scientific reports

Check for updates

OPEN

Structural insight into *Pichia pastoris* fatty acid synthase

Joseph S. Snowden, Jehad Alzahrani, Lee Sherry, Martin Stacey, David J. Rowlands, Neil A. Ranson✉ & Nicola J. Stonehouse✉

Type I fatty acid synthases (FASs) are critical metabolic enzymes which are common targets for bioengineering in the production of biofuels and other products. Serendipitously, we identified FAS as a contaminant in a cryoEM dataset of virus-like particles (VLPs) purified from *P. pastoris*, an important model organism and common expression system used in protein production. From these data, we determined the structure of *P. pastoris* FAS to 3.1 Å resolution. While the overall organisation of the complex was typical of type I FASs, we identified several differences in both structural and enzymatic domains through comparison with the prototypical yeast FAS from *S. cerevisiae*. Using focussed classification, we were also able to resolve and model the mobile acyl-carrier protein (ACP) domain, which is key for function. Ultimately, the structure reported here will be a useful resource for further efforts to engineer yeast FAS for synthesis of alternate products.

Fatty acid synthases (FASs) are critical metabolic enzymes for the endogenous biosynthesis of fatty acids in a diverse range of organisms. Through iterative cycles of chain elongation, FASs catalyse the synthesis of long-chain fatty acids that can produce raw materials for membrane bilayer synthesis, lipid anchors of peripheral membrane proteins, metabolic energy stores, or precursors for various fatty acid-derived signalling compounds[1]. In addition to their key physiological importance, microbial FAS systems are also a common target of metabolic engineering approaches, usually with the aim of generating short chain fatty acids for an expanded repertoire of fatty acid-derived chemicals, including chemicals with key industrial significance such as α-olefins[2–6]. A comprehensive structural and functional understanding of a diverse range of FAS systems is therefore paramount to the success of such complex rational engineering approaches.

Unlike the dissociated systems of individual enzymes that make up type II FAS systems in plants and bacteria[7], Type I FAS systems in fungi and animals are large complexes that integrate the key enzymatic activities required for fatty acid biosynthesis[8,9]. In particular, yeast type I FASs are 2.6 MDa hetero-dodecameric complexes with an $\alpha_6\beta_6$ configuration, which form a cage-like structure comprising two dome-shaped reaction chambers on either side of a central platform[10–14]. Along with structural support regions, enzymatic domains form the walls of the reaction chambers, with most active sites facing the chamber interior. The mobile acyl-carrier protein (ACP) domain tethers the growing fatty acid chain and carries it between active sites, so each can act on this in turn[14–16]. The phosphopantetheine transferase (PPT) domain, which catalyses the attachment of a phosphopantetheinyl group to the active site serine residue of the ACP (necessary for ACP activity), is also an integral part of yeast FAS complexes, and is located on the exterior of the cage[13,17,18].

An extensive body of work underlies our current structural understanding of yeast FAS, although FAS structure determination to date has been challenging. A cryoEM study of *S. cerevisiae* FAS revealed a number of differences with X-ray crystal structures of fungal FASs, suggesting that crystal contacts had altered the conformation of FAS from its unconstrained structure in solution[14]. A more recent attempt to determine the solution structure of *S. cerevisiae* FAS at a resolution sufficient for molecular modelling was hindered by partial denaturation induced by interactions with the air–water interface, an effect that was ameliorated by coating grids with hydrophilized graphene, ultimately leading to a 3.1-Å resolution structure when sample preparation was optimised[19,20]. More recently, the structure of *S. cerevisiae* FAS was determined at 2.8 Å resolution in complex with a previously unknown regulatory subunit[21].

While our understanding of yeast FAS structure is growing, our library of structures remains incomplete. The vast majority of published solution structures of FAS are from *S. cerevisiae*[14,19–22]. To date, no structures are available for *Pichia pastoris* FAS from either cryoEM or X-ray crystallography. This represents an important gap in our knowledge, because *P. pastoris* (*Komagataella spp.*) is a methylotrophic yeast which is used extensively for recombinant protein production[23]. In recent years, it has also been the target of metabolic engineering approaches[24]. *P. pastoris* has a major advantage over other yeast strains like *S. cerevisiae*, as it produces relatively

Astbury Centre for Structural Molecular Biology, School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK. ✉email: n.a.ranson@leeds.ac.uk; n.j.stonehouse@leeds.ac.uk

little ethanol, making the maintenance of productive, high-cell density fermenter cultures more tractable[25]. It has also been suggested that *P. pastoris* cells, in contrast to cells of the plant *Arabidopsis* sp., are highly tolerant of free fatty acids[26].

Here, we report the 3.1-Å resolution solution structure of *P. pastoris* FAS, determined by cryoEM. We initially identified FAS as a contaminant in a *P. pastoris*-derived virus-like particle (VLP) preparation being used for structural studies, and processed these data to yield detailed structural information on FAS, including density for the mobile ACP domain. Comparison of our atomic model with that of *S. cerevisiae* revealed interesting structural differences throughout the complex within both structural and enzymatic domains. This included movement of the malonyl/palmitoyl transferase domain at the level of domain packing, and alternative rotamers in the catalytic centres of the acetyl transferase and enoylreductase domains.

Results

CryoEM structure determination of *Pichia pastoris* fatty acid synthase.

Virus-like particles (VLPs) composed of hepatitis B virus tandem core proteins[27] with a SUMO-binding affimer inserted into the major immunodominant region (MIR)[28] were expressed in transformed *P. pastoris* cells for structural and vaccine studies (construct available upon request). Expression was induced with methanol and the VLPs purified from the yeast lysate by sucrose gradient centrifugation. Image analysis of cryoEM data revealed, in addition to the VLPs, non-VLP-like particles which were recognised as yeast FAS complexes due to their characteristic morphology—two hollow chambers on either side of a central platform. Further inspection of individual micrographs revealed that FAS particles were well-separated from VLPs, indicating that FAS was not co-purified as a result of an association with the VLPs (Fig. 1a).

In contrast to a recent study that reported significant air–water interface-induced denaturation of FAS particles by cryoEM (which could be solved with deposition of hydrophilized graphene on the grid surface)[19], we saw no signs of denaturation in our data set, judged by 2D class averages (Fig. 1b). While some particles were discarded after the initial 2D classification, ~ 67% of these were actually VLPs or areas of carbon film that were inadvertently selected during automated particle picking, and the remaining discarded classes mainly comprised apparently intact FAS particles over thick carbon. It is therefore difficult to estimate how many particles were excluded because of damage that was not visible in class averages. There was evidence of some preferential orientation effects (perhaps induced by adsorption to the carbon film), but this was not a significant problem, as a reasonable coverage of potential orientations within the asymmetric unit of the D3 symmetric structure (i.e., a single α and β subunit) was achieved (Fig. 1c).

Classes corresponding to FAS were taken forward and processed separately, yielding a 3.1 Å reconstruction from ~ 37,000 particles, with parts of the central platform resolved to ~ 2.7 Å (Fig. 1d, Supplementary Fig S1). In comparison, ~ 28,000 particles of *S. cerevisiae* FAS were used to generate a 4.0-Å resolution structure[19]. More recently, ~ 15,000 particles of *S. cerevisiae* FAS led to a 3.1-Å resolution map, although in this case the FAS was incubated with NADPH and malonyl-CoA before vitrification in order to drive synthesis of fatty acids to completion, increasing particle homogeneity[20].

The 3.1-Å resolution solution structure of *P. pastoris* FAS.

The overall structure of the *P. pastoris* FAS reported here was typical of a yeast FAS. Density was present for most domains of the α- and β-subunits, including acetyl transferase (AT), ketoacyl reductase (KR), ketoacyl synthase (KS), enoylreductase (ER), dehydratase (DH) and malonyl/palmitoyl transferase (MPT) domains (Fig. 2a,b). ACP density was weak (only becoming fully apparent when inspecting the map at a threshold of approx. 1 σ) and poorly resolved, as was also the case for previous EM density maps of *S. cerevisiae* FAS (e.g., EMD-10420, EMD-0178). However, the approximate position of the ACP could be identified within the reaction chamber. There was no density apparent for the PPT region, as discussed further below. Furthermore, we saw no indication of a γ subunit, as recently reported for *S. cerevisiae* FAS[21].

The map was sufficiently resolved to allow us to build an atomic model of *P. pastoris* FAS. An initial homology model for the asymmetric unit was generated using SWISS-MODEL[29] and rigid-body fitted into the density map. The fit of model to density was visually inspected and corrected, before the model was symmetrised and refined to improve its fit to the density and geometry. The α subunit is mostly resolved, though several segments of the peptide sequence (including the ACP and PPT domains) are missing (residues 96–322, 535–598 and 1752–1879). The β subunit is practically complete, lacking only a small number of residues at the N- and C-termini (residues 1–9 and 2064–2069). We also observed density for a ligand bound within the active site of the ER domain (for reference, see Fig. 4), presumably a molecule of FMN.

In an attempt to improve the quality of the cryoEM density corresponding to the ACP domain, we performed focussed 3D classification within the reaction chamber of the complex (Fig. 2c). While classification with a regularisation parameter (‘T’ number) of 20 did not reveal any meaningful classes, using a higher value of 40 (i.e., putting a higher relative weighting on the experimental data than on the prior/reference) yielded a class containing relatively well resolved ACP density (Supplementary Fig S2a). Particles from this class were taken forward for asymmetric reconstruction, yielding a complete FAS map of 3.1-Å resolution (Fig. 2d). In the latter map, the additional density is consistent with the ACP domain of the α subunit of FAS, given the fit-to-density of an ACP domain homology model. This additional density was largely missing in the original map (Supplementary Fig S2b,c). Ultimately, the new density was of sufficient quality to build an atomic model of the ACP domain (Fig. 2e). Interestingly, ACP domain residue S178 (which tethers the growing fatty acid chain) is located ~ 18 Å away from the KS active site, which is in agreement with the length of the phosphopantetheine arm (18 Å)[30].

Some classes contained additional low-resolution density at the edge of the masked region used for focussed classification (Supplementary Fig S2a). To examine whether this corresponded to ACP domains positioned

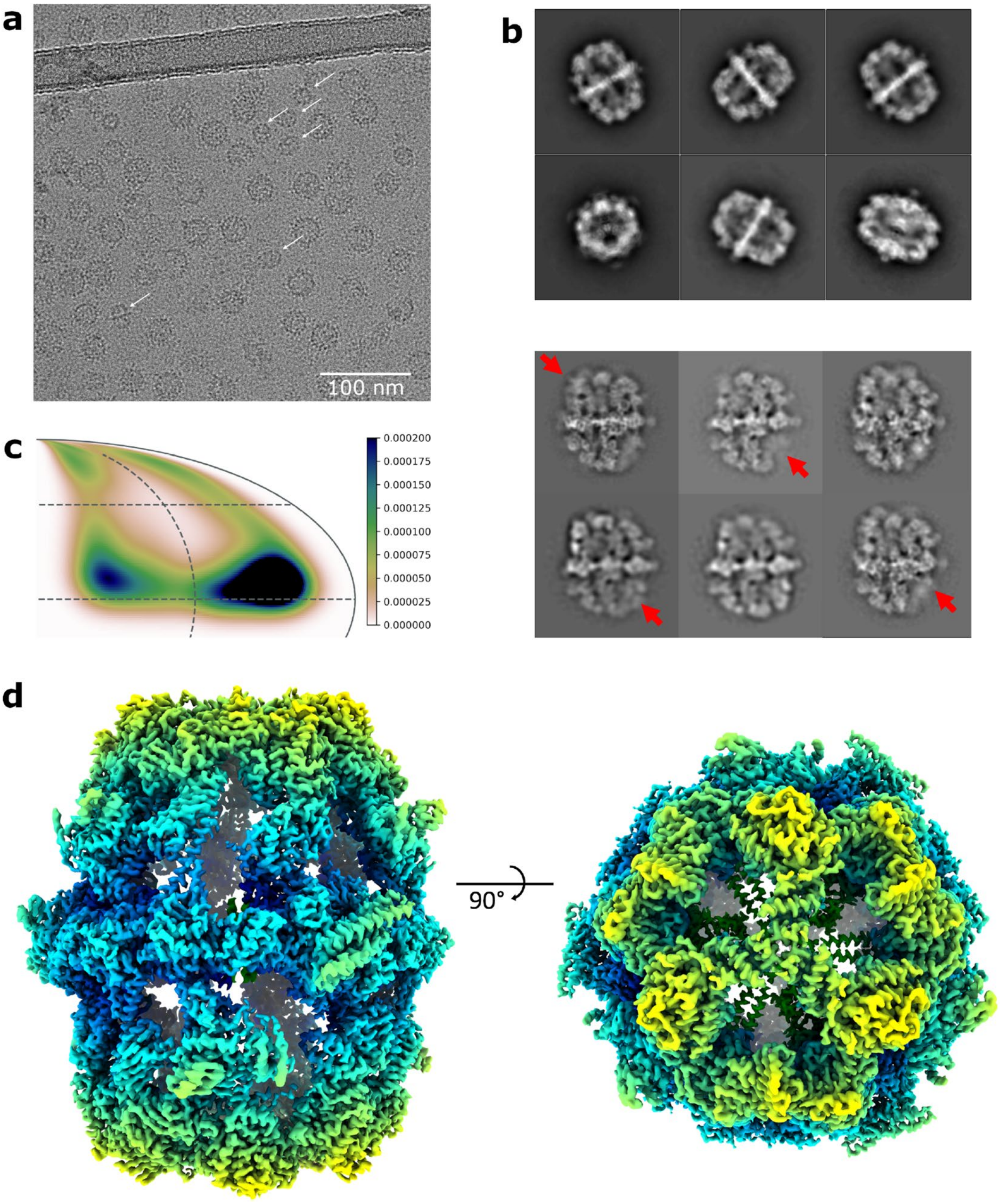

Figure 1. Determination of *P. pastoris* FAS structure by cryoEM. (**a**) Raw micrograph showing VLPs with contaminating FAS particles. FAS particles are indicated by white arrows. Scale bar shows 100 nm. (**b**) The most populated classes from 2D classification of *P. pastoris* FAS using a box size of 460 pixels (after excluding 'junk' classes of FAS over thick carbon support) (upper) compared with 2D class averages of *S. cerevisiae* FAS showing partial denaturation (indicated by red arrows) from D'Imprima et al. (2019)[19] (licensed under CC BY 4.0) (lower). (**c**) Orientation distribution map of particles contributing to the final reconstruction of the FAS asymmetric unit. The map shows the probability that any particle was assigned the orientation indicated, according to the scale shown. (**d**) CryoEM density map of *P. pastoris* FAS, coloured radially.

differently within the complex interior that were not fully contained within the masked region, focussed classification was repeated with an expanded mask (Supplementary Fig S2d). In addition to a class containing well-resolved ACP density, multiple classes containing low-resolution ACP-sized density in different positions were observed (Fig. 2f). However, these classes comprised relatively few subparticles and asymmetric reconstruction only resulted in poor resolution (~ 8 Å).

Focussed classification was also performed in an attempt to resolve the PPT domain. While a low-resolution class was identified as containing PPT density (Supplementary Fig S3), further reconstruction did not yield density sufficient for model building.

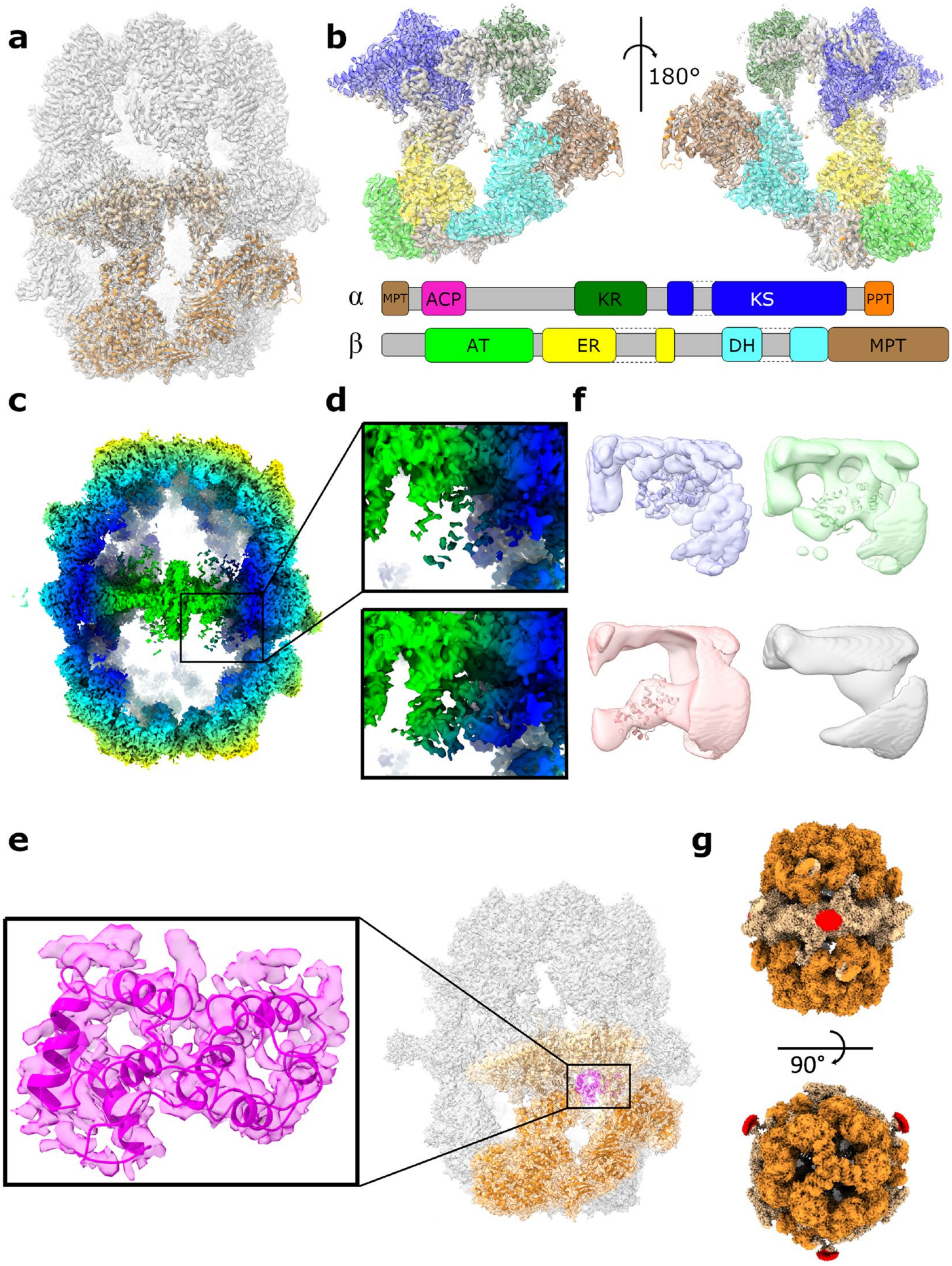

Figure 2. The molecular structure of *P. pastoris* FAS. (**a**) Atomic coordinates for the asymmetric unit of *P. pastoris* FAS, comprising α (light orange) and β (dark orange) subunits, with the density map overlaid. (**b**) Enlarged view of the asymmetric unit, highlighting different enzymatic domains according to the colour scheme indicated in the domain organisation schematic (below). Grey-coloured regions are non-enzymatic, structural domains. Domain boundaries were assigned based on equivalent domain boundaries from *S. cerevisiae* FAS. MPT—malonyl/palmitoyl transferase, ACP—acyl carrier protein, KR—ketoacyl reductase, KS—ketoacyl synthase, PPT—phosphopantetheine transferase, AT—acetyl transferase, ER—enoyl reductase, DH—dehydratase. (**c**) Clipped view of the sharpened FAS density map, coloured according to radius. The box highlights the region of the reaction chamber selected for the first focussed 3D classification. (**d**) Enlarged image of the region depicted by the box in (**c**) (upper) compared with the sharpened, asymmetric reconstruction of FAS derived from an ACP density-containing class from the first focussed 3D classification (lower). Further details (including all classes) from the first focussed classification are given in Supplementary Fig S2a-c. (**e**) Atomic coordinates for the ACP domain of FAS fitted into ACP density from the map shown in (**d**, bottom). ACP is also shown in the context of the full FAS complex. (**f**) Representative classes from the second focussed 3D classification (with expanded masked region) of the complex interior. Atomic coordinates for ACP are shown fitted into the three classes containing putative ACP density to give an indication of size. One class shows no putative density for ACP (lower right). All maps are shown at the same contour level. All classes from the second focussed classification are given in Supplementary Fig S2d. (**g**) Density map of FAS shown at a low contour level (i.e., including weak density), coloured according to subunit (α—light orange, β—dark orange). Low-resolution density corresponding to the flexible dimerization module 2 (DM2) element is highlighted in red.

At low contour levels, we also observed additional density on the exterior of the FAS complex corresponding to dimerization module 2 (DM2), a flexible structural element that contributes to the stability of the complex but is often missing from FAS structures[12,31] (Fig. 2g). This density was weaker than for the rest of the complex and was not well resolved, confirming that DM2 is flexible. Given the poor resolution of this region of the map, we were unable to include DM2 in our atomic model.

Distinct structural features of *P. pastoris* FAS. To identify unique structural features of *P. pastoris* FAS, we carried out a comparative structural analysis. While PDBeFold[32] suggested that the closest structural matches to *P. pastoris* FAS were cryoEM structures of *C. albicans* FAS (PDB 6U5W, PDB 6U5V)[22], which shares 73% (α) and 68% (β) sequence identity with *P. pastoris* FAS, we chose to compare our structure with *S. cerevisiae* FAS, a better characterised FAS with greater relevance to metabolic engineering attempts. At the peptide sequence level, *P. pastoris* FAS α and β subunits show 69% and 63% identity with their equivalents in *S. cerevisiae*. We compared our structure with a recently published 3.1-Å resolution structure of FAS from *S. cerevisiae* (PDB 6TA1)[20]. After aligning the two structures, the RMSD value between equivalent Cα atom pairs was calculated as 7.1 Å, indicating a notable level of difference between the two models (Fig. 3a). A similar level of variation was observed between *P. pastoris* FAS and a range of other *S. cerevisiae* FAS EM and X-ray crystal structures (PDB 6QL6[21], 6QL9[21], 2UV8[30]) (Fig. 3b, Supplementary Fig S4). However, aligning the models (*P. pastoris* FAS and PDB 6TA1) by individual subunits gave RMSD values of 1.7 Å (α–α), and 2.2 Å (β–β), suggesting that this may be partially attributable to differences in the spatial relationship between the α and β subunits in each model.

Visual inspection confirmed that several domains appeared shifted, particularly the MPT domain. Here, the ferredoxin-like subdomain is positioned further away from the α/β-hydrolase fold in *P. pastoris* FAS, opening up a cleft between the two subdomains close to the active site (Fig. 4). The *P. pastoris* FAS ACP domain is positioned very similarly to the *S. cerevisiae* FAS ACP domain (PDB: 6TA1)[20], with a root-mean-square deviation (RMSD) value of 2.1 Å between equivalent Cα atom pairs in the ACP domains after alignment based on the KS domains. This is in contrast to *C. albicans* FAS (PDB: 6U5V), where the ACP domain was observed to localise to the ER domain[22].

Closer examination revealed many differences in the path traced by the peptide backbone. One instance of this is a loop in the MPT domain, which is shifted by 16 Å at its turning point (D1761) (Fig. 3c). Interestingly, when displaying density at a low threshold so that even weak density is visible (0.5 σ), we also saw density tracing the alternative loop conformation seen in the *S. cerevisiae* yeast structure. Other loops and linker regions throughout the β subunit also showed obvious differences in conformation, including in the DH domain (F1283-V1289, E1505-I1511), MPT domain (A1859-Y1870), and structural domains/linker regions (G75-N78, A914-G919) (Fig. 3d). Furthermore, a small sequence insertion of three amino acids introduced a kink into the outer strand of a β-sheet within the DH domain (A1553-L1558) (Fig. 3e). Of particular interest is a sequence of residues (F1113-T1135) linking the ER domain to the DH-adjacent structural domain, which was mostly unmodeled in the *S. cerevisiae* FAS structure (Fig. 3f). For the structure reported here, this linker region was well-resolved and differed significantly from the terminal parts of the linker that are resolved in PDB 6TA1.

There were also subtle rotameric differences in some of the enzymatic domain active sites. In particular, Q171 was shifted within the active site of the AT domain (Fig. 4). H749 of the ER domain was also pointing in the opposite direction to the equivalent histidine in *S. cerevisiae* FAS (Fig. 4). Interestingly, the conformation of H749 reported here was more similar to that seen in a structure of *C. albicans* FAS in the KS-stalled state, with the equivalent *S. cerevisiae* FAS histidine showing more similarity to one of the two conformations observed for *C. albicans* FAS in the apo state (Supplementary Fig S5)[22].

Discussion

We serendipitously identified fatty acid synthase, which is constitutively expressed in yeast[33], as a contaminant in a cryoEM data set of a hepatitis B core VLP preparation purified from *P. pastoris*. Subsequent analysis of raw micrographs revealed that FAS was readily separated from the VLPs. As such, we think it unlikely that FAS co-purified as a result of a specific interaction with the VLPs. Rather, its presence in the VLP preparation was probably as a result of its mass being sufficiently similar to co-purify on sucrose density gradients (FAS 2.6 MDa; $T=3$ VLP 4.7 MDa; $T=4$ VLP 6.3 MDa). By processing this 'contaminant' data and using focussed 3D classification techniques, we were able to determine the structure of the *P. pastoris* FAS to 3.1 Å resolution, and build an atomic model for the complex, including a model for the mobile ACP domain.

Interestingly, we did not observe signs of partial denaturation of FAS in 2D class averages, as was seen previously for a different yeast FAS structure[19]. It is possible that the presence of VLPs in the sample affected the ice thickness or crowded the air–water interface (which is known to induce denaturation[19]), reducing any interactions between the interface and FAS. However, we suspect that the lack of partial denaturation is more likely a result of the continuous thin carbon film used for making these cryoEM grids, which may trap FAS particles away from the air–water interface. While the presence of a thin carbon film adds a background that reduces the signal-to-noise ratio of particles (and can therefore limit the resolution), we were able to resolve our map to 3.1 Å, equivalent to that of the structure reported by Joppe et al. (2020)[20] and sufficient to build a molecular model for FAS. Therefore, using a continuous thin carbon film may be a simple alternative to coating grids with hydrophilized graphene, when the size of the target protein (and hence its signal strength) is similar to FAS.

The overall structure was typical of a yeast FAS. As has previously been observed for yeast FAS structures, the PPT domain was missing[19] and the quality of the ACP domain density was relatively poor. In this case, density corresponding to the ACP domain was initially of insufficient quality for molecular modelling, but this was resolved by focussed classification. Given the known mobile nature of the ACP domain[14], this was unsurprising. Focussed classification led to an improvement in the strength and quality of the ACP density that was sufficient

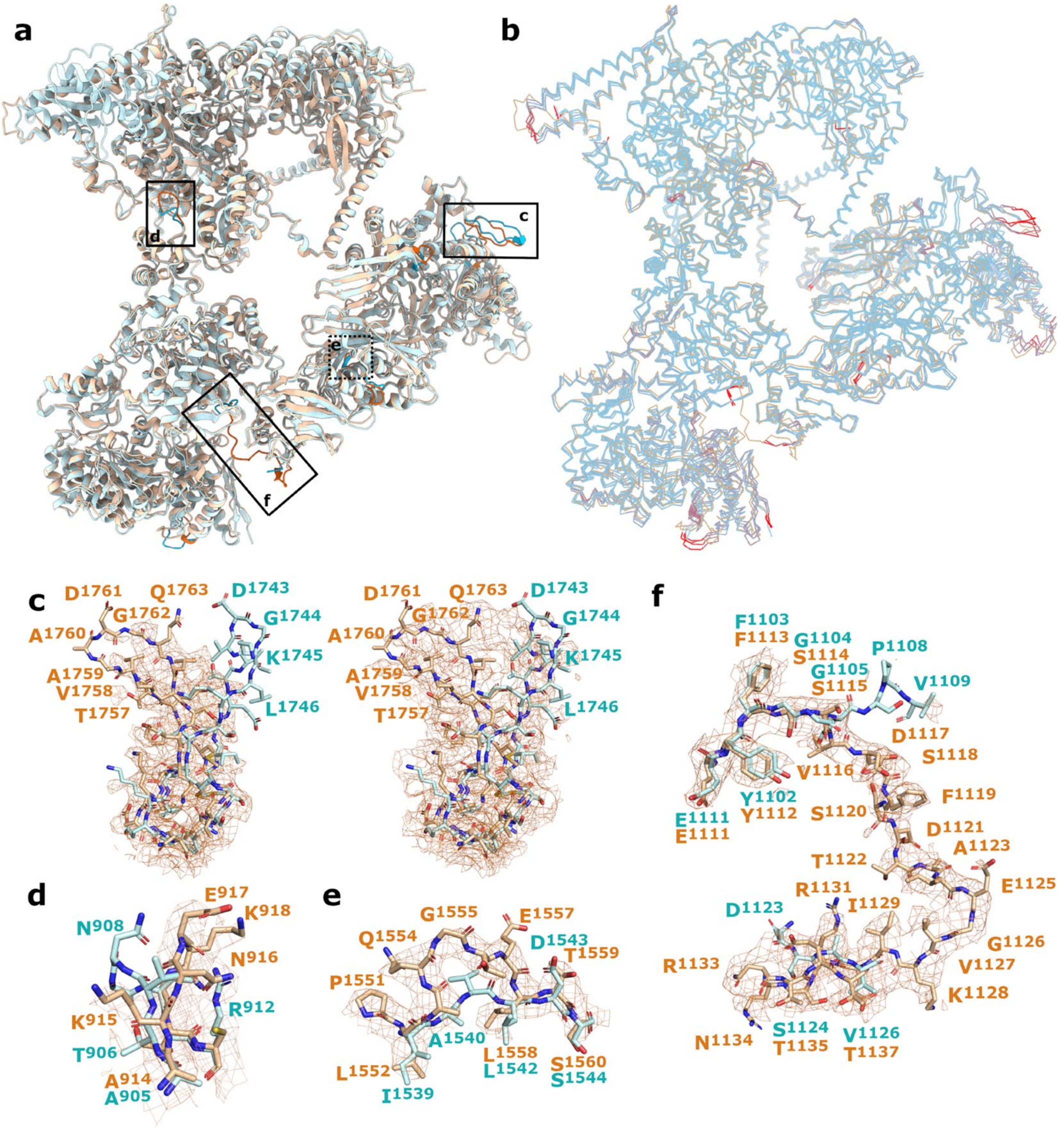

Figure 3. Unique structural features of *P. pastoris* FAS. (**a**) Overlay of *P. pastoris* FAS (orange) and *S. cerevisiae* FAS (PDB 6TA1, blue) from Joppe et al. (2020)[20], with differences highlighted by stronger colours. Enlarged images are indicated on the map with boxes. (**b**) Wire representation of atomic model for *P. pastoris* FAS (orange) overlaid with a range of models of *S. cerevisiae* FAS (PDB 6TA1[20], 6QL9[21], 6QL6[21], 2UV8[30]; blue), with regions of high deviation (RMSD ≥ 15 Å) coloured red for *S. cerevisiae* FAS models. Segments of the backbone that were obviously misaligned (e.g., due to missing residues) were removed for clarity. (**c**) Loop region comprising residues 1753–1768, highlighting the different positioning compared to the equivalent region from *S. cerevisiae* FAS. The *P. pastoris* density map is overlaid at high contour level (left) and low contour level (right). (**d**–**f**) Other regions from different parts of the complex are highlighted and compared with equivalent regions of *S. cerevisiae* FAS, including the loop comprising residues 914–919 (**d**), the beta strand comprising residues 1553–1558 (**e**) and the additional resolved linking region comprising residues 1113–1135 (**f**).

to allow atomic modelling, probably as result of grouping together only subparticles with ACP domains in the same relative position (adjacent to the KS domain, as has been observed in several other yeast FAS structures[20,30]) at the time of vitrification. Use of an expanded mask for focussed classification revealed classes with putative ACP density in other positions within the complex interior. These classes contained relatively few subparticles, suggesting that the vast majority of ACP domains were adjacent to the KS domain, with the remainder either missing or occupying other positions. Notably, the positioning of the ACP domain was different to that observed in FAS structures from *C. albicans* (PDB: 6U5V)[22] and *C. thermophilum* (EMDB: 3757)[34], where the ACP domain localised to the ER domain. This was suspected to be a result of weakened negative surface charge in the KS domain, given that the ACP-KS interaction is thought to be driven by surface electrostatics[22,35]. Interestingly, while *P. pastoris* FAS retains some of the key acidic residues involved in the ACP-KS interaction that are altered in *C. albicans* FAS, other key acidic residues conserved in both *S. cerevisiae* FAS and *C. albicans* FAS are replaced

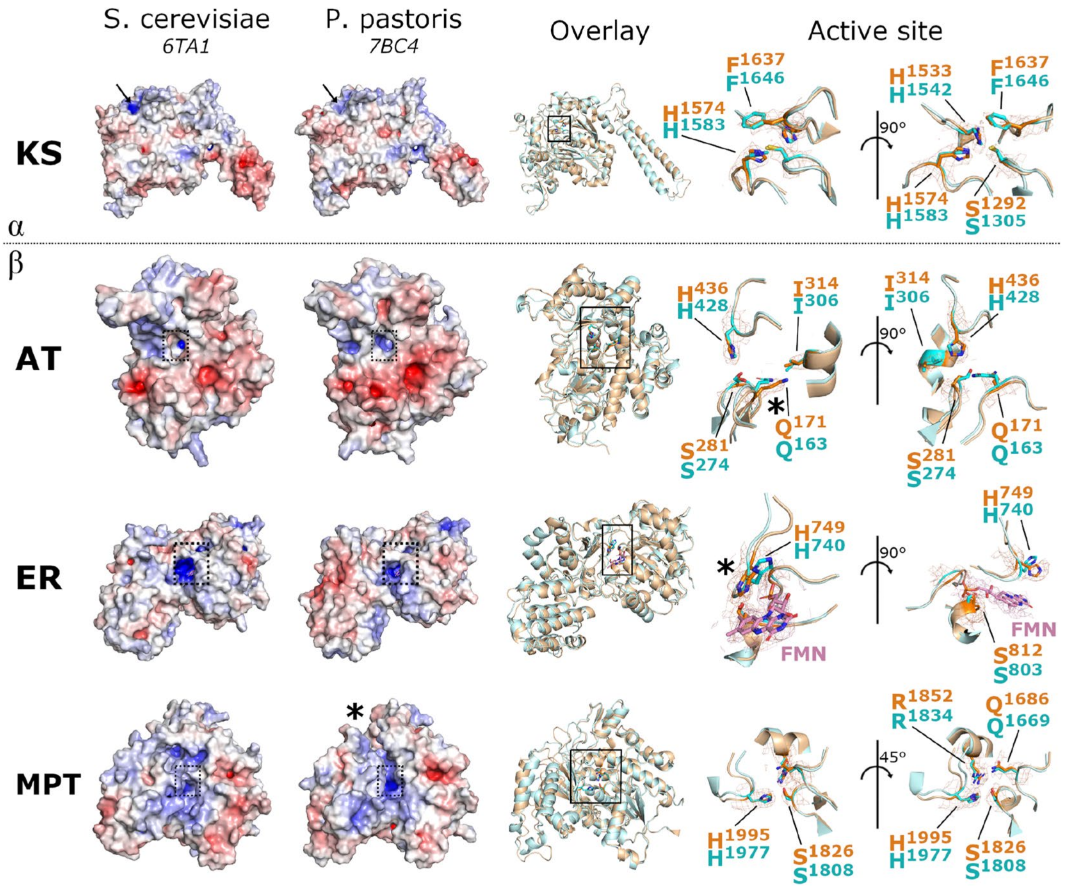

Figure 4. Structural details of key FAS enzymatic domains. For each key enzymatic domain (KS—ketoacyl synthase, AT—acetyl transferase, ER—enoyl reductase, MPT—malonyl/palmitoyl transferase), surfaces are shown for both *P. pastoris* FAS and *S. cerevisiae* FAS (PDB 6TA1[20]), coloured according to electrostatic charge (blue—positive, red—negative). The locations of the active sites are highlighted by dashed boxes (AT/ER/MPT) or black arrows (KS). An overlay of the atomic models is also shown (*P. pastoris* FAS—orange, *S. cerevisiae* FAS—blue) with enlarged images of key active site residues (located within the solid boxes). Density for *P. pastoris* FAS is shown around highlighted active site residues. FMN and corresponding EM density is shown in pink. Key differences are highlighted by asterisks (Q171 in the AT domain, H749 in the ER domain, MPT subdomain arrangement).

with non-charged residues (Supplementary Fig S6). The observation that the ACP domain still localises to the KS domain in *P. pastoris* FAS despite a similarly weakened overall negative surface charge suggests that certain acidic residues at the interface (such as E1107 and E1115) may be more important for driving the ACP-KS interaction.

The lack of density for the PPT domain may result from partial denaturation of this domain in most complexes, as has previously been suggested[19,20]. Since our initial aim was to purify VLPs, not FAS, no special care was taken to ensure the complex was kept intact (although the ER domain cofactor FMN was clearly present in the density map, providing evidence that the complex was still likely to have been active). While focussed classification was not sufficient to resolve the PPT domain, one class did show weak density corresponding to the PPT domain, suggesting that it was intact in at least some particles.

We identified a number of structural differences between the *P. pastoris* FAS structure reported here and a recent 3.1-Å resolution structure from *S. cerevisiae* (PDB 6TA1)[20], another member of the yeast family *Saccharomycetaceae*. The MPT domain was distinctly shifted relative to its position in *S. cerevisiae* FAS, including a change in the positioning of the ferredoxin-like subdomain relative to the α/β-hydrolase fold. This is of particular importance, because this domain is the target of one approach to engineer FAS complexes to produce short chain fatty acids[4]. Many intradomain differences affected structural (non-enzymatic) components of the complex, including a loop within the MPT domain from S1752 – K1768. This loop was particularly interesting, as very weak density tracing the backbone of the alternative (*S. cerevisiae*) conformation also became visible at low contour thresholds. This is indicative of a mixed population of conformations in *P. pastoris* FAS, with most complexes sampling the conformation modelled, and a smaller subpopulation sampling the conformation seen for *S. cerevisiae* FAS. This loop may plausibly be dynamic, an idea reinforced by (i) other EM and X-ray crystal structures of *S. cerevisiae* FAS showing the equivalent loop in 'intermediate' positions (Fig. 3b), and (ii) the local resolution of the loop being poorer than average, at around 3.7–4.1 Å resolution (Supplementary Fig S1).

We also identified differences in the active sites of two enzymatic domains. In the AT domain, Q171 forms part of a tight hydrogen bond network around the active serine (S282) and its backbone amide also contributes to an oxyanion hole, which plays a role in promoting the transfer of the acetyl moiety to the thiol group of the ACP[7,13]. While the alternative rotamer observed here is unlikely to affect the oxyanion hole, the structure suggests a subtle alteration in the arrangement/spacing of the hydrogen bond network at the catalytic centre of the AT domain. Aside from Q171, the catalytic histidine of the ER domain (H749) was seen to be in a 'flipped' conformation relative to its equivalent in *S. cerevisiae* FAS[13,20]. Interestingly, the conformation reported here is much closer to that seen in a crystal structure of *S. pneumoniae* FabK in its active state (PDB 2Z6I), a homologous enzyme that forms part of its type II FAS system[36]. Whether these changes to active site residues have functional significance is unclear. However, it should be noted that the *S. cerevisiae* FAS was incubated with NADPH and malonyl-CoA prior to vitrification in order to drive FA synthesis to completion[20]. The *P. pastoris* FAS reported here received no such treatment and as such, reaction status was likely less uniform and potentially incomplete.

In conclusion, despite non-optimised sample preparation and using just ~ 37,000 particles, we were able to determine the structure of *P. pastoris* FAS to 3.1-Å resolution and identify structural differences with a related yeast FAS. As well as highlighting the importance of considering 'contaminant data' in cryoEM data sets, the structure reported here could prove a useful resource for future efforts to engineer yeast FAS.

Methods

Cells.

FAS was purified from PichiaPink™ yeast strain one (Invitrogen, USA), which was grown according to the manufacturer's instructions.

Yeast transformation and induction.

While the focus of this work was FAS, which is constitutively expressed in yeast[33], our initial aim was to use yeast to express and purify VLPs composed of hepatitis B virus tandem core proteins[27] with a SUMO-binding affimer[28] inserted into the major immunodominant region (MIR) (construct available upon request) for structural characterisation. As such, VLP expression was induced in the transformed yeast prior to purification, as described previously[37]. Briefly, glycerol stocks known to exhibit good levels of expression were cultured in yeast extract-peptone-dextrose (YPD), supplemented with 50 µg/mL ampicillin, at 28 °C, 250 rpm for 48 h. 2 ml of the high-density starter culture was then transferred to 200 mL YPD and incubated for another 24 h under the same conditions. To induce expression of VLPs, cells were pelleted (1500 × *g*, 20 min) and resuspended in 200 mL yeast extract-peptone-methanol (YPM; 0.5% [v/v] methanol) for a further 48-h incubation, with methanol added to a final concentration of 0.5% (v/v) 24 h post-induction. Cells were subsequently collected by pelleting at 2000×*g* (20 min), resuspended in 30 mL breaking buffer (50 mM sodium phosphate, 1 mM phenylmethylsulfonyl fluoride [PMSF], 1 mM EDTA, 5% glycerol, pH 7.4) and stored at – 20 °C prior to purification.

Co-purification of FAS with VLPs.

Though our aim was to purify VLPs (as described previously[37]) for structural characterisation, FAS co-purified with the particles. Frozen *P. pastoris* stocks that had been induced to express VLPs were thawed and incubated with 0.1% (v/v) Triton-X100 for 30 min. Cells were lysed using a CF-1 cell disruptor at ~ 275 MPa and cooled to 4 °C, before the lysate was centrifuged first at 4000 rpm (30 min, 4 °C), then at 10,000×*g* (30 min, 4 °C) to remove insoluble material. The supernatant was supplemented with 2 mM MgCl and treated with HS-nuclease (2BScientific) for 1.5 h at room temperature, then mixed with saturated ammonium sulphate solution (40% [v/v]) and incubated overnight at 4 °C for precipitation. Precipitated protein was recovered by pelleting at 4000 rpm (30 min, 4 °C) and solubilised by resuspending in PBS (supplemented with 0.1% [v/v] Triton-X100). After further centrifugation at 4000 rpm (30 min, 4 °C), then 10,000×*g* (30 min, 4 °C) to remove any remaining insoluble material, the supernatant was pelleted through a 30% sucrose cushion (w/v) at 151,000×*g* (3.5 h, 4 °C, SW-32 Ti). The pellet was resuspended in PBS (supplemented with 1% [v/v] NP-40 and 0.5% [w/v] sodium deoxycholate), then clarified at 10,000×*g* (10 min, 4 °C) before loading onto a 15–45% (w/v) sucrose gradient for ultracentrifugation at 151,000 × *g* (3 h, 4 °C). 2 ml fractions were collected and those containing VLPs were identified by western blot. A centrifugal filter (100 K, 5 mL, Pall Life Sciences) was used according to the manufacturer's instructions to remove sucrose and concentrate the peak fraction to 0.5 mL.

Electron microscopy.

CryoEM grids were prepared by applying 3 µL of a VLP preparation with contaminating FAS complex in PBS to lacey carbon 400-mesh copper grids coated with a < 3-nm continuous carbon film (Agar Scientific, Stansted, UK), after glow-discharging the grids in air (10 mA, 30 s) using a PELCO easiGlow™ Glow Discharge System (Ted Pella). Following application of sample, grids were incubated at 8 °C, 80% relative humidity for 30 s, then blotted (for 1.0–4.0 s) to remove excess liquid and vitrified in liquid nitrogen-cooled liquid ethane using a LEICA EM GP plunge freezing device (Leica Microsystems, Wetzlar, Germany). Blotting time parameters were varied to increase the likelihood of recovering a grid with optimal ice thickness. Following plunge freezing, grids were stored in liquid nitrogen. Grids were imaged using an FEI Titan Krios transmission electron microscope (ABSL, University of Leeds, Leeds, UK) operating at 300 kV. A magnification of 75,000 × was used for a calibrated object sampling of 1.065 Å/pixel. Detailed information on data collection parameters is provided in Supplementary Table S1.

Image processing.

The RELION 3.0 pipeline[38] was used for image processing. Raw micrographs were first motion corrected using MotionCor2[39], then CTF parameters were estimated using Gctf[40]. Particles were selected from a subset of micrographs by manual picking and were then differentiated by 2D classification. High quality 2D classes were used as templates for autopicking the complete data set. A low picking threshold was specified for autopicking to minimize the number of 'missed' VLPs. Picked particles were 2 × down-sampled, then subject to 2D classification (with CTFs ignored until the first peak), leading to the identification of FAS-containing classes. 2D classes containing FAS were taken forward for initial model generation and used as templates for FAS-specific autopicking. Autopicked FAS particles were extracted without down-sampling and subjected to two rounds of 2D classification (first with CTFs ignored until the first peak, then without). Particles in high quality classes were taken forward for 3D refinement (with D3 symmetry applied) with subsequent masking and sharpening. Following this, several cycles of CTF refinement (with per-particle astigmatism correction and beamtilt estimation), Bayesian polishing and 3D refinement (with masking and use of solvent-flattened FSCs) were performed to improve the resolution of the map. After sharpening, the resolution of the final map was determined using the 'gold standard' Fourier shell correlation criterion (FSC = 0.143) (Supplementary Table S1, Supplementary Fig S1). RELION was used to estimate local resolution and generate a local resolution-filtered map.

To resolve the flexible ACP, focussed 3D classification (without signal subtraction) was performed as described previously[16,41–44]. Briefly, a cylindrical mask was generated in SPIDER[45] and resampled onto the D3-symmetric density map of FAS, such that it was positioned over the weak ACP density. The relion_symmetry_expand tool was used to assign 6 symmetrically-redundant orientations to each particle contributing to a symmetric reconstruction, such that the ACP domain in each symmetry-related position would be included in the classification. These symmetry-expanded particles were then subject to masked 3D classification without re-alignment using a regularisation parameter ("T" number) of 40. Given that an ACP density-containing class was resolved with a "T" number of 40, greater values of "T" number were not tested due to the heavy computational resources required for focussed classification. Particles contributing to the ACP density-containing class were taken forward for asymmetric reconstruction using the relion_reconstruct tool.

Model building and refinement. To generate a preliminary model for the asymmetric unit of FAS, the peptide sequences of *P. pastoris* FAS α (NCBI: XP_002490414.1) and β (NCBI: XP_002489642.1) subunits were used to generate a homology model using the SWISS-MODEL web server[29]. The homology model was rigid-body fitted into the sharpened FAS density map using UCSF Chimera[46], then refined in Coot[47]. The refined model was symmetrised in UCSF Chimera, then subject to real space refinement in Phenix[48] to improve simultaneously the fit of the model to the density map and model geometry. Several iterations between refinement in Coot and refinement in Phenix were performed before validating the model with Molprobity[49] and EMRinger[50].

Visualization and structure comparison. Density map and atomic model visualization was performed and figures were generated using UCSF Chimera[46], UCSF Chimera X[51] and PyMOL (The PyMOL Molecular Graphics System, Version 2.1 Schrödinger, LLC). Surface charge colouring was performed using PDB2PQR[52] and APBS[53]. Structural alignment and RMSD calculations were performed using the 'MatchMaker' tool with default settings in UCSF Chimera. PDBeFold (the protein structure comparison service at EBI [http://www.ebi.ac.uk/msd-srv/ssm])[32] was used to probe structural similarity between the FAS structure reported here and other molecular structures within the PDB. Primary sequence alignments were performed using Clustal Omega with default parameters[54]. Particle orientation distribution was visualized using an adapted version[55] of a script from Naydenova & Russo (2017)[56].

Quantification and statistical analysis. Where appropriate, statistical details are given in the Method Details section. Supplementary Table S1 contains quantitative parameters related to data collection and image processing. Supplementary Table S2 contains validation statistics related to model building.

Data availability

The *P. pastoris* FAS cryoEM map was deposited in the Electron Microscopy Data Bank (EMD-12138) along with the reconstructed ACP domain-containing map from focussed classification (EMD-12139). The atomic coordinates for the FAS asymmetric unit (PDB-7BC4) and the ACP domain (PDB-7BC5) were deposited in the Protein Data Bank.

Received: 12 January 2021; Accepted: 19 April 2021
Published online: 07 May 2021

References

1. Tehlivets, O., Scheuringer, K. & Kohlwein, S. D. Fatty acid synthesis and elongation in yeast. *Biochim. Biophys. Acta* **1771**, 255–270 (2007).
2. Heil, C. S., Wehrheim, S. S., Paithankar, K. S. & Grininger, M. Fatty acid biosynthesis: Chain-length regulation and control. *ChemBioChem* **20**, 2298–2321 (2019).
3. Gajewski, J. *et al.* Engineering fatty acid synthases for directed polyketide production. *Nat. Chem. Biol.* **13**, 363–365 (2017).
4. Gajewski, J., Pavlovic, R., Fischer, M., Boles, E. & Grininger, M. Engineering fungal de novo fatty acid synthesis for short chain fatty acid production. *Nat. Commun.* **8**, 14650 (2017).
5. Beld, J., Lee, D. J. & Burkart, M. D. Fatty acid biosynthesis revisited: Structure elucidation and metabolic engineering. *Mol. BioSyst.* **11**, 38–59 (2015).
6. Zhu, Z. *et al.* Expanding the product portfolio of fungal type I fatty acid synthases. *Nat. Chem. Biol.* **13**, 360–362 (2017).
7. White, S. W., Zheng, J., Zhang, Y. M. & Rock, C. O. The structural biology of type II fatty acid biosynthesis. *Annu. Rev. Biochem.* **74**, 791–831 (2005).
8. Maier, T., Leibundgut, M., Boehringer, D. & Ban, N. Structure and function of eukaryotic fatty acid synthases. *Q. Rev. Biophys.* **43**, 373–422 (2010).
9. Schweizer, E. & Hofmann, J. Microbial type I fatty acid synthases (FAS): Major players in a network of cellular FAS systems. *Microbiol. Mol. Biol. Rev.* **68**, 501–517 (2004).

10. Stoops, J. K., Kolodziej, S. J., Schroeter, J. P., Bretaudiere, J. P. & Wakil, S. J. Structure-function relationships of the yeast fatty acid synthase: Negative-stain, cryo-electron microscopy, and image analysis studies of the end views of the structure. *Proc. Natl. Acad. Sci. USA* **89**, 6585–6589 (1992).
11. Jenni, S., Leibundgut, M., Maier, T. & Ban, N. Architecture of a fungal fatty acid synthase at 5 Å resolution. *Science* **311**, 1263–1267 (2006).
12. Jenni, S. *et al.* Structure of fungal fatty acid synthase and implications for iterative substrate shuttling. *Science* **316**, 254–261 (2007).
13. Lomakin, I. B., Xiong, Y. & Steitz, T. A. The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. *Cell* **129**, 319–332 (2007).
14. Gipson, P. *et al.* Direct structural insight into the substrate-shuttling mechanism of yeast fatty acid synthase by electron cryomicroscopy. *Proc. Natl. Acad. Sci. USA* **107**, 9164–9169 (2010).
15. Crosby, J. & Crump, M. P. The structural role of the carrier protein—Active controller or passive carrier. *Nat. Prod. Rep.* **29**, 1111–1137 (2012).
16. Lou, J. W. & Mazhab-Jafari, M. T. Steric occlusion regulates proximal interactions of acyl carrier protein domain in fungal fatty acid synthase. *Commun. Biol.* **3**, 274 (2020).
17. Fichtlscherer, F., Wellein, C., Mittag, M. & Schweizer, E. A novel function of yeast fatty acid synthase. Subunit α is capable of self-pantetheinylation. *Eur. J. Biochem.* **267**, 2666–2671 (2000).
18. Johansson, P. *et al.* Multimeric options for the auto-activation of the Saccharomyces cerevisiae FAS type I megasynthase. *Structure* **17**, 1063–1074 (2009).
19. D'Imprima, E. *et al.* Protein denaturation at the air–water interface and how to prevent it. *Elife* **8**, e42747 (2019).
20. Joppe, M. *et al.* The resolution revolution in cryoEM requires high-quality sample preparation: a rapid pipeline to a high-resolution map of yeast fatty acid synthase. *IUCrJ* **7**, 220–227 (2020).
21. Singh, K. *et al.* Discovery of a regulatory subunit of the yeast fatty acid synthase. *Cell* **180**, 1130-1143.e20 (2020).
22. Lou, J. W., Iyer, K. R., Hasan, S. M. N., Cowen, L. E. & Mazhab-Jafari, M. T. Electron cryomicroscopy observation of acyl carrier protein translocation in type I fungal fatty acid synthase. *Sci. Rep.* **9**, 12987 (2019).
23. Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. Protein expression in Pichia pastoris: Recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* **98**, 5301–5317 (2014).
24. Yang, Z. & Zhang, Z. Engineering strategies for enhanced production of protein and bio-products in Pichia pastoris: A review. *Biotechnol. Adv.* **36**, 182–195 (2018).
25. Cereghino, J. L. & Cregg, J. M. Heterologous protein expression in the methylotrophic yeast Pichia pastoris. *FEMS Microbiol. Rev.* **24**, 45–66 (2000).
26. Meesapyodsuk, D., Chen, Y., Ng, S. H., Chen, J. & Qiu, X. Metabolic engineering of Pichia pastoris to produce ricinoleic acid, a hydroxy fatty acid of industrial importance. *J. Lipid Res.* **56**, 2102–2109 (2015).
27. Peyret, H. *et al.* Tandem Fusion of hepatitis B core antigen allows assembly of virus-like particles in bacteria and plants with enhanced capacity to accommodate foreign proteins. *PLoS ONE* **10**, e0120751 (2015).
28. Tiede, C. *et al.* Affimer proteins are versatile and renewable affinity reagents. *Elife* **6**, e24903 (2017).
29. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
30. Leibundgut, M., Jenni, S., Frick, C. & Ban, N. Structural basis for substrate delivery by acyl carrier protein in the yeast fatty acid synthase. *Science* **316**, 288–290 (2007).
31. Fischer, M. *et al.* Analysis of the co-translational assembly of the fungal fatty acid synthase (FAS). *Sci. Rep.* **10**, 895 (2020).
32. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268 (2004).
33. Schüller, H. J., Schütz, A., Knab, S., Hoffmann, B. & Schweizer, E. Importance of general regulatory factors Rap1p, Abf1p and Reb1p for the activation of yeast fatty acid synthase genes FAS1 and FAS2. *Eur. J. Biochem.* **225**, 213–222 (1994).
34. Kastritis, P. L. *et al.* Capturing protein communities by structural proteomics in a thermophilic eukaryote. *Mol. Syst. Biol.* **13**, 936 (2017).
35. Anselmi, C., Grininger, M., Gipson, P. & Faraldo-Gómez, J. D. Mechanism of substrate shuttling by the acyl-carrier protein within the fatty acid mega-synthase. *J. Am. Chem. Soc.* **132**, 12357–12364 (2010).
36. Saito, J. *et al.* Crystal structure of enoyl-acyl carrier protein reductase (FabK) from Streptococcus pneumoniae reveals the binding mode of an inhibitor. *Protein Sci.* **17**, 691–699 (2008).
37. Sherry, L. *et al.* Comparative molecular biology approaches for the production of poliovirus virus-like particles using Pichia pastoris. *mSphere* **5**, e00838-19 (2020).
38. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
39. Zheng, S. Q. *et al.* MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
40. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
41. Scheres, S. H. W. Processing of structurally heterogeneous cryo-EM data in RELION. *Methods Enzymol.* **579**, 125–157 (2016).
42. Conley, M. J. *et al.* Calicivirus VP2 forms a portal-like assembly following receptor engagement. *Nature* **565**, 377–381 (2019).
43. Snowden, J. S. *et al.* Dynamics in the murine norovirus capsid revealed by high-resolution cryo-EM. *PLoS Biol.* **18**, e3000649 (2020).
44. Elad, N. *et al.* Structure of Type-I Mycobacterium tuberculosis fatty acid synthase at 3.3 Å resolution. *Nat. Commun.* **9**, 3886 (2018).
45. Frank, J. *et al.* SPIDER and WEB: Processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199 (1996).
46. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
47. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
48. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
49. Chen, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
50. Barad, B. A. *et al.* EMRinger: Side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12**, 943–946 (2015).
51. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
52. Dolinsky, T. J. *et al.* PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35**, W522–W525 (2007).
53. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**, 10037–10041 (2001).
54. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

55. Klebl, D. P. *et al.* Need for speed: Examining protein behavior during cryoEM grid preparation at different timescales. *Structure* **28**, 1238-1248.e4 (2020).
56. Naydenova, K. & Russo, C. J. Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nat. Commun.* **8**, 629 (2017).

Acknowledgements

We are grateful to the Wellcome Trust for PhD Studentship support to JSS (102174/B/13/Z) and to the Saudi Ministry of Education for studentship support to JA. LS is funded by the WHO (2019/883397-O, "Generation of virus free polio vaccine – phase IV"). Electron microscopy was performed in the Astbury Biostructure Laboratory, which was funded by the University of Leeds and the Wellcome Trust (108466/Z/15/Z). We also wish to thank David Klebl (University of Leeds) for assistance with identifying FAS, and we are very grateful to Glyn Hemsworth (University of Leeds) and Alison Baker (University of Leeds) for critical reading of the manuscript.

Author contributions

J.A. and L.S. maintained *P. pastoris*; J.A. purified FAS; J.S.S. prepared samples for cryoEM, carried out cryoEM data collection and performed image processing; J.S.S., J.A., D.J.R., N.A.R. and N.J.S. analysed data; J.S.S., N.A.R. and N.J.S. wrote the paper; L.S., M.S., D.J.R., N.A.R. and N.J.S. provided supervision.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-89196-2.

Correspondence and requests for materials should be addressed to N.A.R. or N.J.S.

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

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021