Article

Structural basis for enzymatic terminal C–H bond functionalization of alkanes

Jin Chai [1,3], Gongrui Guo [1,2,3], Sean M. McSweeney[2], John Shanklin [1] ✉ & Qun Liu [1,2] ✉

Alkane monooxygenase (AlkB) is a widely occurring integral membrane metalloenzyme that catalyzes the initial step in the functionalization of recalcitrant alkanes with high terminal selectivity. AlkB enables diverse microorganisms to use alkanes as their sole carbon and energy source. Here we present the 48.6-kDa cryo-electron microscopy structure of a natural fusion from *Fontimonas thermophila* between AlkB and its electron donor AlkG at 2.76 Å resolution. The AlkB portion contains six transmembrane helices with an alkane entry tunnel within its transmembrane domain. A dodecane substrate is oriented by hydrophobic tunnel-lining residues to present a terminal C–H bond toward a diiron active site. AlkG, an [Fe–4S] rubredoxin, docks via electrostatic interactions and sequentially transfers electrons to the diiron center. The archetypal structural complex presented reveals the basis for terminal C–H selectivity and functionalization within this broadly distributed evolutionary class of enzymes.

Alkanes are saturated hydrocarbons that constitute 20–50% of crude oils. Due to their abundance and low cost, alkanes are attractive starting materials for producing chemical feedstocks and value-added products[1,2]. Alkanes contain hydrocarbons with fully occupied orbitals for their C–H bonds and similar bond energies, rendering them almost chemically inert and challenging to functionalize selectively. Alkane monooxygenase (AlkB) is a transmembrane metalloenzyme that catalyzes terminal hydroxylation reactions of alkanes with broad chain-length specificity and exclusive terminal C–H bond selectivity[3,4]. AlkB can be engineered to allow it to derivatize abundant alkanes to alcohols, aldehydes, carboxylic acids and epoxides[5,6].

AlkB was originally identified in *Pseudomonas oleovorans*, which thrives in oil-rich environments where it uses alkanes as its sole carbon and energy source[7–10]. The first and key step in alkane utilization is its hydroxylation, which requires an electron transfer and redox system of AlkBGT, consisting of a catalytic membrane-bound monooxygenase AlkB, a soluble rubredoxin AlkG and a soluble rubredoxin reductase AlkT[11,12]. AlkB, AlkG and AlkT can be encoded as separate or fusion proteins[13–15] (Extended Data Fig. 1). Intense research over the past half century has been devoted to characterizing and engineering the AlkBGT system for alkane C–H functionalization[10,16–21]. The structure

of this electron transfer complex and the molecular determinants of alkane terminal selectivity have nonetheless remained elusive.

Results

Structure determination

We screened 21 AlkB homologs for production in *Escherichia coli* and identified a natural AlkB-AlkG fusion from *Fontimonas thermophila* (*Ft*AlkBG) that was expressed well and could be purified using the detergent *n*-dodecyl-β-D-maltoside (DDM) (Extended Data Fig. 2a). We reconstituted the protein in amphipol PMAL-C8 nanodiscs and further purified it using size-exclusion chromatography (Extended Data Fig. 2b). The molecular mass of *Ft*AlkBG on SDS-PAGE is less than 50 kDa (Extended Data Fig. 2a), which is relatively small for structure determination by cryo-electron microscopy (cryo-EM). To enhance the signals of cryo-EM images, we used a camera pixel size of 0.333-Å resolution under super-resolution mode, with an energy filter slit width at 15 eV, and collected data from holes with the thinnest-possible ice capable of encapsulating particles. Under these conditions, we were able to obtain well-identifiable particles on micrographs (Extended Data Fig. 2c). Two-dimensional (2D) class averaging showed recognizable secondary structural features and

[1]Biology Department, Brookhaven National Laboratory, Upton, NY, USA. [2]NSLS-II, Brookhaven National Laboratory, Upton, NY, USA. [3]These authors contributed equally: Jin Chai, Gongrui Guo. ✉e-mail: shanklin@bnl.gov; qunliu@bnl.gov

Table 1 | Cryo-EM data collection, 3D reconstruction and refinement statistics

	FtAlkBG (EMD-28890, PDB 8F6T)
Data collection	
Microscope	Titan Krios G3i
Magnification	×130,000
Stage type	Autoloader
Voltage (kV)	300
Detector	Gatan K3
Energy filter (eV)	15
Acquisition mode	Super-resolution
Physical pixel size (Å)	0.666
Defocus range (μm)	0.7–2.5
Electron exposure (e⁻ Å⁻²)	60
Reconstruction	
Software	Relion v.3.08, CryoSPARC v.2.15
Particles picked (no.)	2,950,051
Particles final (no.)	46,953
Extraction box size (pixels)	128
Rescaled box size (pixels)	64
Final pixel size (Å)	1.332
Symmetry imposed	C1
Map resolution (Å)	2.76
FSC threshold	0.143
Map resolution range (Å)	341–2.72
Map sharpening B factor (Å²)	66
Model refinement	
Software	PHENIX
Refinement algorithm	Real Space
Clipped box size (pixels)	None
Residues (no.)	428
Iron (no.)	3
Ligand (no.)	1
R.m.s. deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.573
Validation	
MolProbity clashscore	4.79
Rotamer outliers (%)	0.0
Cβ deviations (%)	0.0
Ramachandran plot	
Favored (%)	94.58
Allowed (%)	5.42
Outliers (%)	0

amphipol densities around the protein transmembrane region (Extended Data Fig. 2d).

We determined the structure of FtAlkBG, using the workflow illustrated in Extended Data Fig. 3, and achieved a cryo-EM map at a resolution of 2.76 Å (Table 1). FtAlkBG is wrapped tightly by amphipols that cover approximately half of the protein (Extended Data Fig. 4a–c). Amphipols appear to form a shield that stabilizes the transmembrane region of FtAlkBG during sample vitrification. In addition, the relatively small size of the PMAL-C8 amphipol may be well suited for cryo-EM structure determination of small membrane proteins. The final reconstruction used only 1.6% (46,953) of the total picked particles to reach a resolution of 2.76 Å, estimated using the gold-standard Fourier shell correlation at 0.143 (Extended Data Fig. 4d). These particles have well-distributed orientation angles, with only slightly more particles viewed from the periplasmic side (Extended Data Fig. 4e). The cryo-EM map is of high quality and allowed the building and refinement of an atomic model (Extended Data Fig. 5).

Overall structure

The solved FtAlkBG structure is a monomer (Fig. 1a), consisting of both AlkB (residues 13–387) and AlkG (residues 415–467) domains (Fig. 1b,c), with a total molecular mass of 48.6 kDa. The FtAlkB structure is mostly α-helical, with six transmembrane α-helices (TM1–6), two partially membrane-embedded short α-helices α2 and α3 (Fig. 1d), a long membrane-associated amphiphilic helix α4 (Fig. 1b,d) and six other α-helices outside the membrane (Fig. 1c,d). TM2 is kinked within the membrane, making an angle of about 100°. The protein is narrower on the periplasmic side than on the cytoplasmic side, forming a wedge-shaped structure that may be responsible for bending the membranes to facilitate the formation of membrane vesicles[22]. In terms of sequence, TM1–4 are clustered and separated from TM5 and TM6 by α2–α4. The soluble domain follows TM6 and connects to the FtAlkG domain through a 27-residue, disordered linker (residues 388–414).

The diiron active site

FtAlkB binds two iron ions, which form a diiron center that is essential for C−H bond activation (Fig. 2)[23]. The diiron center is coordinated by nine conserved histidines and one carboxylate residue (Extended Data Fig. 6) with His137, His141, His167, His172 and His314 coordinating Fe1 and four histidines (His171, His272, His311 and His315) and glutamate Glu271 coordinating Fe2. The coordination geometry of Fe1 is octahedral, such that four histidine nitrogen atoms are coplanar with Fe1, analogous to the heme in cytochrome P450 enzymes that can also catalyze C−H bond hydroxylation (Extended Data Fig. 7)[24]. The octahedral geometry of Fe1 allows a fifth ligand on one side of the plane. His141 is on the other side of the plane to further stabilize Fe1. Fe2 is 6.1 Å away from Fe1, in line with Fe1 and His141, but too far to coordinate with Fe1. The four histidines (His171, His272, His311 and His315) are on the same side of the Fe2 site, thus allowing additional coordination to Fe2.

Additional density was present in the active site of the cryo-EM structure, allowing a dodecane substrate (C-12) to be modeled with its terminal methyl group in close proximity of the diiron center (Fig. 2b and Extended Data Fig. 8). Experimentally, FtAlkBG hydroxylates dodecane, as demonstrated by the decrease in NADH fluorescence compared to a nonsubstrate control (Extended Data Fig. 9). No alkanes were provided during protein production or sample preparation. Therefore, the dodecane was likely derived from the E. coli which was used as the expression host for FtAlkBG. The linear dodecane is oriented by hydrophobic residues, such that its terminal methyl group (C(sp^3)−H bonds) is 5.3 Å from Fe1 and 4.8 Å from Fe2. The methyl group is also located 3.2 Å away from Glu271, suggesting a possible role for Glu271 in hydroxylation of the C−H bond (Fig. 2a). The methyl group is not in line with Fe1 and His141. Instead, it is closer to Fe2 and lies on the noncoordinated side of Fe2. The proximity of the substrate terminal methyl group to Fe2 suggests the Fe2 center may be more directly involved in C−H bond activation.

Substrate selectivity

FtAlkBG has a long, hydrophobic substrate binding channel starting from near the middle of transmembrane helices TM1 and TM2 and ending at the diiron center (Fig. 2b). The hydrophobic residues lining the channel show increasing degrees of conservation the nearer they are
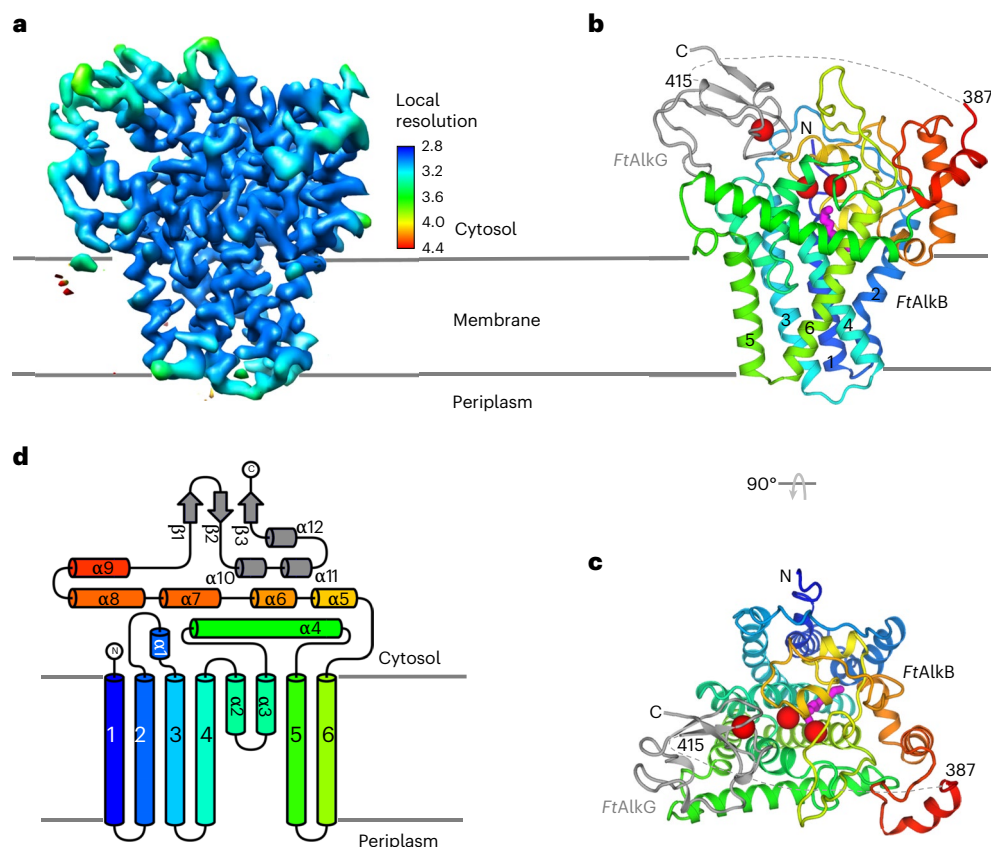
Fig. 1 | Overall structure. a, Cryo-EM density map colored with local resolutions. **b,c**, Two views of the overall structure: side view (**b**) and top view with a 90° rotation from **b** (**c**). *Ft*AlkB is shown as rainbow cartoons, with colors ranging from blue for N terminus to red for C terminus. *Ft*AlkG is shown in gray. Three irons, two in *Ft*AlkB and one in *Ft*AlkG, are shown as red spheres. The substrate dodecane is drawn as magenta sticks. The six transmembrane helices are labeled by numbers 1–6. **d**, Secondary structure topology of the structure. The coloring is the same as in **b**. Six transmembrane helices are indicated by numbers 1–6.

to the diiron site. Dodecane occupies roughly half the tunnel length, suggesting the enzyme could accommodate longer alkanes. Because the tunnel entrance is narrower than the dodecane-occupied region, binding of longer substrates would presumably be accommodated by minor conformational changes. The side chains of residues Ile27, Ile54 and Ala127 form a restriction site. Sequence alignment shows that Ile54 in *Ft*AlkBG is equivalent to Trp55 in AlkB from *P. oleovorans* (*Po*AlkB) (Extended Data Fig. 6). Trp55 in *Po*AlkB determines alkane chain-length selectivity because mutating it to a serine or cysteine extended its alkane chain-length selectivity from C-10 to C-13 (ref. 25). Conversely, AlkB in *Mycobacterium tuberculosis* has a leucine at position 69 (equivalent to Trp55 in *Po*AlkB and Ile54 in *Ft*AlkBG); mutating it to a bulky side chain residue phenylalanine or tryptophan reduces the chain-length selectivity from C-16 to C-11 (ref. 25). Similarly, in AlkBG from *Dietzia cinnamea*, a V91W mutant (equivalent to Ile54 in *Ft*AlkBG) rendered the enzyme less active for alkanes longer than C-9 (ref. 26). Collectively, these hydrophobic residues appear to mediate alkane chain-length selectivity.

Electron transfer mechanism

C–H bond activation requires the transfer of electrons to the Fe–Fe diiron center in AlkB by AlkG[11]. *Ft*AlkG has an [Fe–4S] cluster coordinated by Cys418, Cys421, Cys451 and Cys454 (Extended Data Fig. 10a). *Ft*AlkG sits on a positively charged surface of *Ft*AlkB, adjacent to the diiron center (Fig. 3a). On the surface of AlkG are four negatively charged residues (that is, Glu434, Asp453, Asp458 and Asp461; Extended Data Fig. 10b) that form electrostatic interactions with six positively charged residues (Arg143, Arg148, Arg169, Arg215, Arg284 and Arg287) on the

surface of *Ft*AlkB. However, we did not observe specific salt-bridge interactions between carboxylate–guanidinium pairs in the cryo-EM density. We interpret this to mean that nonspecific electrostatic interactions may facilitate fast on and off rates of *Ft*AlkG association with *Ft*AlkB that would act as a shuttle to relay pairs of electrons one at a time from *Ft*AlkT to *Ft*AlkB. This mechanism is also consistent with the disordering of the linker between *Ft*AlkG and *Ft*AlkB (Fig. 1b).

AlkG alone cannot supply electrons; it functions as a shuttle of electrons originating from its electron donor NADH through NADH oxidation by AlkT, a rubredoxin reductase[12]. Electron transfer between AlkT and AlkG has been proposed on the basis of the structure of a complex between rubredoxin reductase and rubredoxin from *P. aeruginosa* (*Pa*AlkT and *Pa*AlkG, respectively) (Extended Data Fig. 10c)[27]. We superimposed the *Ft*AlkBG and *Pa*AlkG-*Pa*AlkT structures on the basis of AlkG orientation (Extended Data Fig. 10d). We found that AlkB and AlkT bind to the same face on AlkG (Fig. 3a,b), indicating AlkG must dissociate from AlkB to associate with AlkT to enable its reduction and subsequently transfer the electron to AlkB, thus forming a sequential electron transfer model shuttling one electron at a time.

Although there are two irons in *Ft*AlkB, the [Fe–4S] cluster in *Ft*AlkG is closer to Fe1, that is, at 13.6 Å, than to Fe2, at 17.1 Å (Fig. 3c). The unequal distance raises a question as to which iron ion the electrons are transferred. In natural redox proteins, physiological tunneling electron transfer can happen 6–14 Å between redox centers[28]. Therefore, we suggest that, in *Ft*AlkBG, an electron is transferred from the [Fe–4S] cluster to Fe1. Considering the close distance of 6.1 Å between Fe1 and Fe2, Fe1 may transfer an electron to Fe2. Fe1 may function as a redox center to relay electrons from [Fe–4S] in *Ft*AlkG to Fe2 in *Ft*AlkB.
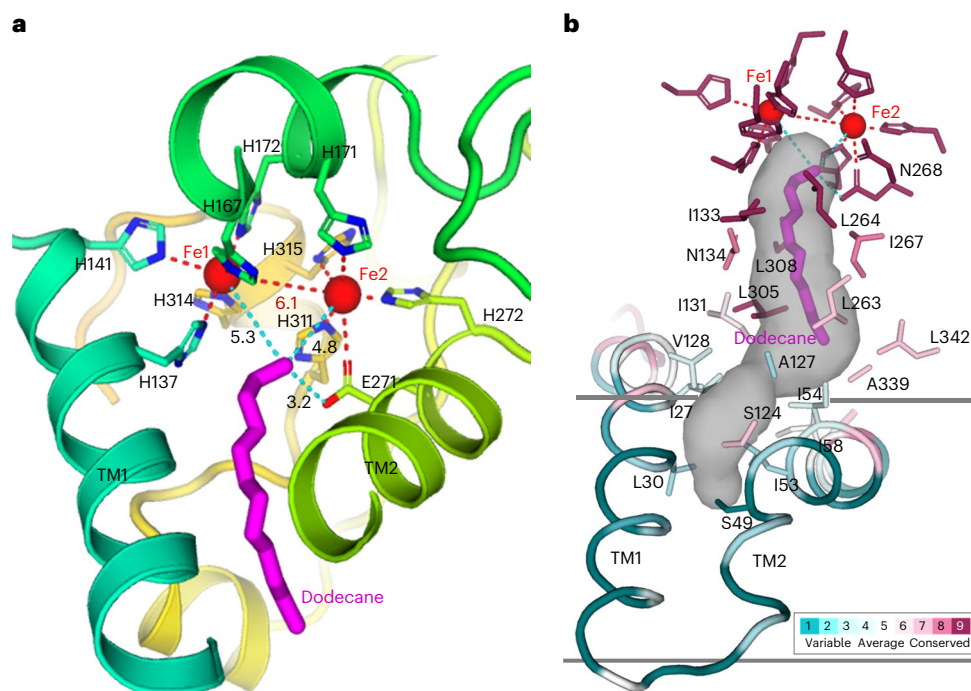
Fig. 2 | Active site structure. a, Structure of the diiron center with a substrate, dodecane, bound. The diiron center (Fe1 and Fe2, red spheres) interacts with nine histidines and one glutamate (sticks) and forms the diiron-center structure. Substrate is shown as magenta sticks, and its distance to the diiron and Glu271 are shown as dashed lines in cyan. **b**, Substrate binding site and substrate entry tunnel. Residues forming the tunnel are shown as sticks and colored by degree of conservation. The substrate, dodecane, is shown as magenta sticks. The entrance of the tunnel is in the membrane between TM1 and TM2.
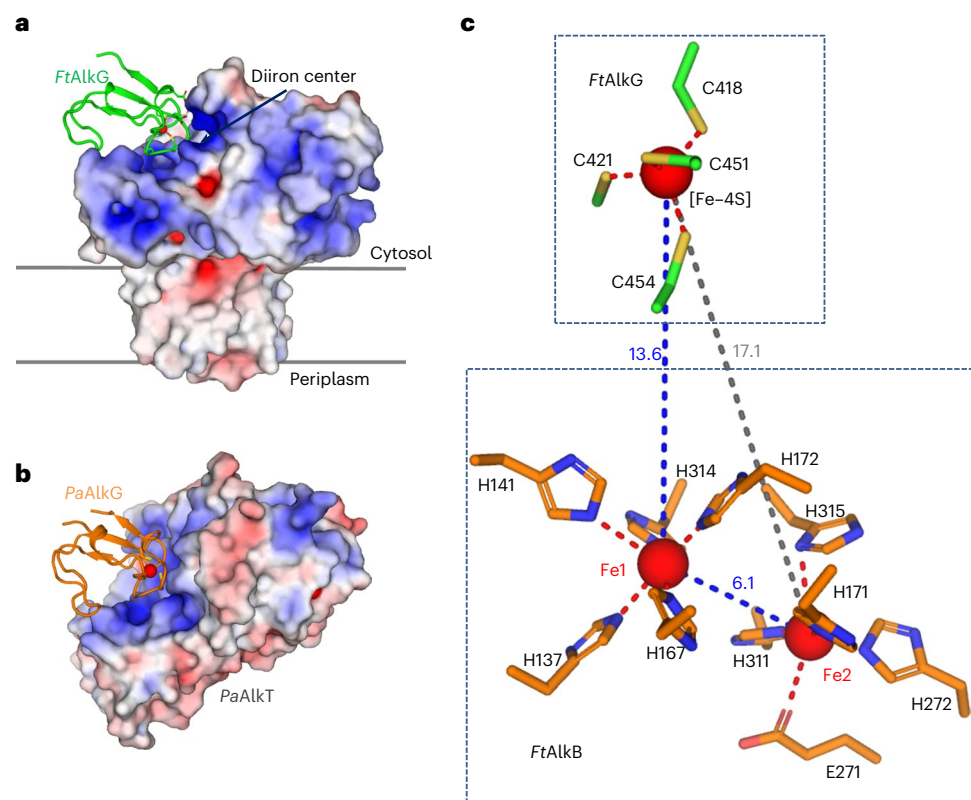
Fig. 3 | Electron transfer mechanism. a, Positioning of *Ft*AlkG relative to *Ft*AlkB for electron transfer. *Ft*AlkB is shown as electrostatic potential surface. The electrostatics was calculated using the program APBS[36] and plotted at the level of ±5 *kT/e*. **b**, Positioning of AlkG relative to AlkT from *Pseudomonas aeruginosa* (*Pa*AlkT and *Pa*AlkG, respectively; PDB 2V3B)[27]. *Pa*AlkT is shown as electrostatic potential surface. *Pa*AlkG is shown as a ribbon with its iron center shown as a red sphere. The electrostatics was calculated using the program APBS and plotted at the level of ±5 *kT/e*. **c**, Geometry of electron transfer between *Ft*AlkG and *Ft*AlkB. The distances between the [Fe–4S] cluster to the two irons are indicated as blue and gray dashed lines. The distance between Fe1 and Fe2 is indicated as a blue dashed line. Residues forming the diiron center site are shown as sticks.
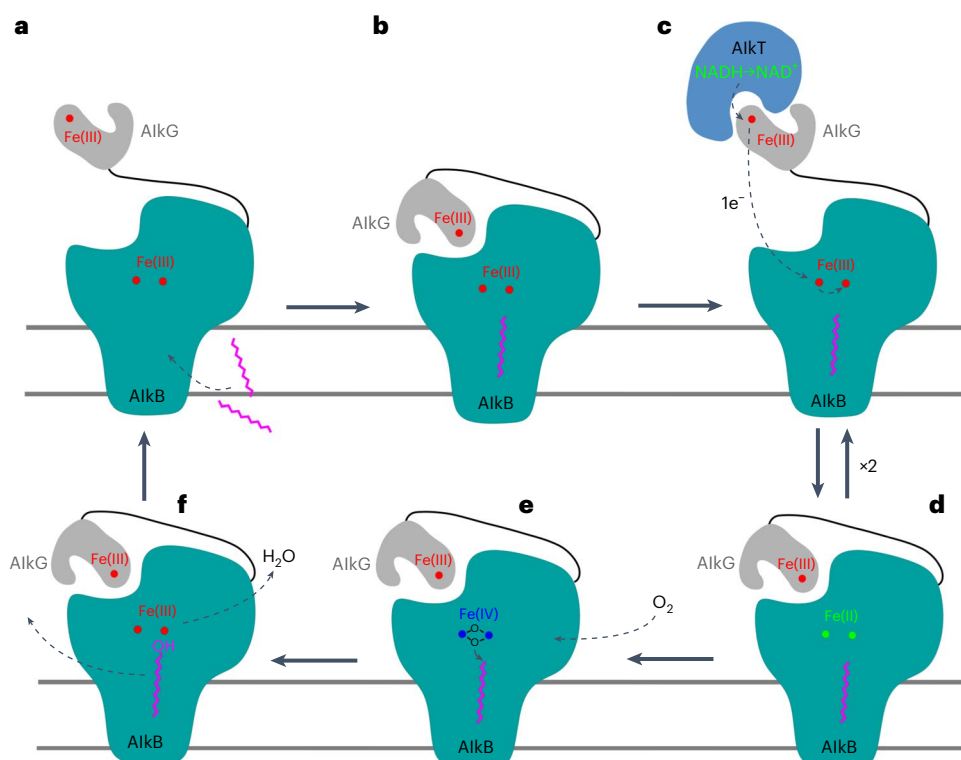
Fig. 4 | A proposed model of alkane terminal C–H functionalization. a, In ambient atmospheric oxygen conditions, both AlkB and AlkG are in an oxidized state with AlkG being dislodged. Fe(III)–Fe(III) is shown as red spheres. **b**, Substrate alkane can diffuse into the substrate entry tunnel in the membranes, and its terminal C–H is oriented toward to the diiron center in AlkB. **c,d**, When AlkT and NADH are available, AlkG shuttles two electrons, one at a time, from oxidation of NADH in AlkT to the diiron center in AlkB. Fe(III)-Fe(III) in oxidized AlkB (**c**) is reduced to Fe(II)–Fe(II) (green spheres) (**d**). **e**, Reduced diiron center activates oxygen to produce an active high-valent diiron-oxo intermediate Fe(IV)–Fe(IV) (blue spheres) that attacks the terminal C–H bond. **f**, The abstraction of hydrogen followed by addition of an OH group completes the two-electron hydroxylation process. The release of product will likely involve conformational changes in AlkB and the release of AlkG.

Discussion

Alkanes are highly hydrophobic and are almost immiscible with water, but they can diffuse into hydrophobic membranes. A structural feature of the transmembrane protein AlkB is its alkane-accessible tunnel with an entrance in the transmembrane region, allowing substrate access. We expressed the protein in the absence of reductant and in the presence of supplemented ferric chloride and purified it under ambient atmospheric oxygen conditions. Our previous Mossbauer spectroscopy findings identified that the 'as-isolated AlkB' from *P. oleovorans* was in the Fe(III)-Fe(III) state[22]; therefore, we postulate that, in an isolated state, the iron ions in the *Ft*AlkBG structure are also in the oxidized Fe(III)-Fe(III) state (Fig. 4a), consistent with the observed long Fe–Fe at 6.1 Å (Fig. 2a). Alkanes diffuse into the membranes where they interact with the tunnel-forming hydrophobic residues that lead toward the diiron center. The shape of the tunnel means that only its terminal C–H bonds can reach the cavity adjacent to the diiron center (Figs. 2b and 4b). Binding of a substrate to the active site triggers conformational changes that induce the docking of AlkG to AlkB (Fig. 4b). When AlkT and NADH are available, AlkG will be reduced and will transfer two electrons derived from oxidation of NADH in AlkT to the diiron center in AlkB. The process requires two single electron transfer cycles from AlkT to AlkG and from AlkG to AlkB (Fig. 4c,d). The fully oxidized Fe(III)-Fe(III) diiron center is not functional because of its long Fe–Fe distance of 6.1 Å. For AlkB to be functional, its diiron site needs to be reduced to Fe(II)-Fe(II) to activate a molecular oxygen (O_2)[22,29,30]. Binding of O_2 to the reduced Fe(II)-Fe(II) diiron site facilitates O_2 activation and, upon cleavage of the O–O bond, an as-yet-uncharacterized high-valent Fe–O intermediate species is created[31]. In the high-valent intermediate, the

Fe–Fe distance would be substantially shorter to form the hypothesized Fe(IV)-Fe(IV)-O_2 diamond-core intermediate structure[32] (Fig. 4e). The diamond core is reactive and abstracts a hydrogen from the terminal C–H bond followed by oxygen rebound[33] to form the terminal hydroxyl group. One of the two oxygen atoms is utilized for the production of the hydroxyl group during substrate oxidation, whereas the other is likely protonated by Glu271 to form a water molecule (Figs. 2a and 4f).

Biocatalysts provide environmentally benign alternatives to high-temperature and high-pressure precious metal-based catalysis for alkane C–H bond activation and functionalization[6,34]. The *Ft*AlkBG structure may provide a structural platform, in a similar way to the P450 system[35], for creating novel biocatalysts for super-selective production of chemical feedstocks and value-added products from abundant alkanes.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-023-00958-0.

References

1. Labinger, J. A. & Bercaw, J. E. Understanding and exploiting C–H bond activation. *Nature* **417**, 507–514 (2002).
2. Hartwig, J. F. & Larsen, M. A. Undirected, homogeneous C–H bond functionalization: challenges and opportunities. *ACS Cent. Sci.* **2**, 281–292 (2016).

3. van Beilen, J. B., Wubbolts, M. G. & Witholt, B. Genetics of alkane oxidation by Pseudomonas oleovorans. *Biodegradation* **5**, 161–174 (1994).

4. Nie, Y., Liang, J., Fang, H., Tang, Y.-Q. & Wu, X.-L. Two novel alkane hydroxylase-rubredoxin fusion genes isolated from a *Dietzia* bacterium and the functions of fused rubredoxin domains in long-chain *n*-alkane degradation. *Appl. Environ. Microbiol.* **77**, 7279–7288 (2011).

5. Witholt, B. et al. Bioconversions of aliphatic compounds by *Pseudomonas oleovorans* in multiphase bioreactors: background and economic potential. *Trends Biotechnol.* **8**, 46–52 (1990).

6. Bordeaux, M., Galarneau, A. & Drone, J. Catalytic, mild, and selective oxyfunctionalization of linear alkanes: current challenges. *Angew. Chem. Int. Ed. Engl.* **51**, 10712–10723 (2012).

7. Rojo, F. Degradation of alkanes by bacteria. *Environ. Microbiol.* **11**, 2477–2490 (2009).

8. van Beilen, J. B. et al. Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the *alk* genes. *Microbiology* **147**, 1621–1630 (2001).

9. van Beilen, J. B., Penninga, D. & Witholt, B. Topology of the membrane-bound alkane hydroxylase of *Pseudomonas oleovorans*. *J. Biol. Chem.* **267**, 9194–9201 (1992).

10. Eggink, G. et al. Alkane utilization in *Pseudomonas oleovorans*. Structure and function of the regulatory locus alkR. *J. Biol. Chem.* **263**, 13400–13405 (1988).

11. Staijen, I. E., van Beilen, J. B. & Witholt, B. Expression, stability and performance of the three-component alkane mono-oxygenase of *Pseudomonas oleovorans* in *Escherichia coli*. *Eur. J. Biochem.* **267**, 1957–1965 (2000).

12. Lee, H. J., Basran, J. & Scrutton, N. S. Electron transfer from flavin to iron in the *Pseudomonas oleovorans* rubredoxin reductase–rubredoxin electron transfer complex. *Biochemistry* **37**, 15513–15522 (1998).

13. Kok, M. et al. The *Pseudomonas oleovorans* alkane hydroxylase gene. Sequence and expression. *J. Biol. Chem.* **264**, 5435–5441 (1989).

14. Owen, D. J. et al. Physical structure, genetic content and expression of the alkBAC operon. *Mol. Gen. Genet.* **197**, 373–383 (1984).

15. Nie, Y. et al. Diverse alkane hydroxylase genes in microorganisms and environments. *Sci. Rep.* **4**, 4968 (2014).

16. Glieder, A., Farinas, E. T. & Arnold, F. H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nat. Biotechnol.* **20**, 1135–1139 (2002).

17. Nieboer, M., Kingma, J. & Witholt, B. The alkane oxidation system of *Pseudomonas oleovorans*: induction of the *alk* genes in *Escherichia coli* W3110(pGEc47) affects membrane biogenesis and results in overexpression of alkane hydroxylase in a distinct cytoplasmic membrane subtraction. *Mol. Microbiol.* **8**, 1039–1051 (1993).

18. Baptist, J. N., Gholson, R. K. & Coon, M. J. Hydrocarbon oxidation by a bacterial enzyme system. I. Products of octane oxidation. *Biochim. Biophys. Acta* **69**, 40–47 (1963).

19. Kusunose, M., Coon, M. J. & Kusunose, E. Enzymatic ω-oxidation of fatty acids. I. Products of octanoate, decanoate, laurate oxidation. *J. Biol. Chem.* **239**, 1374–1378 (1964).

20. Kusunose, M., Kusunose, E. & Coon, M. J. Enzymatic ω-oxidation of fatty acids. II. Substrate specificity and other properties of the enzyme system. *J. Biol. Chem.* **239**, 2135–2139 (1964).

21. Koch, D. J., Chen, M. M., Beilen, J. B. V. & Arnold, F. H. In vivo evolution of butane oxidation by terminal alkane hydroxylases AlkB and CYP153A6. *Appl. Environ. Microbiol.* **75**, 337–344 (2009).

22. Shanklin, J., Achim, C., Schmidt, H., Fox, B. G. & Münck, E. Mössbauer studies of alkane ω-hydroxylase: evidence for a diiron cluster in an integral-membrane enzyme. *Proc. Natl Acad. Sci. USA* **94**, 2981–2986 (1997).

23. Shanklin, J., Whittle, E. & Fox, B. G. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* **33**, 12787–12794 (1994).

24. Williams, P. A. et al. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* **305**, 683–686 (2004).

25. van Beilen, J. B. et al. Identification of an amino acid position that determines the substrate range of integral membrane alkane hydroxylases. *J. Bacteriol.* **187**, 85–91 (2005).

26. Williams, S. C. et al. Investigation of the prevalence and catalytic activity of rubredoxin-fused alkane monooxygenases (AlkBs). *J. Inorg. Biochem.* **219**, 111409 (2021).

27. Hagelueken, G. et al. Crystal structure of the electron transfer complex rubredoxin–rubredoxin reductase of *Pseudomonas aeruginosa*. *Proc. Natl Acad. Sci. USA* **104**, 12276–12281 (2007).

28. Page, C. C., Moser, C. C., Chen, X. & Dutton, P. L. Natural engineering principles of electron tunnelling in biological oxidation–reduction. *Nature* **402**, 47–52 (1999).

29. Shanklin, J. & Whittle, E. Evidence linking the *Pseudomonas oleovorans* alkane ω-hydroxylase, an integral membrane diiron enzyme, and the fatty acid desaturase family. *FEBS Lett.* **545**, 188–192 (2003).

30. Smits, T. H. M., Balada, S. B., Witholt, B. & van Beilen, J. B. Functional analysis of alkane hydroxylases from gram-negative and gram-positive bacteria. *J. Bacteriol.* **184**, 1733–1742 (2002).

31. Cooper, H. L. R. et al. Parallel and competitive pathways for substrate desaturation, hydroxylation, and radical rearrangement by the non-heme diiron hydroxylase AlkB. *J. Am. Chem. Soc.* **134**, 20365–20375 (2012).

32. Shu, L. J. et al. An $Fe_2^{IV}O_2$ diamond core structure for the key intermediate Q of methane monooxygenase. *Science* **275**, 515–518 (1997).

33. Groves, J. T. & McClusky, G. A. Aliphatic hydroxylation via oxygen rebound. Oxygen transfer catalyzed by iron. *J. Am. Chem. Soc.* **98**, 859–861 (1976).

34. Urlacher, V. B. & Girhard, M. in *Enzyme Catalysis in Organic Synthesis* 3rd edn (eds Drauz, K., Gröger, H. & May, O.) Ch. 1 (Wiley, 2012).

35. Peters, M. W., Meinhold, P., Glieder, A. & Arnold, F. H. Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3. *J. Am. Chem. Soc.* **125**, 13442–13450 (2003).

36. Jurrus, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128 (2018).

Methods

Protein expression and purification

The gene encoding the full length of *Ft*AlkBG was synthesized and cloned into the pET16b expression vector by GenScript. The vector contains an N-terminal 10×His tag followed by a tobacco etch virus (TEV) cleavage site. Protein expression was performed in *Escherichia coli* BL21-Gold (DE3) pLysS cells (Agilent) growing in LB medium with an induction of 0.4 mM IPTG at 28 °C for 4.5 h. One milliliter of 0.1 M $FeCl_3$ in 0.12 M HCl was added to every 1 L of culture 15 min before induction. Cells were harvested by centrifugation at 5,000*g* for 10 min at 4 °C, and were resuspended in lysis buffer containing 30 mM HEPES, pH 7.5, 150 mM NaCl and 2 mM $MgCl_2$. Cells were lysed using an EmulsiFlex-C3 homogenizer (Avestin) at 15,000 psi. Cell lysates were cleaned by centrifuging at 25,000*g* for 25 min at 4 °C, and the supernatants were collected for pelleting membranes by ultracentrifugation at 360,000*g* for 2 h at 4 °C. The membranes were frozen in liquid nitrogen and stored at −80 °C until further use.

Membranes were resuspended in solubilization buffer (30 mM HEPES, pH 7.5, 150 mM NaCl, 20% glycerol (wt vol⁻¹), 0.5 mM Tris(2-carboxyethyl)phosphine (TCEP), 8% DDM (wt vol⁻¹) and Roche proteases inhibitor cocktail). After solubilization for 4 h at 4 °C, insoluble components were removed by ultracentrifugation at 360,000*g* for 30 min at 4 °C. The supernatant was applied to pre-equilibrated Ni^{2+}-nitrilotriacetic acid (NTA) column (Anatrace) and washed by 8 column volumes of wash buffer (30 mM HEPES, pH 7.5, 150 mM NaCl, 20% glycerol, 0.25 mM TCEP, 50 mM imidazole, 0.05% DDM). The protein was eluted by 500 mM imidazole and digested by TEV protease overnight at 8 °C. The His tag-removed protein was clarified by ultracentrifugation at 190,000*g* for 30 min at 4 °C, desalted by Econo-Pac 10DG desalting column (Bio-Rad) and passed through the pre-equilibrated Ni^{2+}-NTA column. The flow-through was concentrated to 10 mg ml⁻¹ and further purified by size-exclusion chromatography using Superdex 200 Increase (GE Healthcare) in buffer containing 20 mM HEPES, pH 7.5, 100 mM NaCl, 5% glycerol, 0.25 mM TCEP and 0.01% DDM. Fractions containing the protein were pooled and concentrated to 10 mg ml⁻¹ using a 50-kDa molecular cutoff concentrator (Millipore Sigma).

Amphipol reconstitution preparation

Reconstitution of detergent-solubilized *Ft*AlkBG (10 mg ml⁻¹) into amphipol nanodiscs was performed on the basis of the previously published method[37]. Purified protein in a buffer (20 mM HEPES, pH 7.5, 100 mM NaCl, 5% glycerol, 0.25 mM TCEP, 0.01% DDM (wt vol⁻¹)) was mixed with amphipol PMAL-C8 (catalog no. P5008, Anatrace) in a protein:amphipol mass ratio of 1:5 (wt wt⁻¹). The mixture was incubated at 4 °C for 20 h and concentrated using a 50-kDa molecular cutoff concentrator. Then, SM-2 Bio-Beads (40 mg wet; catalog no. 1523920, Bio-Rad) were added to the concentrated mixture, and the mixture was nutated at 4 °C for 3 h. Before use, Bio-Beads were soaked with buffer containing 20 mM HEPE, pH 7.6, 100 mM NaCl overnight at 4 °C. The Bio-Beads were removed by passing through a Micro Bio-Spin chromatography column. The flow-through containing the reconstituted particles was centrifuged at 20,000*g* for 1 h at 4 °C; the supernatant was used for size-exclusion chromatography through a Superdex 200 Increase column (GE Healthcare) in a buffer containing 20 mM HEPES, pH 7.6, 100 mM NaCl. The fractions containing the *Ft*AlkBG–PMAL-C8 nanodiscs were collected and concentrated to 5 mg ml⁻¹ for cryo-EM analyses.

Cryo-EM sample preparation and data collection

Three microliters of the *Ft*AlkBG–PMAL-C8 nanodiscs were applied to a glow-discharged (15 mA current for 15 s), 300-mesh *R* 0.6/1 UltrAuFoil Holey Gold grid (catalog no. Q350AR1A, Electron Microscopy Sciences). After waiting for 60 s, vitrification was performed using a Mark IV Vitrobot (Thermo Fisher Scientific) with a blotting condition of 5 s blot time, 0 blot force and 100% humidity at 6 °C.

Single-particle cryo-EM data were collected at the Laboratory for Biomolecular Structure (LBMS) facility at Brookhaven National Laboratory using a Thermo Fisher Titan Krios electron microscope (G3i) equipped with a Gatan K3 camera and a BioQuantum energy filter. With a physical pixel size of 0.666 Å (0.333 Å in super-resolution mode), a total dose of 60 e⁻ Å⁻² was fractioned to 50 frames under the super-resolution mode using the Thermo Fisher data acquisition program EPU. A total of 11,757 movies were collected with an energy filter width of 15 eV throughout the data acquisition. Data collection statistics are listed in Table 1.

Cryo-EM data processing

Dose-fractioned movies were corrected and averaged using Motion-Corr2 (ref. 38) with a bin factor of two. Averaged movie frames were further corrected by contrast transfer function estimation using Gctf[39]. Micrographs with an estimated resolution better than 4.5 Å were selected for particle picking. Particles were initially picked and cleaned up by 2D class averaging in Relion3 (ref. 40). The cleaned particles were then used for training and particle picking using Kpicker[41]. A total of 2,950,051 particles were picked, extracted at 256 pixels and binned to 64 pixels with a pixel size of 2.664 Å.

We used CryoSPARC[42] for additional 2D class averaging, which resulted in 1,738,105 selected particles. With four classes each, selected particles were used for three cycles of ab initio reconstruction followed by three-dimensional (3D) heterogenous refinements with a particle size of 64 pixels (2.664 Å). The process selected 369,194 particles, which were centered and re-extracted at 256 pixels and binned to 128 pixels with a pixel size of 1.332 Å (Extended Data Fig. 3).

Re-extracted particles were auto-refined to convergence in Relion3, followed by nonalignment 3D classification into four classes (Extended Data Fig. 3). Particles from the 3D class with the best structural feature (α-helices and side chains) as visualized in Chimera[43] were selected. These particles (13%) from two such classes were selected for contrast transfer function refinement and Bayesian polishing in Relion3, followed by 2D classification and nonuniform refinement in CryoSPARC to reach a refined reconstruction at 2.76 Å. Local resolutions were estimated using BlocRes[44]. Reconstruction statistics are listed in Table 1.

Model building and refinement

To boost side chain features of the map, the masked and filtered cryo-EM map was sharpened using PHENIX[45] with a *B* factor of −66 Å². COOT[46] was used to build the atomic model, including irons and a substrate dodecane. All refinements were performed in real space in PHENIX. The refined model was validated using MolProbity[47] and the refinement statistics are listed in Table 1.

Model visualization

The cryo-EM density map and atomic models were visualized using either PyMOL (http://www.pymol.org/) or Chimera[43].

Expression of AlkT

AlkT is a flavin adenine dinucleotide-dependent rubredoxin reductase that transfers electrons from NADH to *Ft*AlkBG for C−H activation and functionalization[10,12]. The gene encoding AlkT from *P. oleovorans* (*Po*AlkT) was amplified using PCR and cloned into the pETDuet expression vector. Protein expression was performed in *E. coli* BL21-Gold (DE3) pLysS cells (Agilent) growing in TB medium with an induction of 0.4 mM IPTG at 16 °C for 17 h. Cells were harvested by centrifugation at 5,000*g* for 10 min at 4 °C, and were then resuspended in lysis buffer containing 50 mM Tris-HCl, pH 8.0, 100 mM KCl, 16% glycerol, 1 mM TCEP, 2 mM K_3PO_4 and 2 mM $MgCl_2$. Cells were lysed by passing through a French press twice at 13,000 psi. Cell lysates were cleaned up by centrifugation at 25,000*g* for 30 min at 4 °C, and the supernatants were applied to pre-equilibrated Ni^{2+}-NTA resin in a buffer containing

50 mM Tris-HCl, pH 8.0, 100 mM KCl, 16% glycerol, 0.5 mM TCEP, 2 mM K_3PO_4 and 20 mM imidazole. After a wash with 10 column volume of wash buffer (50 mM Tris-HCl, pH 8.0, 100 mM KCl, 16% glycerol, 0.5 mM TCEP, 2 mM K_3PO_4, 5 mM ATP, 10 mM $MgCl_2$ and 50 mM imidazole), the protein was eluted with the lysis buffer supplemented with 500 mM imidazole. The protein buffer was changed to 50 mM Tris-HCl, pH 8.0, 200 mM KCl, 16% glycerol, 2 mM K_3PO_4, 0.5 mM EDTA and 1 mM TCEP using a concentrator. Purified protein was frozen in liquid nitrogen and stored at −80 °C until further use.

Enzyme activity

For the activity assay, *Ft*AlkBG was overexpressed in *E. coli* BL21 Star cells (Thermo Fisher Scientific). Cells grew in LB medium supplemented with 100 μM $FeCl_3$. After growing for 14–19 h at 16 °C, cells were harvested by centrifugation at 5,000g for 10 min at 4 °C. Cells were resuspended in lysis buffer (30 mM Tris-HCl, pH 7.75, 150 mM NaCl, 0.2 mM PMSF, 100 μM ferrous ammonium sulfate hexahydrate and 100 μM sodium hydrosulfite) and lysed using a French press at 13,000 psi. Lysates were centrifuged at 10,000g for 25 min at 4 °C. Pellets were resuspended in lysis buffer without PMSF, frozen in liquid nitrogen and stored at −80 °C until further use.

The dodecane assay was performed on the basis of a modified protocol from McKenna and Coon[31,48] using a Spark 20 M multimode microplate reader (Tecan). In a reaction volume of 100 μl, we added 20 μg *Ft*AlkBG, 10 μg *Po*AlkT and 500 μM NADH in a buffer of 50 mM Tris-HCl, pH 7.5, 100 mM NaCl. The reaction was initiated by addition of 400 μM dodecane dissolved in acetone at 24 °C. The reaction was monitored by measuring the decrease of NADH fluorescence (excitation/emission at 360/460 nm with a bandwidth of 10 nm). Consumption of NADH, that is, decreased NADH fluorescence, is proportional to dodecane hydroxylation to dodecanol. For the control experiment, we added all components except we used acetone instead of dodecane.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The 3D cryo-EM density map has been deposited in the Electron Microscopy Data Bank under the accession number EMD-28890. Atomic coordinates have been deposited in the Protein Data Bank under the accession number PDB 8F6T. Source data are provided with this paper.

References

37. Sauer, D. B. et al. Structural basis for the reaction cycle of DASS dicarboxylate transporters. *eLife* **9**, e61350 (2020).
38. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
39. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
40. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
41. McSweeney, D. M., McSweeney, S. M. & Liu, Q. A self-supervised workflow for particle picking in cryo-EM. *IUCrJ* **7**, 719–727 (2020).
42. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
43. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
44. Vilas, J. et al. Local resolution estimates of cryoEM reconstructions. *Curr. Opin. Struct. Biol.* **64**, 74–78 (2020).
45. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
46. Casañal, A., Lohkamp, B. & Emsley, P. Current developments in *Coot* for macromolecular model building of electron cryo-microscopy and crystallographic data. *Protein Sci.* **29**, 1055–1064 (2020).
47. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
48. Mckenna, E. J. & Coon, M. J. Enzymatic ω-oxidation. IV. Purification and properties of ω-hydroxylase of *Pseudomonas oleovorans*. *J. Biol. Chem.* **245**, 3882–3889 (1970).

Author contributions

Q.L. and J.S. designed the study and experiments. J.C., G.G. and Q.L. performed the experiments. J.C., S.M.M., J.S. and Q.L. analyzed the data. Q.L. and J.S. wrote the paper with help from other coauthors.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41594-023-00958-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-023-00958-0.

Correspondence and requests for materials should be addressed to John Shanklin or Qun Liu.

Peer review information *Nature Structural & Molecular Biology* thanks Wolf-Dieter Schubert and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Florian Ullrich and Katarzyna Ciazynska, in collaboration with the *Nature Structural & Molecular Biology* team.

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

Extended Data Fig. 1 | Organization of the components of the AlkBGT biocatalysts. The three components (AlkB, AlkG, and AlkT) in the alkane hydroxylation biocatalyst are encoded as separate genes in *P. oleovorans*. AlkG is an Fe-sulfur cluster-containing redox protein that transfers electrons to the diiron center in AlkB. AlkT is a flavin adenine dinucleotide (FAD)-dependent rubredoxin reductase that transfers two electrons from NADH singly to the Fe-sulfur cluster of AlkG, which then relays two of them singly to AlkB to enable biocatalysis[10,12]. The membrane-bound AlkB is a nonheme diiron-containing monooxygenase that specifically oxidizes the terminal methyl group of alkanes to produce the corresponding primary alcohols[8,10,13,14,17,22,30]. There are microbes that have evolved to have the AlkB and AlkG fused together, presumably for improved electron transfer[15,26]. Additionally, in some species (Leptospira, Limnobacter, and Polaromonas), AlkB is fused with multiples proteins of an N-terminal ferredoxin and a ferredoxin reductase[15].

a

b

c

d
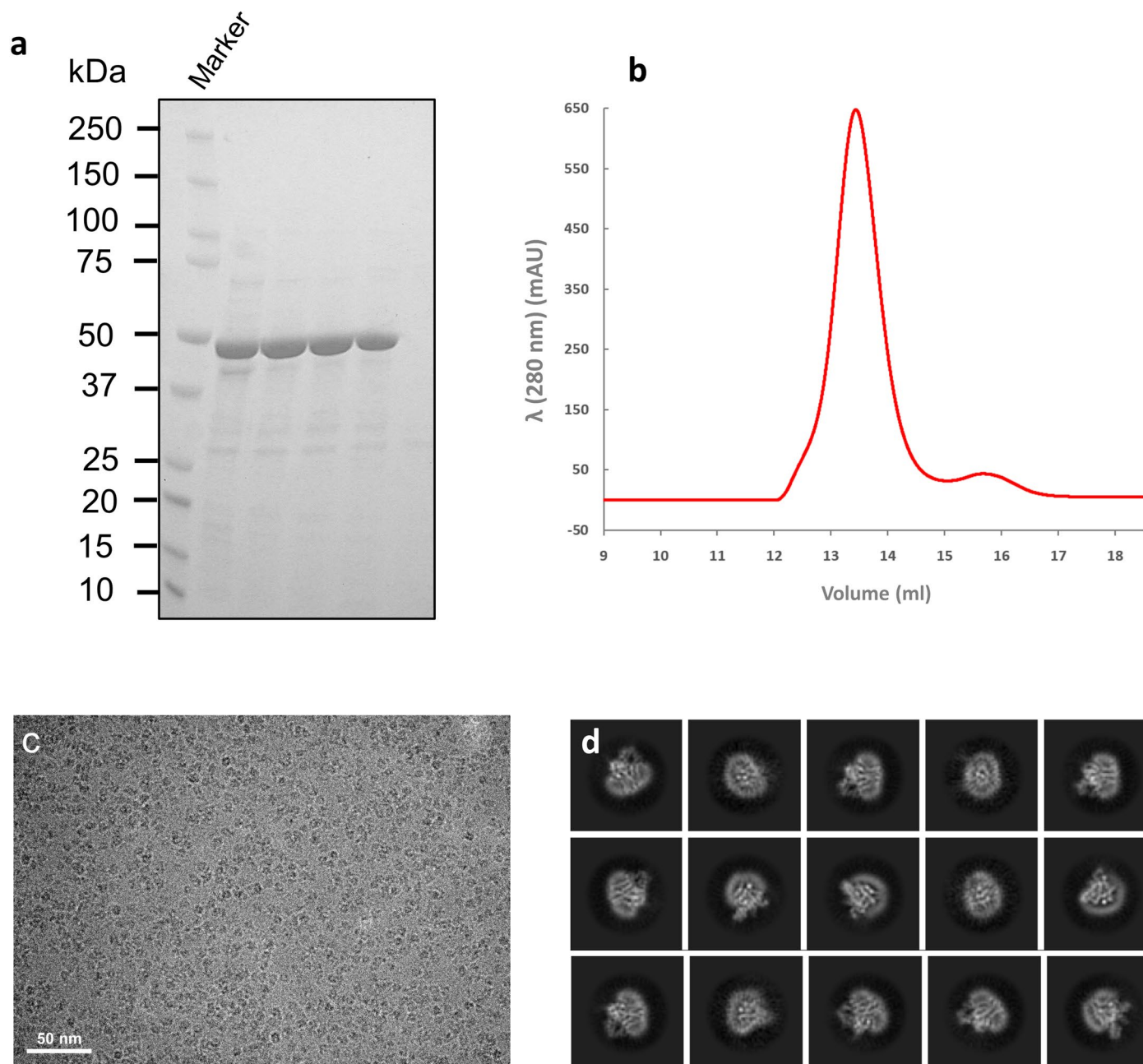
Extended Data Fig. 2 | Production and analysis of reconstituted *Ft*AlkBG particles. a, SDS-PAGE gel analysis of purified *Ft*AlkBG in detergent DDM. This experiment was reproducible and was repeated for at least three times. **b**, Size-exclusion chromatography profile of *Ft*AlkBG reconstituted into PMAL-C8 nanodiscs. **c**, A typical motion-corrected cryo-EM micrograph. We have collected a total of 11,757 cryo-EM micrographs of similar particle size and distribution. **d**, 2D class averages showing secondary structure features of *Ft*AlkBG with its transmembrane region wrapped by PMAL-C8.
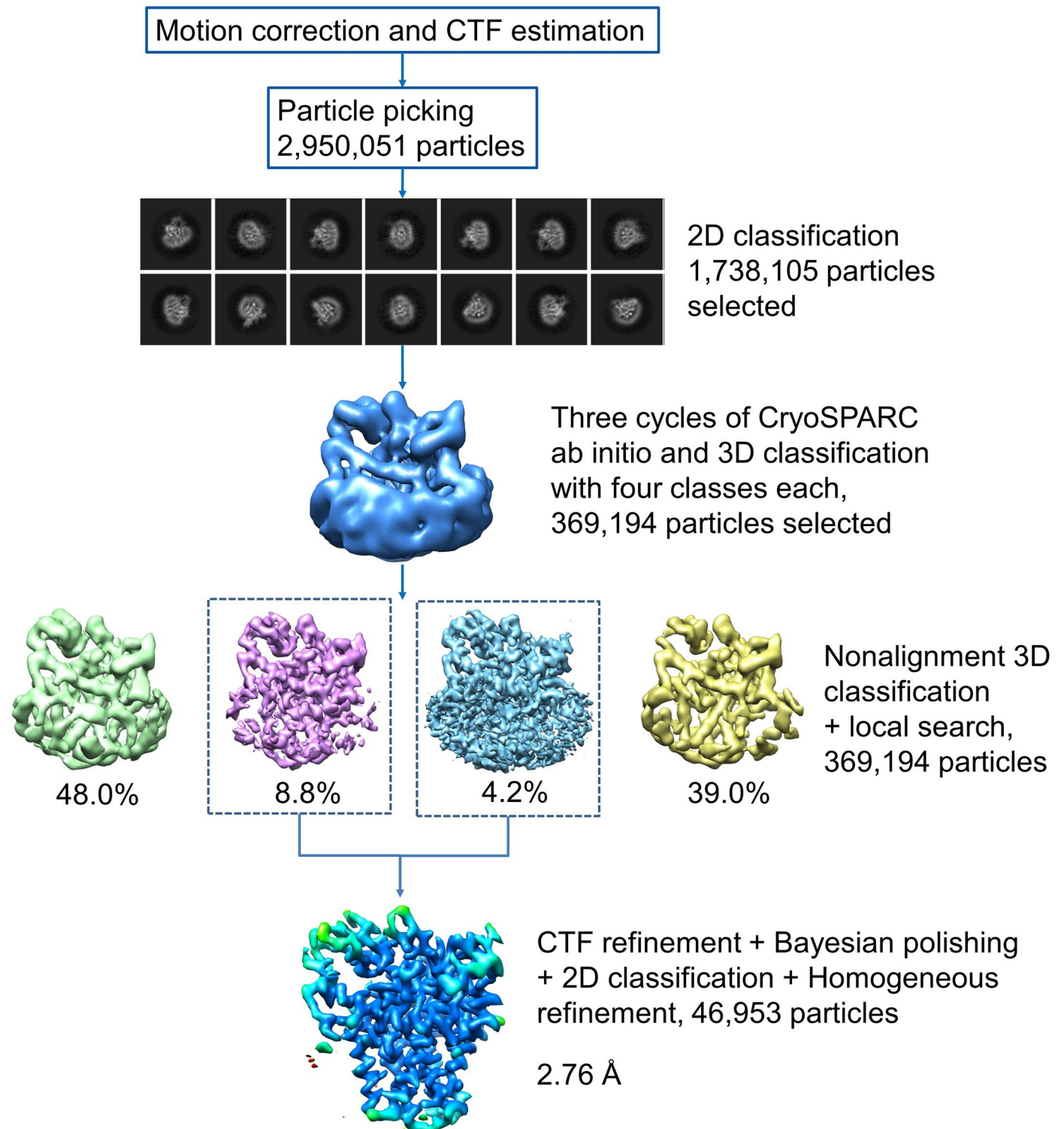
Extended Data Fig. 3 | Flow chart of cryo-EM data processing. The flow chart includes motion correction and CTF estimation, 2D and 3D class averages, and 3D refinement.
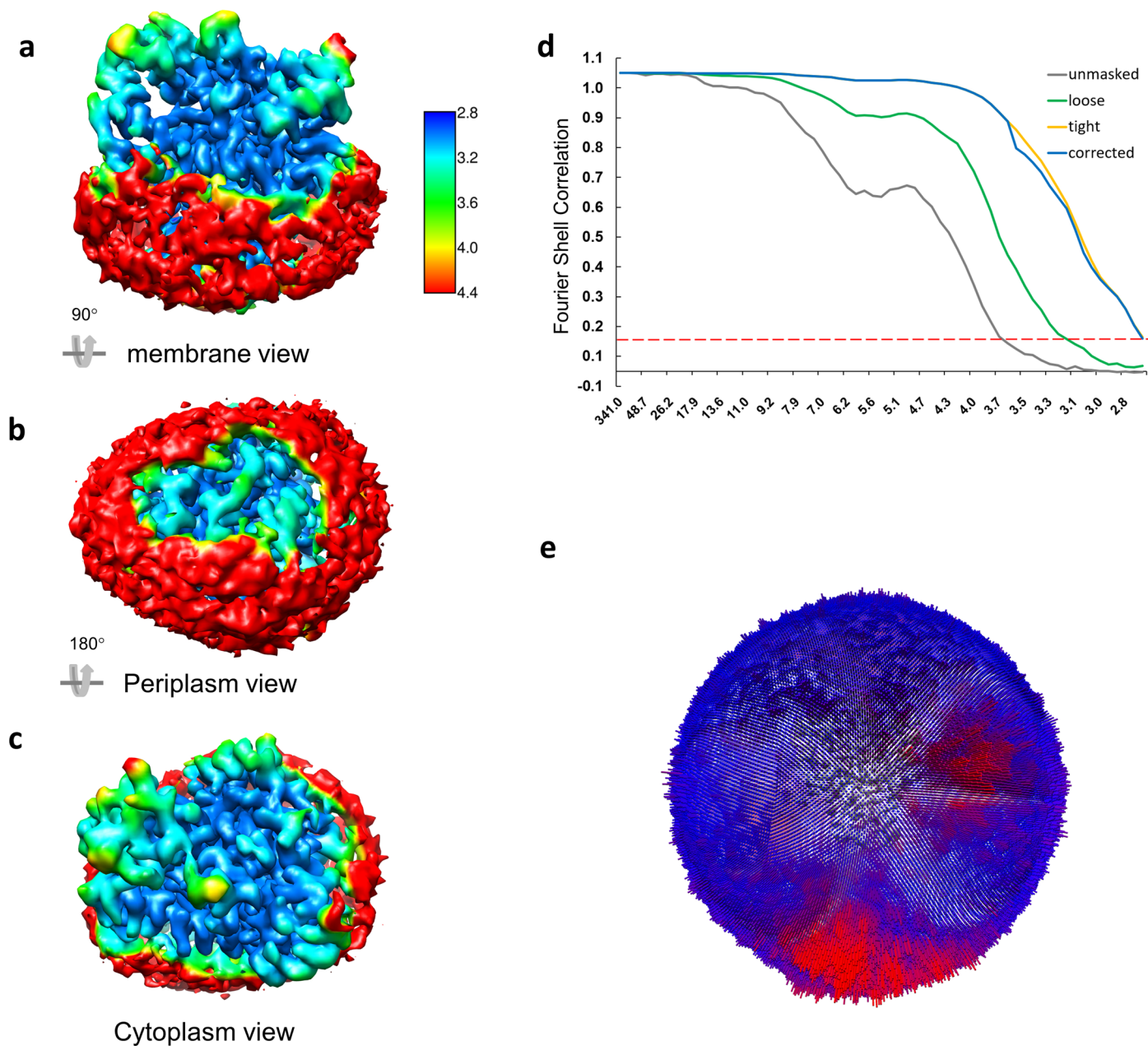
Extended Data Fig. 4 | 3D reconstruction of *Ft*AlkBG. a-c, Three views for the reconstructed map colored with local resolutions. Red belts are amphipol densities around the transmembrane regions of the complex. **d**, Fourier Shell Correlation (FSC) curves for the 3D reconstruction. **e**, Angular distribution for particles used in the final 3D reconstruction.
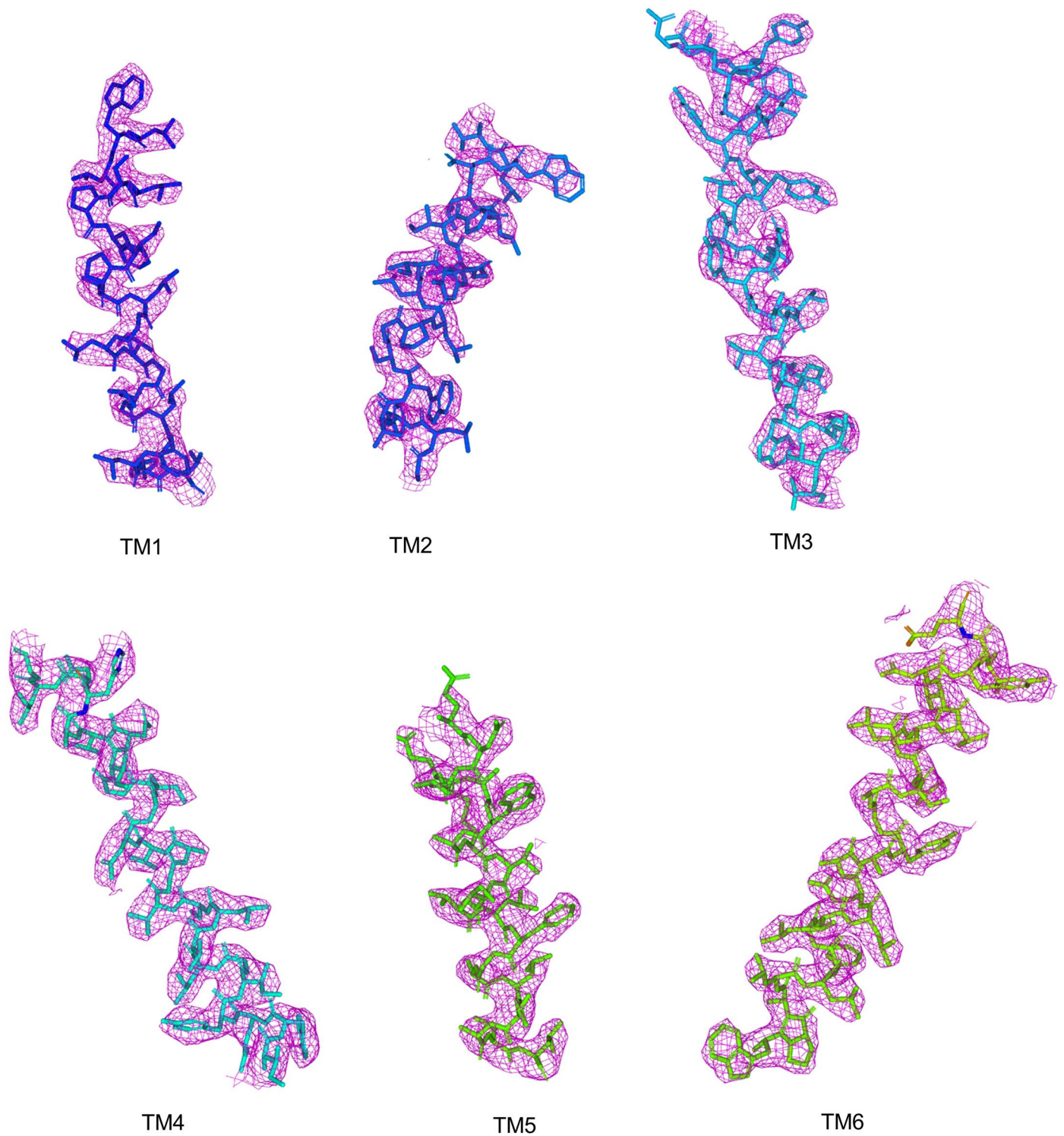
Extended Data Fig. 5 | Quality of the cryo-EM maps. Cryo-EM densities for the six transmembrane regions (TM1-6) are shown as magenta isomeshes. The six transmembrane helices are shown as sticks.
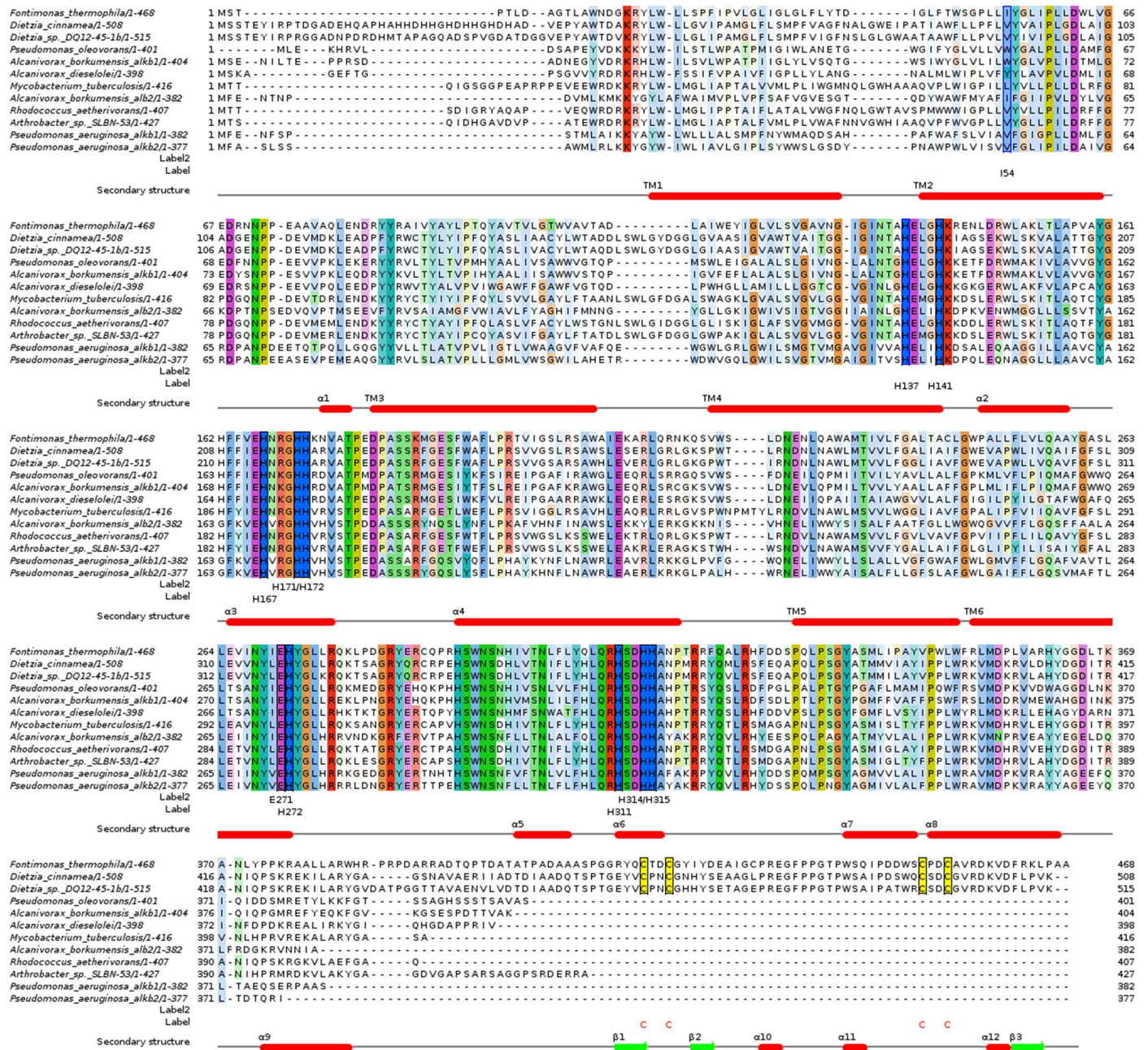
Extended Data Fig. 6 | Structure-based sequence alignments of AlkBs. Nine conserved histidine and one glutamate residues are indicated. Residue I54 and its equivalent residues in other AlkBs are also indicated. Conserved cysteine residues in three AlkBG fusions are highlighted in yellow.
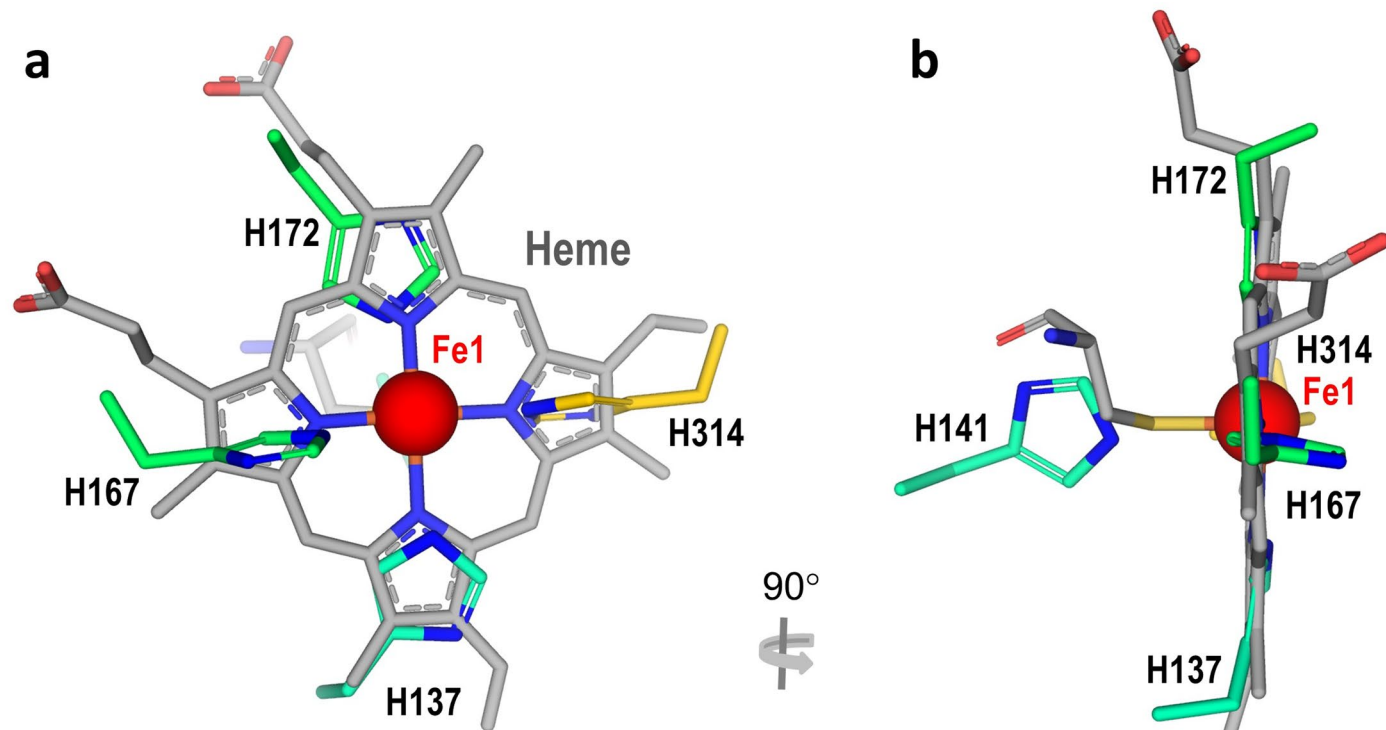
Extended Data Fig. 7 | Superimposition of the Fe1 site with heme in cytochrome P450. a, The Fe1-center was superimposed with the heme (gray sticks) in a cytochrome P450 monooxygenase (PDB code 6H1T). The four nitrogen atoms on four histidines are in the same plane as the heme and align well with the four porphyrin inner nitrogen atoms. **b**, A view of 90° rotation of (**a**). His141 is behind the Fe1-hisitine plane to further stabilize Fe1.
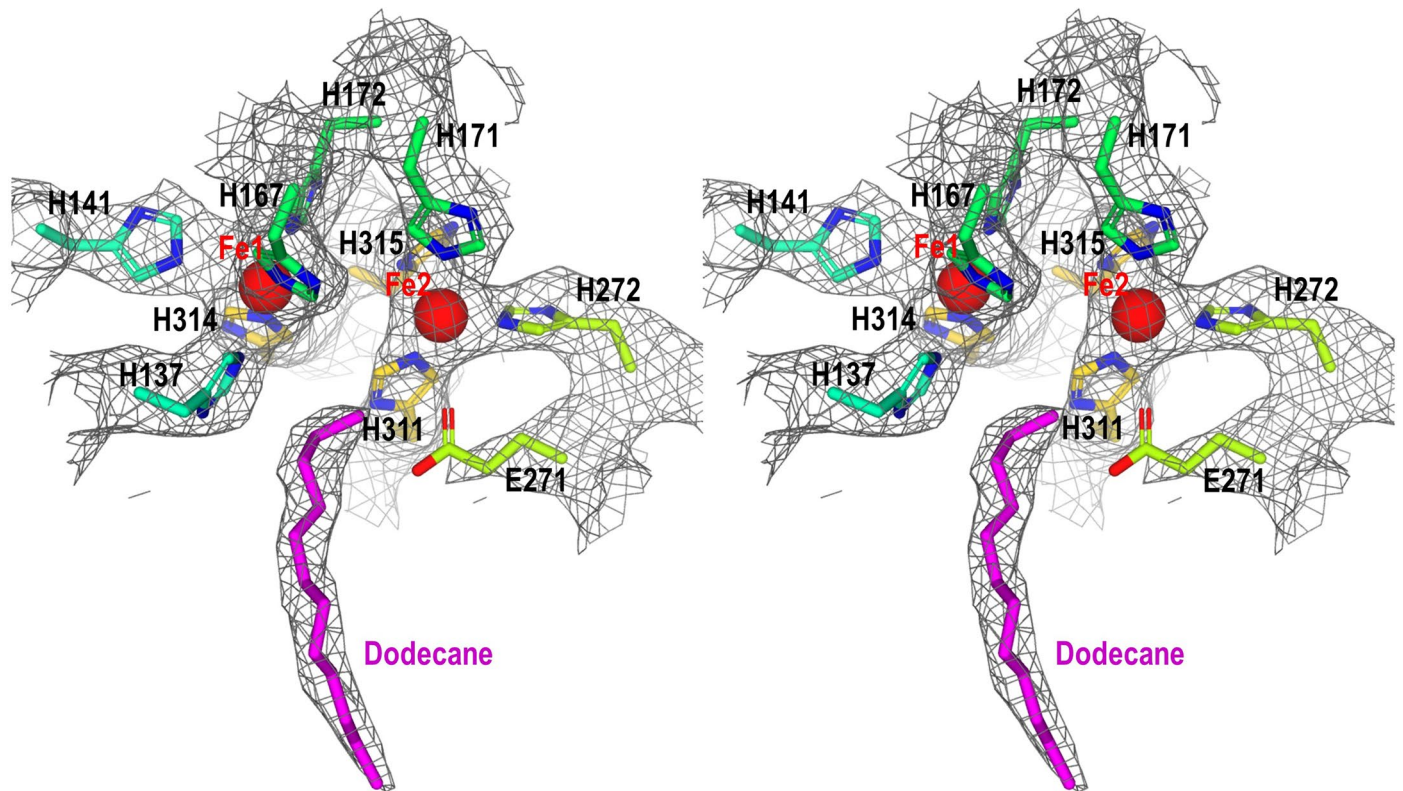
Extended Data Fig. 8 | A stereo view of cryo-EM densities for the active site. The densities are shown as gray isomeshes covering two irons (red spheres), substrate (magenta sticks), nine histidines (sticks) and one glutamate (sticks).
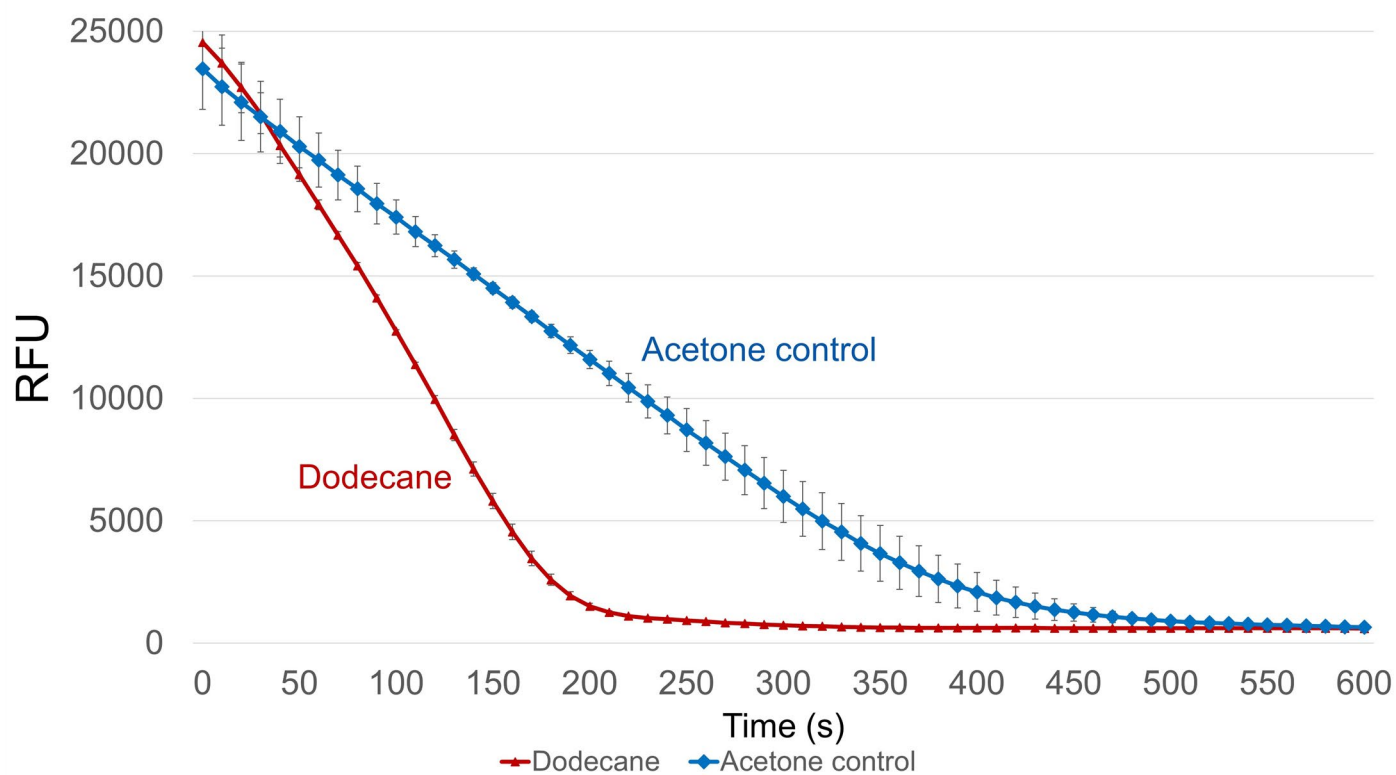
Extended Data Fig. 9 | *Ft*AlkBG activity measured by NADH fluorescence intensity. The consumption of NADH was monitored by NADH fluorescence with ex/em of 360/460 nm using a plate reader. Conversion of dodecane to dodecanol resulted in the increased consumption of NADH, while the addition of acetone control decreased the NADH fluorescence intensity slowly. RFU is relative fluorescence unit. The error bar is the standard deviation (s.d.) of the mean from three independent experiments.
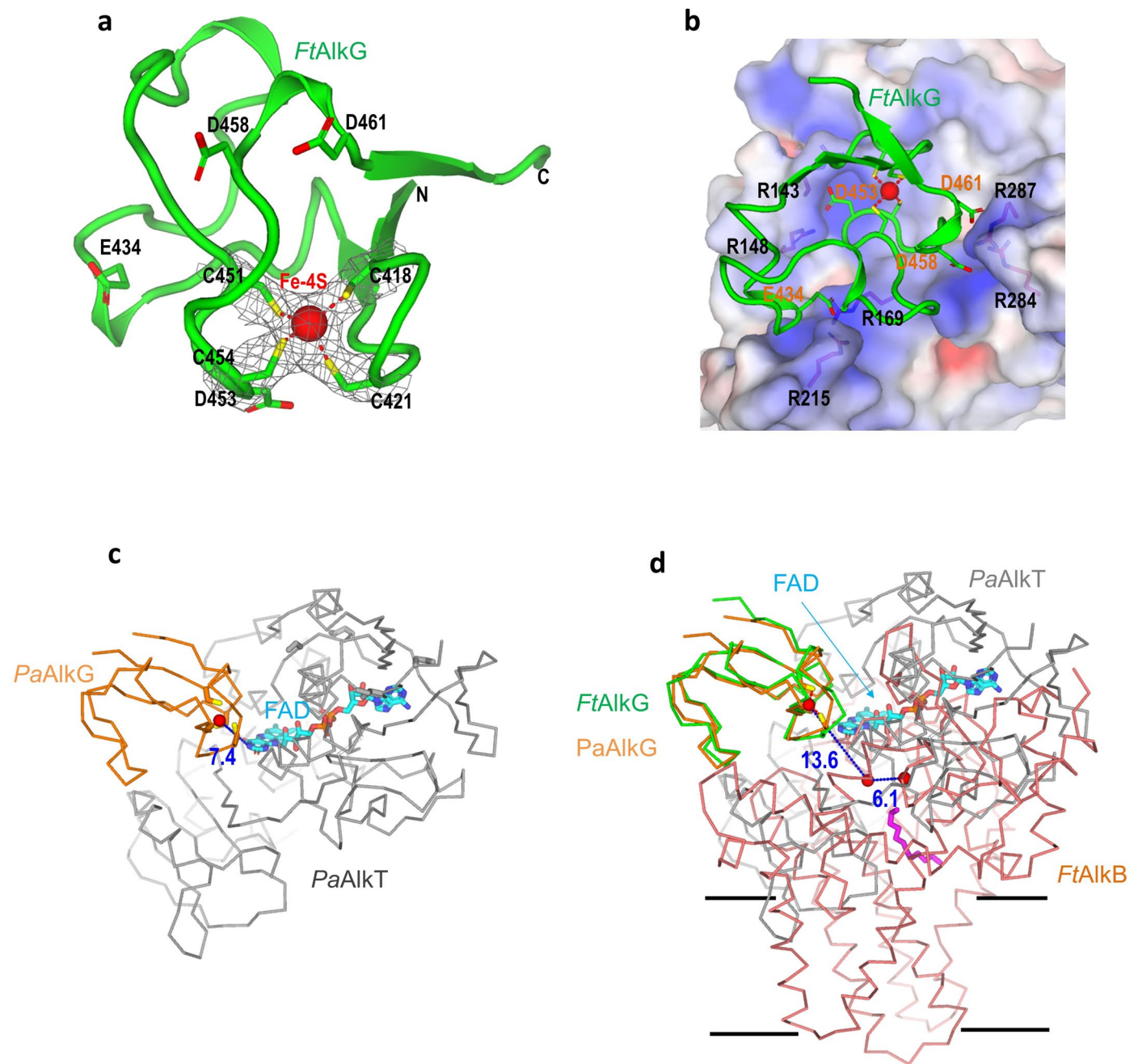
Extended Data Fig. 10 | AlkG in electron transfer. a, *Ft*AlkG structure and its [Fe-4S] iron-sulfur cluster. The iron is shown as a red sphere and the four cysteine residues are shown as sticks. Four negatively charged residues are also shown as sticks. Cryo-EM density for the iron-sulfur cluster was shown as gray isomeshes. **b**, Electrostatic interactions between *Ft*AlkB and *Ft*AlkG. Positively charged arginine residues on AlkB and negatively charged glutamate or aspartate residues are shown as sticks. **c**, Electron transfer between *Pa*AlkT and *Pa*AlkG at a distance of 7.4 Å from FAD to the iron center. **d**, AlkG on and off from AlkT and

AlkB for electron transfer. We superimposed the *Ft*AlkBG structure with *Pa*AlkG-AlkT complex. The sequence identity between *Ft*AlkG and *Pa*AlkG is 54.5%. The superimposition was performed on the *Ft*AlkG and *Pa*AlkG with an R.M.S.D. of 0.932 Å for 49 aligned Cα atoms. In the superimposed structure, AlkT and AlkB bind to the same AlkG site to transfer and receive an electron, respectively. *Ft*AlkB, salmon; *Ft*AlkG, green; *Pa*AlkT, gray; *Pa*AlkG, orange. FAD molecule in *Pa*AlkT is drawn as sticks. The three irons in the *Ft*AlkBG are shown as red spheres. Substrate dodecane is shown as magenta sticks.

Corresponding author(s): Qun Liu

Last updated by author(s): Feb 5, 2023

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	ThermoFisher EPU 1.8
Data analysis	Relion 3.08, CryoSPARC 2.15, MotionCorr2 1.4.4, Gctf v1.18, Coot 0.9.5, Chimera 1.14, Phenix 1.18.2, PyMOL 2.6.0a0, Molprobity 1.18.2-3874

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All relevant data are available from the authors and/or are included with the manuscript or Supplementary Information. Atomic coordinates and the cryo-EM density map have been deposited in the Protein Data Bank (PDB) (access code 8F6T) and the Electron Microscopy Data Bank (EMDB) (accession code EMD28890), respectively.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. For enzyme activity assays, the sample sizes were chosen so that the maximum relative fluorescence unit (RFU) is less than 30000. For cryo-EM, sample sizes were determined by the number of micrographs and particle densities. The rationale here is to have sufficient samples that allowed the reconstruction of a 3D cryo-EM map at a resolution better than 4.0 angstrom. A total of 2,950,051 particles were picked automatically and based on 2D and 3D classifications and refinements, finally 46,953 particles were used for final 3D reconstruction.
Data exclusions	Cryo-EM data were processed in Relion and CryoSPARC. After 2D and 3D classifications, particles that do not belong to classes with high-resolution features were rejected from further analysis.
Replication	For enzyme activity assays, all measurements were performed in triplicate. All attempts at replication were successful.
Randomization	Randomization is not applicable because there are not groups to allocate.
Blinding	Blinding is not applicable because there are no groups to allocate.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology and archaeology
☒ ☐	Animals and other organisms
☒ ☐	Human research participants
☒ ☐	Clinical data
☒ ☐	Dual use research of concern

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging