicosanoyl (C20:0)-CoA
444 reaction product was carried out on a Phenomenex Kinetex 5 μ m EVO C18 column (100 mm
445 \times 3.0 mm, 100 Å) using a flow rate of 2 mL/min in a 3 min gradient elution, with 20 μ L

446 injections. The mobile phase was a mixture of 93% H₂O, 5% acetonitrile, and 2% of 0.5 M
447 ammonium acetate adjusted to pH 6 with glacial acetic acid (solvent A) and 18% H₂O, 80%
448 acetonitrile, and 2% of 0.5 M ammonium acetate adjusted to pH 6 with glacial acetic acid
449 (solvent B). Gradient elution was as follows: 95:5 (A/B) 0.35 min, 95:5 (A/B) to 5:95 (A/B)
450 over 1 min, 5:95 (A/B) over 0.75 min, and then reversion back to 95:5 (A/B) over 0.1 min and
451 95:5 (A/B) over 0.8 min. A blank injection of MeOH (20 µL) was included between runs to
452 avoid any carry-over from the previous sample.

453 Samples were analysed by mass spectrometry in negative ESI single ion scans set for
454 mass ions of 852.4 m/z (ESI-) (malonyl-CoA, Sigma-Aldrich, Cat. No. M4263), 1032.9 m/z
455 (ESI-) (C18:0-CoA, Sigma-Aldrich, Cat. No. S0802) and 1074.9 m/z (ESI-) (3-keto-
456 eicosanoyl-CoA). The retention times for these compounds were 0.23, 1.60 and 1.63 mins,
457 respectively. The capillary voltage was set to 3.00 kV, the cone voltage was set to 30 V.
458 Nitrogen was used as desolvation gas at flow rate of 500 L/h. Desolvation temperature was set
459 to 250 °C.

460 **Size exclusion chromatography with multiangle light scattering (SEC-MALS)**

461 SEC-MALS analysis was performed by injecting 100 µl of purified protein at 1 mg/mL
462 onto a Superdex 200 10/300 Increase GL column (GE Healthcare), pre-equilibrated in 20 mM
463 HEPES pH 7.5, 200 mM NaCl, 0.1% w/v OGNG / 0.01% w/v CHS, using an OMNISEC
464 RESOLVE system (Malvern). Light scattering and refractive index changes were monitored
465 using an integrated OMNISEC REVEAL multi-detector module (Malvern), and the data was
466 analysed using OMNISEC 5.1 software. To enable calculation of the protein molecular weight
467 within the protein-detergent complex, the detergent dn/dc was experimentally determined by
468 generating a standard curve of refractive index peak area vs. detergent concentration. The slope

469 of this curve, corresponding to the detergent dn/dc , was 0.127 ml/g. The protein dn/dc was
470 assumed to be 0.185 ml/g.

471

472

473

474 **Data Availability**

475 Atomic coordinates and structure factors for the reported crystal structure are deposited
476 in the protein databank (PDB) under accession code 6Y7F.

477

478 **References**

- 479 1. Leonard, A.E., Pereira, S.L., Sprecher, H. & Huang, Y.-S. Elongation of long-chain
480 fatty acids. *Progress in Lipid Research* **43**, 36-54 (2004).
- 481 2. Sassa, T. & Kihara, A. Metabolism of very long-chain Fatty acids: genes and
482 pathophysiology. *Biomol Ther (Seoul)* **22**, 83-92 (2014).
- 483 3. Ohno, Y. et al. ELOVL1 production of C24 acyl-CoAs is linked to C24 sphingolipid
484 synthesis. *Proc Natl Acad Sci U S A* **107**, 18439-44 (2010).
- 485 4. Imgrund, S. et al. Adult ceramide synthase 2 (CERS2)-deficient mice exhibit myelin
486 sheath defects, cerebellar degeneration, and hepatocarcinomas. *J Biol Chem* **284**,
487 33549-60 (2009).
- 488 5. Isokawa, M., Sassa, T., Hattori, S., Miyakawa, T. & Kihara, A. Reduced chain length
489 in myelin sphingolipids and poorer motor coordination in mice deficient in the fatty acid
490 elongase Elov11. *FASEB Bioadv* **1**, 747-759 (2019).
- 491 6. Sassa, T. et al. Impaired epidermal permeability barrier in mice lacking *elov11*, the gene
492 responsible for very-long-chain fatty acid production. *Mol Cell Biol* **33**, 2787-96 (2013).

- 493 7. Li, W. et al. Depletion of ceramides with very long chain fatty acids causes defective
494 skin permeability barrier function, and neonatal lethality in ELOVL4 deficient mice. *Int*
495 *J Biol Sci* **3**, 120-8 (2007).
- 496 8. Harkewicz, R. et al. Essential role of ELOVL4 protein in very long chain fatty acid
497 synthesis and retinal function. *J Biol Chem* **287**, 11469-80 (2012).
- 498 9. Pewzner-Jung, Y. et al. A critical role for ceramide synthase 2 in liver homeostasis: II.
499 insights into molecular changes leading to hepatopathy. *J Biol Chem* **285**, 10911-23
500 (2010).
- 501 10. Zhang, K. et al. A 5-bp deletion in ELOVL4 is associated with two related forms of
502 autosomal dominant macular dystrophy. *Nat Genet* **27**, 89-93 (2001).
- 503 11. Cadieux-Dion, M. et al. Expanding the clinical phenotype associated with ELOVL4
504 mutation: study of a large French-Canadian family with autosomal dominant
505 spinocerebellar ataxia and erythrokeratoderma. *JAMA Neurol* **71**, 470-5 (2014).
- 506 12. Di Gregorio, E. et al. ELOVL5 mutations cause spinocerebellar ataxia 38. *Am J Hum*
507 *Genet* **95**, 209-17 (2014).
- 508 13. Moon, Y.A., Hammer, R.E. & Horton, J.D. Deletion of ELOVL5 leads to fatty liver
509 through activation of SREBP-1c in mice. *J Lipid Res* **50**, 412-23 (2009).
- 510 14. Matsuzaka, T. et al. Crucial role of a long-chain fatty acid elongase, Elovl6, in obesity-
511 induced insulin resistance. *Nat Med* **13**, 1193-202 (2007).
- 512 15. Tamura, K. et al. Novel lipogenic enzyme ELOVL7 is involved in prostate cancer
513 growth through saturated long-chain fatty acid metabolism. *Cancer Res* **69**, 8133-40
514 (2009).
- 515 16. Tolkach, Y. et al. Signatures of Adverse Pathological Features, Androgen Insensitivity
516 and Metastatic Potential in Prostate Cancer. *Anticancer Res* **35**, 5443-51 (2015).
- 517 17. Zhang, X. & Wang, Y. Identification of hub genes and key pathways associated with
518 the progression of gynecological cancer. *Oncol Lett* **18**, 6516-6524 (2019).

- 519 18. Li, G. et al. Association of GALC, ZNF184, IL1R2 and ELOVL7 With Parkinson's
520 Disease in Southern Chinese. *Front Aging Neurosci* **10**, 402 (2018).
- 521 19. Parisi, L.R. et al. Membrane Disruption by Very Long Chain Fatty Acids during
522 Necroptosis. *ACS Chem Biol* **14**, 2286-2294 (2019).
- 523 20. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res*
524 **38**, W545-9 (2010).
- 525 21. Naganuma, T., Sato, Y., Sassa, T., Ohno, Y. & Kihara, A. Biochemical characterization
526 of the very long-chain fatty acid elongase ELOVL7. *FEBS Letters* **585**, 3337-3341
527 (2011).
- 528 22. Hernandez-Buquer, S. & Blacklock, B.J. Site-directed mutagenesis of a fatty acid
529 elongase ELO-like condensing enzyme. *FEBS Lett* **587**, 3837-42 (2013).
- 530 23. Wang, H. et al. Crystal structure of human stearyl-coenzyme A desaturase in complex
531 with substrate. *Nat Struct Mol Biol* **22**, 581-5 (2015).
- 532 24. Rottig, A. & Steinbuchel, A. Acyltransferases in bacteria. *Microbiol Mol Biol Rev* **77**,
533 277-321 (2013).
- 534 25. Paiva, P., Sousa, S.F., Ramos, M.J. & Fernandes, P.A. Understanding the Catalytic
535 Machinery and the Reaction Pathway of the Malonyl-Acetyl Transferase Domain of
536 Human Fatty Acid Synthase. *ACS Catalysis* **8**, 4860-4872 (2018).
- 537 26. Trent, M.S., Worsham, L.M.S. & Ernst-Fonberg, M.L. HlyC, the Internal Protein
538 Acyltransferase That Activates Hemolysin Toxin: Roles of Various Conserved
539 Residues in Enzymatic Activity As Probed by Site-Directed Mutagenesis. *Biochemistry*
540 **38**, 9541-9548 (1999).
- 541 27. Dodds, A.W., Ren, X.D., Willis, A.C. & Law, S.K. The reaction mechanism of the
542 internal thioester in the human complement component C4. *Nature* **379**, 177-9 (1996).
- 543 28. Gadjeva, M. et al. The covalent binding reaction of complement component C3. *J*
544 *Immunol* **161**, 985-90 (1998).

- 545 29. Elliott, R.J., Bennet, A.J., Braun, C.A., MacLeod, A.M. & Borgford, T.J. Active-site
546 variants of *Streptomyces griseus* protease B with peptide-ligation activity. *Chem Biol* **7**,
547 163-71 (2000).
- 548 30. Bolon, D.N. & Mayo, S.L. Enzyme-like proteins by computational design. *Proc Natl*
549 *Acad Sci U S A* **98**, 14274-9 (2001).
- 550 31. Burke, A.J. et al. Design and evolution of an enzyme with a non-canonical
551 organocatalytic mechanism. *Nature* **570**, 219-223 (2019).
- 552 32. Aldahmesh, M.A. et al. Recessive mutations in ELOVL4 cause ichthyosis, intellectual
553 disability, and spastic quadriplegia. *Am J Hum Genet* **89**, 745-50 (2011).
- 554 33. Giroux, J.M. & Barbeau, A. Erythrokeratoderma with ataxia. *Arch Dermatol* **106**, 183-
555 8 (1972).
- 556 34. Herron, B.J., Bryda, E.C., Heverly, S.A., Collins, D.N. & Flaherty, L. Scraggly, a new
557 hair loss mutation on mouse chromosome 19. *Mamm Genome* **10**, 864-9 (1999).
- 558 35. Westerberg, R. et al. Role for ELOVL3 and fatty acid chain length in development of
559 hair and skin function. *J Biol Chem* **279**, 5621-9 (2004).
- 560 36. Engelen, M., Kemp, S. & Poll-The, B.-T. X-Linked Adrenoleukodystrophy:
561 Pathogenesis and Treatment. *Current Neurology and Neuroscience Reports* **14**, 486
562 (2014).
- 563 37. Sassa, T., Wakashima, T., Ohno, Y. & Kihara, A. Lorenzo's oil inhibits ELOVL1 and
564 lowers the level of sphingomyelin with a saturated very long-chain fatty acid. *J Lipid*
565 *Res* **55**, 524-30 (2014).
- 566 38. Winter, G. et al. DIALS: implementation and evaluation of a new integration package.
567 *Acta Crystallographica Section D* **74**, 85-97 (2018).
- 568 39. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta*
569 *Crystallographica Section D* **67**, 235-242 (2011).

- 570 40. Evans, P. Scaling and assessment of data quality. *Acta Crystallographica Section D-*
571 *Biological Crystallography* **62**, 72-82 (2006).
- 572 41. Tickle, I.J. et al. STARANISO. (Global Phasing Ltd, Cambridge, UK, 2018).
- 573 42. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular
574 structure solution. *Acta Crystallographica Section D-Biological Crystallography* **66**,
575 213-221 (2010).
- 576 43. Terwilliger, T. SOLVE and RESOLVE: automated structure solution, density
577 modification and model building. *Journal of Synchrotron Radiation* **11**, 49-52 (2004).
- 578 44. Cowtan, K.D., Zhang, K.Y.J. & Main, P. *DM/DMMULTI software for phase*
579 *improvement by density modification*, (International Union of Crystallography, 2012).
- 580 45. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein
581 chains. *Acta Crystallographica Section D* **62**, 1002-1011 (2006).
- 582 46. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot.
583 *Acta Crystallographica Section D-Biological Crystallography* **66**, 486-501 (2010).
- 584 47. Bricogne, G. et al. BUSTER version 2.10.3. (Global Phasing Ltd, Cambridge, UK,
585 2017).
- 586 48. Smart, O.S. et al. Exploiting structure similarity in refinement: automated NCS and
587 target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* **68**, 368-80
588 (2012).
- 589 49. Glaser, F. et al. ConSurf: identification of functional regions in proteins by surface-
590 mapping of phylogenetic information. *Bioinformatics* **19**, 163-4 (2003).

591

592 **Acknowledgements**

593 L.N., A.C.W.P., G.F.R., A.C., V.C., S.R.B., D.S., T.M., L.S., S.M.M.M., N.B.-B., E.P.C. were
594 members of the SGC, a registered charity (number 1097737) that receives funds from AbbVie,

595 Bayer Pharma AG, Boehringer Ingelheim, the Canada Foundation for Innovation, Genome
596 Canada, Janssen, Merck KGaA, Merck & Co., Novartis, the Ontario Ministry of Economic
597 Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP and Takeda, as
598 well as the Innovative Medicines Initiative Joint Undertaking ULTRA-DD grant 115766 and
599 the Wellcome grant no 106169/Z/14/Z. T.C.P. is supported by a Wellcome PhD studentship.
600 A.Q. is supported by Wellcome grant no 20289/Z16/Z. The authors would like to thank
601 Diamond Light Source for beam time (BAG proposal mx19301), and the I24 beamline staff for
602 assistance with beam time, crystal screening and data collection. We thank Oliver Smart and
603 Clemens Vornrhein at Global Phasing Ltd for assistance with generating covalent ligand
604 restraints. We acknowledge the use of the UCSF Chimera package from the Resource for
605 Biocomputing, Visualisation, and Informatics at the University of California, San Francisco
606 (supported by NIGMS P41-GM103311). We are grateful to the Membrane Protein Laboratory
607 (Wellcome grant ref (20289/Z16/Z)) at the Research Complex at Harwell for access to SEC-
608 MALS equipment and assistance with these experiments.

609

610 **Author Contributions**

611 L.N. purified protein, carried out biophysical characterisations, assisted by V.C. and D.S., and
612 obtained crystals that diffracted to beyond 4 Å resolution. L.N. and A.C.W.P. collected X-ray
613 diffraction data and solved and built the structure. T.C.P. purified protein samples, performed
614 mass spectrometry adduct analysis and SEC-MALS experiments. A.Q. provided access to
615 equipment and assisted in SEC-MALS. G.F.R. assisted with design and execution of the rapid
616 fire mass spectrometry activity assay, supervised by P.E.B.. T.M. assisted with mass
617 spectrometry methods optimisation. S.R.B. and A.C. were involved in the early stages of the
618 project, including design of constructs, optimisation of the protein purification, production and
619 screening of initial crystals and initial MS studies. J.D.L. assisted with early stages of the

620 project, including testing constructs and providing materials. Constructs were screened for
621 expression by L.S. and large scale insect cell expressions were produced by S.M.M.M.,
622 supervised by N.B.B. Data were analysed and the paper was written by L.N., T.C.P., A.C.W.P.
623 G.F.R., P.E.B. and E.P.C.. E.P.C. supervised all aspects of the project.

624

625 **Author Information**

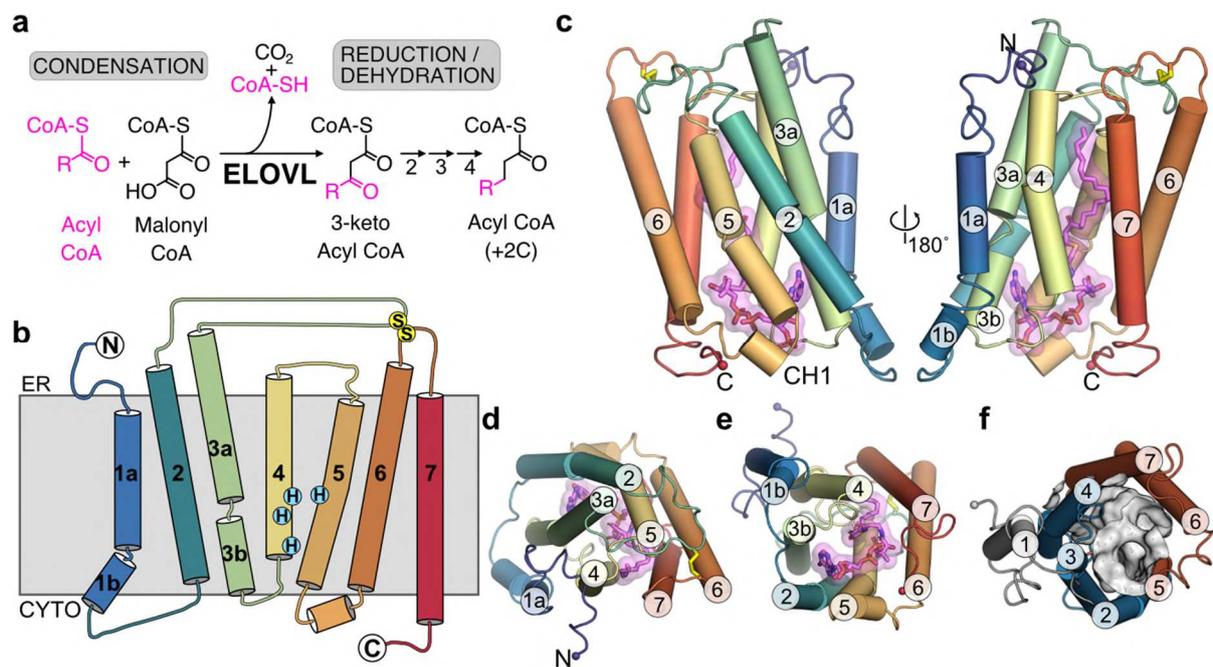
626 Reprints and permissions information are available at www.nature.com/reprints.

627 The other authors declare no competing interests.

628 Correspondence and requests for materials should be addressed to E.P.C.

629 (liz.carpenter@cmd.ox.ac.uk).

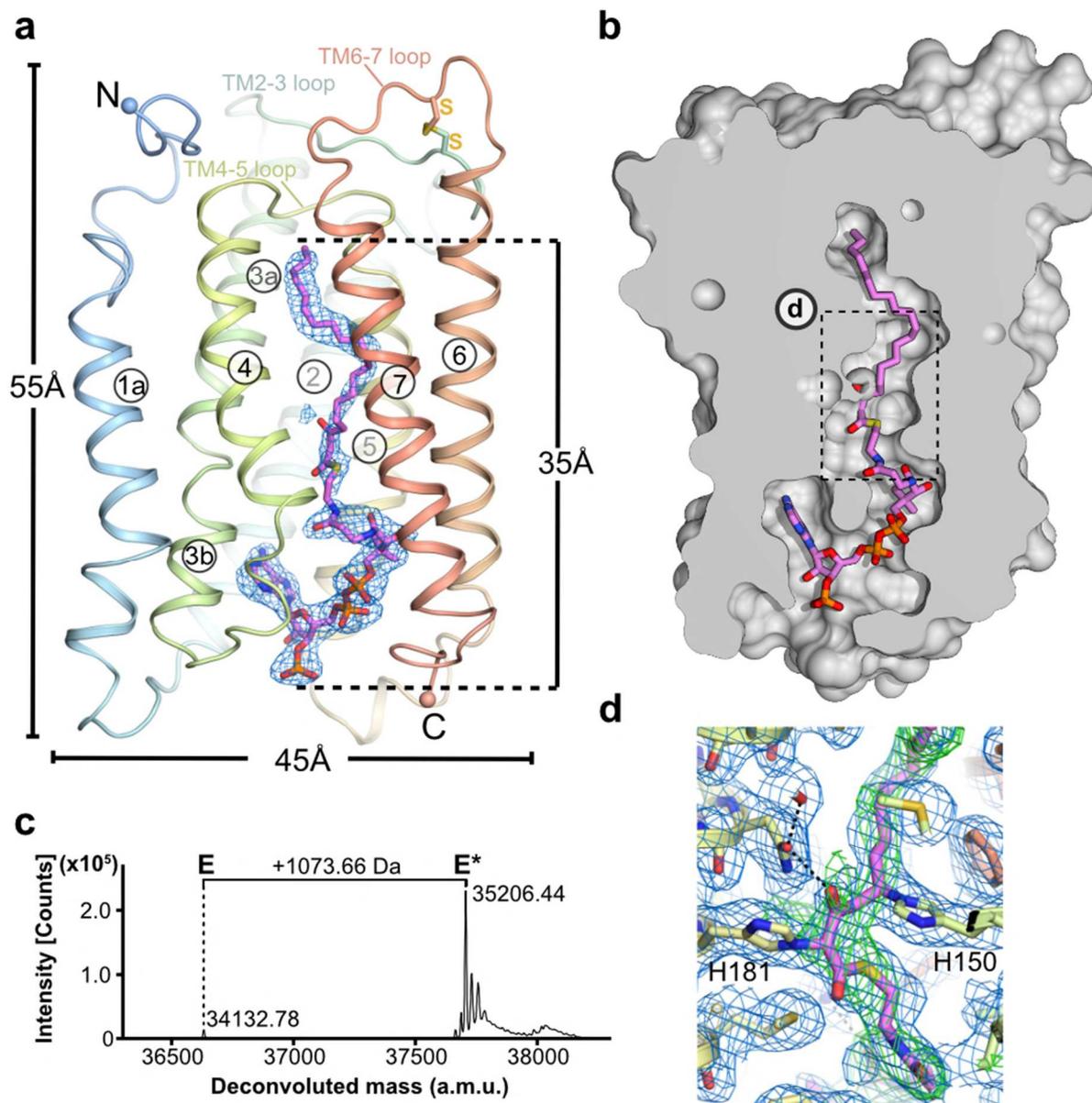
630 **Figures**



631

632

Fig. 1. Overall structure of ELOVL7. a, FA elongase reaction sequence. **b**, Schematic representation of the seven TM transmembrane topology. **c-e**, Cartoon of ELOVL7 structure viewed **c**, parallel to the membrane plane and from either **d**, the ER or **e**, the cytoplasmic faces. **f**, ELOVL7 6-TM inverted barrel organisation viewed from the cytoplasmic face.



634

635 **Fig. 2. Heterologously expressed ELOVL7 is covalently bound to a 3-keto acyl-CoA. a,**

636 Cartoon showing the covalently bound 3-keto eicosanoyl CoA along with the FoFc omit

637 electron density map (blue mesh, contoured at 3sigma). **b,** Cutaway molecular surface

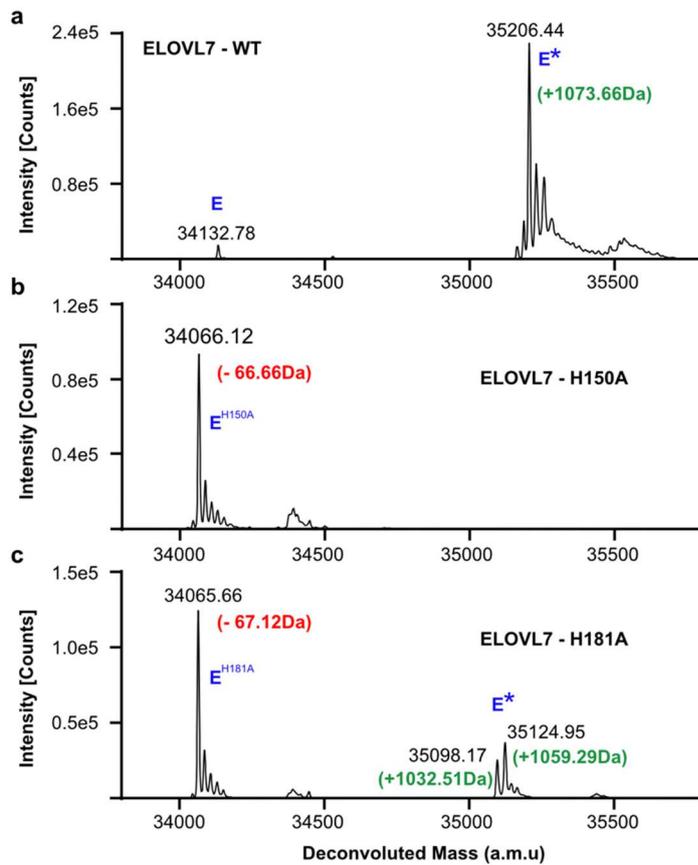
638 representation showing the extent of the enclosed central active site tunnel. **c,** Intact mass

639 analysis of purified protein (E) highlighting adduct (E*; +1074Da). **d,** Electron density in region

640 around covalent linkages to active site histidines. Final BUSTER 2mFo-DFc (blue mesh,

641 contoured at 1sigma) and omit mFo-DFc (green mesh, contoured at 2.5sigma) electron density
642 maps are overlaid on the final model.

643

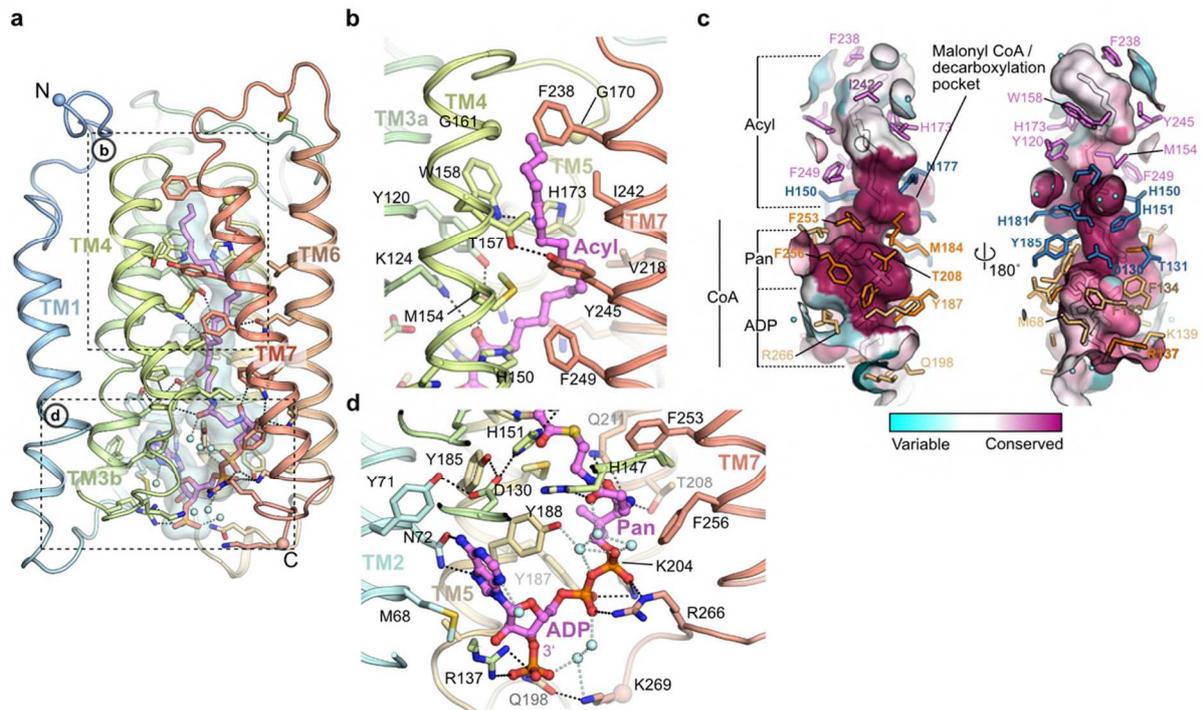


644

645

646 **Fig. 3. Intact mass analysis of WT ELOVL7, H150A and H181A mutants.** a-c, Comparison
647 of the deconvoluted intact mass spectra for the untagged a, WT enzyme b, H150A and c, H181A
648 mutant. The expected mass decrease of a His-to-Ala mutation is 66.06 Da. *In vivo* modification
649 (E*) is observed for the WT and H181A mutant but not for H150A.

650



651

652 **Fig. 4. Acyl chain and CoA binding sites. a**, Overview of the central acyl-CoA binding tunnel.

653 **b**, Acyl chain binding pocket. **c**, Conservation of active site tunnel. Molecular surface

654 representation is coloured by amino acid conservation score calculated by CONSURF⁴⁹

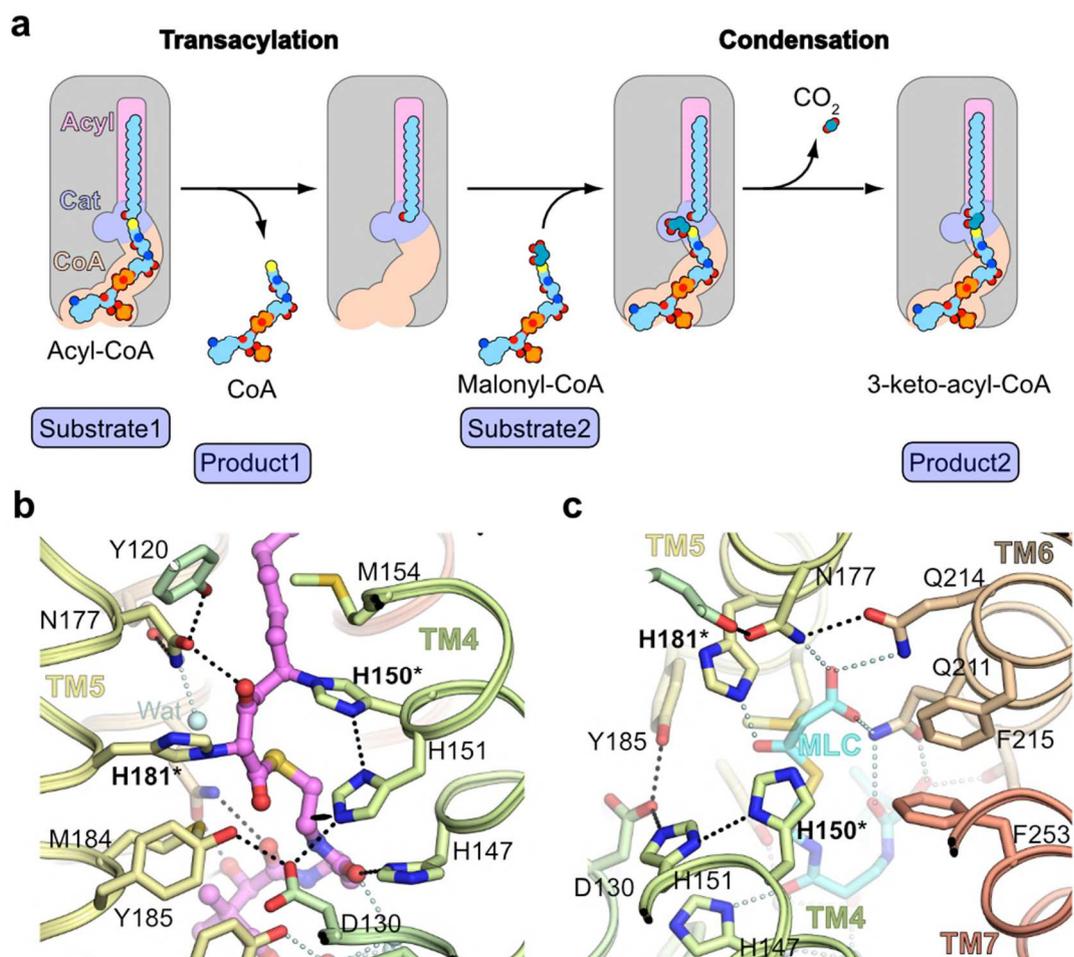
655 analysis of a diverse set of ELOV1-7 family members. The various subregions of the tunnel are

656 indicated (ADP / Pan from CoA and Acyl chain). Amino acid residues that form the binding

657 tunnel are coloured according to region (pink, acyl; blue, catalytic site; orange CoA binding).

658 **d**, Cytoplasmic-facing CoA 3'-phospho ADP/pantothenic (Pan) binding pocket.

659



660

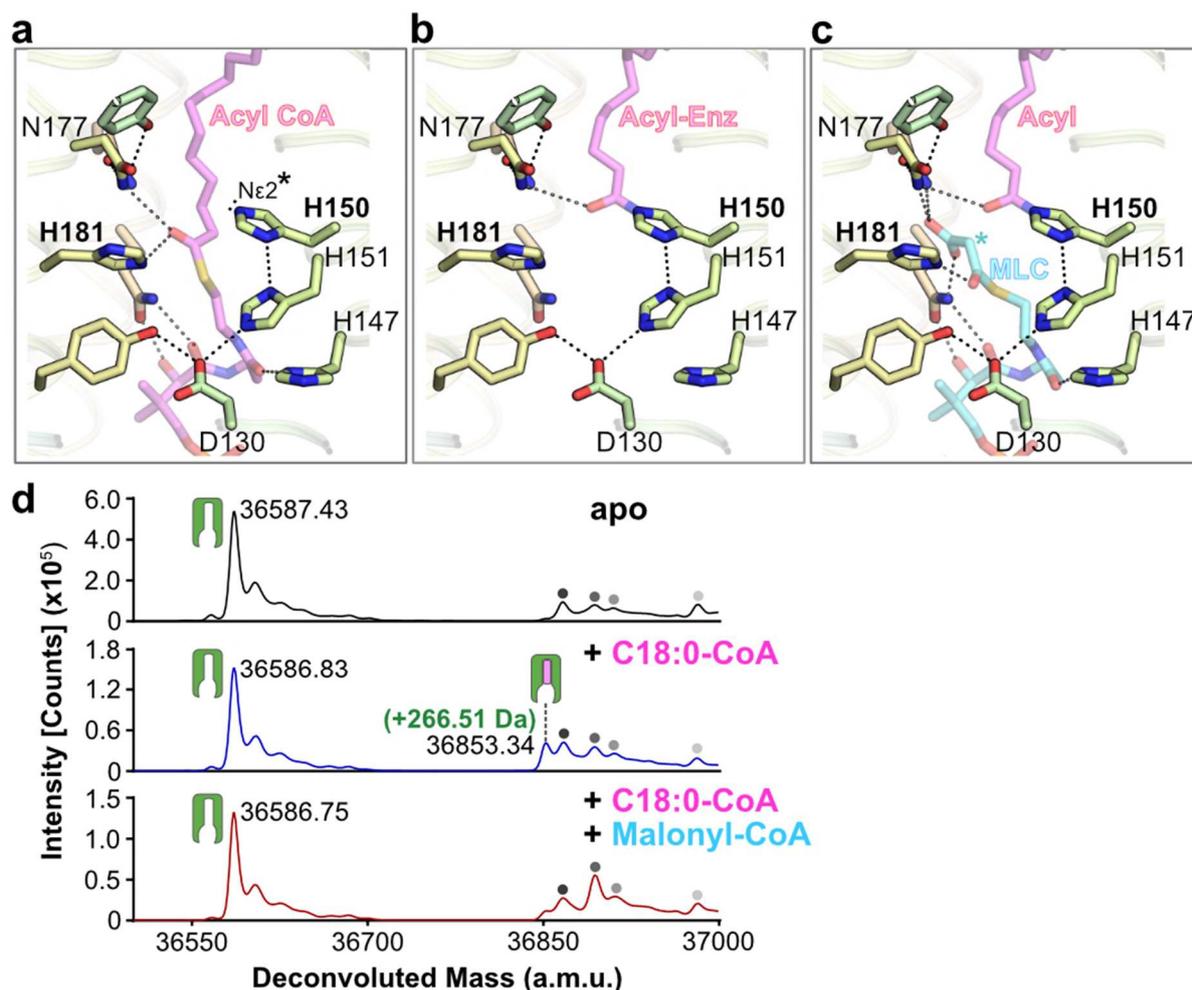
661 **Fig. 5. Active site. a**, Schematic outlining proposed ELOVL ping-pong mechanism. **b-c**,

662 Detailed view of **b**, central catalytic site around conserved histidine motif and **c**, putative

663 decarboxylation pocket. Covalently modified histidines (H150/H181) are highlighted by

664 asterisks. Bound malonyl-CoA (cyan; MLC) has been modelled based on the coordinates of

665 the acyl-CoA.



667

668 **Fig. 6. Proposed steps in catalytic mechanism and evidence for a covalent acyl-imidazole**
 669 **intermediate. a-c**, Proposed catalytic steps for **a**, acyl-CoA binding; **b**, transacylation to form
 670 an acyl-imidazole intermediate. **c**, Proposed malonyl-CoA (MLC) condensation step with acyl
 671 intermediate. Ligand complexes are models based on product/inhibitor crystal structure. The
 672 unprotonated imidazole nitrogen of H150 involved in the transacylation step and the location
 673 of the enolate carbon nucleophile formed after decarboxylation of MLC are highlighted with
 674 asterisks. **d**, Intact mass analysis of protein only (black), protein + C18:0 acyl-CoA (blue) and
 675 protein + C18:0 acyl-CoA + malonyl-CoA (red). Peaks are indicated as follows: protein (green
 676 icon), acyl enzyme intermediate (green icon+pink oblong), background species in all traces
 677 (grey circles).

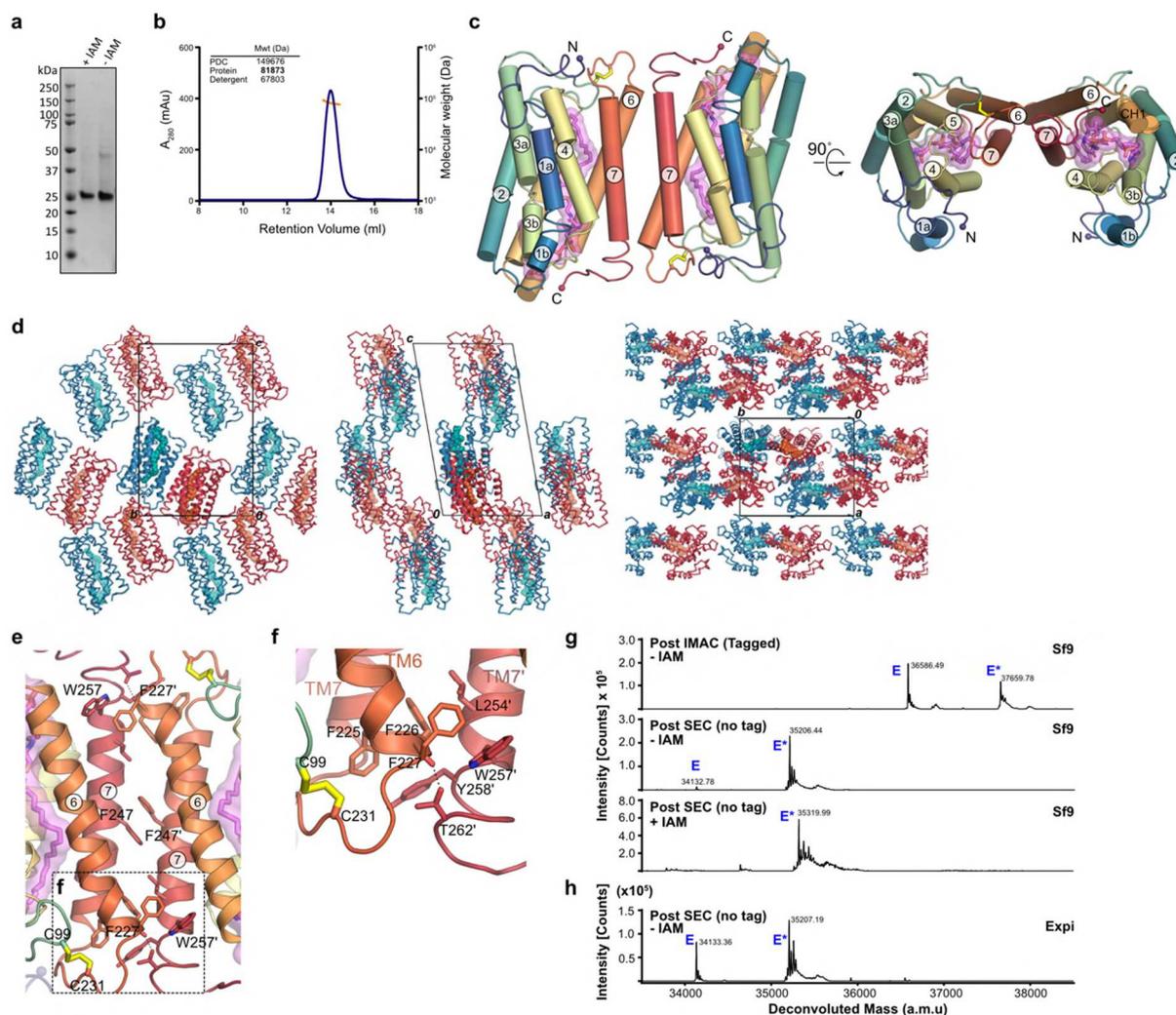
678

679 **Table 1 X-ray data collection, refinement & validation statistics.**

	Hg-peak	WT-native	STARANISO
680			
681	Data collection		
	Space group	<i>P</i> 2 ₁	<i>P</i> 2 ₁
			-
682	Cell dimensions		
	<i>a</i> , <i>b</i> , <i>c</i> (Å)	62.34, 72.42, 111.68	63.23, 72.46, 112.04
683	α , β , γ (°)	90, 100.14, 90	90, 100.03, 90
	Resolution limits [Å] ¹	4.6, 3, 3.2	3.5, 2.05, 2.65
684		(4.4, 3, 3)	(3.24, 2.05, 2.22)
	Nominal Resolution [Å]²	3.2	2.6
			-
685	Resolution [Å]	58 - 3.0 (3.0-3.18) ³	62 - 2.05 (2.05-2.10) ³
	<i>R</i> _{meas}	0.396 (2.516) ³	0.119 (0.839) ³
686	<i>R</i> _{pim}	0.093 (0.588) ³	0.061 (1.534) ³
	<i>I</i> / σ <i>I</i>	6.2 (1.5) ³	4.9 (0.4) ³
687	<i>CC</i> _{1/2}	0.996 (0.838) ³	0.985 (0.473) ³
	Completeness [%]	99.9 (99.6) ³	99.5 (96.6) ³
688	Redundancy	17.6 (18) ³	6.9 (6.9) ³
689	<i>CC</i> _{anom}	0.251 (0.0) ³	6.9 (6.2) ³
	Refinement		
690	Resolution (Å)	-	28.68 – 2.05
	No. reflections (free)	-	36690 (1765)
691	<i>R</i> _{work} / <i>R</i> _{free}	-	21.04 / 22.61
	No. atoms		
692	Protein	-	4260
	Ligand	-	140
693	Water	-	112
	Other	-	153
694	<i>B</i> -factors (Å ²)		
	Protein	-	44
	Ligand	-	43
695	Water	-	37
	Other	-	55
696	r.m.s. deviations		
	Bond lengths (Å)	-	0.008
	Bond angles (°)	-	0.87
697	¹ Anisotropic resolution limits along each of the three principal directions as defined by AIMLESS based on Mn (<i>I</i> /sd(<i>I</i>)) > 2. Values in parentheses are resolution limits in each direction based on half dataset correlation > 0.5 (<i>CC</i> _{1/2}).		
698	² Nominal resolution is defined based on overall Mn (<i>I</i> /sd(<i>I</i>)) > 2 as estimated by AIMLESS.		
699	³ Values in parentheses are for highest resolution shell		

700

701 Extended Data Figures



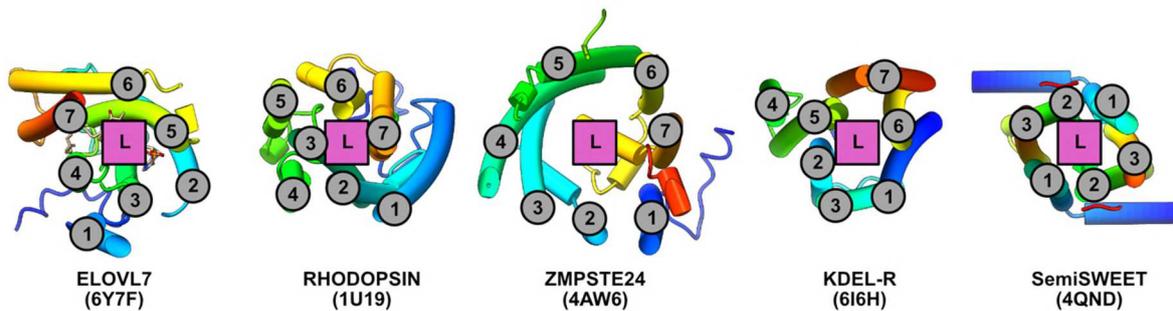
702

703 Extended Data Fig. 1 Properties of purified ELOVL7.

704 **a**, SDS-PAGE gel of purified ELOVL7. **b**, SEC profile and MALLS analysis showing that
705 OGNG-solubilised protein exists as a dimer in solution. **c**, Representation of head-to-tail dimer
706 present in the crystal. **d**, ELOVL7 head-to-tail dimer packing within the crystal lattice. **e, f**
707 Details of dimer interface interactions between TM6 and TM7 of each molecule. **g-h**, Intact
708 mass analysis of ELOVL7 protein at various stages during purification. Deconvoluted mass
709 spectra are shown for ELOVL7 protein purified from (**g**) insect and (**h**) Expi293F cells. For
710 protein purified after expression in insect cells, the samples are shown after immobilized metal
711 affinity chromatography (IMAC), after cleavage of the tag and size exclusion chromatography

712 (SEC) and after treatment with iodoacetamine (IAM). The expected mass of the untagged 'apo'
713 enzyme (E) based on the sequence is 34222 Da. The observed mass peak (34133 Da; -89 Da)
714 corresponds to the loss of the N-terminal methionine (-131.21 Da) and acetylation of the
715 resulting new N-terminus (+42 Da). All samples were run in their reduced state. The modified
716 material (E*) appears as an adduct with an average mass shift of +1073.6 Da. The addition of
717 113.55 Da upon treatment with IAM suggests modification of two cysteine residues.

718



719

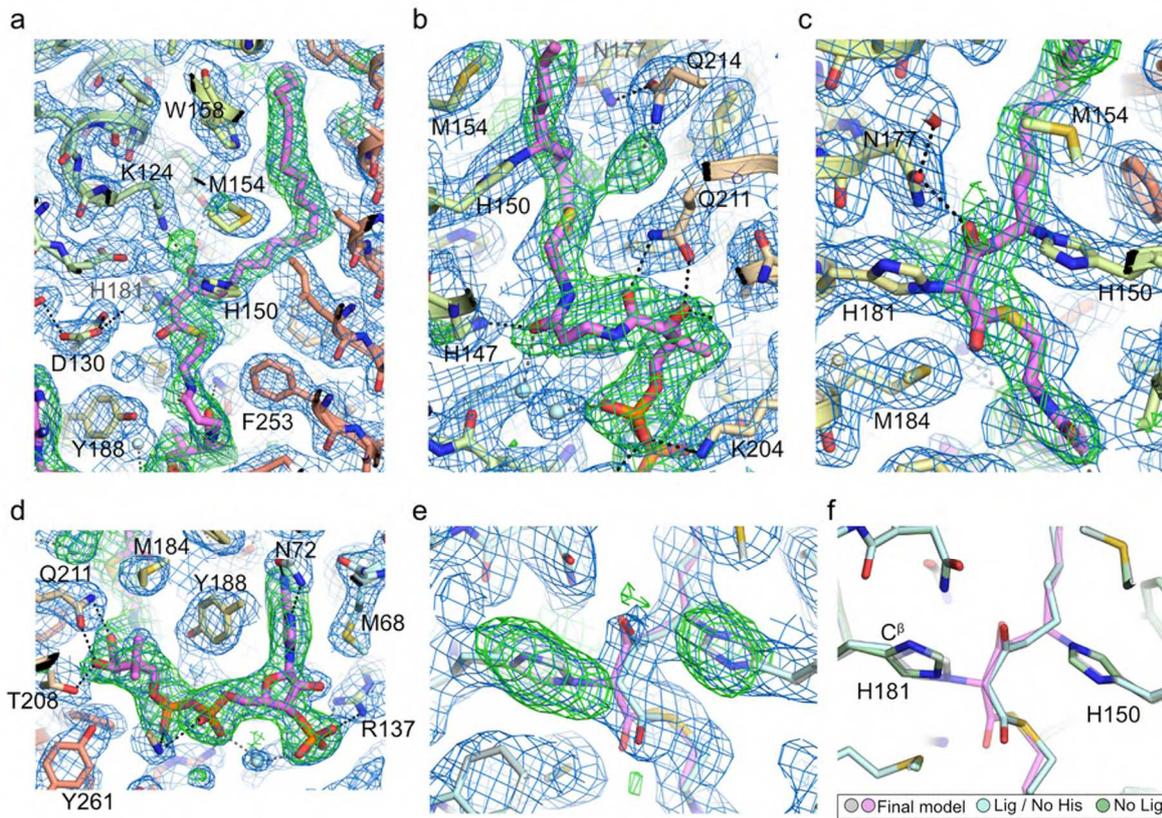
720 **Extended Data Fig. 2 TM helix topology of ELOVL7.**

721 TM helical topology of ELOVL7 is compared with other six and seven membered TM bundles.

722 TM helices are numbered and location of substrate/ligand site marked. Underlying cartoon

723 representations of each structure are coloured from blue to red from the N- to C-termini

724 respectively. PDB accession codes are shown in parentheses.

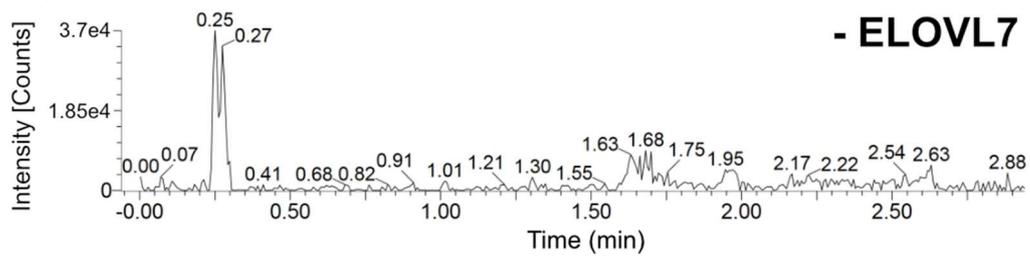


725

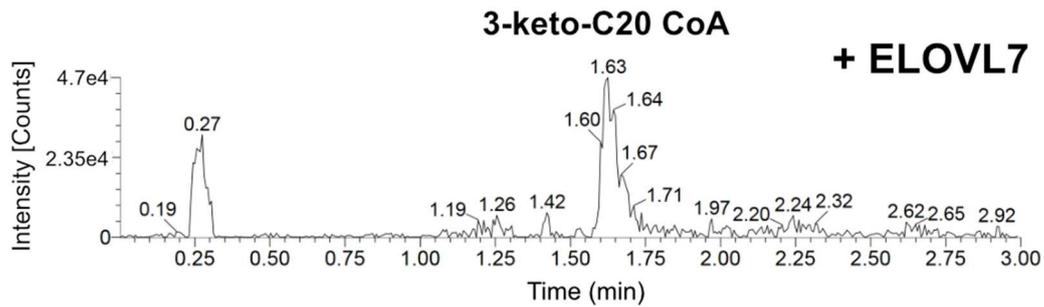
726 **Extended Data Fig. 3 Electron density clearly shows covalently bound 3-keto-eicosanoyl**
 727 **CoA.**

728 **a-d**, Electron density running along the catalytic tunnel. Final BUSTER 2mFo-DFc (blue mesh,
 729 contoured at 1sigma) and omit mFo-DFc (green mesh, contoured at 2.5sigma) electron density
 730 maps are overlaid on the final model. **e**, Comparison of a test refinement in which the imidazole
 731 groups of H150 and H181 were removed from the model (grey carbon protein atoms / palecyan
 732 ligand carbon atoms) and the final model (palecyan protein carbons / violet ligand carbon
 733 atoms). The BUSTER 2mFo-DFc (blue mesh, contoured at 1sigma) and mFo-DFc (green mesh,
 734 contoured at 3sigma) maps for the refined histidine-truncated model / unlinked acyl-CoA are
 735 shown. **f**, Comparison of various refined models (green carbons - protein only; palecyan -
 736 protein without H150/181 sidechain plus ligand; grey/violet - final model with covalently
 737 attached ligand).

a



b

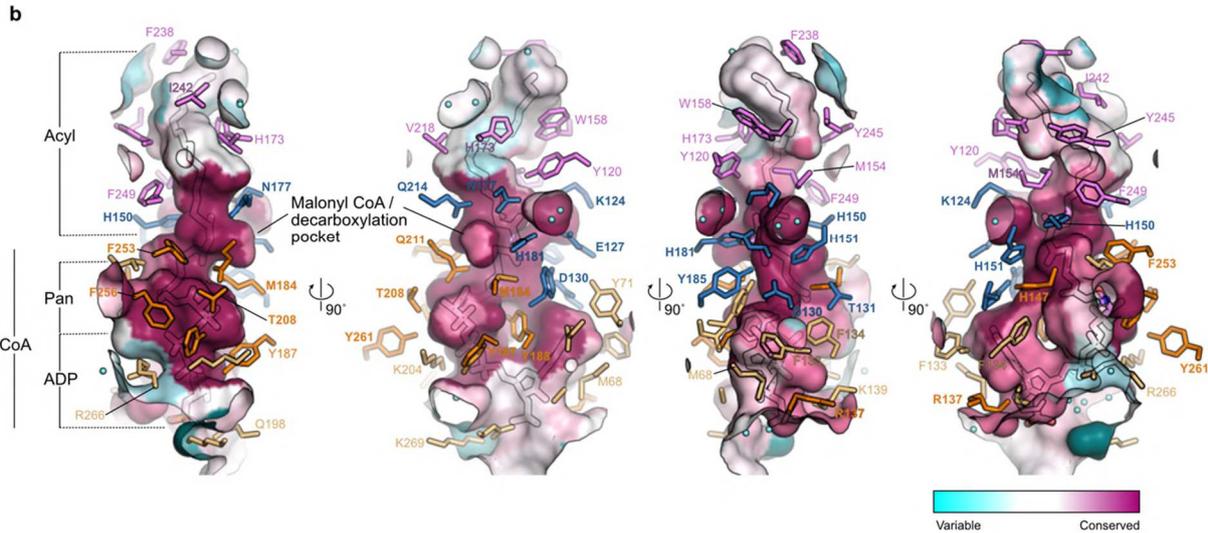
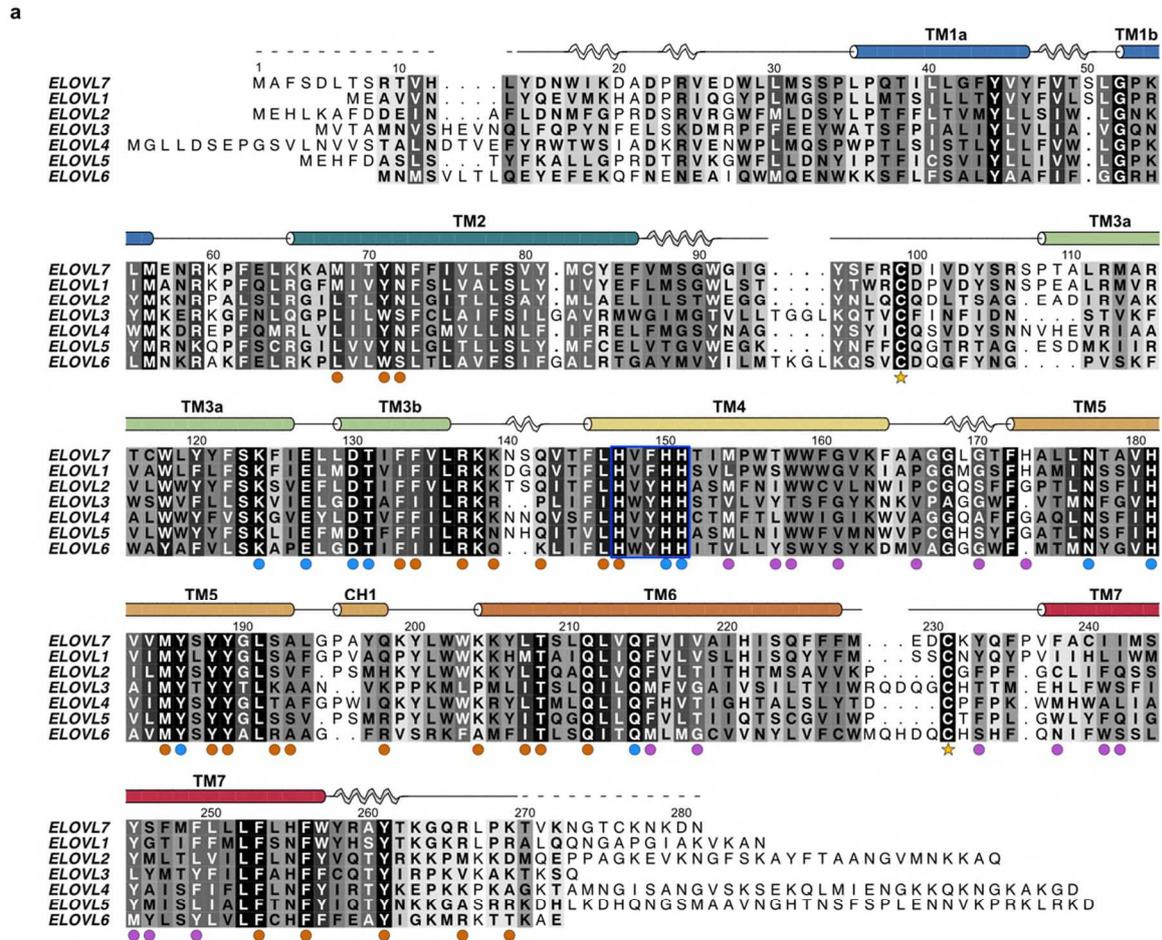


738

739 **Extended Data Fig. 4 WT ELOVL7 activity**

740 **a,b** Activity of residual WT enzyme on incubation with palmitoyl-CoA (C18:0) and malonyl-
741 CoA. Selected ion recording is shown for **a**, reaction mixture without added enzyme and **b**,
742 reaction mixture after 3hr incubation with ELOVL7 enzyme. Ion peak at 1.61 mins corresponds
743 to the expected 3-keto-eicosanoyl (C20)-CoA product of the elongation reaction.

744

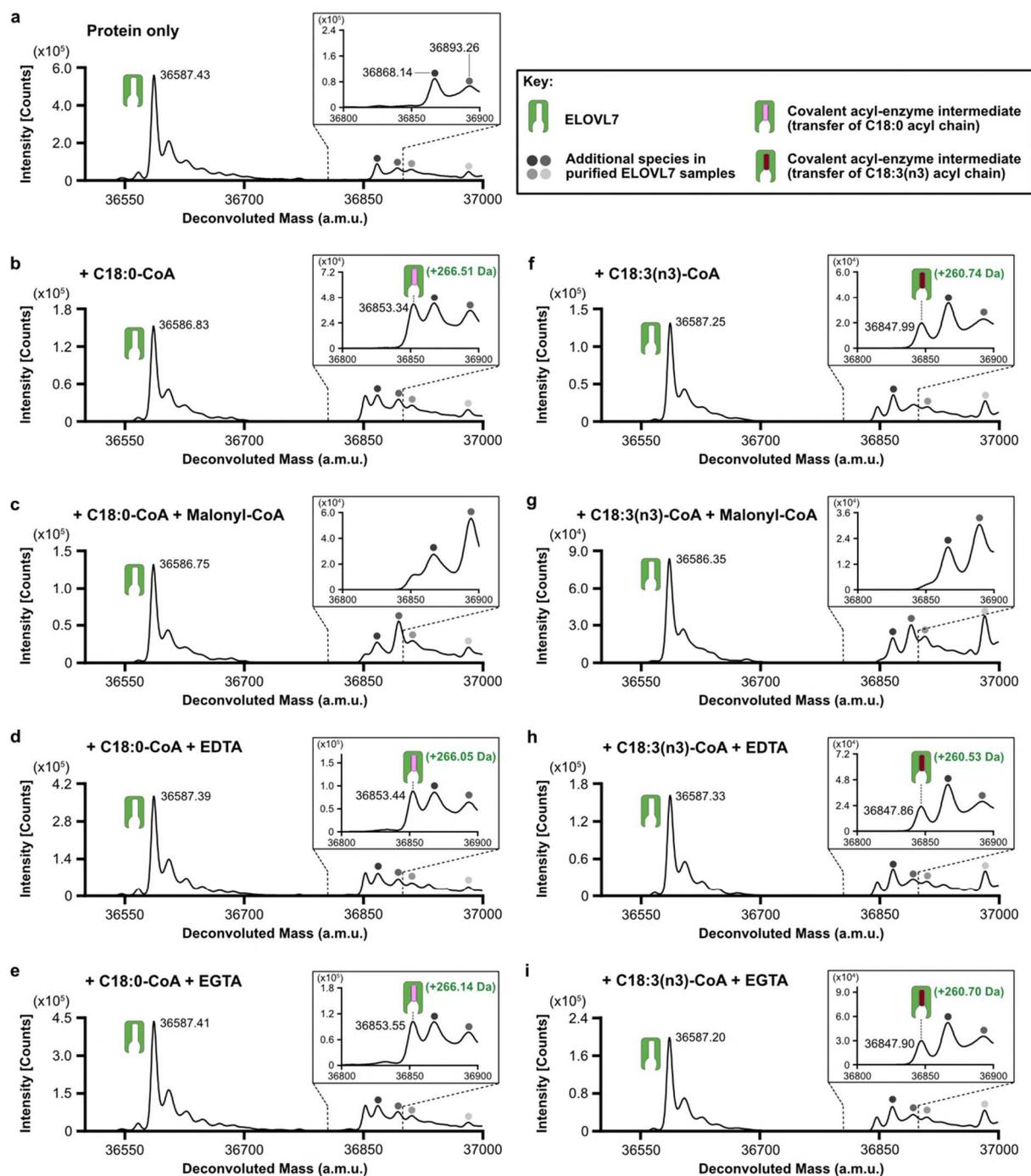


745

746 **Extended Data Fig. 5 Sequence alignment and active site conservation of human ELOVL**
 747 **family members.**

748 **a**, Sequence alignment of human ELOVL1-7. The conserved histidine box (147 HxxHH 151) is
 749 highlighted by a blue box. Filled circles below alignment indicate residues with a proposed

750 catalytic role (blue) and residues interacting with either the CoA (orange) or acyl (plum) portion
751 of the substrate. Cysteines that form the disulphide bridge (C99-C231) between the TM2/3 and
752 TM6/7 loops are indicated by stars. **b**, Conservation of active site tunnel. Molecular surface
753 representation is coloured by amino acid conservation score calculated by CONSURF
754 analysis⁴⁹ of a diverse set of ELOVL1-7 family members. The various subregions of the tunnel
755 are indicated (ADP / Pan from CoA and Acyl chain). Amino acid residues that form the binding
756 tunnel are coloured according to region (pink, acyl; blue, catalytic site; orange CoA binding).



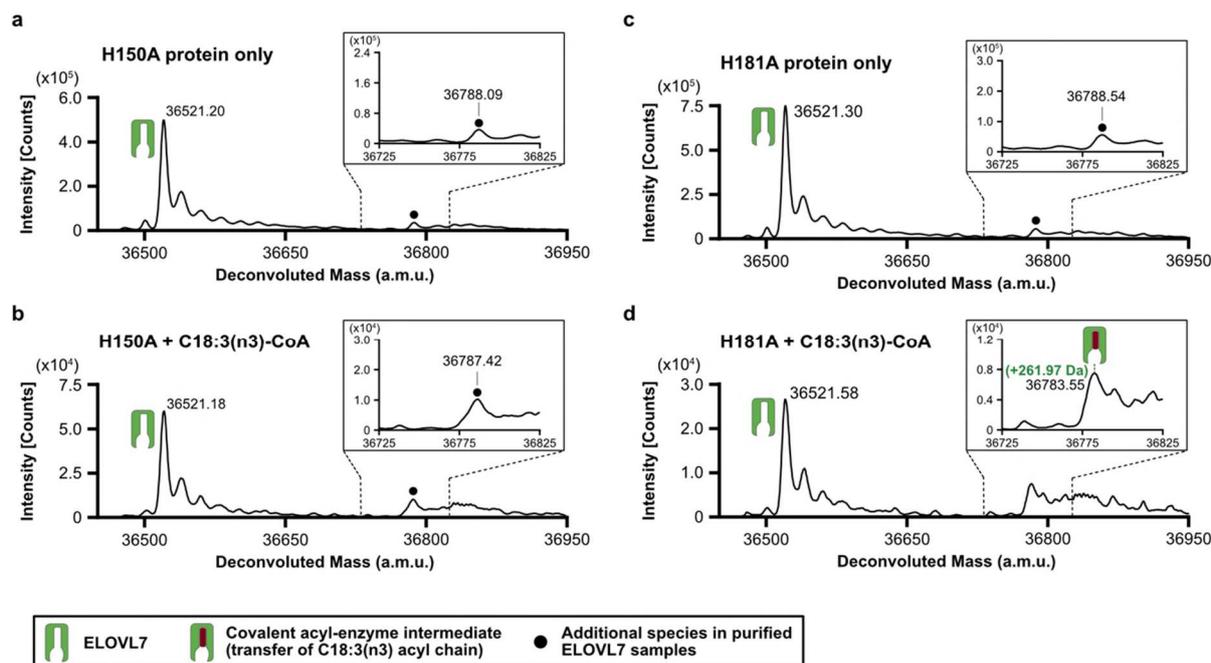
757

758 **Extended Data Fig. 6 Identification of a covalent acyl-enzyme intermediate of ELOVL7.**

759 Purified, tagged, wild-type ELOVL7 was incubated for 2h at 37°C in the presence and absence
 760 of known substrates and metal-chelating agents prior to LC-ESI-MS intact mass analysis.
 761 Deconvoluted intact mass spectra for ELOVL7 incubated **a**, in the absence of substrates. **b**,
 762 with 100µM C18:0-CoA. Expected mass addition for acyl intermediate upon reaction with
 763 C18:0-CoA: +266.47 Da. **c**, Incubation of ELOVL7 with 100µM C18:0-CoA and 100µM

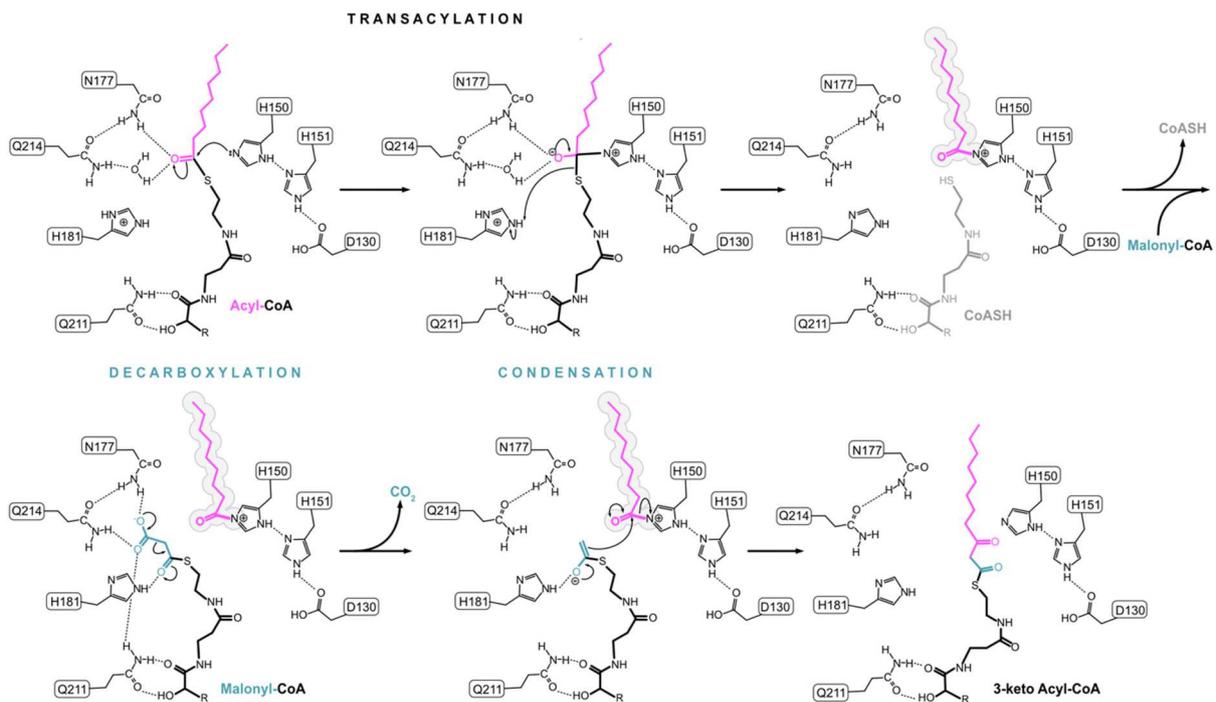
764 malonyl-CoA prevents covalent acyl-enzyme intermediate accumulation. **d-e**, ELOVL7
765 incubated with 100 μ M C18:0-CoA in the presence of **d**, 1mM EDTA or **e**, 1mM EGTA,
766 showing that intermediate formation is not inhibited by metal-chelating agents. **f**, ELOVL7
767 incubated with 100 μ M C18:3(n3)-CoA. Expected mass addition for acyl intermediate upon
768 reaction with C18:3(n3)-CoA: +260.42 Da. **g**, Incubation of ELOVL7 with 100 μ M C18:3(n3)-
769 CoA and 100 μ M malonyl-CoA prevents covalent acyl-enzyme intermediate accumulation. **h-i**,
770 ELOVL7 incubated with 100 μ M C18:3(n3)-CoA in the presence of **h**, 1mM EDTA or **i**, 1mM
771 EGTA. All experiments were repeated independently twice with similar results (n=2 biological
772 repeats, see Supplementary Information).

773



774 **Extended Data Fig. 7 Covalent acyl-enzyme intermediate is formed upon substrate**
 775 **reaction at His150.**

776 LC-ESI-MS intact mass analysis of mutant proteins upon incubation at 37°C for 2h. **a-b**,
 777 His150Ala mutant protein incubated **a**, in the absence of substrate or **b**, in the presence of 100
 778 μM C18:3(n3) CoA. No acyl-enzyme intermediate formation could be detected with the
 779 His150Ala mutant protein. **c-d**, His181Ala mutant protein incubated **c**, in the absence of
 780 substrate or **d**, in the presence of 100 μM C18:3(n3) CoA. Expected mass shift upon reaction
 781 with C18:3(n3) CoA: +260.42 Da. The presence of a background peak at ~36788 Da precluded
 782 testing with C18:0 CoA. All experiments were repeated twice with similar results (n=2
 783 biological repeats, see Supplementary Information).



784 **Extended Data Fig. 8 Proposed ping-pong reaction mechanism for ELOVL7.**

785 **Extended Data Table 1. Mutagenesis primer sequences.**

786

Primer purpose		Primer sequence (5' to 3')
WT ELOVL7	Forward	TTAAGAAGGAGATATACTATGGCCTTCAGTGATCTTAC
	Reverse	GATTGGAAGTAGAGGTTCTCTGCATTATCTTTGTTTTTGCAAG
H150A	Forward	TCATGTATTCGCTCATACCATCATGCCG
	Reverse	AGGAAAGTCACTTGGCTATTTTTC
H151A	Forward	TGTATTCCATGCTACCATCATGCC
	Reverse	TGAAGGAAAGTCACTTGG
H181A	Forward	TACAGCTGTAGCTGTAGTCATGTATTCC
	Reverse	TTTAGAAGGGCATGGAATG

787

Figures

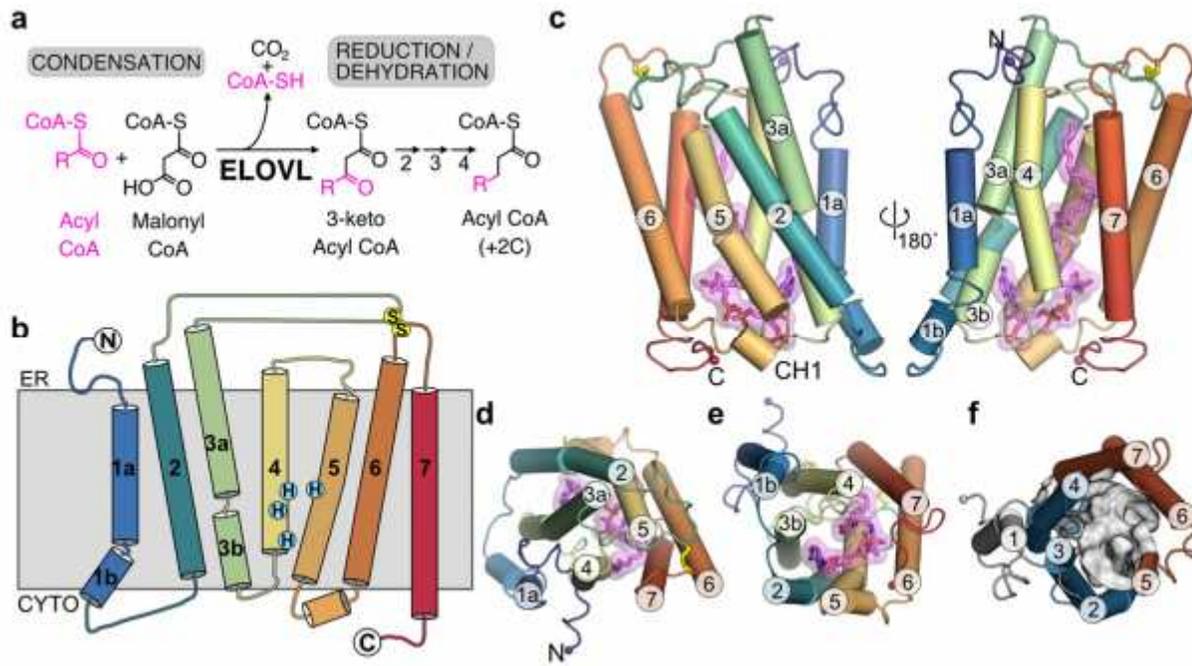


Figure 1

[Please see the manuscript file to view the figure caption.]

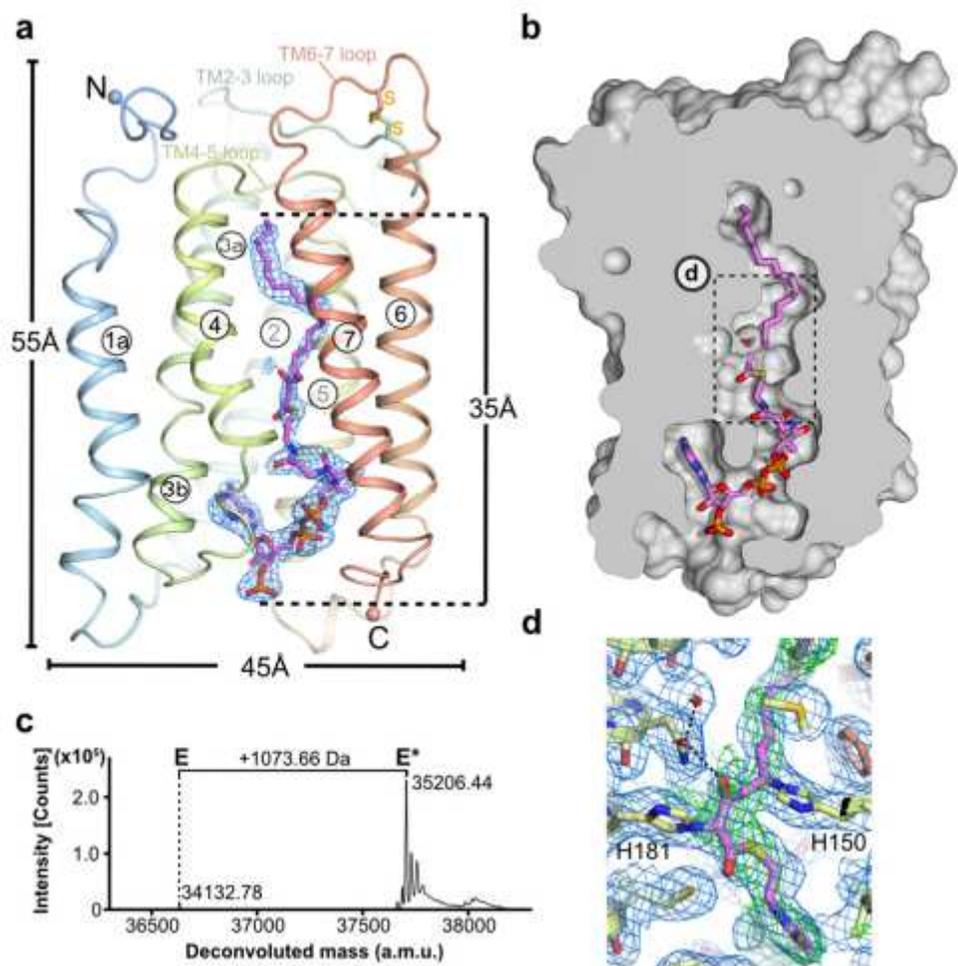


Figure 2

[Please see the manuscript file to view the figure caption.]

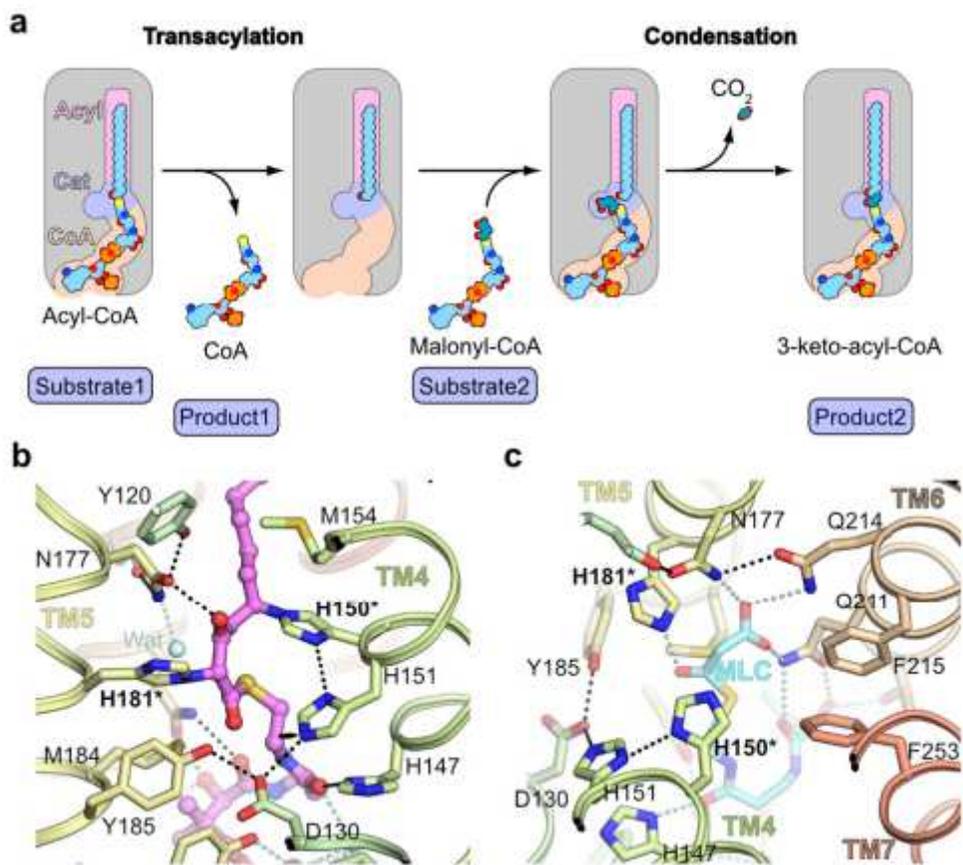


Figure 5

[Please see the manuscript file to view the figure caption.]

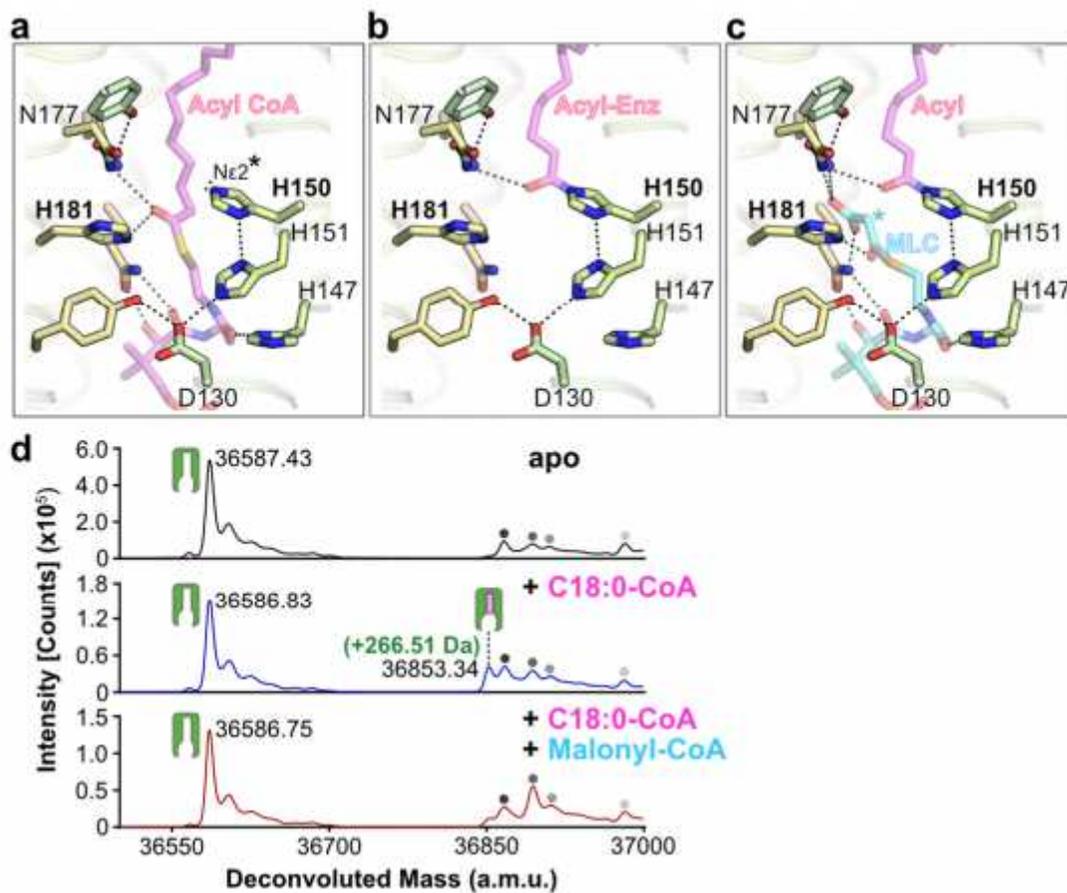


Figure 6

[Please see the manuscript file to view the figure caption.]

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [NieetalELOVL7SupplementaryInformation.pdf](#)