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], Z. Hong Zhou[1,2,3,*]

1 Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA 90095, USA

2 California NanoSystems Institute, UCLA, Los Angeles, CA 90095, USA

3 Molecular Biology Institute, UCLA, Los Angeles, CA 90095, USA

# Contributed equally

* Corresponding Author

Email: hong.zhou@ucla.edu; Phone: 310-694-7527

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 their native states remain unknown. Here, we report the cryo electron microscopy (cryoEM) structures of the endogenous cubic 2-oxoglutarate dehydrogenase complex (OGDC) and icosahedral pyruvate dehydrogenase complex (PDC) cores from bovine kidney determined at 3.5 Å and 3.8 Å resolution, respectively. The structures of multiple protein 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, native structures are decorated with their peripheral E1 and E3 subunits. Asymmetric sub-particle reconstructions support heterogeneity in the arrangements of these peripheral subunits. Additionally, 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 two-fold related intertrimer interface identified secondary structural differences and chemical interactions between icosahedral and cubic geometries of the core. Taken together, our study provides 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]. Genomic 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].

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]. Bovine E2 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). 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 mechanism[11].

Although each E1, E2, and E3 subunit in an α-keto acid dehydrogenase is highly similar both in sequence and overall structure to its counterpart in the other classes of the family[12,13], there are notable differences in the architecture of how these components assemble. E1p is a

52 heterotrimer composed two alpha and two beta subunits, whereas E1o is a homodimer where
53 the alpha and beta subunits are a singular polypeptide[14]. In bacteria, PDC, OGDC, and BCKDC
54 cores all have a 24-mer cubic architecture. As an exception, actinobacteria distinctly lack
55 assembled cores of E2p and possess a unique OdhA protein that fuses the E2o catalytic
56 domain to E1o[15]. However, in eukaryotes, PDC cores adopt a different 60-mer icosahedral
57 architecture despite OGDC and BCKDC cores retaining a cubic geometry[13,16–20]. Additionally,
58 eukaryotic PDC have an additional noncatalytic E3-binding protein (E3BP) that specifically binds
59 E3 to the core and is structurally similar to E2[13,16,21–27]. In mammals, E3BP substitutes for E2p in
60 the core scaffold at possible stoichiometric ratios of 48:12 or 40:20 E2p:E3BP in an unknown
61 arrangement[23,26]. Nonetheless, the subunits of each α-keto acid dehydrogenase complex self-
62 assemble into their respective complexes[28].

63 Much of our current structural understanding of α-keto acid dehydrogenase complexes is
64 derived from crystal structures of recombinantly expressed proteins. However, the flexible
65 assembly of intact α-keto acid dehydrogenase complexes is unamenable to crystallography,
66 and components were often crystallized individually. Structures from these individual proteins
67 are unable to fully capture how multiple components come together and function as a complex.
68 Additionally, crystallization may impose artificial order on protein side chains[29]. In the absence
69 of the structure of native complexes, we have an insufficient understanding of how these
70 subunits are organized into functional units. Recently, structures of native PDC from fungus and
71 bacteria became available from cryo electron microscopy (cryoEM)[13,20,27,30,31]. In the intact
72 fungus PDC structures, E3BP was observed as an additional component appended to the
73 interior of the core instead of substituting for an E2p subunit of the core like in mammals and
74 contains a conserved fold found in other E2 proteins[13,27,30]. From bacteria cultured in minimal
75 media to minimize the activity of PDC, cryoEM density of the lipoyl domains interacting with the
76 PDC IC domains could be observed. Despite these exciting progresses, no such endogenous,
77 intact α-keto acid dehydrogenase complexes have been determined for mammals.

78 In this study, we report the cryoEM 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 two-fold related intertrimer interface, which may direct self-assembly within a milieu of similar components.

Results

Extraction, cryoEM, 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. 1a). 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 cryoEM micrographs (Supplementary Fig. S3a), and smeared densities are present at the periphery of well-resolved cores in the cryoEM 2D classes (Supplementary Fig. S3b). Notably, additional classes belonging to other unknown protein species were obtained (Supplementary Fig. S2, S3b), thus highlighting the capability of cryoEM for the study of molecular sociology. CryoEM 3D reconstruction yielded cubic and icosahedral cores surrounded by peripheral densities (Fig. 1b, 2a, 3a).

To structurally characterize the putative OGDC and PDC in these two samples, we obtained 3D reconstructions of the cubic and icosahedral cores of the complexes by single-

106 particle cryoEM at 3.5 Å and 3.8 Å resolution, respectively (Fig. 2b, 3b; Supplementary Fig. S3;
107 Supplementary Table S1). We initially identified PDC based on its distinct icosahedral
108 architecture. This is confirmed by atomic model building of the bovine E2p sequence into the
109 icosahedral cryoEM density map, which showed good agreement between the sequence and
110 the respective side chain densities. FindMySequence results also support the identity
111 assignment of PDC[32] (E-value 4.80e-74, 3.30e-19, and 2.40e-15 for E2p, E2o, and E2b
112 sequences, respectively).

113 Although PDC is distinct for its icosahedral architecture, both OGDC and BCKDC share
114 cubic architecture, which introduces ambiguity to the identification of cubic α-keto acid
115 dehydrogenase complexes. To confirm whether our cubic complex was OGDC, BCKDC, or a
116 hybrid of both, we created a homology model by mutating a pre-existing human E2o model to
117 the bovine E2o sequence[19]. After fitting and real-space refinement of the model into our cubic
118 cryoEM density map, the side chains for the E2o model fit the density well (Fig. 2e). Sequence
119 alignment of bovine E2o and E2b reveals primary structure features in E2b that cannot
120 correspond to side chain densities in our map. Fitting of a pre-existing E2b model into our map
121 also showed discrepancies between the modeled side chains and the density[17].
122 FindMySequence and checkMySequence results also support the identity assignment of
123 OGDC[32,33] (E-value 5.30e-89 and 6.30e-11 for E2o and E2b sequences, respectively). The
124 unambiguous identification of the cubic complex as OGDC is also reflected in the low presence
125 of BCKDC components in the mass spectrometry data.

126 **Heterogeneous arrangement of peripheral subunits around the native E2 cores**

127 Unmasked 3D reconstructions have external densities suggestive of E1 or E3 around the core
128 (Fig. 2a, 3a). Both OGDC and PDC have a large density over each core face. In PDC, thin
129 densities from the trimer vertices extend toward the density above the pentameric face. OGDC
130 has additional smaller densities along each cubic edge with thin densities of putative linker
131 regions connecting to the E2o trimer vertices. Previous models based on hydrogen/deuterium
132 exchange and cross-linking mass spectrometry studies propose that E1o binds at the edges

133 and E3 binds at the faces[19,34]. The opposite orientation of the small connecting densities
134 between OGDC and PDC suggests different predominant states of activity and a large range of
135 motion for the flexible linker regions. However, enforced symmetry during reconstruction
136 restricts the interpretability of these external densities for the structures and locations of E1 and
137 E3.

138 Peripheral subunit densities were further evaluated by symmetry expanded C_1 sub-
139 particle reconstructions (Supplementary Fig. S4a, b). A 20 Å low pass filter and color zone of 8
140 Å around the E2 core model were applied to the sub-particles for greater interpretability of low-
141 resolution features (Supplementary Fig. S4c, d). Although the resolution is too low for
142 unambiguous positioning of pre-existing E1 or E3 models within the external densities, we
143 utilized the asymmetric sub-particle reconstructions to narrow down their potential localization.
144 Thin densities corresponding to N-terminal linkers originate from both the two-fold related
145 intertrimer interface and a previously proposed LD binding site in E2o and only the LD binding
146 site in E2p/E3BP[31]. In both PDC and OGDC, extra densities are located near the LD binding
147 sites not adjacent to the face of interest, and fitting of *E. coli* E2p with bound LDs into the E2o
148 sub-particle supports binding of the LD to these sites[31]. In PDC, the extra density at the LD
149 binding site protrudes outward and some have thin extensions (Supplementary Fig. S4d, e).
150 These densities have varying prominence and are notably absent at some IC domains
151 (Supplementary Fig. S4e), suggesting a heterogeneous binding of LDs.

152 The facial density of OGDC localizes around a single trimer vertex and is connected to
153 the E2o IC domains by the linker originating from the LD binding sites (Supplementary Fig.
154 S4c). Additional smaller densities protrude from the linkers originating from the two-fold related
155 interface. In PDC, only a large facial density is observed (Supplementary Fig. S4d). Similar to
156 that of OGDC, the facial density is localized near a trimer vertex and is connected to the IC
157 domain at the LD binding site. Linkers from the two-fold related interface of E2p/E3BP like those
158 in E2o were not observed. Trajectories of the linker between the IC domain and PSBD from the
159 two-fold related interface and along the three-fold axis have been previously observed but only

160 mutually exclusively[13,15,15,17–20,25,28,31,35]. In particular, the N-terminal linker of native *E. coli* E2p in
161 the resting state folds back towards the three-fold axis of the trimer after reaching the two-fold
162 related interface[31]. Our asymmetric sub-particle reconstructions suggests that the fold of the E2
163 N-terminal linker is associated with the activity or reaction step of the complex, and both
164 conformations of the PSBD possibly coexist based on the binding of E1 or E3 (Supplementary
165 Fig. S4c). It remains unknown whether the linker folds back to position the PSBD and LD(s)
166 against its own IC domain or continues past the three-fold axis to a neighboring IC domain.

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

169 For E2o, the backbone of residues 220-455 are fully traceable. For E2p, the backbone of
170 residues 420-647 are fully traceable, except for residues 519 and 520, which lack densities.
171 Both E2o and E2p share conserved secondary structure features: six α-helices (H1-6), a short
172 C-terminal 3_{10}-helix (H7), and ten β-strands (βA-J) (Fig. 2c, 3c). The secondary structure
173 elements described here are labeled consistently with *E. coli* E2o, bovine E2b, and human
174 E2p[17,25,36].

175 The E2 IC domains of OGDC and PDC are arranged as trimers (Fig. 2d, 3d) with 51
176 interdomain hydrogen bonds and an average buried surface area of 4957 $Å^2$ (standard deviation
177 (SD) = 3.47) buried surface in E2o and an average of 42.75 (SD = 2.36) interdomain hydrogen
178 bonds and an average buried surface area of 4397 $Å^2$ (SD = 22) in E2p. Although E2o and E2p
179 share secondary structure features and overall fold, they have notable differences in certain
180 regions and in the interactions that stabilize their respective trimers (Fig. 4). The N-terminus
181 extends at a more elevated angle towards the neighboring IC domain at the clockwise position
182 in E2o than in E2p (Fig. 4a). The N-terminus wraps around a turn between βC′ and βD′ (prime
183 symbol is used to denote the neighboring IC domain at the clockwise position from the exterior
184 view), which is longer in E2o than in E2p. The longer βC-βD turn in E2o enables an electrostatic
185 interaction between the negatively charged surface of the turn and the positively charged
186 surfaces of the N-terminus that is not present in E2p (Fig. 4c). This positions βA in a β-sheet

187 with βD′ and βC′, and the increased stability allowed for modeling of additional N-terminal
188 residues in E2o. The extended βC-βD turn is also present in human E2o and not E2p, but it is
189 lacking in bacteria (Fig. 4b**)**. Additionally, E2p possesses a longer interior hairpin (Fig. 4a),
190 possibly for engaging in interactions in the larger volume of the icosahedral interior of PDC
191 compared to the cubic interior of OGDC[37].

192 Within the IC domain trimer, an E2 subunit binds to its clockwise adjacent neighbor
193 through a short β-β motif between βH and βB′. In E2p, this binding is described by three main
194 chain-to-main chain hydrogen bonds (Supplementary Fig. S5). In E2o, this binding is described
195 by two main chain-to-main chain hydrogen bonds, a main chain-to-side chain hydrogen bond,
196 and a slipped-parallel π-stacking interaction[38] (Phe383-Phe253′). The slipped-parallel π-
197 stacking interaction of E2o occurs near the three-fold axis of the trimer and may form a
198 delocalized π-electron network as previously suggested in *C. thermophilum* E2p between
199 Arg384[30]. By sequence and structure alignment, Phe383 in bovine E2o is the equivalent residue
200 of Arg384 in *C. thermophilum* E2p. Notably, in bovine E2o, the delocalized π-electron network is
201 extended within the three-fold axis and encompasses Phe253, Phe383, and Tyr419 (Fig. 4d).
202 This delocalized π-electron network is also present in human E2o but not in bovine E2p at this
203 structurally analogous position because the equivalent residues Asn575 are too distant from
204 each other[19] (Fig. 4d). Like mismatched puzzle pieces, the different interfaces of E2o and E2p
205 may prevent incorrect assembly due to co-existence of their components within the cell.

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

207 In the cryoEM density maps of E2p, a region of flexibility has been identified: a β-turn
208 connecting βE and βF near the interdomain active site (Fig. 5a). Due to missing density at the
209 βE-βF β-turn, two residues (Ala519 and Gly520) are unmodelled in our E2p model (Fig. 5b**)**.
210 Although the corresponding density is not missing in E2o, it is still weaker relative to that of the
211 adjacent residues. This missing or weak density has also been observed in previous cryoEM
212 density maps[13,15,19,20,31]. The density of the adjacent residue (Leu521 in E2p and Leu329 in E2o)
213 is well-defined. This leucine is a conserved gatekeeper residue for the binding of the lipoyl

214 moiety[17,25]. This gatekeeper residue is proximal to the interdomain active site, near the catalytic
215 residues Ser566 and His620′ in E2p, and positions the dihydrolipoamide (DHLA) into the
216 catalytic site upon binding of CoA, as suggested previously[17].

217 In the E2p cryoEM density map, a putative density for the lipoyl moiety was identified
218 near the catalytic residues Ser566 and His620′ and the gatekeeper residue Leu521 (Fig. 5b).
219 When the contour level is decreased, the density forms a thin extension through the active site
220 channel and accommodates a model of DHLA[35] (adopted from PDB ID: 1EAE) (Fig. 5c). The
221 E2o cryoEM density map also contains a putative lipoyl moiety density (Supplementary Fig.
222 S6a). Following the trajectory of these densities, the dihydrolipoyllysine enters the active site
223 from the exterior of the E2 core at an entrance at the end of a cleft between H1 and a turn
224 connecting βH and βI1, as previously depicted in *E. Coli* E2p in the resting state[31]. Densities of
225 the lipoyl moiety are not present in cryoEM density maps of native *C. thermophilum* E2p and
226 E2o and *N. crassa* E2p[13,20,30]. In the E2o density map at low contour level, weak cylindrical
227 densities reminiscent of β-strands from the LD or helices from the PSBD are observed to the left
228 and to the right of the entrance of the cleft. The left densities are located similarly to that of the
229 LD in native *E. coli* E2p in the resting state[31]. (Supplementary Fig. S6b). These additional
230 densities extend along the cleft towards the active site entrance. We attempted to fit an LD
231 model and PBSD models from NMR (PDB ID: 1LAC; 1W4H) and Alphafold prediction[39–42] (AF-
232 P11179-F1; AF-P11181-F1), but there are mismatches between the cylindrical densities and the
233 LD models when the dihydrolipoyllysine-containing loop is positioned adjacent to the active site
234 channel. These observations combined with structural variation between LD domain models
235 suggest that the β-sheets of the LD may be conformationally variable and undergo
236 conformational and positional changes during catalysis.

237 To search for E3BP, we expanded the symmetry of the trimer vertex and pentameric
238 face sub-particles and reconstructed them asymmetrically (C_1) (Supplementary Fig. S7a). Due
239 to low resolution of the C_1 reconstructions, we are unable to utilize side chain densities to
240 distinguish sequence identity. Sequence alignment between E2p and E3BP and fitting of

241 predicted E3BP structures (PDB ID: 6H60; AF-P22439-F1) show two notable regions of
242 difference in E3BP within the IC domain[26,40,41]: a three-residue shorter linker between H2 and
243 H3 and a three-residue longer interior hairpin between βI2 and βJ (Supplementary Fig. S7b).
244 We searched the classes of the trimer vertex and pentameric C_1 sub-particles but were unable
245 to identify any differences corresponding to the presence of E3BP (Supplementary Fig. S7c).
246 E3BP is not evenly distributed in the trimers[26], and current 3D classification methods may place
247 too much weight on the larger features in the conserved fold of E2p and E3BP, which may
248 reduce the quality of asymmetric reconstructions of the interior hairpin. Heterogeneity of E3BP
249 distribution precludes identification and experimental structural determination of E3BP from
250 native mammalian PDC.

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

253 E2o and E2p have near-identical structures, yet they assemble into core scaffolds with different
254 geometry. The cubic and icosahedral geometries are related by principles of quasi-equivalence
255 and Euclidean geometry[28]. Each IC domain trimer vertex of the core is bound to its neighboring
256 trimers through a palindromic two-fold related intertrimer interface[19,25] (Fig. 6a; Supplementary
257 Fig. S8a). The hydrophobic C-terminal 3_{10}-helix (H7) of each IC domain binds to a hydrophobic
258 pocket formed by residues of H2*, H7*, and the N-terminal end of H4* (asterisk is used to
259 denote the partner IC domain across the two-fold related interface) on the opposite IC domain in
260 a knob-and-socket interaction (Supplementary Fig. S8b). This binding is further stabilized by
261 electrostatic interactions from oppositely charged surfaces lining the exteriors of H7 and the
262 hydrophobic pocket (Supplementary Fig. S8c). In actinobacteria E2p, where H7 folds back
263 against rather than away from the IC domain, the knob-and-socket interaction cannot occur, so
264 the trimers are unable assemble into a core[15].

265 The symmetries for bovine E2o and E2p are obtained from a different combination of
266 interactions (Supplementary Table S2). The intertrimer interface has an average of 18.7 (SD =
267 0.49) hydrogen bonds and an average buried surface area of 1102 $Å^2$ (SD = 0.46) in E2o and

268 an average of 5.6 (SD = 0.63) hydrogen bonds and an average buried surface area of 825 $Å^2$
269 (SD = 14.18) in E2p. The exterior of the E2o socket is more charged than that of E2p
270 (Supplementary Fig. S8c), which leads to the greater number of hydrogen bonds at the E2o
271 intertrimer interface. Additionally, the exterior surface of the socket in E2o and E2p are
272 oppositely charged from one another. Although the residues comprising the hydrophobic
273 sockets between the two species are similar, the shapes of the sockets are different to
274 accommodate likewise different knobs. The E2p knob has one hydrophobic surface that binds a
275 single hydrophobic pocket in the socket. However, the E2o knob has an additional hydrophobic
276 surface on the opposite side of the knob, and the E2o socket contains two pockets to
277 accommodate this difference (Supplementary Fig. S8b). Similar to the purpose of the
278 interdomain interface differences to prevent heterologous assembly of E2 subunits, the
279 oppositely charged exterior surfaces and differently shaped knob-and-socket of the intertrimer
280 interface between E2o and E2p suggest a means of preventing heterologous assembly of E2
281 trimers during core formation.
282 To compare the relative positions of trimers at the intertrimer interface between E2o and
283 E2p, the IC domains of the two-fold related interface were aligned by superimposition at one
284 trimer vertex (Fig. 6b). The fold of H2 and the position of H7 differ between E2o and E2p. H7 of
285 E2o (Pro448-Leu453) has a length of six residues and has two additional residues (Asp454 and
286 Leu455) appended at the C-terminus. H7 of E2p (Pro642-Leu646) has a length of five residues
287 and one additional residue (Leu647) at the C-terminus. H7 is positioned more closely to H4* in
288 E2p than in E2o, which is enabled in E2p by an additional hydrogen bond between H7 and H4*
289 and fewer hydrogen bonds between H7 and H2*. (Supplementary Table S2). In E2o, H2 has a
290 length of 18 residues (Ser259-His276) and is comprised of a helix-kink-helix with the kink
291 starting at residue 268. In E2p, H2 has a length of 14 residues (Gly455-Leu468) and is a single,
292 uninterrupted helix (Fig. 6c). The H2 kink in E2o positions the C-terminal end of the helix more
293 closely against H7* for hydrogen bonds between Lys275 and His276 with Asp454* and
294 Asp454*, respectively. The different interactions at the knob-and-socket stabilize E2 IC trimers

295 at angles that enable the Euclidean geometry relationship between a cube and icosahedron.

296

297 **Discussion**

298 Advances in cryoEM and its integration with mass spectrometry analysis have enabled the
299 structural characterization of intact, native complexes[32,43,44]. Endogenous methods have recently
300 been applied to determine structures of PDC in fungi and bacteria[13,20,27,30,31]. Here, we extended
301 these methods to the study of mammalian tissues and determined structures of native bovine
302 OGDC and PDC at sufficient resolutions for model building of their E2 IC domains. Analysis of
303 OGDC and PDC from the same source enables direct comparison between the two protein
304 species. Although E2 IC domains share a nearly identical fold to conduct analogous reactions,
305 nuanced differences in their structures highlight the divergent evolution of α-keto acid
306 dehydrogenase complexes to meet the requirements of different organisms with varied cellular
307 contents and metabolic environments.

308 Besides the structures reported here, native structures of α-keto acid dehydrogenase
309 complexes from the same endogenous source only became available early this year for the
310 fungus *C. thermophilum*, where both PDC and OGDC were structurally characterized[20]. In
311 contrast to our bovine E2o and E2p structures, the back-folded N-terminal region comprised of
312 βA and H1 is absent in *C. thermophilum* E2p. Interestingly, H1 is still present in *C. thermophilum*
313 E2o from the same size-exclusion chromatography fraction[20], and is also present in native *N.*
314 *crassa* and *E. coli* E2p[13,31]. Given that the structured conformation of the N-terminal region is
315 stabilized by electrostatic interactions between βA and H1 with the βC-βD turn (Fig. 4c), the
316 back-folded conformation may exist transiently depending on the activity state of the complex.

317 The unfolded state has been suggested to be necessary for unhindered insertion of the
318 lipoyl moiety[30]. For *E. coli* E2p in the absence of pyruvate, H1 is folded and binding of the N-
319 terminus extends towards the three-fold axis, which was suggested to position and stabilize the
320 E2-LD complex in a resting state[31]. Our sub-particle reconstructions show the N-terminal region
321 extending outwards from both the two-fold related interface and about the three-fold axis

322 simultaneously (Supplementary Fig. S4c), which indicates that multiple conformations of the
323 PSBD linker are present. The conformation of the linker may correspond to the transient binding
324 of E1 and E3. There could be a potential mixture of activity or assembly states in different lysate
325 fractions when complexes are sorted by size, and our single-particle cryoEM analysis of the
326 entire core might have only reconstructed the most stable or populous conformation or a
327 potential hybrid of the active and resting states.

328 The interdomain interactions between an E2 IC domain and its neighbor are not
329 conserved between E2p and E2o. Bovine and human E2o contain an extensive delocalized π-
330 electron network at the three-fold axis (Fig. 4d) that is absent in their E2p counterparts[19,25]. The
331 differences in interdomain interfaces between E2p and E2o within an organism may enable
332 correct self-assembly within a milieu of components in the mitochondrial matrix. *C.*
333 *thermophilum* E2o may also have a similar delocalized π-electron network, but instead with
334 Met350 at the analogous position of Phe383 in E2o[20]. Met-aromatic motifs can provide
335 additional stabilization compared to purely hydrophobic interactions[45]. Rather than a delocalized
336 π-electron network of aromatic residues, *C. thermophilum* E2p instead possesses a potential
337 stabilizing arginine cluster that is absent in both its E2o counterpart, mammalian E2p, and
338 mesophilic fungus *N. crassa* E2p[13,20,30]. The greater interdomain interactions could contribute
339 towards increased thermal stability needed for thermophiles compared to mesophiles as a result
340 of divergent evolution.

341 While there are pre-existing structures of α-keto acid dehydrogenase complexes, the
342 diversity of interactions within these complexes between different organisms occludes direct
343 comparison for determinants of assembly. Although *E.coli* E2p and bovine E2b are both cubic
344 like bovine E2o, their interdomain interface interactions differ greatly. The knob of *E. coli* E2p
345 has a semicylindrical hydrophobic surface that complements a narrow, elongated hydrophobic
346 socket[31]. Bovine E2b has a double-sided knob like bovine E2o[17], but the hydrophobic surfaces
347 are rotated about and have different sizes. Although α-keto acid dehydrogenase complexes can
348 share the same geometry, the interdomain interactions that hold the complex cores together

349 differ between both family members and model organisms. Thus, structural characterization of
350 samples from the same organism are needed to describe distinct interactions.
351 In mammals, E3BP substitutes for a corresponding E2 in the core scaffold as
352 determined by biochemical analysis instead of being an additional subunit as it is in
353 fungus[13,23,26,27], and no internal density was observed for our native bovine PDC structure.
354 Because of the heterogeneous arrangement of E3BP and peripheral E1 and E3 subunits, PDC
355 and other α-keto acid dehydrogenase complex may need to be studied at the individual level.
356 Tomography could identify peripheral subunits and avoid issues of classification for E3BP if
357 sufficient resolution could be reached in the future.
358 Advancements in cryoEM have enabled atomic modeling of native α-keto acid
359 dehydrogenase complexes and other endogenous complexes. Looking forward, faster direct
360 electron detectors and improved cryoEM grid preparation methods will improve acquisition of
361 particles for low population species for unambiguous identification and atomic modeling.
362 Heterogeneous reconstruction could enable visualization of distinct protein conformations[46,47].
363 Machine-learning based methods in particle picking[48], structure prediction[40], and identification
364 can resolve unknown identities in a society of proteins[32]. Structural determination of native
365 geometrically variable complexes from similar folds, exemplified by the cubic and icosahedral
366 complexes presented here, should not only provide new biological insight but also inform protein
367 engineering applications by providing an opportunity for comparison within the same
368 environment to identify the determinants of protein architecture.

369 ## Materials and methods

370 ### Preparation of bovine mitochondria

371 Bovine mitochondria were prepared from *Bos taurus* kidneys as described with modifications[49].
372 6 kg of bovine kidneys were collected and chilled on ice immediately after slaughter. Cortical
373 tissues were cut into small slices (0.5-1 cm) and soaked in 4 L water for 1 h and then washed
374 with 2 L kidney buffer (20 mM potassium phosphate pH 7.6, 250 mM sucrose, 1 mM EDTA).
375 The slices were passed through an electric meat grinder, and the ground meat was suspended
376 in kidney buffer with 10 mM 2-mercaptoethanol and diluted to a final volume of 12 L. The
377 suspension was homogenized using an Ultra-Turrax (IKA) homogenizer (40 s), filtered through 8
378 layers of cheesecloth accompanied with sieve, and diluted to a total volume of 18 L using kidney
379 buffer. The resulting supernatant was centrifuged at 2,000-3,000 × g for 10 min. The
380 supernatant was decanted and further centrifuged at 5,600 × g for 25 min. The lysosome-
381 enriched fluffy layer was carefully removed from the mitochondrial pellet. The pellets were
382 resuspended in a total of 8 L kidney buffer and filtered again through 8 layers of cheesecloth.
383 The suspension was homogenized using an Ultra-Turrax homogenizer (30 s), diluted to 12 L
384 with kidney buffer, and centrifuged at 5,600 × g for 25 min. The pellets were resuspended with
385 kidney buffer with 0.2 mM PMSF, 0.25 μg/ml aprotinin, 0.14 μg/ml pepstatin, and 1 μg/ml
386 leupeptin in a total volume of 4 L. The suspension was treated with 0.01% digitonin for 15 min to
387 remove outer mitochondrial membranes and diluted to 12 L using kidney buffer. The resulting
388 mitoplasts were concentrated by centrifugation at ~25,600 × g for 20 min followed by removal of
389 excess kidney buffer. The pellet of the final step was flash frozen in liquid nitrogen and stored at
390 −80 °C.

391 ### Isolation of bovine PDC and OGDC complexes

392 100 g of frozen mitoplasts were thawed in 150 mL lysis buffer [20 mM potassium phosphate pH
393 7.6, 50 mM NaCl, 0.02 mM thiamin pyrophosphate, 20 mM $MgCl_2$, 2 mM EGTA, protease
394 inhibitors (2 mM PMSF, 2.5 μg/mL aprotinin, 1.4 μg/mL pepstatin, and 10 μg/mL leupeptin)].
395 The sample was diluted to 337.5 mL with lysis buffer. 37.5 mL Triton X-100 buffer [lysis buffer

396 components in addition with 16% (v/v) Triton X-100] was then added, and the sample was
397 stirred for 15 min at 4 °C. The suspension was centrifuged (SLA-3000, 13,000 rpm, 20 min,
398 4 °C), and the supernatant was PEG precipitated in 5% (w/v) PEG 10,000 for 15 min. The
399 precipitate was collected by centrifugation (2,500 × g, 7 min, 4 °C) and re-suspended in 35 mL
400 lysis buffer. The suspension was homogenized using a Dounce homogenizer and centrifuged
401 (SW28 Ti rotor, 28,000 rpm, 17 min, 4 °C). The resulting supernatant was loaded onto 50% (w/v)
402 sucrose cushions (15 ml) and centrifuged (SW41 rotor, 40,000 rpm, 24 h, 4 °C). Pellets were
403 each dissolved in 0.5 mL sample buffer (50 mM MOPS-KOH buffer pH 7.0, 0.02 mM thiamin
404 pyrophosphate, 10 mM $MgCl_2$, 0.1 mM EGTA, protease inhibitors, 1mM DTT) by shaking at
405 230 rpm for 1 h and clarified by centrifugation (SLA-3000, 16,000 rpm, 20 min, 4 °C). The
406 sample was loaded onto 10–40% (w/v) sucrose gradients (1 mL per gradient) and centrifuged
407 (SW41 Ti, 26,000 rpm, 12 h, 4 °C). The gradients were fractionated and assessed by SDS-
408 PAGE and western blot using a PDC complex antibody cocktail (Abcam) that contains four
409 different mAbs reacting specifically with E1α, E1β, and E2/E3BP. The fractions containing PDC
410 and OGDC complexes were pooled, dialyzed to final buffer (50 mM MOPS-KOH buffer pH 7.0,
411 0.02 mM thiamin pyrophosphate, 1 mM $MgCl_2$, 1 mM NAD+, 0.1 mM EGTA, protease inhibitors,
412 1 mM DTT), and concentrated for cryoEM analysis. Liquid chromatography-tandem mass
413 spectrometry was performed on the sucrose gradient fractions. Gradient fractions corresponding
414 to OGDC were subjected to gel filtration on a Superose 6 column (GE Healthcare) to remove
415 ferritin.

416 **CryoEM sample preparation and image acquisition**

417 CryoEM grids were prepared by using an FEI Mark IV Vitrobot (Thermo Fisher Scientific). 3 μL
418 of sample was applied onto a glow-discharged lacey carbon copper grid with a thin, continuous
419 carbon film (Ted Pella). After waiting for 30 s, the grid was blotted (8°C, 100% humidity, 10 s
420 blotting time, 1 blotting force) and plunge-frozen in liquid ethane. Grids were loaded into a Titan
421 Krios (Thermo Fisher Scientific) equipped with a Gatan Imaging Filter (GIF) Quantum LS and a
422 Gatan K2 Summit direct electron detector. Movies were recorded with SerialEM at super-

423 resolution mode[50]. The nominal magnification was 105,000 ×, corresponding to a calibrated
424 pixel size of 0.68 Å at the specimen level. The total exposure time for each movie was set to 8 s
425 and fractionated equally into 40 frames, leading to a total dosage of ~45 electrons/Å^2. Defocus
426 was set to -1.8 to -2.6 μm.

427 **Image processing**

428 The movies were motion corrected with MotionCor2 (Zheng, 2017). 6,959 good images were
429 selected from a total of 9,029 by manual screening. After defocus determination by
430 CTFFIND4[51], particles were automatically picked using Gautomatch (https://www2.mrc-
431 lmb.cam.ac.uk/research/locally-developed-software/zhang-software/). 33,138 particles for PDC
432 complex and 41,165 particles for OGDC complex were obtained by 2D and 3D classification in
433 RELION-3[52]. FSC calculations were performed in RELION-3.

434 For PDC complex, all particles from the previous step were recentered and extracted
435 with a box size of 576 pixels. These particles were refined with icosahedral symmetry, resulting
436 in a map with a resolution of 3.8 Å. To get better density for the peripheral E1 and E3 subunits,
437 a structure-guided sub-particle strategy was used. First, the STAR file from refinement was
438 expanded with icosahedral symmetry, leading to a total of 1,988,280 particles. Second, the
439 center of the expanded particles was shifted to one trimer or pentameric face. Finally, 3D
440 classification and refinement focused on the putative E1 and E3 region was performed.

441 For OGDC complex, all particles from the previous step were recentered and extracted
442 with a box size of 384 pixels. These particles were refined with octahedral symmetry and a
443 mask around the E2 core, resulting in a map with a resolution of 3.5 Å. To get better density for
444 the peripheral E1 and E3 subunits, a structure-guided sub-particle strategy was used. First, the
445 STAR file from refinement was expanded with octahedral symmetry, leading to a total of
446 987,960 particles. Second, the center of the expanded particles was shifted to one face of the
447 cubic core, and particles were extracted in RELION with a box size of 208. Finally, 3D
448 classification and refinement of the external area was performed.

449 **Atomic modeling**

450 For model building, human E2o (PDB ID: 6H05) and human E2p (PDB ID: 6CT0) were used as
451 templates for bovine E2o and E2p IC domains, respectively[19,25]. Each monomeric subunit model
452 was fitted into the respective cryoEM density map of the core scaffold using ChimeraX[53],
453 mutated and manually refined in Coot[54], and real-space refined in Phenix[55]. The monomer
454 model was then duplicated to the appropriate stoichiometry and fit into the respective map and
455 refined iteratively using ISOLDE and Phenix[56]. FindMySequence and checkMySequence were
456 later used to confirm the identity of OGDC[32,33]. Electrostatic potential maps were calculated in
457 ChimeraX. Sequence alignments were performed [Sequence alignments performed with Clustal
458 Omega and EMBOSS Needle on the EMBL-EBI server[57–59]. Structure predictions were obtained
459 from Alphafold using Colabfold or the EBI database[40,41,60].
460
461 **Data availability** Mass spectrometry data and source gel/blot images are provided with this
462 paper. CryoEM density maps of the icosahedral bovine PDC core and cubic bovine OGDC core
463 have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-
464 XXXX and EMD-XXXX, respectively. The coordinates of the bovine E2p IC domain and E2o IC
465 domain have been deposited in the Protein Data Bank under accession numbers XXXX and
466 XXXX, respectively.
467 **Acknowledgments** This research was supported in part by grants from NSF (DMR-1548924)
468 and NIH (R01GM071940). J.Z. acknowledges support from an Interdisciplinary Training in
469 Virology and Gene Therapy training grant (NIH 5T32AI060567). We acknowledge the use of
470 resources in the Electron Imaging Center for Nanomachines supported by UCLA and grants
471 from NIH (S10RR23057, S10OD018111, and U24GM116792) and NSF (DBI-1338135). We
472 thank the UCLA Proteome Research Center for assistance in mass spectrometry.
473 **Contributions** Z.H.Z. conceived the project and supervised the research. S.L designed the
474 experimental protocols. S.L. and X.X. prepared samples, collected cryoEM images, and
475 determined the 3D structures. Z.L. and J.Z. built atomic models and made figures. J.Z.
476 interpreted the results and wrote the manuscript; all authors reviewed and approved the paper.

Conflict of interest The authors declare no competing interests.

479 **References**

480 1. Patel, M. S. & Roche, T. E. Molecular biology and biochemistry of pyruvate dehydrogenase
481 complexes. *FASEB J.* **4**, 3224–3233 (1990).
482 2. Reed, L. J. A Trail of Research from Lipoic Acid to α-Keto Acid Dehydrogenase Complexes.
483 *J. Biol. Chem.* **276**, 38329–38336 (2001).
484 3. Ævarsson, A. *et al.* Crystal structure of human branched-chain α-ketoacid dehydrogenase
485 and the molecular basis of multienzyme complex deficiency in maple syrup urine disease.
486 *Structure* **8**, 277–291 (2000).
487 4. Roche, T. E. & Hiromasa, Y. Pyruvate dehydrogenase kinase regulatory mechanisms and
488 inhibition in treating diabetes, heart ischemia, and cancer. *Cell. Mol. Life Sci.* **64**, 830 (2007).
489 5. Shi, Q. *et al.* Inactivation and Reactivation of the Mitochondrial α-Ketoglutarate
490 Dehydrogenase Complex. *J. Biol. Chem.* **286**, 17640–17648 (2011).
491 6. Burrage, L. C., Nagamani, S. C. S., Campeau, P. M. & Lee, B. H. Branched-chain amino
492 acid metabolism: from rare Mendelian diseases to more common disorders. *Hum. Mol.*
493 *Genet.* **23**, R1–R8 (2014).
494 7. Whitley, M. J. *et al.* Pyruvate dehydrogenase complex deficiency is linked to regulatory loop
495 disorder in the αV138M variant of human pyruvate dehydrogenase. *J. Biol. Chem.* **293**,
496 13204–13213 (2018).
497 8. McFate, T. *et al.* Pyruvate Dehydrogenase Complex Activity Controls Metabolic and
498 Malignant Phenotype in Cancer Cells. *J. Biol. Chem.* **283**, 22700–22708 (2008).
499 9. Tran, Q., Lee, H., Park, J., Kim, S.-H. & Park, J. Targeting Cancer Metabolism - Revisiting
500 the Warburg Effects. *Toxicol. Res.* **32**, 177–193 (2016).
501 10. Chen, J. *et al.* Compartmentalized activities of the pyruvate dehydrogenase complex sustain
502 lipogenesis in prostate cancer. *Nat. Genet.* **50**, 219–228 (2018).
503 11. Perham, R. N. Swinging Arms and Swinging Domains in Multifunctional Enzymes: Catalytic
504 Machines for Multistep Reactions. *Annu. Rev. Biochem.* **69**, 961–1004 (2000).

505 12. Yu, X. *et al.* Structures of the Human Pyruvate Dehydrogenase Complex Cores: A Highly
506 Conserved Catalytic Center with Flexible N-Terminal Domains. *Structure* **16**, 104–114
507 (2008).
508 13. Forsberg, B. O., Aibara, S., Howard, R. J., Mortezaei, N. & Lindahl, E. Arrangement and
509 symmetry of the fungal E3BP-containing core of the pyruvate dehydrogenase complex. *Nat.*
510 *Commun.* **11**, 4667 (2020).
511 14. Frank, R. A. W., Price, A. J., Northrop, F. D., Perham, R. N. & Luisi, B. F. Crystal Structure
512 of the E1 Component of the Escherichia coli 2-Oxoglutarate Dehydrogenase Multienzyme
513 Complex. *J. Mol. Biol.* **368**, 639–651 (2007).
514 15. Bruch, E. M. *et al.* Actinobacteria challenge the paradigm: A unique protein architecture for a
515 well-known, central metabolic complex. *Proc. Natl. Acad. Sci.* **118**, (2021).
516 16. Stoops, J. K. *et al.* On the Unique Structural Organization of the Saccharomyces cerevisiae
517 Pyruvate Dehydrogenase Complex. *J. Biol. Chem.* **272**, 5757–5764 (1997).
518 17. Kato, M. *et al.* A synchronized substrate-gating mechanism revealed by cubic-core structure
519 of the bovine branched-chain α-ketoacid dehydrogenase complex. *EMBO J.* **25**, 5983–5994
520 (2006).
521 18. Andi, B. *et al.* Structure of the dihydrolipoamide succinyltransferase catalytic domain from
522 Escherichia coli in a novel crystal form: a tale of a common protein crystallization
523 contaminant. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **75**, 616–624 (2019).
524 19. Nagy, B. *et al.* Structure of the dihydrolipoamide succinyltransferase (E2) component of the
525 human alpha-ketoglutarate dehydrogenase complex (hKGDHc) revealed by cryo-EM and
526 cross-linking mass spectrometry: Implications for the overall hKGDHc structure. *Biochim.*
527 *Biophys. Acta BBA - Gen. Subj.* **1865**, 129889 (2021).
528 20. Skalidis, I. *et al.* Cryo-EM and artificial intelligence visualize endogenous protein community
529 members. *Structure* doi:10.1016/j.str.2022.01.001 (2022).

530 21. Marcucci, O. D. & Lindsay, J. G. Component X. An immunologically distinct polypeptide
531 associated with mammalian pyruvate dehydrogenase multi-enzyme complex. *Eur. J.*
532 *Biochem.* **149**, 641–648 (1985).
533 22. Jilka, J. M., Rahmatullah, M., Kazemi, M. & Roche, T. E. Properties of a newly characterized
534 protein of the bovine kidney pyruvate dehydrogenase complex. *J. Biol. Chem.* **261**, 1858–
535 1867 (1986).
536 23. Hiromasa, Y., Fujisawa, T., Aso, Y. & Roche, T. E. Organization of the Cores of the
537 Mammalian Pyruvate Dehydrogenase Complex Formed by E2 and E2 Plus the E3-binding
538 Protein and Their Capacities to Bind the E1 and E3 Components. *J. Biol. Chem.* **279**, 6921–
539 6933 (2004).
540 24. Brautigam, C. A. *et al.* Structural Insight into Interactions between Dihydrolipoamide
541 Dehydrogenase (E3) and E3 Binding Protein of Human Pyruvate Dehydrogenase Complex.
542 *Structure* **14**, 611–621 (2006).
543 25. Jiang, J. *et al.* Atomic Structure of the E2 Inner Core of Human Pyruvate Dehydrogenase
544 Complex. *Biochemistry* **57**, 2325–2334 (2018).
545 26. Prajapati, S. *et al.* Structural and Functional Analyses of the Human PDH Complex Suggest
546 a “Division-of-Labor” Mechanism by Local E1 and E3 Clusters. *Structure* **27**, 1124-1136.e4
547 (2019).
548 27. Kyrilis, F. L. *et al.* Integrative structure of a 10-megadalton eukaryotic pyruvate
549 dehydrogenase complex from native cell extracts. *Cell Rep.* **34**, 108727 (2021).
550 28. Izard, T. *et al.* Principles of quasi-equivalence and Euclidean geometry govern the assembly
551 of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *Proc. Natl. Acad.*
552 *Sci.* **96**, 1240–1245 (1999).
553 29. Bartesaghi, A., Matthies, D., Banerjee, S., Merk, A. & Subramaniam, S. Structure of β-
554 galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad.*
555 *Sci.* **111**, 11709–11714 (2014).

30. Tüting, C. *et al.* Cryo-EM snapshots of a native lysate provide structural insights into a metabolon-embedded transacetylase reaction. *Nat. Commun.* **12**, 6933 (2021).

31. 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).

32. Chojnowski, G. *et al.* findMySequence: a neural-network-based approach for identification of unknown proteins in X-ray crystallography and cryo-EM. *IUCrJ* **9**, (2022).

33. Chojnowski, G. *Sequence assignment validation in cryo-EM models with checkMySequence*. *bioRxiv* doi:10.1101/2022.01.04.474974 (2022).

34. 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).

35. Mattevi, A. *et al.* Atomic Structure of the Cubic Core of the Pyruvate Dehydrogenase Multienzyme Complex. *Science* **255**, 1544–1550 (1992).

36. 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).

37. 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).

38. 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).

580 39. Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. Three-dimensional Structure of the
581 Lipoyl domain from Bacillus stearothermophilus Pyruvate Dehydrogenase Multienzyme
582 Complex. *J. Mol. Biol.* **229**, 1037–1048 (1993).
583 40. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**,
584 583–589 (2021).
585 41. Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural
586 coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**,
587 D439–D444 (2022).
588 42. Ferguson, N. *et al.* Ultra-fast Barrier-limited Folding in the Peripheral Subunit-binding
589 Domain Family. *J. Mol. Biol.* **353**, 427–446 (2005).
590 43. Ho, C.-M. *et al.* Bottom-up structural proteomics: cryoEM of protein complexes enriched
591 from the cellular milieu. *Nat. Methods* 1–7 (2019).
592 44. Ho, C.-M. *et al.* Native structure of the RhopH complex, a key determinant of malaria
593 parasite nutrient acquisition. *Proc. Natl. Acad. Sci.* **118**, (2021).
594 45. Valley, C. C. *et al.* The Methionine-aromatic Motif Plays a Unique Role in Stabilizing Protein
595 Structure. *J. Biol. Chem.* **287**, 34979–34991 (2012).
596 46. Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of
597 heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18**, 176–185
598 (2021).
599 47. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete
600 heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
601 48. Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-
602 electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
603 49. Pettit, F. H. & Reed, L. J. [65] Pyruvate dehydrogenase complex from bovine kidney and
604 heart. in *Methods in Enzymology* vol. 89 376–386 (Academic Press, 1982).

605 50. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of
606 specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
607 51. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron
608 micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
609 52. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in
610 RELION-3. *eLife* **7**, e42166 (2018).
611 53. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and
612 analysis. *Protein Sci.* **27**, 14–25 (2018).
613 54. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr.*
614 *D Biol. Crystallogr.* **60**, 2126–2132 (2004).
615 55. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and
616 electrons: recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877
617 (2019).
618 56. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution
619 electron-density maps. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 519–530 (2018).
620 57. Rice, P., Longden, I. & Bleasby, A. EMBOSS: The European Molecular Biology Open
621 Software Suite. *Trends Genet.* **16**, 276–277 (2000).
622 58. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence
623 alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
624 59. McWilliam, H. *et al.* Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **41**,
625 W597–W600 (2013).
626 60. Mirdita, M. *et al.* ColabFold - Making protein folding accessible to all. *bioRxiv*
627 doi:10.1101/2021.08.15.456425v3 (2022).
628

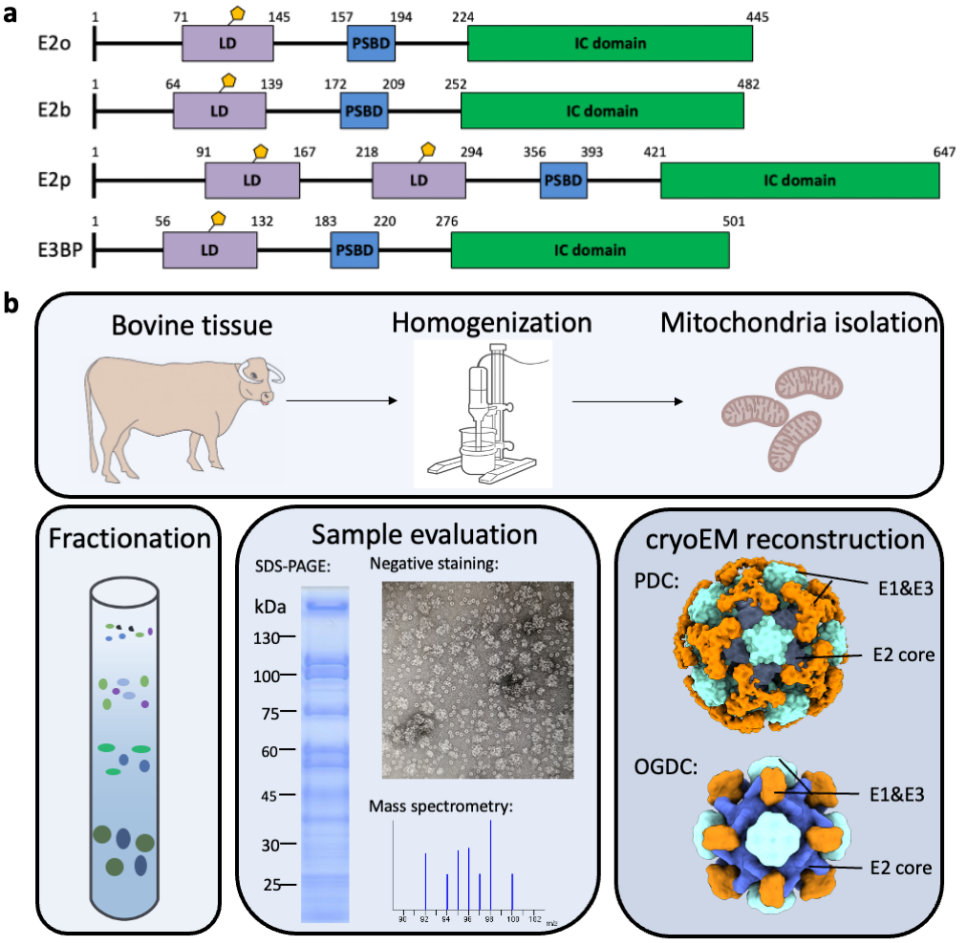


Figure 1. Preparation of bovine OGDC and PDC from native tissue lysate. (a) Organization of E2 and E3BP domains. **(b)** Workflow for extraction of PDC and OGDC from bovine tissue lysate. Bovine kidney was homogenized to isolate mitochondria. Mitochondrial lysates were then fractionated by sucrose gradient. Samples were evaluated for α-keto acid dehydrogenase complex by SDS-PAGE, negative stain, and mass spectrometry. CryoEM reconstruction of these samples yielded icosahedral and cubic cores surrounded by peripheral densities.

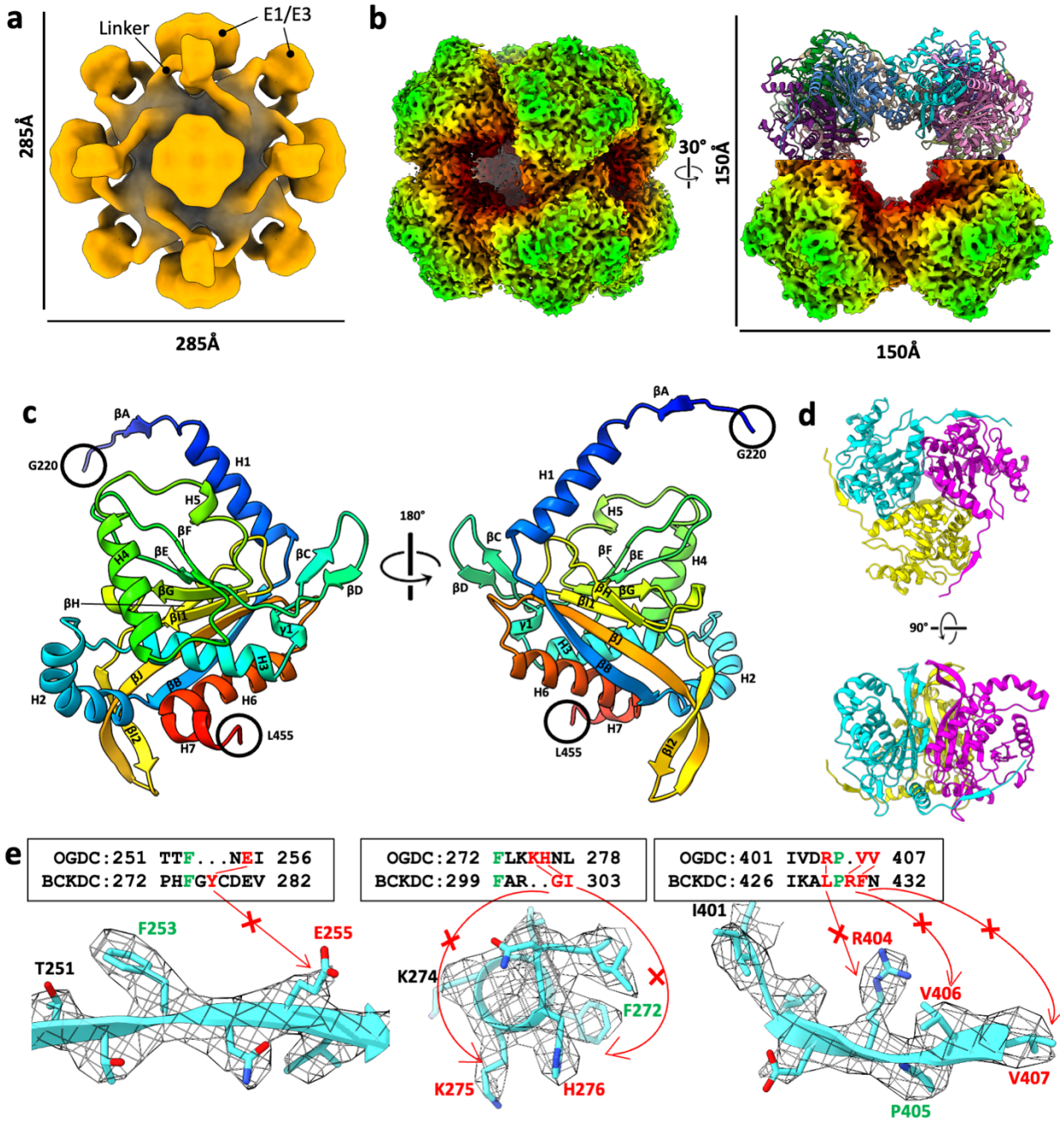


Figure 2. CryoEM reconstruction and identification of bovine OGDC core. (a) 3D reconstruction of the full OGDC with octahedral symmetry showing external densities (orange) around the core density (grey). Linkers extend from each E2o subunit towards a density along the edges of the core. Additional large densities are present atop the faces of the core. **(b)** Two views of the cryoEM density map of E2o core colored by radius, with half of our E2o atomic model assembly in the right. **(c)** Rainbow-colored ribbon of E2o IC domain atomic model with secondary structures labeled. Resolvable N-terminal residues (blue) and C-terminus (red) are indicated. **(d)** Structure of the E2o IC domain trimer, colored by subunits. **(e)** Sequence alignment of E2o and E2b with sidechain fit of corresponding E2o sequence into cryoEM density map. Green letters indicate conserved residues between E2o and E2b that establish landmark features in cryoEM density. Red letters and arrows indicate mismatch of residue identity in E2b with corresponding cryoEM density, thus establishing the map to be that of E2o.

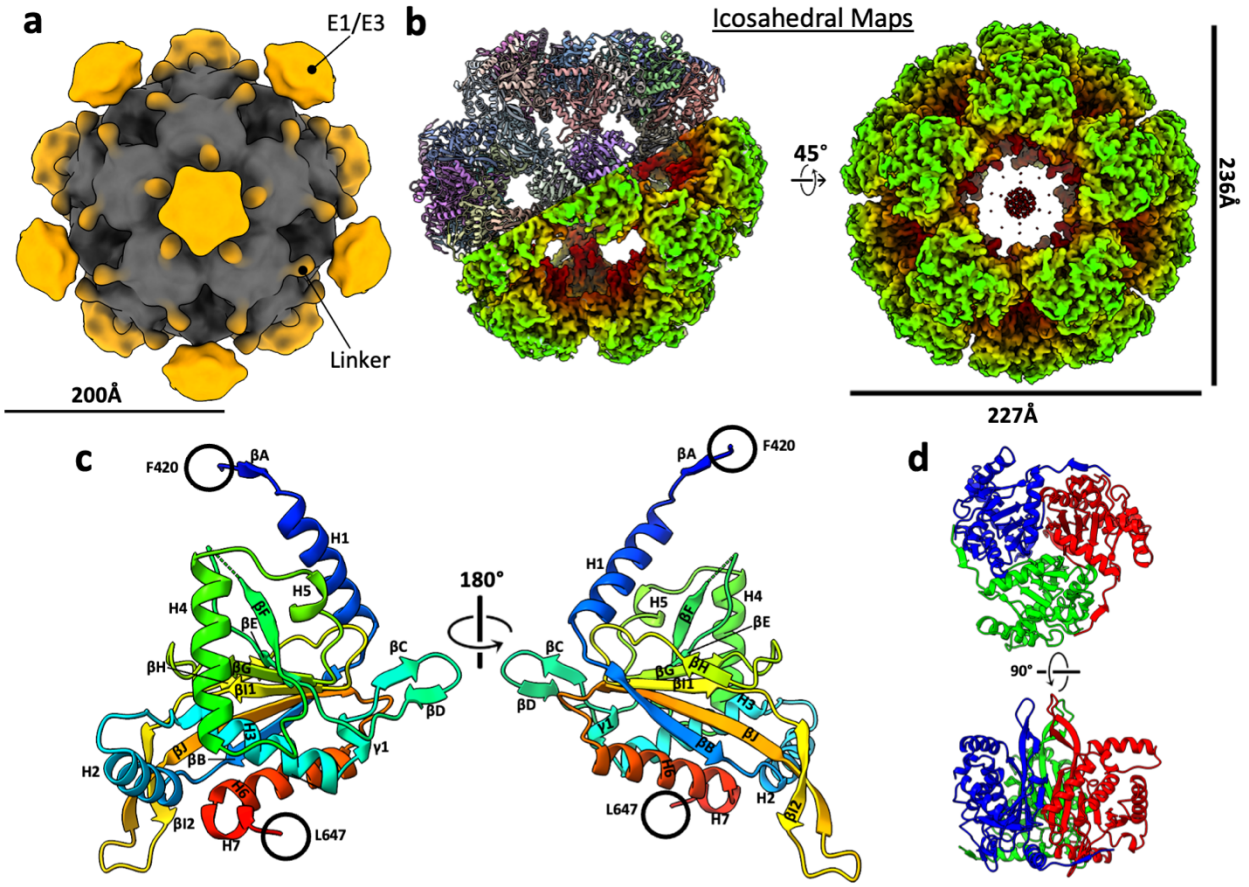


Figure 3. CryoEM reconstructions of bovine PDC and atomic structure of native E2p IC domain. (a) 3D reconstruction of the full PDC with icosahedral symmetry shows external densities (orange) around the core density (grey). **(b)** Two views of the cryoEM density map of E2p core colored by radius, with our E2p atomic model assembly of a hemisphere shown in the left panel. **(c)** Rainbow-colored ribbon of E2p IC domain atomic model with secondary structures labeled. Resolvable N-terminal residues (blue) and C-terminus (red) are indicated. **(d)** Two views of the atomic model of the E2p IC domain trimer, colored by subunits.

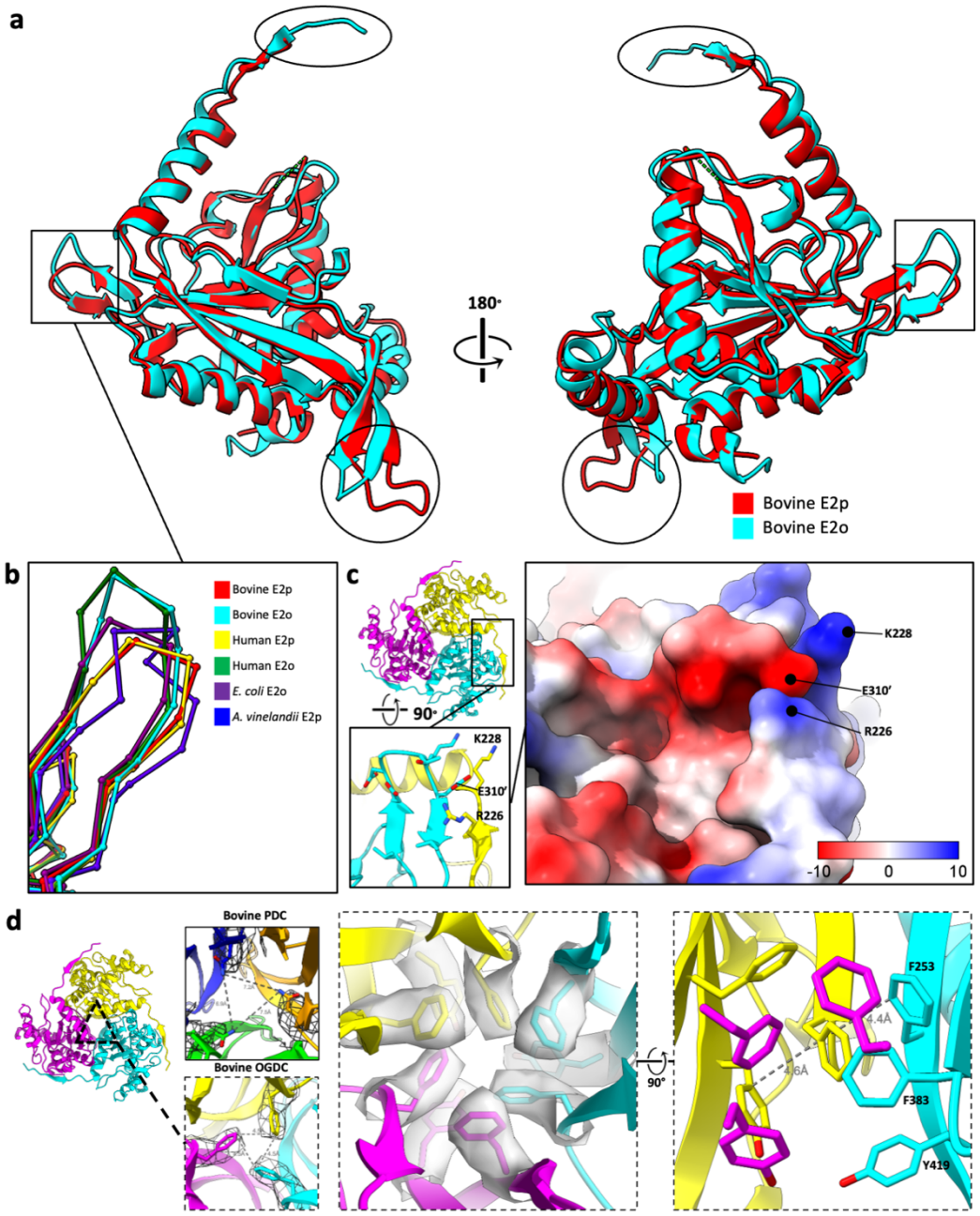


Figure 4. Structure comparison between PDC and OGDC E2 IC domains. (a) Superimposition of E2p and E2o IC domains with the N-terminal linker, βC-βD turn, and interior βI2-βJ hairpin indicated by an oval, box, and circle, respectively. E2p has a longer interior hairpin than that of E2o (circle). N-terminal tail diverges at different angles between endogenous recombinant E2o (oval). E2o has a more extended βC-βD turn than that of E2p. (box). **(b)** Comparison of βC-βD turn from E2o and E2p of various organisms (left). **(c)** Location of N-terminal linker and βC-βD turn contact in the E2o trimer. Lower inset shows residue interactions between R226 and K228 of N-terminal linker from the yellow subunit with E310′ of βC-βD turn from the cyan subunit. Right inset: electrostatic surface potential of the same area as that in the lower inset. The positively charged N-terminal linker (primarily blue) associates with the negatively charged βC-βD turn (primarily red). **(d)** π-stacking interactions present at the three-fold axis of E2o trimer but absent in E2p trimer. Expanded views along the three-fold axis of E2o shows an extended π-electron network formed by F253, F383, and Y419.

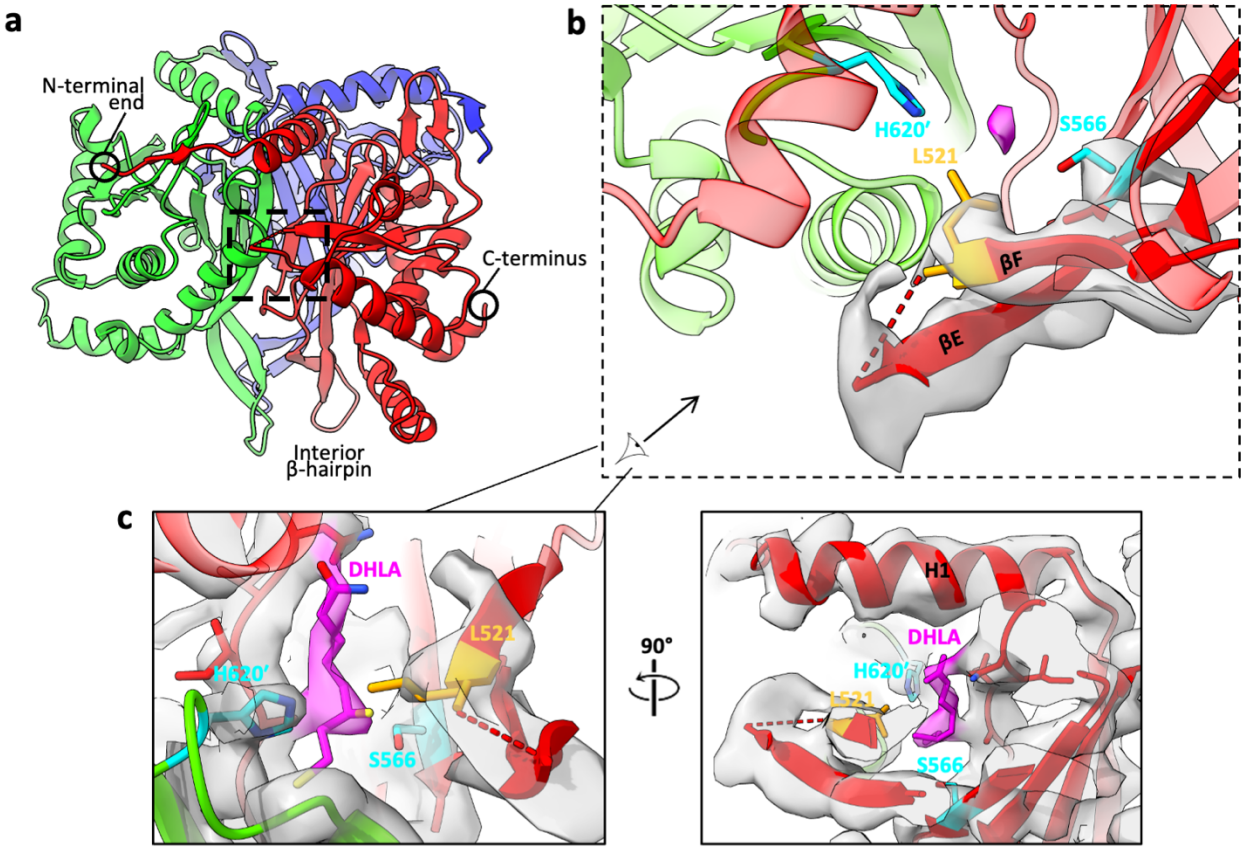


Figure 5. Flexible elements within the bovine PDC E2 IC domain trimers. (a) Ribbon-model of E2p trimer (colored by subunits) with boxes highlighting the relative locations of the flexible βE-βF turn (black box). **(b)** Enlarged views of the βE-βF turn near interdomain active site relative to conserved catalytic residues (teal), Ser566 from the red subunit and His620 from the green subunit. Two residues (Ala519 and Gly520) are unmodeled in the turn due to absence of continuous cryoEM density (semi-transparent gray; for clarity, only those for βE-turn-βF are shown), suggesting flexibility of the residues in the turn. The conserved leucine gatekeeper (orange) is next to these unmodeled residues. A weak density (pink) bordered by catalytic and gatekeeper residues within active site is visible. **(c)** CryoEM density of the same region of (B) shown at a decreased threshold as semi-transparent pink for the lipoyl moiety or gray for the surrounding region. Atomic models for the lipoyl moiety (from PDB ID: 1EAE) and the surrounding area are shown as sticks and ribbon, respectively.

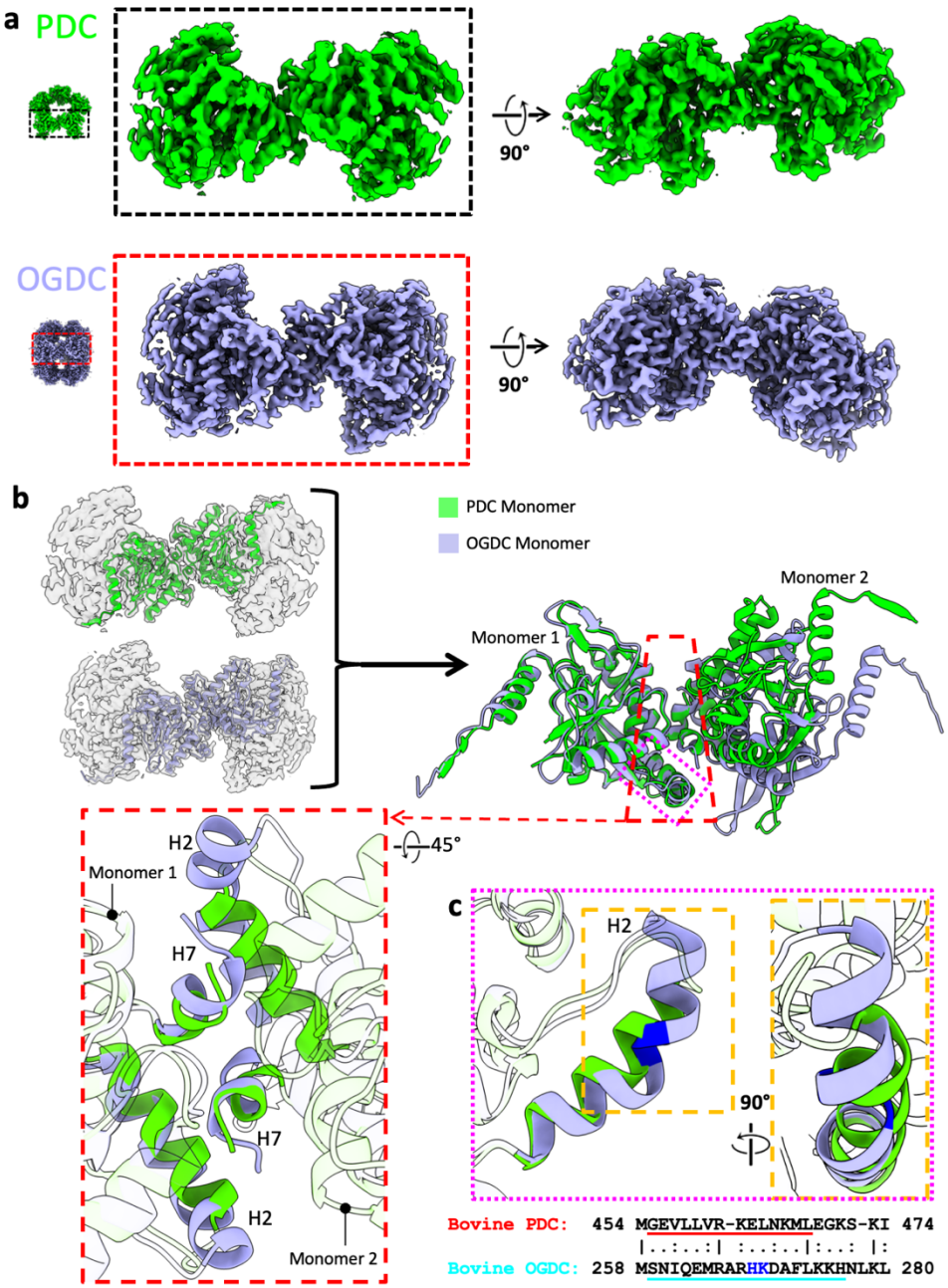
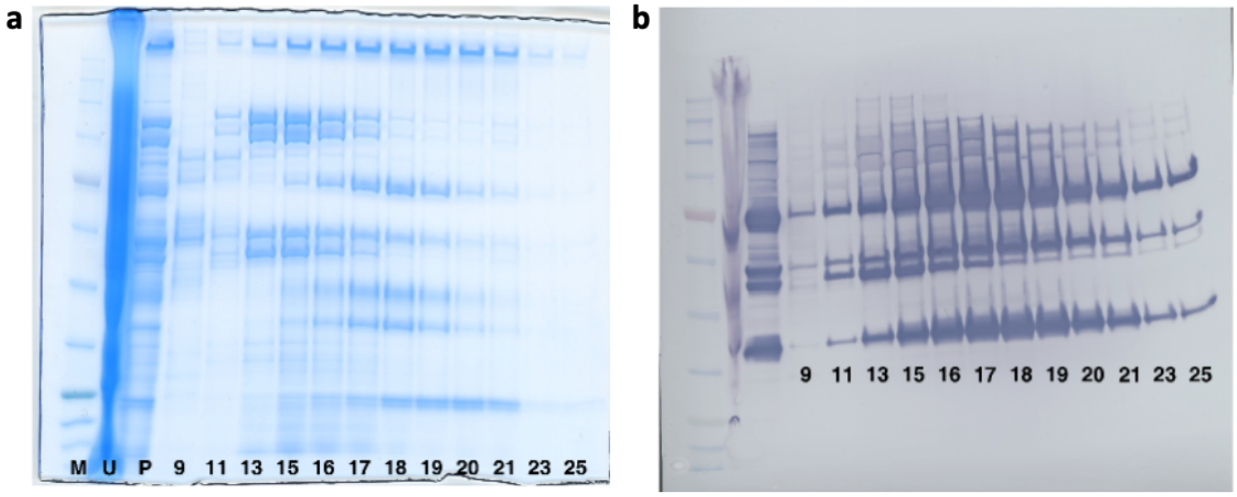
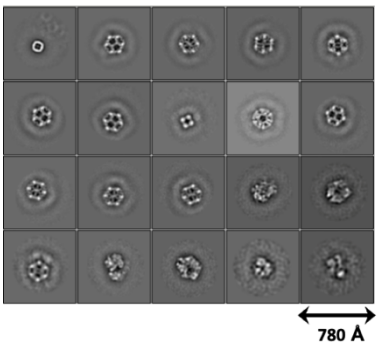


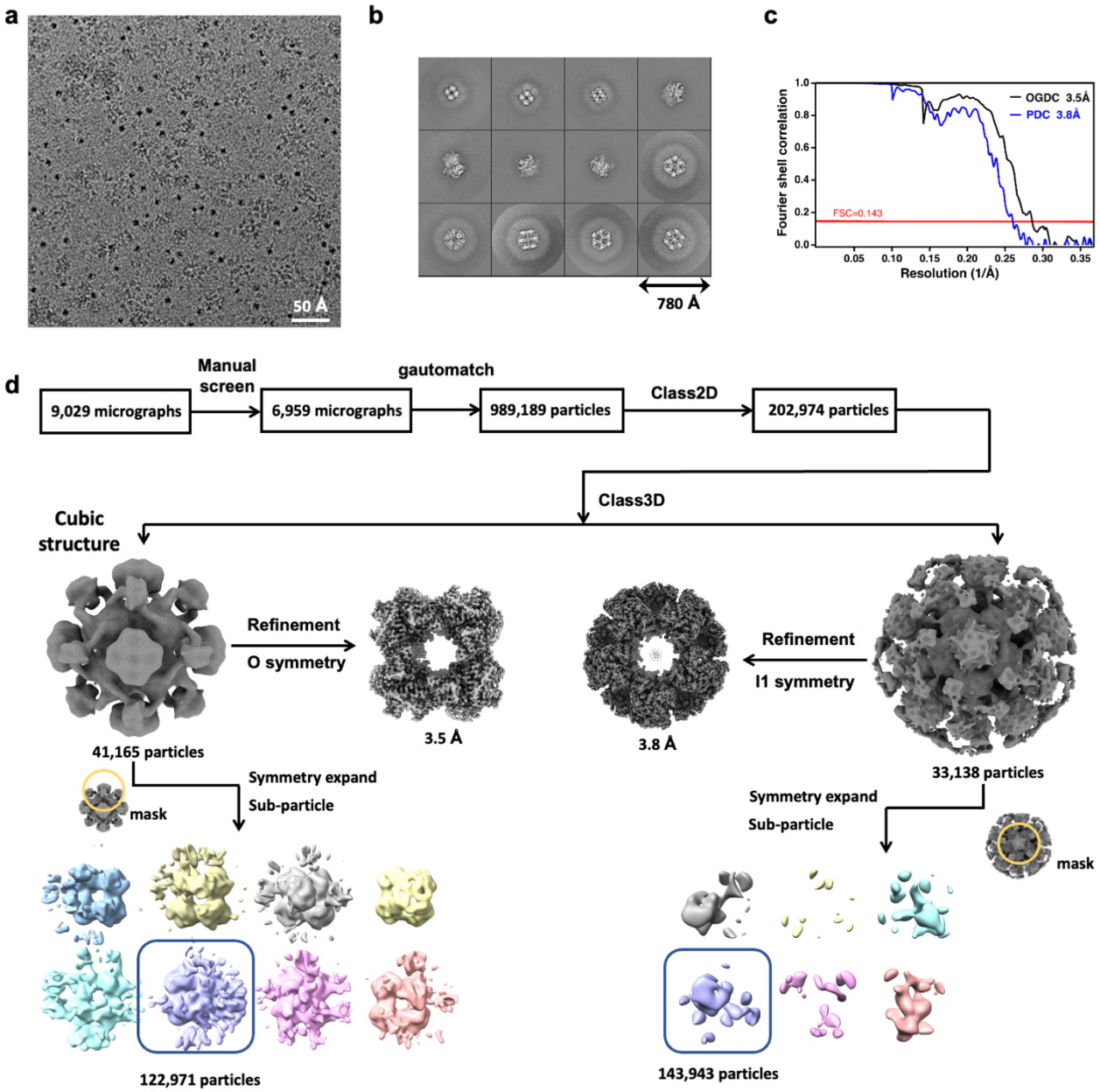
Figure 6. Differences in the E2 inter-trimer interactions between PDC and OGDC. (a) Two orthogonal views (perpendicular and parallel to the 2-fold axis) of interacting E2 trimers from the cryoEM density maps of PDC (green) and OGDC (blue). **(b)** Comparison of the intertrimer interface in PDC (green) and OGDC (blue). Interactions involve only two subunits, one from each trimer, as apparent from the left panel where the atomic model (ribbons) of the interacting subunits are superposed in their respective densities. The two interacting subunits shown in (A) are aligned through their left subunit (monomer 1) and shown together to identify differences at the interacting interfaces with boxed regions shown as insets to highlight differences of their corresponding H2 helix. The inset in red box shows the knob-and-socket interaction of H2 and H7 (colored; rest of ribbon transparent for clarity) at two-fold interface in E2p and E2o. **(c)** Superimposition and sequence alignment of H2 in E2p and E2o. The helix-kink-helix (kink highlighted in blue) is present in E2o but absent in E2p. Inset in the orange box: Enlarged orthogonal view of the diverging H2 trajectories.



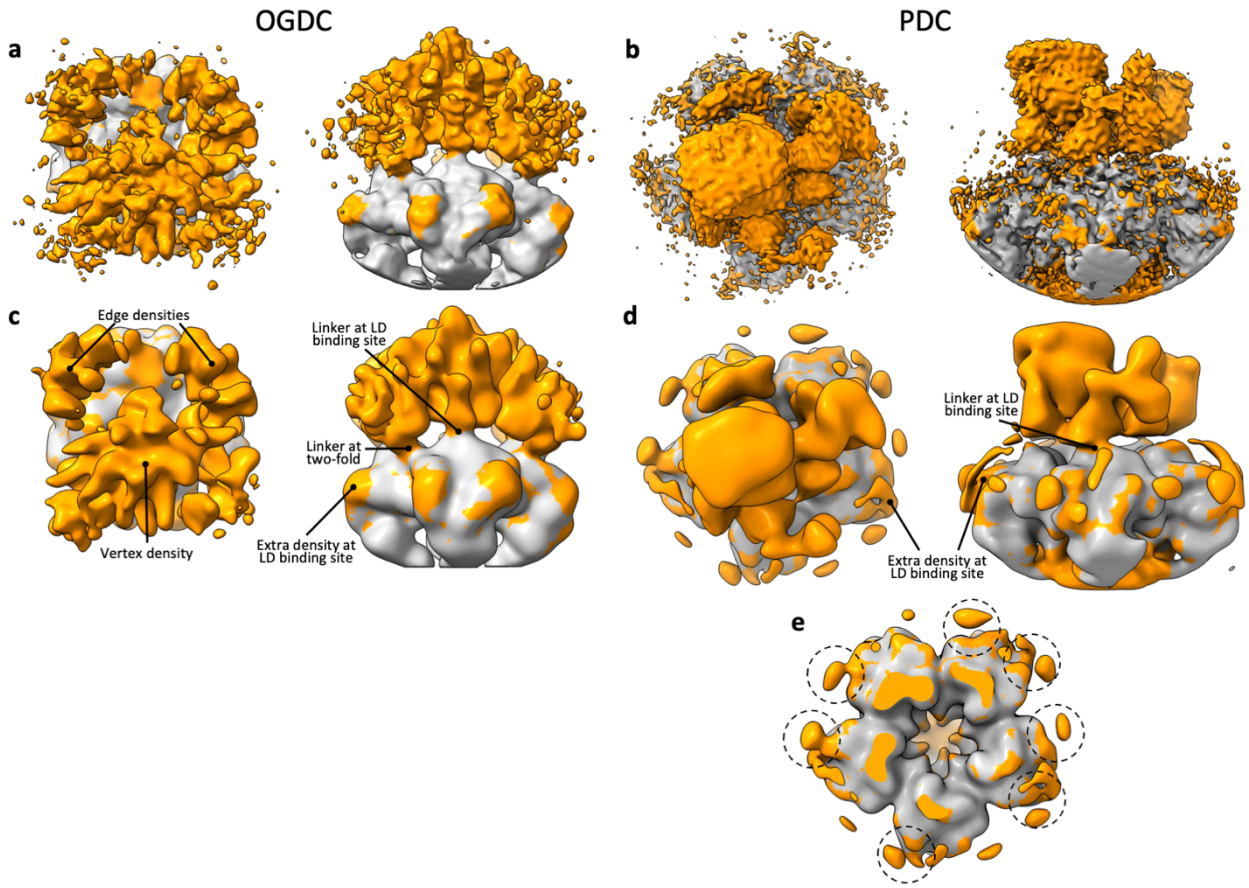
Supplementary Fig. 1. Raw images of gels and blots. (a) SDS-PAGE. **(b)** Immunoblot. Numbers below lanes indicate sucrose gradient fraction.



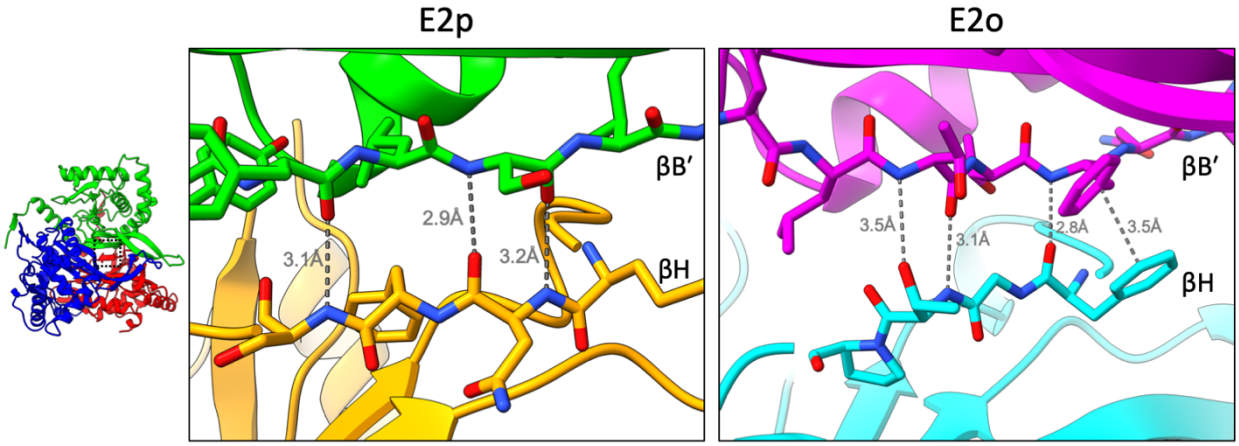
Supplementary Fig. 2. Negative stain 2D class averages. Negative stain 2D class averages show cubic and icosahedral cores with peripheral density. Additional classes of unknown protein species were also observed.



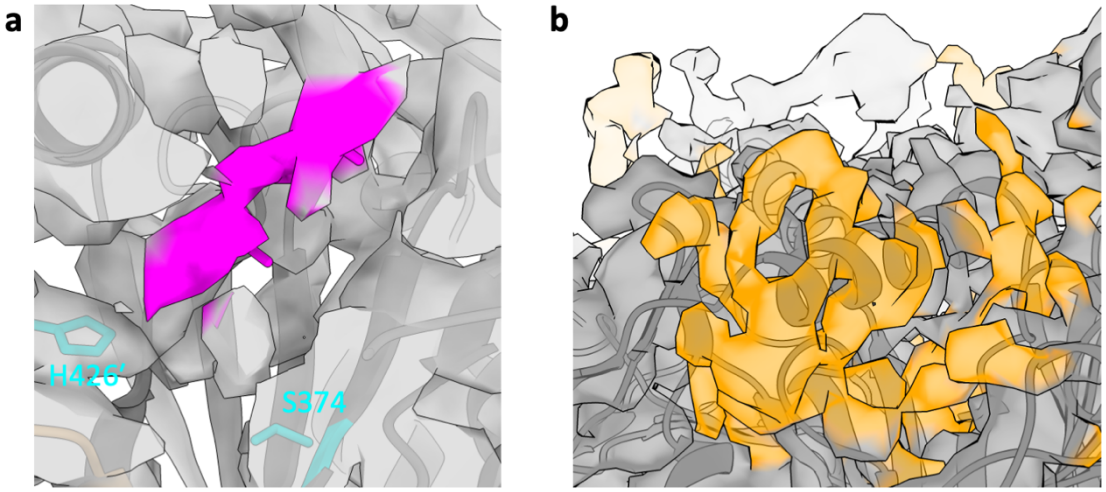
Supplementary Fig. 3. CryoEM data processing workflow. (a) Representative cryoEM micrograph of lysate containing PDC and OGDC. **(b)** CryoEM 2D class averages depicting icosahedral and cubic cores with smeared peripheral density. Classes of additional unidentified protein were also obtained. **(c)** FSC of E2p and E2o core. **(d)** Data processing workflow for reconstruction of E2o and E2p cores and sub-particles.



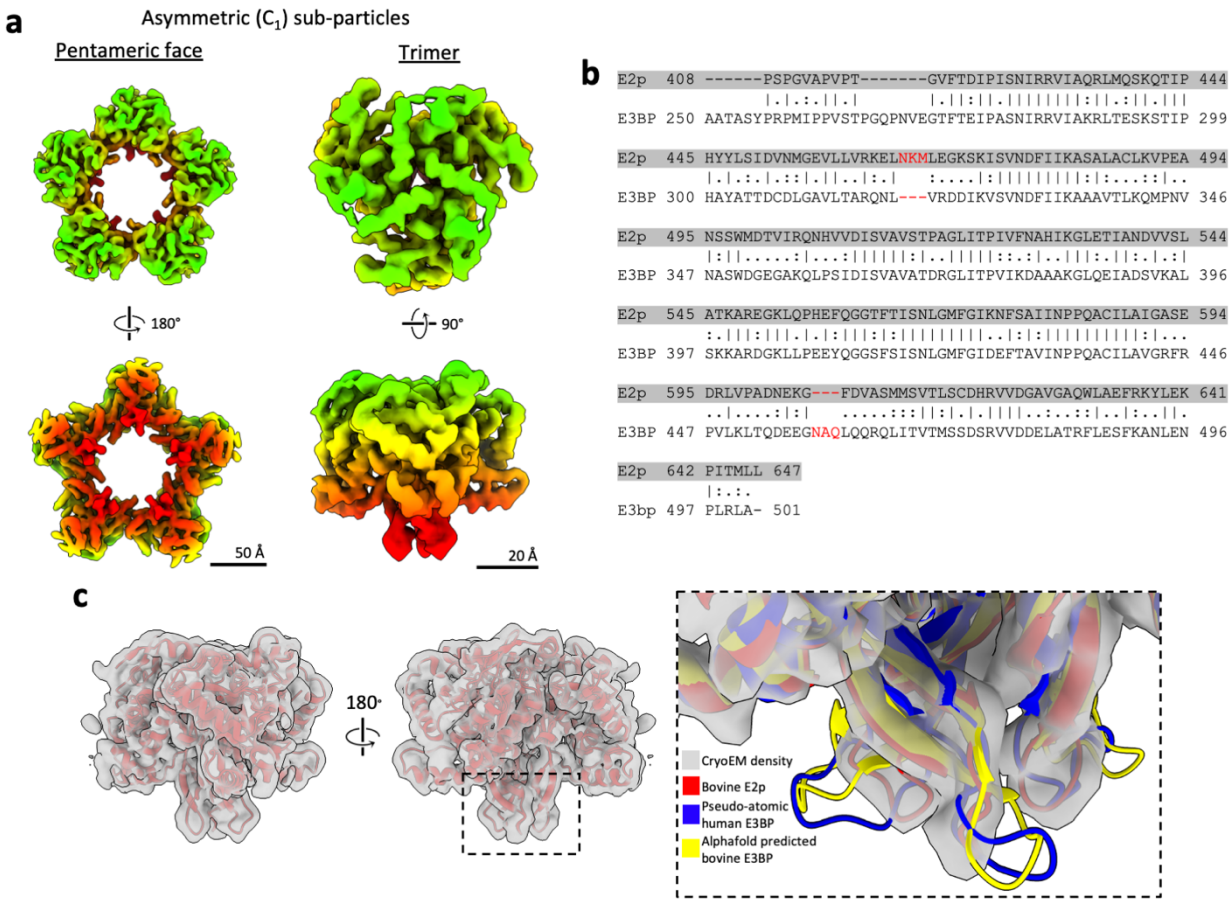
Supplementary Fig. 4. Symmetry expansion and sub-particle reconstruction of the PDC and OGDC cores. (a) Top (left) and side (right) views of sub-particle reconstruction with C_1 symmetry of cubic face of OGDC. **(b)** Top (left) and side (right) views of sub-particle reconstruction with C_1 symmetry of pentameric face of PDC. **(c)** Low-pass filtered (20 Å) OGDC sub-particle map with density external to fitted E2o model in orange. **(d)** Low-pass filtered (20 Å) PDC sub-particle map with density external to fitted E2p model in orange. **(e)** Sliced view of PDC sub-particle at LD binding site. Additional densities at LD binding site are circled.



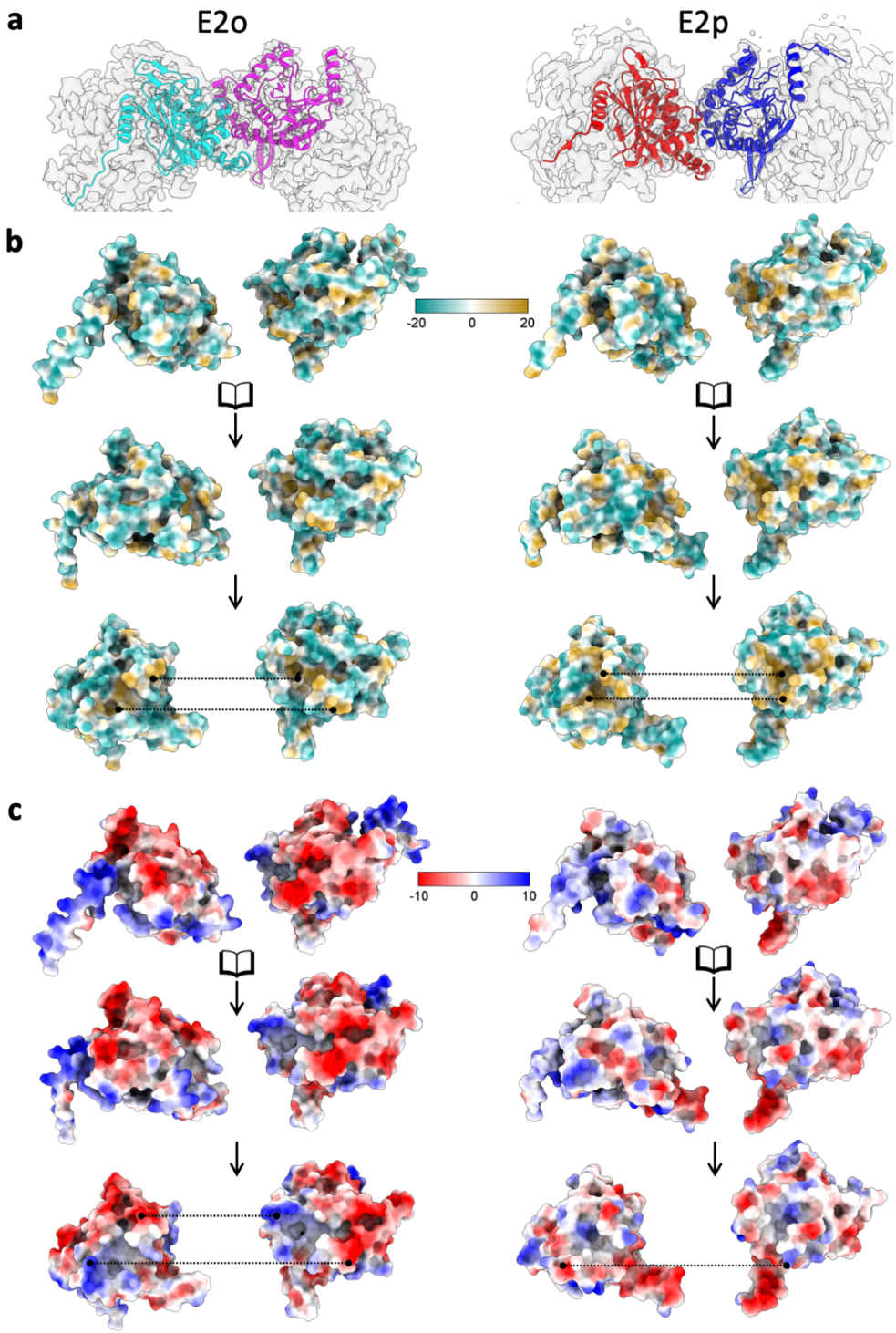
714 **Supplementary Fig. 5. βH-βB′ interdomain interactions.** View of the βH-βB′ interface within
715 the interdomain trimer for E2p (left) and E2o (right). βH (orange; cyan) forms three hydrogen
716 bonds with βB′ (green; magenta) in both E2p and E2o. The E2o interface has an additional
717 slipped-parallel π-stacking interaction between Phe383 and Phe253′.



Supplementary Fig. 6. Lipoyl moiety and tentative lipoyl domain density in native E2o. (a) Weak, thin density within active site channel of E2o corresponding to putative lipoyl moiety (magenta). Catalytic residues (teal) indicated for reference. **(b)** Extra, cylindrical density (orange) located near previously proposed LD binding site.



Supplementary Fig. 7. Search for E3BP in asymmetric sub-particle reconstructions. (a) Representative pentameric face (left) and trimer (right) sub-particle reconstructions with C_1 symmetry. **(b)** Sequence alignment between bovine E2p and E3BP. Three-residue differences are indicated in red. **(c)** Atomic model of E2p IC domain fit into trimer sub-particle reconstruction with C_1 symmetry. The interior loop in E3BP is extended compared to that of E2p and is a key feature for distinguishing the two components. All cryoEM densities of interior loops from each trimer class appear identical and fit E2p closely. Enlarged view of the boxed area depicts atomic models fit into cryoEM density of interior loop. The extended loop of predicted human E3BP models (PDB ID: 6H60; AF-P22439-F1) are not accommodated by the cryoEM density.



Supplementary Fig. 8. Electrostatic surface potential and hydrophobicity of E2o two-fold related intertrimer interface. (a) Two E2o (left) and E2p (right) IC domains at two-fold related intertrimer interface superposed in their respective density (transparent grey) and colored by chain. **(b**) Open-book view of hydrophobicity at intertrimer interface. A hydrophobic knob on each subunit binds to a complementary hydrophobic pocket on the partner subunit across the interface. Dotted lines indicate contacting surfaces. **(c)** Open-book view of electrostatic surface potential at intertrimer interface. Oppositely charged surfaces on the exterior mediate electrostatic interactions between the two subunits at the intertrimer interface.

740 **Supplementary Table 1. CryoEM data collection, refinement, and validation statistics**

	#1 OGDC (EMDB-xxxx) (PDB xxxx)	#2 PDC (EMDB-xxxx) (PDB xxxx)
Data collection and processing		
Magnification	105,000 ×	105,000 ×
Voltage (kV)	300	300
Electron exposure (e^{-}/Å^2)	45	45
Defocus range (μm)	-1.8 to -2.6	-1.8 to -2.6
Pixel size (Å)	1.36	1.36
Symmetry imposed	O	I1
Initial particle images (no.)	989,189	989,189
Final particle images (no.)	41,165	33,138
Map resolution (Å)	3.5	3.8
FSC threshold	0.143	0.143
Map sharpening B factor (Å^2)	-152	-150
Refinement		
Model-to-map fit, mask CC	0.88	0.87
Model composition		
Non-hydrogen atoms	44496	104640
Protein residues	5664	13560
Ligands	0	0
Mean Isotropic *B* factors (Å^2)		
Protein	46.08	78.36
Ligand	N/A	N/A
R.m.s. deviations		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.944	0.940
Validation		
MolProbity score	1.13	1.39
Clashscore	3.05	7.12
Rotamer outliers (%)	0.5	0.09
Ramachandran plot		
Favored (%)	97.86	98.54
Allowed (%)	2.14	1.46
Outliers (%)	0.00	0.00

741

742 **Supplementary Table 2. Interactions at the intertrimer interface**

Interactions	Cubic E2o core	Icosahedral E2p core
Hydrogen bonds (donor-acceptor)		
H2 and H7 interaction		
	Arg267-Asp447	(Lys641-Glu463)
	His268-Asp447	
	Lys275-Asp454	
	(Lys275-Asp455)	
	His276-Asp454	
	Arg449-Asn260	
	Arg449-Glu263 [×3]	
H4 and H7 interaction		
	Tyr342-Leu452	Leu534-Leu646
		Glu535-Leu647
Hydrophobic residues composing knob		
	Leu452	Ile643
	Leu453	Leu646
	Leu455	
Hydrophobic residues lining socket		
	Met264	Val460
	Phe272	Leu464
	Leu280	Phe479
	Tyr342	Leu534
	Pro448	Pro642
	Leu451	Met645

Duplicate two-fold interactions not listed. Parentheses indicate interface-variable interaction. Equivalent interactions indicated by matching colors. Arg449-Glu263 share three hydrogen bonds.

743