

ARTICLE

Open Access

Structures and comparison of endogenous 2-oxoglutarate and pyruvate dehydrogenase complexes from bovine kidney

Shiheng Liu^{1,2}, Xian Xia^{1,2}, James Zhen^{1,2,3}, Zihang Li^{1,2} and Z. Hong Zhou^{1,2,3}✉

Abstract

The α -keto acid dehydrogenase complex family catalyzes the essential oxidative decarboxylation of α -keto acids to yield acyl-CoA and NADH. Despite performing the same overarching reaction, members of the family have different component structures and structural organization between each other and across phylogenetic species. While native structures of α -keto acid dehydrogenase complexes from bacteria and fungi became available recently, the atomic structure and organization of their mammalian counterparts in native states remain unknown. Here, we report the cryo-electron microscopy structures of the endogenous cubic 2-oxoglutarate dehydrogenase complex (OGDC) and icosahedral pyruvate dehydrogenase complex (PDC) cores from bovine kidney determined at resolutions of 3.5 Å and 3.8 Å, respectively. The structures of multiple proteins were reconstructed from a single lysate sample, allowing direct structural comparison without the concerns of differences arising from sample preparation and structure determination. Although native and recombinant E2 core scaffold structures are similar, the native structures are decorated with their peripheral E1 and E3 subunits. Asymmetric sub-particle reconstructions support heterogeneity in the arrangements of these peripheral subunits. In addition, despite sharing a similar monomeric fold, OGDC and PDC E2 cores have distinct interdomain and intertrimer interactions, which suggests a means of modulating self-assembly to mitigate heterologous binding between mismatched E2 species. The lipoyl moiety lies near a mobile gatekeeper within the interdomain active site of OGDC E2 and PDC E2. Analysis of the twofold related intertrimer interface identified secondary structural differences and chemical interactions between icosahedral and cubic geometries of the core. Taken together, our study provides a direct structural comparison of OGDC and PDC from the same source and offers new insights into determinants of interdomain interactions and of architecture diversity among α -keto acid dehydrogenase complexes.

Introduction

The pyruvate dehydrogenase complex (PDC; components suffixed by 'p') is an essential element of life that catalyzes the oxidative decarboxylation of pyruvate to acetyl-coenzyme A (acetyl-CoA). This key metabolic reaction links glycolysis to oxidative phosphorylation and

the Krebs cycle. PDC is a member of the α -keto acid dehydrogenase complex family, alongside 2-oxoglutarate dehydrogenase complex (OGDC; 'o') and branched-chain α -keto acid dehydrogenase complex (BCKDC; 'b'), which all perform analogous reactions in central metabolism^{1,2}. Genetic studies have identified mutations in components of these multienzyme complexes linked to severe clinical consequences, including metabolic acidosis and neurodegeneration^{3–7}. PDC also attracts interest in cancer biology for its role in modulating the Warburg effect to promote tumor anabolism^{4,8–10}.

Correspondence: Z. Hong Zhou (hongzhou@ucla.edu)

¹Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA, USA

²California NanoSystems Institute, UCLA, Los Angeles, CA, USA

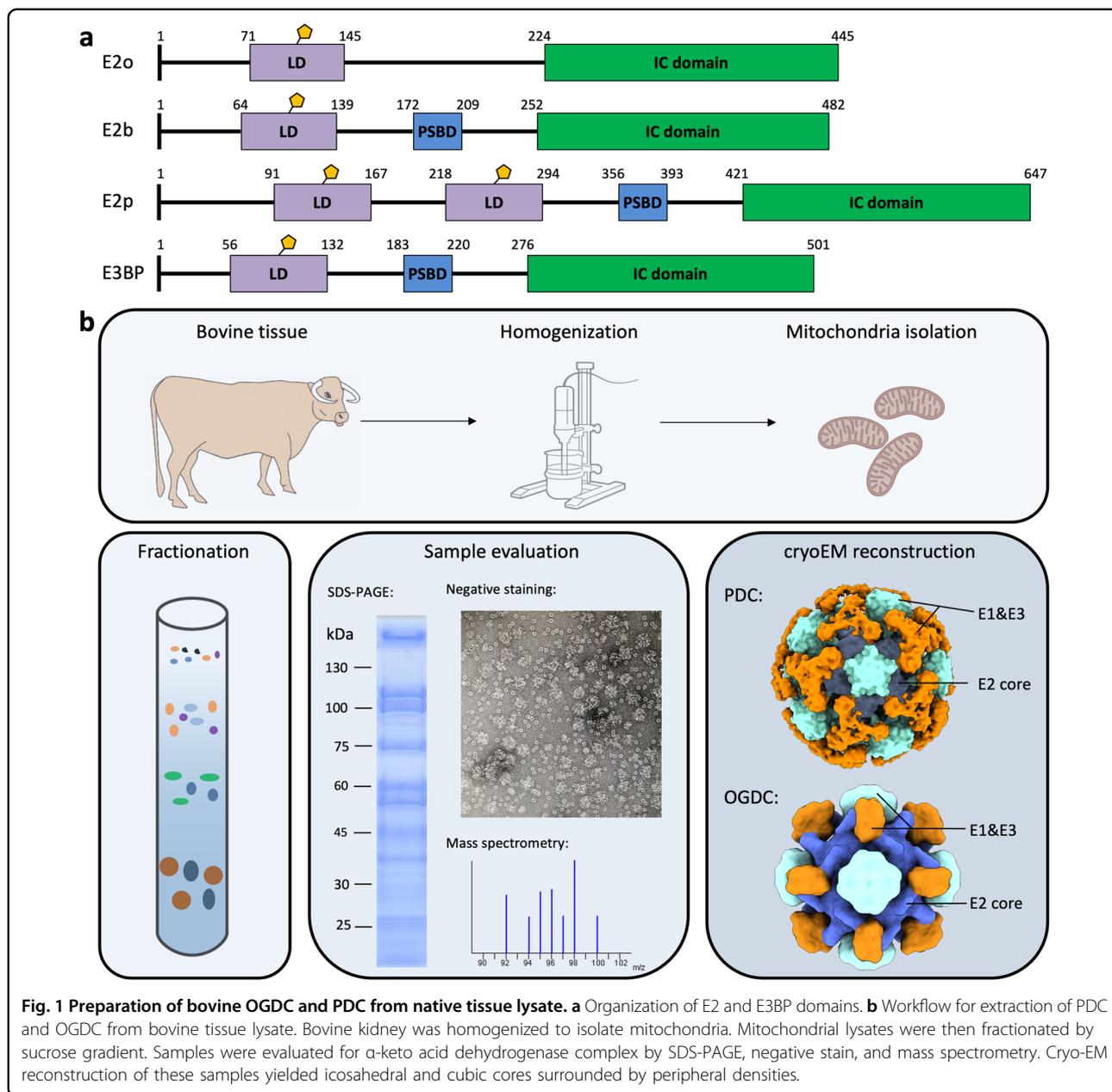
Full list of author information is available at the end of the article

These authors contributed equally: Shiheng Liu, Xian Xia, James Zhen

© The Author(s) 2022



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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.



PDC, OGDC, and BCKDC are each comprised of multiple copies of a substrate-specific dihydrolipoamide acetyltransferase (E2p), dihydrolipoamide succinyltransferase (E2o), or dihydrolipoamide acyltransferase (E2b) inner catalytic (IC) core surrounded by multiple copies of the respective α -keto acid dehydrogenase (E1p, E1o, or E1b) and universal dihydrolipoamide dehydrogenase (E3) components^{1,2}. Mammalian E2p or E2b is comprised of one or two lipoyl domains (LDs) followed by a peripheral subunit-binding domain (PSBD) and an IC domain that are connected by flexible linkers (Fig. 1a). Mammalian E2o has a similar composition but lacks the classical PSBD of E2p and E2b^{11–13}. Each α -keto acid

dehydrogenase complex proceeds through a three-step mechanism, starting with the decarboxylation of the α -keto acid substrate by E1 and transfer of an acetyl, acyl, or succinyl functional group to lipoate in the LD of E2. Then, the functionalized LD localizes to the IC core, where the functional group is transferred from the lipoyl moiety to CoA. Lastly, E3 reoxidizes the lipoyl moiety, which enables the process to repeat with the production of NADH. The LD translocates between these active sites through a flexible swinging arm mechanism¹⁴.

Although each E1 and E2 subunit in an α -keto acid dehydrogenase complex is highly similar in both sequence and overall structure to its counterpart in the other classes

of the family and the same E3 protein is shared by all classes within a tissue^{1,15,16}, there are notable differences in the architecture of how these components assemble. E1o is a homodimer, and E1p is either a homodimer or a heterotetramer in which each polypeptide of the homodimer is divided into an alpha and a beta subunit^{17,18}. In the gram-negative bacteria *Escherichia coli* and *Azotobacter vinelandii*, PDC, OGDC, and BCKDC cores all have a 24-mer cubic architecture. However, in eukaryotes and gram-positive bacteria, PDC cores adopt a different 60-mer icosahedral architecture despite OGDC and BCKDC cores retaining a cubic geometry^{16,18–24}. As an exception, actinobacteria distinctly lack assembled cores of E2p and also possess a unique OdhA protein that fuses the E2o catalytic domain to E1o²⁵. Eukaryotic PDC has an additional noncatalytic E3-binding protein (E3BP) that specifically binds E3 to the core and is structurally similar to E2^{16,19,26–32}. In mammals, E3BP substitutes for E2p in the core scaffold at possible stoichiometric ratios of 48:12 or 40:20 E2p:E3BP in an unknown arrangement^{28,31}. Although peripheral E1 and E3 subunits are bound to the core by a discrete PSBD located in E2 or E3BP, eukaryotic E2o distinctly lacks a discrete PSBD for the binding of E1 and E3^{11–13}. Instead, E3 binds an N-terminal region of E1o that shares antisera reactivity with E2p and E3BP and that is similar to them in sequence^{13,33,34}. In turn, E1o binds to E2o at a region comprised of residues from the IC domain and the preceding linker^{13,22}. Nonetheless, the subunits of each α -keto acid dehydrogenase complex self-assemble into their respective complexes²⁴.

Much of our current structural understanding of α -keto acid dehydrogenase complexes is derived from crystal structures of recombinantly expressed proteins. However, the flexible assembly of intact α -keto acid dehydrogenase complexes is unamenable to crystallography, and components were often crystallized individually. Structures from these individual proteins are unable to fully capture how multiple components come together and function as a complex. In addition, crystallization may impose artificial order on protein side chains³⁵. In the absence of the structure of native complexes, we have an insufficient understanding of how these subunits are organized into functional units. Recently, structures of native PDC from fungus and bacteria became available from cryo-electron microscopy (cryo-EM)^{16,23,32,36,37}. In the intact fungus PDC structures, E3BP was observed as an additional component appended to the interior of the core instead of substituting for an E2p subunit of the core like in mammals and contains a conserved fold found in other E2 proteins^{16,32,36}. From bacteria cultured in minimal media to minimize the activity of PDC, cryo-EM density of the lipoyl domains interacting with the PDC IC domains could be observed. Despite these exciting progresses, no

such endogenous, intact α -keto acid dehydrogenase complexes have been determined for mammals.

In this study, we report the cryo-EM structures of endogenous OGDC and PDC E2 IC domain cores extracted from bovine kidney tissue at 3.5 Å and 3.8 Å resolution, respectively. We observe peripheral subunits around the core and the lipoyl moiety substrate within the E2 active site for both complexes. By comparing structures of PDC and OGDC from the same source, we identify distinct interactions in the IC domain trimer and in the twofold related intertrimer interface, which may direct self-assembly within a milieu of similar components.

Results

Extraction, cryo-EM analysis, and identification of native, intact OGDC and PDC

To generate samples of intact α -keto acid dehydrogenase complexes, we extracted mitochondrial lysates from bovine (*Bos taurus*) kidney tissue. We then used sucrose gradient fractionation to separate and enrich the protein species present in the mitochondrial lysates (Fig. 1b). Fractions were evaluated by SDS-PAGE, western blot, and mass spectrometry to confirm the presence of PDC and OGDC (Fig. 1b; Supplementary Fig. S1a and Data S1). The mass spectrometry data indicated BCKDC to be present only as a minor species compared to PDC and OGDC. Fractions containing PDC and OGDC were pooled together into a single lysate sample for EM studies. Negative stain 2D class average evaluation of the sample on a single grid showed both icosahedral classes and cubic classes (Supplementary Fig. S2). Negative stain classes also show peripheral densities around the cores, suggesting that E1 or E3 are present and that intact α -keto acid dehydrogenase complexes were recovered. Higher-order assemblies are visible in cryo-EM micrographs (Supplementary Fig. S3a), and smeared densities are present at the periphery of well-resolved cores in the cryo-EM 2D classes (Supplementary Fig. S3b). Notably, additional classes belonging to other unknown protein species were obtained (Supplementary Figs. S2, S3b), thus highlighting the capability of cryo-EM for the study of molecular sociology. Cryo-EM 3D reconstruction yielded cubic and icosahedral cores surrounded by peripheral densities (Figs. 1b, 2a, 3a).

To structurally characterize the putative OGDC and PDC in the pooled lysate, we obtained 3D reconstructions of the cubic and icosahedral cores of the complexes by single-particle cryo-EM at 3.5 Å and 3.8 Å resolution, respectively (Figs. 2b, 3b; Supplementary Fig. S3 and Table S1). We initially identified PDC based on its distinct icosahedral architecture. This is confirmed by atomic model building of the bovine E2p sequence into the icosahedral cryo-EM density map, which showed good

and endogenous OGDC^{30,31} (Supplementary Fig. S4c), which suggests that the central density is unique to native PDC and does not belong to E2 nor E3BP.

Peripheral subunit densities were further evaluated by symmetry expanded C_1 sub-particle reconstructions. A 20 Å low-pass filter and color zone of 8 Å around the E2 core model were applied to the sub-particles for greater interpretability of low-resolution features (Supplementary Fig. S4e, f). Although the resolution is too low for unambiguous identification or positioning of E1, E3, LD, or PBSB within the external densities, we utilized the asymmetric sub-particle reconstructions to narrow down their potential localization. Thin densities corresponding to N-terminal linkers originate from both the twofold related intertrimer interface and the N-terminal end of H1 (Supplementary Fig. S4e) in E2o, but the linkers originate from only the H1 site in E2p/E3BP (Supplementary Fig. S4f). A region extending along the N-terminal end of H1 through the preceding linker (residues 224 to 240) has previously been identified as the PBSB in mammalian E2o^{13,22}. In native *E. coli* E2p during the resting state, the LD binds to a cleft flanked by the N-terminal end of H1 and the C-terminal end of H4³⁷. In both PDC and OGDC, extra densities are located near the H1 sites of E2/E3BP subunits not lining the inner face of interest, and fitting of *E. coli* E2p with bound LDs into the E2o and E2p/E3BP sub-particles supports binding of the LD to these sites³⁷. Due to the conflicting role of the H1 site as the PBSB in E2o, the H1 site of E2o may have a dual function as both the LD-binding site and the PBSB, where the binding as either role is mutually exclusive. In PDC, the extra densities at the H1 sites protrude outward and some have thin extensions (Supplementary Fig. S4f, g). These densities have varying prominence and are notably absent at some IC domains (Supplementary Fig. S4g), suggesting a heterogeneous binding of LDs or E1/E3.

The facial density of OGDC localizes around a single trimer vertex and is connected to the E2o IC domains by the linker originating from the H1 sites (Supplementary Fig. S4e). Additional smaller edge densities protrude from the linkers originating from the twofold related interface. In PDC, only a large facial density is observed (Supplementary Fig. S4f). Similar to that of OGDC, the facial density is localized near a trimer vertex and is connected to the IC domain at the H1 site. Linkers from the twofold related interface of E2p/E3BP like those in E2o were not observed. Trajectories of the linker between the IC domain and PBSB from the twofold related interface and from along the threefold axis have been previously observed but only mutually exclusively^{16,20–25,30,37,41}. In particular, the N-terminal linker of native *E. coli* E2p in the resting state folds back toward the threefold axis of the trimer after reaching the twofold related interface³⁷. Our asymmetric sub-particle reconstructions suggest that

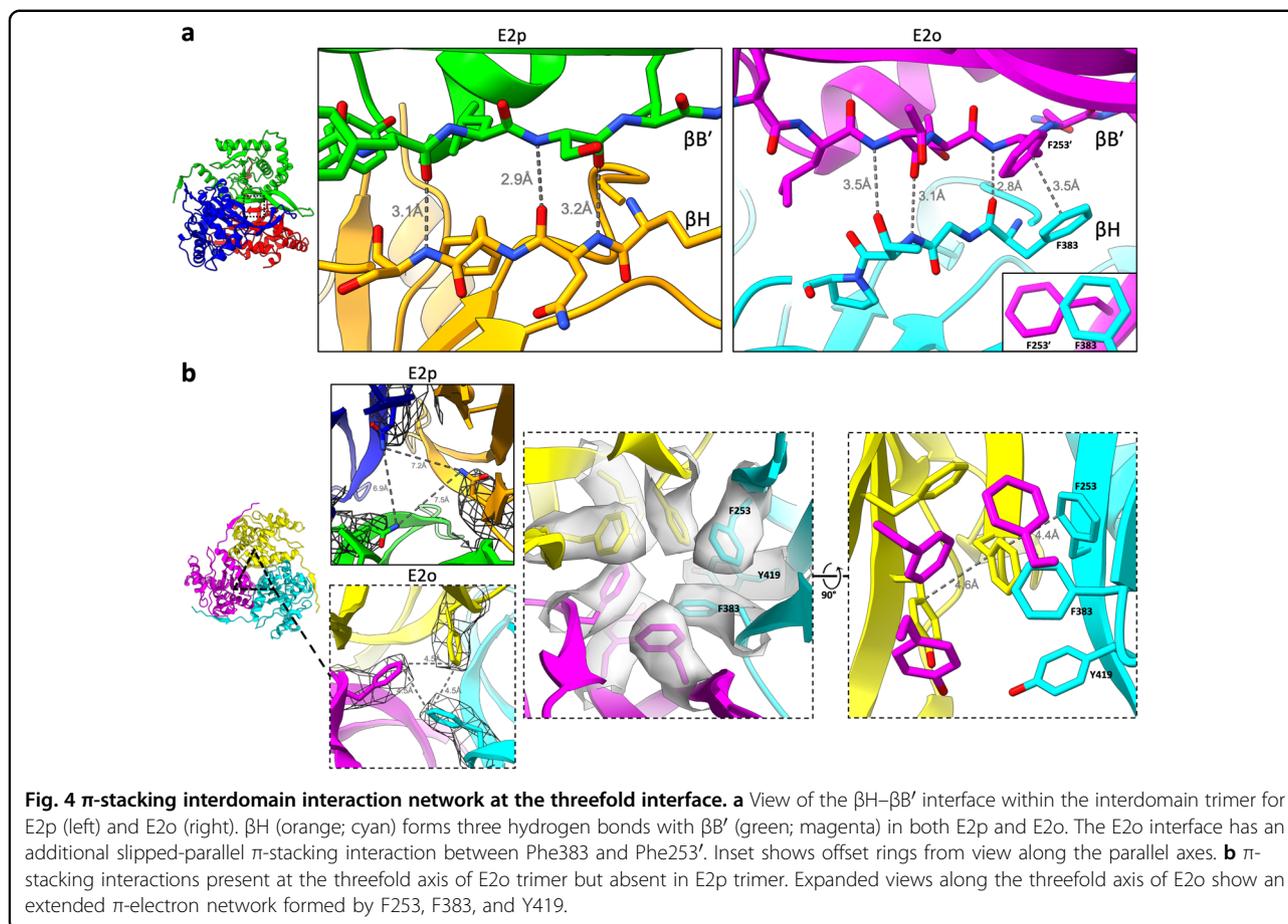
the fold of the E2 N-terminal linker is associated with the activity or reaction step of the complex, and both orientations of the linker possibly coexist based on the binding of LD, E1 or E3 (Supplementary Fig. S4e). It remains unknown whether the linker folds back to position the PBSB and LD(s) against its own IC domain or continues past the threefold axis to a neighboring IC domain.

E2o and E2p utilize distinct interactions to ensure correct self-assembly into E2 trimers and prevent incorrect heterologous binding of different E2 subunits

For E2o, the backbones of residues 220–455 are fully traceable. For E2p, the backbones of residues 420–647 are fully traceable, except for residues 519 and 520, which lack densities. Both E2o and E2p share conserved secondary structure features: six α -helices (H1–H6), a short C-terminal 3_{10} -helix (H7), and ten β -strands (β A– β J) (Figs. 2c, 3c). The secondary structure elements described here are labeled consistently with *E. coli* E2o, bovine E2b, and human E2p^{20,30,42}.

The E2 IC domains of OGDC and PDC are arranged as trimers (Figs. 2d, 3d) with 51 interdomain hydrogen bonds and an average buried surface area of 4957 Å² (standard deviation (SD) = 3.47) in E2o and an average of 42.75 (SD = 2.36) interdomain hydrogen bonds and an average buried surface area of 4397 Å² (SD = 22) in E2p. Although E2o and E2p share secondary structure features and overall fold, they have notable differences in certain regions and in the interactions that stabilize their respective trimers (Supplementary Fig. S5a). The N-terminus extends at a less acute angle toward the neighboring IC domain at the clockwise position in E2o than in E2p (Supplementary Fig. S5b). The N-terminus wraps around a turn between β C' and β D' (prime symbol is used to denote the neighboring IC domain at the clockwise position from the exterior view), which is longer in E2o than in E2p. The longer β C– β D turn in E2o enables an electrostatic interaction between the negatively charged surface of the turn and the positively charged surfaces of the N-terminus that is not present in E2p (Supplementary Fig. 5d). This positions β A in a β -sheet with β D' and β C', and the increased stability allowed for modeling of additional N-terminal residues in E2o. The extended β C– β D turn is also present in human E2o and not E2p, but it is lacking in bacteria (Supplementary Fig. S5c). In addition, E2p possesses a longer interior hairpin (Supplementary Fig. S5a), possibly for engaging in interactions in the larger volume of the icosahedral interior of PDC compared to the cubic interior of OGDC⁴³.

Within the IC domain trimer, an E2 subunit binds to its clockwise adjacent neighbor through a short β – β motif between β H and β B'. In E2p, this binding is described by three main chain-to-main chain hydrogen bonds (Fig. 4a). In E2o, this binding is described by two main chain-to-



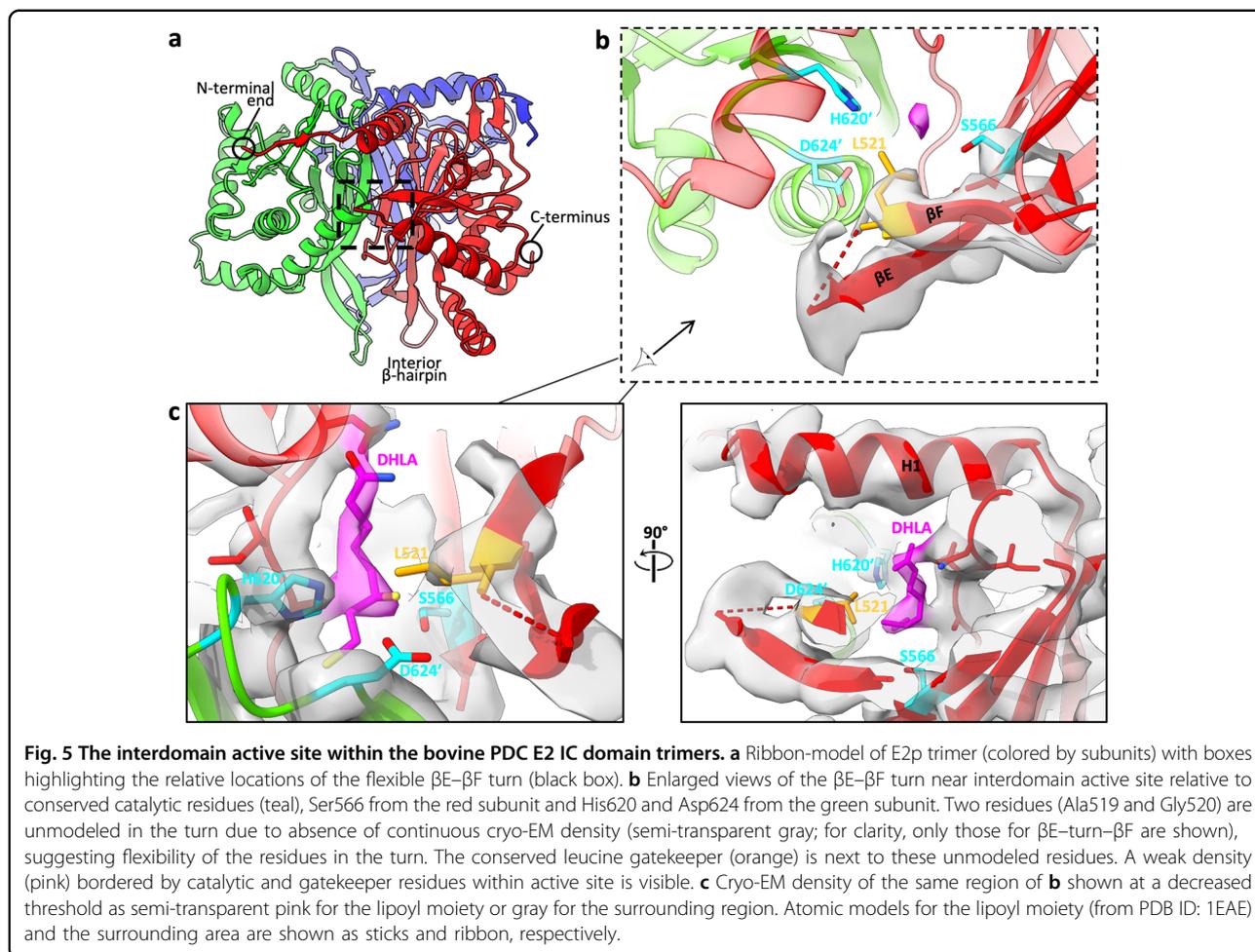
main chain hydrogen bonds, a main chain-to-side chain hydrogen bond, and a slipped-parallel π -stacking interaction⁴⁴ (Phe383–Phe253'). The slipped-parallel π -stacking interaction of E2o occurs near the threefold axis of the trimer and may form a delocalized π -electron network as previously suggested in *Chaetomium thermophilum* E2p between Arg384³⁶. By sequence and structure alignment, Phe383 in bovine E2o is the equivalent residue of Arg384 in *C. thermophilum* E2p. Notably, in bovine E2o, the delocalized π -electron network is extended within the threefold axis and encompasses Phe253, Phe383, and Tyr419 (Fig. 4b). This delocalized π -electron network is also present in human E2o but not in bovine E2p at this structurally analogous position because the equivalent bovine E2p residues Asn575 are too distant from each other²² (Fig. 4b). Like mismatched puzzle pieces, the different interfaces of E2o and E2p may prevent incorrect assembly due to co-existence of their components within the cell.

The lipoyl moiety is positioned near a flexible gatekeeper in the catalytic site

In the cryo-EM density maps of E2p, a region of flexibility has been identified: a β -turn connecting β E and β F

near the interdomain active site (Fig. 5a). Due to missing density at the β E- β F β -turn, two residues (Ala519 and Gly520) are unmodelled in our E2p model (Fig. 5b). Although the corresponding density is not missing in E2o, it is still weaker relative to that of the adjacent residues. This missing or weak density has also been observed in previous cryo-EM density maps^{16,22,23,25,37}. The density of the adjacent residue (Leu521 in E2p and Leu329 in E2o) is well-defined. This leucine is a conserved gatekeeper residue for the binding of the lipoyl moiety^{20,30}. This gatekeeper residue is proximal to the interdomain active site, near the catalytic residues Ser566 and His620' in E2p, and positions the dihydrolipoamide (DHLA) into the catalytic site during a previously suggested CoA-bound state²⁰.

Adjacent to the region of missing density, a putative density for the lipoyl moiety was identified near the catalytic residues Ser566 and His620' and the gatekeeper residue Leu521 (Fig. 5b). When the contour level is decreased, the density forms a thin extension through the active site channel and accommodates a model of DHLA⁴⁵ (adopted from PDB ID: 1EAE) (Fig. 5c). The E2o cryo-EM density map also contains a putative lipoyl



moiety density (Supplementary Fig. S6a). Following the trajectory of these densities, the dihydrolipoylsine enters the active site from the exterior of the E2 core at an entrance at the end of the LD-binding site cleft between the C-terminal end of H1 and a turn connecting β H and β I1, as previously depicted in *E. coli* E2p in the resting state³⁷. Densities of the lipoyl moiety are not present in cryo-EM density maps of native *C. thermophilum* E2p and E2o and *Neurospora crassa* E2p^{16,23,36}.

In the E2o density map at low contour level, weak cylindrical densities reminiscent of β -strands from the LD or helices from the PSBD are observed to the left and to the right of the entrance of the cleft. The left densities are located similarly to that of the LD in native *E. coli* E2p in the resting state³⁷ (Supplementary Fig. S6b). These additional densities extend along the cleft toward the active site entrance. We attempted to fit LD and PSBD models from NMR (PDB IDs: 1LAC, 1W4H) and AlphaFold prediction^{46–49} (AF-P11179-F1, AF-P11181-F1), but there are mismatches between the cylindrical densities and the LD models when the dihydrolipoylsine-containing loop is positioned adjacent to the active site channel. These

observations combined with structural variation between LD domain models suggest that the β -sheets of the LD may be conformationally variable and undergo conformational and positional changes during catalysis.

To search for E3BP, we expanded the symmetry of the trimer vertex and pentameric face sub-particles and reconstructed them asymmetrically (C_1) (Supplementary Fig. S7a). Due to low resolution of the C_1 reconstructions, we are unable to utilize side chain densities to distinguish sequence identity. Sequence alignment between E2p and E3BP and fitting of predicted E3BP structures (PDB ID: 6H60; AF-P22439-F1) show two notable regions of difference in E3BP within the IC domain^{31,47,48}: a three-residue shorter linker between H2 and H3 and a three-residue longer interior hairpin between β I2 and β J (Supplementary Fig. S7b). We searched the classes of the trimer vertex and pentameric C_1 sub-particles but were unable to identify any differences corresponding to the presence of E3BP (Supplementary Fig. S7c). E3BP is not evenly distributed in the trimers³¹, and current 3D classification methods may place too much weight on the larger features in the conserved fold of E2p and E3BP, which may reduce the quality of

asymmetric reconstructions of the interior hairpin. Heterogeneity of E3BP distribution precludes identification and experimental structural determination of E3BP from native mammalian PDC.

Different knob-and-socket intertrimer interactions prevent heterologous self-assembly of E2 trimers

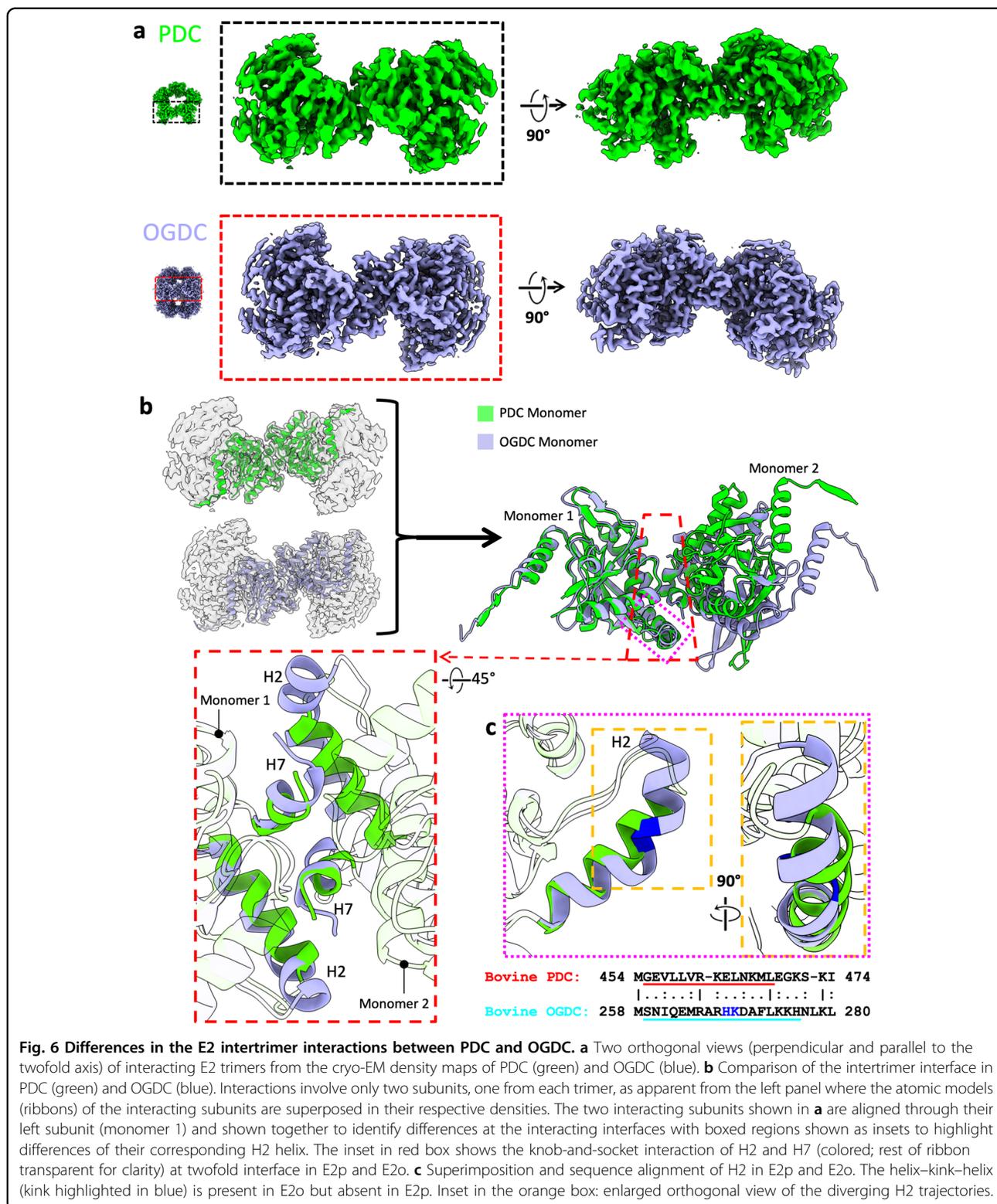
E2o and E2p have near-identical structures, yet they assemble into core scaffolds with different geometry. The cubic and icosahedral geometries are related by principles of quasi-equivalence and Euclidean geometry²⁴. Each IC domain trimer vertex of the core is bound to its neighboring trimers through a palindromic twofold related intertrimer interface^{22,30} (Fig. 6a; Supplementary Fig. S8a). The hydrophobic C-terminal 3_{10} -helix (H7) of each IC domain binds to a hydrophobic pocket formed by residues of H2*, H7*, and the N-terminal end of H4* (the asterisk is used to denote the partner IC domain across the twofold related interface) on the opposite IC domain in a knob-and-socket interaction (Supplementary Fig. S8b). This binding is further stabilized by electrostatic interactions from oppositely charged surfaces lining the exteriors of H7 and the hydrophobic pocket (Supplementary Fig. S8c). In actinobacteria E2p, where H7 folds back against rather than away from the IC domain, the knob-and-socket interaction cannot occur, so the trimers are unable to assemble into a core²⁵.

The symmetries for bovine E2o and E2p are obtained from a different combination of interactions (Supplementary Table S2). The intertrimer interface has an average of 18.7 (SD = 0.49) hydrogen bonds and an average buried surface area of 1102 Å² (SD = 0.46) in E2o and an average of 5.6 (SD = 0.63) hydrogen bonds and an average buried surface area of 825 Å² (SD = 14.18) in E2p. The exterior of the E2o socket is more charged than that of E2p (Supplementary Fig. S8c), which leads to the greater number of hydrogen bonds at the E2o intertrimer interface. In addition, the exterior surface of the socket in E2o and E2p are oppositely charged from one another. Although the residues comprising the hydrophobic sockets between the two species are similar, the shapes of the sockets are different to accommodate likewise different knobs. The E2p knob has one hydrophobic surface that binds a single hydrophobic pocket in the socket. In contrast, the E2o knob has an additional hydrophobic surface on the opposite side of the knob, and the E2o socket contains two pockets to accommodate this difference (Supplementary Fig. S8b). Similar to the purpose of the interdomain interface differences to prevent heterologous assembly of E2 subunits, the oppositely charged exterior surfaces and differently shaped knob-and-socket of the intertrimer interface between E2o and E2p suggest a means of preventing heterologous assembly of E2 trimers during core formation.

To compare the relative positions of trimers at the intertrimer interface between E2o and E2p, the IC domains of the twofold related interface were aligned by superimposition at one trimer vertex (Fig. 6b). The fold of H2 and the position of H7 differ between E2o and E2p. H7 of E2o (Pro448–Leu453) has a length of six residues and has two additional residues (Asp454 and Leu455) appended at the C-terminus. H7 of E2p (Pro642–Leu646) has a length of five residues and one additional residue (Leu647) at the C-terminus. H7 is positioned more closely to H4* in E2p than in E2o, which is enabled in E2p by an additional hydrogen bond between H7 and H4* and fewer hydrogen bonds between H7 and H2* (Supplementary Table S2). In E2o, H2 has a length of 18 residues (Ser259–His276) and is comprised of a helix–kink–helix with the kink starting at residue 268. In E2p, H2 instead has a length of 14 residues (Gly455–Leu468) and is a single, uninterrupted helix (Fig. 6c). The H2 kink in E2o positions the C-terminal end of the helix more closely against H7* for hydrogen bonds between Lys275 and His276 with Asp454* and Asp454*, respectively. The different interactions at the knob-and-socket stabilize E2 IC trimers at angles that enable the Euclidean geometry relationship between a cube and icosahedron.

Discussion

Advances in cryo-EM and its integration with mass spectrometry analysis have enabled the structural characterization of intact, native complexes^{38,50,51}. Endogenous methods have recently been applied to determine structures of PDC in fungi and bacteria^{16,23,32,36,37}. Here, we extended these methods to the study of mammalian tissues and determined structures of native bovine OGDC and PDC at sufficient resolutions for model building of their E2 IC domains (Figs. 2, 3). Despite exhaustive classification and refinement of an enormous data set, structures beyond the E2 cores in our asymmetric sub-particle reconstructions were only resolved at low resolution (Supplementary Fig. S4), indicating heterogeneity in the arrangements of the peripheral subunits attached to the core. Nonetheless, analysis of OGDC and PDC from the same source enables direct comparison between the two protein species. Although the E2 IC domains of PDC and OGDC share a nearly identical fold to conduct analogous reactions, OGDC and PDC cores differ in their interdomain (Fig. 4; Supplementary Fig. S5) and intertrimer (Fig. 6; Supplementary Fig. S8 and Table S2) interactions. These nuanced differences suggest a means of modulating self-assembly to mitigate heterologous binding between mismatched E2 species and highlight the divergent evolution of α -keto acid dehydrogenase complexes to meet the requirements of different organisms with varied cellular contents and metabolic environments. In



addition, native structures retain biologically relevant features, such as the position of the lipoyl moiety within the E2 active site (Fig. 5; Supplementary Fig. S6a), and

present opportunities for discovery in molecular sociology (Supplementary Figs. S4d, S6b) that may be lost in recombinant samples.

Besides the structures reported here, native structures of α -keto acid dehydrogenase complexes from the same endogenous source only became available early this year for the fungus *C. thermophilum*, where both PDC and OGDC were structurally characterized²³. In contrast to our bovine E2o and E2p structures, the back-folded N-terminal region comprised of β A and H1 is absent in *C. thermophilum* E2p. Interestingly, H1 is still present in *C. thermophilum* E2o from the same size-exclusion chromatography fraction²³, and is also present in native *N. crassa* and *E. coli* E2p^{16,37}. Given that the structured conformation of the N-terminal region is stabilized by electrostatic interactions between β A and H1 with the β C– β D turn (Supplementary Fig. S5d), the back-folded conformation may exist transiently depending on the activity state of the complex.

The unfolded state of H1 has been suggested to be necessary for unhindered insertion of the lipoyl moiety³⁶. For *E. coli* E2p in the absence of pyruvate, H1 is folded and binding of the N-terminus extends toward the threefold axis, which was suggested to position and stabilize the E2-LD complex in a resting state³⁷. Our sub-particle reconstruction of E2o shows the N-terminal region extending outwards from both the twofold related interface and about the threefold axis simultaneously (Supplementary Fig. S4e), which indicates that multiple conformations of the PSBD-like H1 and preceding linker of E2o are present. The conformation of the linker may correspond to the transient and mutually exclusive binding of either LD to the E2o IC domain or of E1/E3 to the PSBD.

Native structures of PDC and OGDC have been captured in various activity states. While the tertiary structure of E2 remains mostly identical between structures, the conformation of the aspartate (Asp624 in PDC and Asp430 in OGDC of bovine) that is responsible for stabilizing the protonated catalytic histidine within the active site differs. Previous crystallization experiments of *Azotobacter vinelandii* E2p with reduced forms of CoA or dihydrolipoamide yielded structures with the side chain of the aspartate-equivalent (Asn614) facing that of the catalytic histidine while other states of E2p have the aspartate-equivalent rotated away from the histidine⁴⁵. In our E2o and E2p structures, the stabilizing aspartate faces the catalytic histidine (Fig. 5b, c; Supplementary Fig. S6a), which suggests that substrate is bound in the active site. This asparagine is also facing toward the histidine in the native structure of *N. crassa* E2p¹⁶, but the asparagine is rotated away in the native structures of *C. thermophilum* E2p and E2o and *E. coli* E2p in the resting state^{23,36,37}. In recombinant human E2p and E2o, the asparagine is modeled as facing toward the histidine^{22,30}. However, the EM density for this asparagine side chain is missing, indicating motility of the side chain in the apo state. There

could be a potential mixture of activity or assembly states of OGDC and PDC in different lysate fractions when complexes are sorted by size, and our single-particle cryo-EM analysis of the entire core may have only reconstructed the most stable or populous conformation or a potential hybrid of the active and resting states.

The interdomain interactions between an E2 IC domain and its neighbor are not conserved between E2p and E2o. Bovine and human E2o contain an extensive delocalized π -electron network at the threefold axis (Fig. 4b) that is absent in their E2p counterparts^{22,30}. The differences in interdomain interfaces between E2p and E2o within an organism may enable correct self-assembly within a milieu of components in the mitochondrial matrix. *C. thermophilum* E2o may also have a similar delocalized π -electron network, but instead with Met350 at the analogous position of Phe383 in E2o²³. Met-aromatic motifs can provide additional stabilization compared to purely hydrophobic interactions⁵². Rather than a delocalized π -electron network of aromatic residues, *C. thermophilum* E2p instead possesses a potential stabilizing arginine cluster that is absent in both its E2o counterpart, mammalian E2p, and mesophilic fungus *N. crassa* E2p^{16,23,36}. The greater interdomain interactions could contribute toward increased thermal stability needed for thermophiles compared to mesophiles as a result of divergent evolution.

While there are pre-existing structures of α -keto acid dehydrogenase complexes, the diversity of interactions within these complexes between different organisms occludes direct comparison for determinants of assembly. Although *E. coli* E2p and bovine E2b are both cubic like bovine E2o, their interdomain interface interactions differ greatly. The knob of *E. coli* E2p has a semicylindrical hydrophobic surface that complements a narrow, elongated hydrophobic socket³⁷. Bovine E2b has a double-sided knob like bovine E2o²⁰, but the hydrophobic surfaces are rotated about and have different sizes. Although α -keto acid dehydrogenase complexes can share the same geometry, the interdomain interactions that hold the complex cores together differ between both family members and model organisms. Thus, structural characterization of samples from the same organism is needed to describe distinct interactions.

In mammals, E3BP substitutes for a corresponding E2 in the core scaffold as determined by biochemical analysis instead of being an additional subunit as it is in fungus^{16,28,31,32}. Although there is an internal density for our native bovine PDC structure (Supplementary Fig. S4d), the density is centralized and does not resemble the tetrahedral arrangement of fungal E3BP³⁶. In addition to the subtle differences between E2p and E3BP in the IC domain (Supplementary Fig. S7b), the linker region between the PSBD and IC domain is twice as long in E3BP

compared to that in E2p (Fig. 1a). Due to the short length of the PSBD–IC domain linker in E2p, E1p is expected to be close to the core. Likewise, the longer linker in E3BP enables binding of E3 further away from the core, possibly to better accommodate the larger size of E3 compared to E1p. Because of the heterogeneous arrangement of E3BP and peripheral E1 and E3 subunits, PDC and other α -keto acid dehydrogenase complexes may need to be studied at the individual level. Tomography could identify peripheral subunits and avoid issues of classification for E3BP if sufficient resolution could be reached in the future.

Advancements in cryo-EM have enabled atomic modeling of native α -keto acid dehydrogenase complexes and other endogenous complexes. Looking forward, faster direct electron detectors and improved cryo-EM grid preparation methods will improve acquisition of particles for low-population species for unambiguous identification and atomic modeling. Heterogeneous reconstruction could enable visualization of distinct protein conformations^{53,54}. Machine-learning-based methods in particle picking⁵⁵, structure prediction⁴⁷, and identification can resolve unknown identities in a society of proteins³⁸. Structural determination of native geometrically variable complexes from similar folds, exemplified by the cubic and icosahedral complexes presented here, should not only provide new biological insight but also inform protein engineering applications by providing an opportunity for comparison within the same environment to identify the determinants of protein architecture.

Materials and methods

Preparation of bovine mitochondria

Bovine mitochondria were prepared from *B. taurus* kidneys as described with modifications⁵⁶. 6 kg of bovine kidneys were collected and chilled on ice immediately after slaughter. Cortical tissues were cut into small slices (0.5–1 cm) and soaked in 4 L water for 1 h and then washed with 2 L kidney buffer (20 mM potassium phosphate, pH 7.6, 250 mM sucrose, 1 mM EDTA). The slices were passed through an electric meat grinder, and the ground meat was suspended in kidney buffer with 10 mM 2-mercaptoethanol and diluted to a final volume of 12 L. The suspension was homogenized using an Ultra-Turrax (IKA) homogenizer (40 s), filtered through 8 layers of cheesecloth accompanied with a sieve, and diluted to a total volume of 18 L using kidney buffer. The resulting supernatant was centrifuged at 2000–3000 $\times g$ for 10 min. The supernatant was decanted and further centrifuged at 5600 $\times g$ for 25 min. The lysosome-enriched fluffy layer was carefully removed from the mitochondrial pellet. The pellets were resuspended in a total of 8 L kidney buffer and filtered again through 8 layers of cheesecloth. The suspension was homogenized using an Ultra-Turrax

homogenizer (30 s), diluted to 12 L with kidney buffer, and centrifuged at 5600 $\times g$ for 25 min. The pellets were resuspended with kidney buffer with 0.2 mM PMSF, 0.25 μ g/mL aprotinin, 0.14 μ g/mL pepstatin, and 1 μ g/mL leupeptin in a total volume of 4 L. The suspension was treated with 0.01% digitonin for 15 min to remove outer mitochondrial membranes and diluted to 12 L using kidney buffer. The resulting mitoplasts were concentrated by centrifugation at ~25,600 $\times g$ for 20 min followed by removal of excess kidney buffer. The pellet of the final step was flash frozen in liquid nitrogen and stored at -80°C .

Isolation of bovine PDC and OGDC complexes

100 g of frozen mitoplasts were thawed in 150 mL lysis buffer (20 mM potassium phosphate, pH 7.6, 50 mM NaCl, 0.02 mM thiamin pyrophosphate, 20 mM MgCl_2 , 2 mM EGTA, protease inhibitors (2 mM PMSF, 2.5 μ g/mL aprotinin, 1.4 μ g/mL pepstatin, and 10 μ g/mL leupeptin)). The sample was diluted to 337.5 mL with lysis buffer. 37.5 mL Triton X-100 buffer (lysis buffer components in addition with 16% (v/v) Triton X-100) was then added, and the sample was stirred for 15 min at 4°C . The suspension was centrifuged (SLA-3000, 13,000 rpm, 20 min, 4°C), and the supernatant was PEG precipitated in 5% (w/v) PEG 10,000 for 15 min. The precipitate was collected by centrifugation (2500 $\times g$, 7 min, 4°C) and resuspended in 35 mL lysis buffer. The suspension was homogenized using a Dounce homogenizer and centrifuged (SW28 Ti rotor, 28,000 rpm, 17 min, 4°C). The resulting supernatant was loaded onto 50% (w/v) sucrose cushions (15 mL) and centrifuged (SW41 rotor, 40,000 rpm, 24 h, 4°C). Pellets were each dissolved in 0.5 mL sample buffer (50 mM MOPS-KOH buffer, pH 7.0, 0.02 mM thiamin pyrophosphate, 10 mM MgCl_2 , 0.1 mM EGTA, protease inhibitors, 1 mM DTT) by shaking at 230 rpm for 1 h and clarified by centrifugation (SLA-3000, 16,000 rpm, 20 min, 4°C). The sample was loaded onto 10%–40% (w/v) sucrose gradients (1 mL per gradient) and centrifuged (SW41 Ti, 26,000 rpm, 12 h, 4°C). The gradients were fractionated and assessed by SDS-PAGE and western blot using a PDC complex antibody cocktail (Abcam) that contains four different mAbs reacting specifically with E1 α , E1 β , and E2/E3BP. The fractions containing PDC and OGDC complexes were pooled, dialyzed to final buffer (50 mM MOPS-KOH buffer, pH 7.0, 0.02 mM thiamin pyrophosphate, 1 mM MgCl_2 , 1 mM NAD^+ , 0.1 mM EGTA, protease inhibitors, 1 mM DTT), and concentrated for cryo-EM analysis. Liquid chromatography-tandem mass spectrometry was performed on the sucrose gradient fractions. Gradient fractions corresponding to OGDC were subjected to gel filtration on a Superose 6 column (GE Healthcare) to remove ferritin.

Cryo-EM sample preparation and image acquisition

Cryo-EM grids were prepared by using an FEI Mark IV Vitrobot (Thermo Fisher Scientific). 3 μ L of sample was applied onto a glow-discharged lacey carbon copper grid with a thin, continuous carbon film (Ted Pella). After waiting for 30 s, the grid was blotted (8 °C, 100% humidity, 10 s blotting time, 1 blotting force) and plunge-frozen in liquid ethane. Grids were loaded into a Titan Krios (Thermo Fisher Scientific) equipped with a Gatan Imaging Filter (GIF) Quantum LS and a Gatan K2 Summit direct electron detector. Movies were recorded with SerialEM at super-resolution mode⁵⁷. The nominal magnification was 105,000 \times , corresponding to a calibrated pixel size of 0.68 Å at the specimen level. The total exposure time for each movie was set to 8 s and fractionated equally into 40 frames, leading to a total dosage of \sim 45 electrons/Å². Defocus was set from -1.8 to -2.6 μ m.

Image processing

The movies were motion-corrected with MotionCor2⁵⁸. 6959 good images were selected from a total of 9029 by manual screening. After defocus determination by CTFIND4⁵⁹, particles were automatically picked using Gautomatch (<https://www2.mrc-lmb.cam.ac.uk/research/locally-developed-software/zhang-software/>). 33,138 particles for PDC complex and 41,165 particles for OGDC complex were obtained by 2D and 3D classification in RELION-3⁶⁰. FSC calculations were performed in RELION-3.

For PDC complex, all particles from the previous step were recentered and extracted with a box size of 576 pixels. These particles were refined with icosahedral symmetry, resulting in a map with a resolution of 3.8 Å. To get better density for the peripheral E1 and E3 subunits, a structure-guided sub-particle strategy was used. First, the STAR file from refinement was expanded with icosahedral symmetry, leading to a total of 1,988,280 particles. Second, the center of the expanded particles was shifted to one trimer or pentameric face. Sub-particles were extracted with a box size sufficient for accommodating E1 and E3. Finally, 3D classifications with a local mask focused on either the internal or external region were performed. No classes with identifiable features were obtained for the internal region. For the external region, one class with 13.4% of all particles best displays a connecting density between E2p and the putative E1/E3/LD/PSBD. The particles in this class were subjected to further classifications with different parameters (skip-align classification and local search classification), but they did not yield improved density maps for E1/E3/LD/PSBD.

For OGDC complex, all particles from the previous step were recentered and extracted with a box size of 384 pixels. These particles were refined with octahedral

symmetry and a mask around the E2 core, resulting in a map with a resolution of 3.5 Å. To get better density for the peripheral E1 and E3 subunits, a structure-guided sub-particle strategy was used. First, the STAR file from refinement was expanded with octahedral symmetry, leading to a total of 987,960 particles. Second, the center of the expanded particles was shifted to one face of the cubic core, and particles were extracted in RELION with a box size of 208. Finally, 3D classification with a local mask of the external area was performed. One class with 12.7% of all particles showed the best connections between E2 and the exterior E1/E3. The particles in this class were subjected to further classifications with different parameters (skip-align classification and local search classification). However, no maps with better density of exterior E1/E3/LD were obtained.

Atomic modeling

For model building, human E2o (PDB ID: 6H05) and human E2p (PDB ID: 6CT0) were used as templates for bovine E2o and E2p IC domains, respectively^{22,30}. Each monomeric subunit model was fitted into the respective cryo-EM density map of the core scaffold using ChimeraX⁶¹, mutated and manually refined in Coot⁶², and real-space refined in Phenix⁶³. The monomer model was then duplicated to the appropriate stoichiometry and fitted into the respective map and refined iteratively using ISOLDE and Phenix⁶⁴. FindMySequence and checkMySequence were later used to confirm the identity of OGDC^{38,39}. Electrostatic potential maps were calculated in ChimeraX. Sequence alignments were performed with Clustal Omega and EMBOSS Needle on the EMBL-EBI server^{65–67}. Structure predictions were obtained from AlphaFold using ColabFold or the EBI database^{47,48,68}.

Acknowledgements

This research was supported in part by grants from NSF (DMR-1548924) and NIH (R01GM071940). J.Z. acknowledges support from an Interdisciplinary Training in Virology and Gene Therapy training grant (NIH 5T32AI060567). We acknowledge the use of resources in the Electron Imaging Center for Nanomachines supported by UCLA and grants from NIH (S10RR23057, S10OD018111, and U24GM116792) and NSF (DBI-1338135). We thank the UCLA Proteome Research Center for assistance in mass spectrometry.

Author details

¹Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA, USA. ²California NanoSystems Institute, UCLA, Los Angeles, CA, USA. ³Molecular Biology Institute, UCLA, Los Angeles, CA, USA

Author contributions

Z.H.Z. conceived the project and supervised the research. S.L. designed the experimental protocols. S.L. and X.X. prepared samples, collected cryo-EM images, and determined the 3D structures. Z.L. and J.Z. built atomic models and made figures. J.Z. interpreted the results and wrote the manuscript. All authors reviewed and approved the manuscript.

Data availability

Mass spectrometry data and source gel/blot images are provided with this paper. Cryo-EM density maps of the cubic bovine OGDC core and icosahedral

bovine PDC core have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-26649 and EMD-26650, respectively. The coordinates of the E2 α IC domain and bovine E2 β IC domain have been deposited in the Protein Data Bank under accession numbers 7UOL and 7UOM, respectively.

Conflict of interest

The authors declare no competing interests.

Publisher's note

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

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41421-022-00487-y>.

Received: 22 April 2022 Accepted: 20 October 2022

Published online: 22 November 2022

References

- Patel, M. S. & Roche, T. E. Molecular biology and biochemistry of pyruvate dehydrogenase complexes. *FASEB J.* **4**, 3224–3233 (1990).
- Reed, L. J. A Trail of research from lipoic acid to α -keto acid dehydrogenase complexes. *J. Biol. Chem.* **276**, 38329–38336 (2001).
- Åvarsson, A. et al. Crystal structure of human branched-chain α -ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease. *Structure* **8**, 277–291 (2000).
- Roche, T. E. & Hiromasa, Y. Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer. *Cell. Mol. Life Sci.* **64**, 830 (2007).
- Shi, Q. et al. Inactivation and reactivation of the mitochondrial α -ketoglutarate dehydrogenase complex. *J. Biol. Chem.* **286**, 17640–17648 (2011).
- Burrage, L. C., Nagamani, S. C. S., Campeau, P. M. & Lee, B. H. Branched-chain amino acid metabolism: from rare Mendelian diseases to more common disorders. *Hum. Mol. Genet.* **23**, R1–R8 (2014).
- Whitley, M. J. et al. Pyruvate dehydrogenase complex deficiency is linked to regulatory loop disorder in the α V138M variant of human pyruvate dehydrogenase. *J. Biol. Chem.* **293**, 13204–13213 (2018).
- McFate, T. et al. Pyruvate dehydrogenase complex activity controls metabolic and malignant phenotype in cancer cells. *J. Biol. Chem.* **283**, 22700–22708 (2008).
- Tran, Q., Lee, H., Park, J., Kim, S.-H. & Park, J. Targeting cancer metabolism - revisiting the Warburg Effects. *Toxicol. Res.* **32**, 177–193 (2016).
- Chen, J. et al. Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer. *Nat. Genet.* **50**, 219–228 (2018).
- Nakano, K. et al. Purification and molecular cloning of succinyltransferase of the rat α -ketoglutarate dehydrogenase complex. Absence of a sequence motif of the putative E3 and/or E1 binding site. *J. Biol. Chem.* **266**, 19013–19017 (1991).
- Nakano, K. et al. Isolation, characterization and structural organization of the gene and pseudogene for the dihydrolipoamide succinyltransferase component of the human 2-oxoglutarate dehydrogenase complex. *Eur. J. Biochem.* **224**, 179–189 (1994).
- Zhou, J. et al. A multipronged approach unravels unprecedented protein–protein interactions in the human 2-oxoglutarate dehydrogenase multienzyme complex. *J. Biol. Chem.* **293**, 19213–19227 (2018).
- Perham, R. N. Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* **69**, 961–1004 (2000).
- Yu, X. et al. Structures of the human pyruvate dehydrogenase complex cores: a highly conserved catalytic center with flexible N-terminal domains. *Structure* **16**, 104–114 (2008).
- Forsberg, B. O., Aibar, S., Howard, R. J., Mortezaei, N. & Lindahl, E. Arrangement and symmetry of the fungal E3BP-containing core of the pyruvate dehydrogenase complex. *Nat. Commun.* **11**, 4667 (2020).
- Frank, R. A. W., Price, A. J., Northrop, F. D., Perham, R. N. & Luisi, B. F. Crystal structure of the E1 component of the *Escherichia coli* 2-oxoglutarate dehydrogenase multienzyme complex. *J. Mol. Biol.* **368**, 639–651 (2007).
- Reed, L. J. Multienzyme complexes. *Acc. Chem. Res.* **7**, 40–46 (1974).
- Stoops, J. K. et al. On the unique structural organization of the *Saccharomyces cerevisiae* pyruvate dehydrogenase complex. *J. Biol. Chem.* **272**, 5757–5764 (1997).
- Kato, M. et al. A synchronized substrate-gating mechanism revealed by cubic-core structure of the bovine branched-chain α -ketoacid dehydrogenase complex. *EMBO J.* **25**, 5983–5994 (2006).
- Andi, B. et al. Structure of the dihydrolipoamide succinyltransferase catalytic domain from *Escherichia coli* in a novel crystal form: a tale of a common protein crystallization contaminant. *Acta Crystallogr. F Struct. Biol. Commun.* **75**, 616–624 (2019).
- Nagy, B. et al. Structure of the dihydrolipoamide succinyltransferase (E2) component of the human α -ketoglutarate dehydrogenase complex (hKGDHC) revealed by cryo-EM and cross-linking mass spectrometry: Implications for the overall hKGDHC structure. *Biochim. Biophys. Acta Gen. Subj.* **1865**, 129889 (2021).
- Skalidis, I. et al. Cryo-EM and artificial intelligence visualize endogenous protein community members. *Structure* **30**, 575–589 (2022).
- Izard, T. et al. Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *Proc. Natl. Acad. Sci. USA* **96**, 1240–1245 (1999).
- Bruch, E. M. et al. Actinobacteria challenge the paradigm: A unique protein architecture for a well-known, central metabolic complex. *Proc. Natl. Acad. Sci. USA* **118**, e2112107118 (2021).
- Marcucci, O. D. & Lindsay, J. G. Component X. An immunologically distinct polypeptide associated with mammalian pyruvate dehydrogenase multi-enzyme complex. *Eur. J. Biochem.* **149**, 641–648 (1985).
- Jilka, J. M., Rahmatullah, M., Kazemi, M. & Roche, T. E. Properties of a newly characterized protein of the bovine kidney pyruvate dehydrogenase complex. *J. Biol. Chem.* **261**, 1858–1867 (1986).
- Hiromasa, Y., Fujisawa, T., Aso, Y. & Roche, T. E. Organization of the cores of the mammalian pyruvate dehydrogenase complex formed by E2 and E2 plus the E3-binding protein and their capacities to bind the E1 and E3 components. *J. Biol. Chem.* **279**, 6921–6933 (2004).
- Brautigam, C. A. et al. Structural insight into interactions between dihydrolipoamide dehydrogenase (E3) and E3 binding protein of human pyruvate dehydrogenase complex. *Structure* **14**, 611–621 (2006).
- Jiang, J. et al. Atomic structure of the E2 inner core of human pyruvate dehydrogenase complex. *Biochemistry* **57**, 2325–2334 (2018).
- Prajapati, S. et al. Structural and functional analyses of the human PDH complex suggest a “division-of-labor” mechanism by local E1 and E3 clusters. *Structure* **27**, 1124–1136.e4 (2019).
- Kyriilis, F. L. et al. Integrative structure of a 10-megadalton eukaryotic pyruvate dehydrogenase complex from native cell extracts. *Cell Rep.* **34**, 108727 (2021).
- Rice, J. E., Dunbar, B. & Lindsay, J. G. Sequences directing dihydrolipoamide dehydrogenase (E3) binding are located on the 2-oxoglutarate dehydrogenase (E1) component of the mammalian 2-oxoglutarate dehydrogenase multienzyme complex. *EMBO J.* **11**, 3229–3235 (1992).
- McCartney, R. G. et al. Subunit interactions in the mammalian α -ketoglutarate dehydrogenase complex: evidence for direct association of the α -ketoglutarate dehydrogenase and dihydrolipoamide dehydrogenase components. *J. Biol. Chem.* **273**, 24158–24164 (1998).
- Bartesaghi, A., Matthies, D., Banerjee, S., Merk, A. & Subramaniam, S. Structure of β -galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad. Sci. USA* **111**, 11709–11714 (2014).
- Tütting, C. et al. Cryo-EM snapshots of a native lysate provide structural insights into a metabolon-embedded transacetylase reaction. *Nat. Commun.* **12**, 6933 (2021).
- Škerlová, J., Berndtsson, J., Nolte, H., Ott, M. & Stenmark, P. Structure of the native pyruvate dehydrogenase complex reveals the mechanism of substrate insertion. *Nat. Commun.* **12**, 5277 (2021).
- Chojnowski, G. et al. findMySequence: a neural-network-based approach for identification of unknown proteins in X-ray crystallography and cryo-EM. *IUCr* **9**, 86–97 (2022).
- Chojnowski, G. Sequence-assignment validation in cryo-EM models with checkMySequence. *Acta Crystallogr. D Struct. Biol.* **78**, 806–816 (2022).
- Murphy, G. E. & Jensen, G. J. Electron cryotomography of the *E. coli* pyruvate and 2-oxoglutarate dehydrogenase complexes. *Structure* **13**, 1765–1773 (2005).
- Mattevi, A. et al. Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* **255**, 1544–1550 (1992).
- Knapp, J. E. et al. Crystal structure of the truncated cubic core component of the *Escherichia coli* 2-oxoglutarate dehydrogenase multienzyme complex. *J. Mol. Biol.* **280**, 655–668 (1998).

43. Zhou, Z. H. et al. Direct evidence for the size and conformational variability of the pyruvate dehydrogenase complex revealed by three-dimensional electron microscopy: the “breathing” core and its functional relationship to protein dynamics. *J. Biol. Chem.* **276**, 21704–21713 (2001).
44. Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M. & Tanabe, K. Origin of attraction and directionality of the π/π interaction: model chemistry calculations of benzene dimer interaction. *J. Am. Chem. Soc.* **124**, 104–112 (2002).
45. Mattevi, A., Obmolova, G., Kalk, K. H., Teplyakov, A. & Hol, W. G. J. Crystallographic analysis of substrate binding and catalysis in dihydrolipoyl transacetylase (E2p). *Biochemistry* **32**, 3887–3901 (1993).
46. Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. Three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037–1048 (1993).
47. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
48. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
49. Ferguson, N. et al. Ultra-fast barrier-limited folding in the peripheral subunit-binding domain family. *J. Mol. Biol.* **353**, 427–446 (2005).
50. Ho, C.-M. et al. Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. *Nat. Methods* **17**, 79–85 (2020).
51. Ho, C.-M. et al. Native structure of the RhopH complex, a key determinant of malaria parasite nutrient acquisition. *Proc. Natl. Acad. Sci. USA* **118**, e2100514118 (2021).
52. Valley, C. C. et al. The methionine-aromatic motif plays a unique role in stabilizing protein structure. *J. Biol. Chem.* **287**, 34979–34991 (2012).
53. Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18**, 176–185 (2021).
54. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
55. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
56. Pettit, F. H. & Reed, L. J. [65] Pyruvate dehydrogenase complex from bovine kidney and heart. in *Methods in Enzymology* **89** 376–386 (Academic Press, 1982).
57. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
58. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
59. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
60. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
61. Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
62. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
63. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
64. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018).
65. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European molecular biology open software suite. *Trends Genet.* **16**, 276–277 (2000).
66. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
67. McWilliam, H. et al. Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* **41**, W597–W600 (2013).
68. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).