Edinburgh Research Explorer

Integrative structure of a 10-megadalton eukaryotic pyruvate dehydrogenase complex from native cell extracts

Cell Reports

Integrative structure of a 10-megadalton eukaryotic pyruvate dehydrogenase complex from native cell extracts

Graphical Abstract
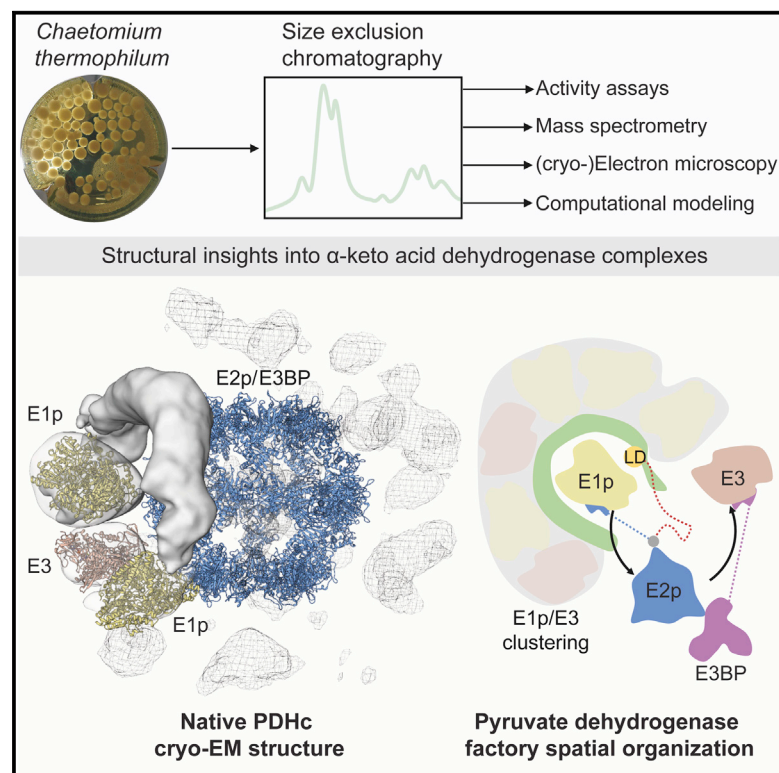
Authors

Fotis L. Kyrilis, Dmitry A. Semchonok, Ioannis Skalidis, ..., Francis J. O'Reilly, Juri Rappsilber, Panagiotis L. Kastritis

Correspondence

panagiotis.kastritis@bct.uni-halle.de

In Brief

Kyrilis et al. combine diverse biochemical and biophysical methods to probe native cell extract organization. They report an asymmetric reconstruction of the active, native pyruvate dehydrogenase complex by cryo-EM and determine the spatial relationships among the protein components in the dynamic assembly. They identify enzyme clusters that form a transient catalytic nanocompartment and define this as the "pyruvate dehydrogenase factory."

Highlights

- 2-oxo-acid dehydrogenase complexes are studied by integrative analysis of cell extracts

- A cryo-EM structure of the active, native pyruvate dehydrogenase complex is resolved

- Asymmetric cryo-EM maps reveal enzyme clustering during pyruvate oxidation

- A transiently formed catalytic chamber is defined as the "pyruvate dehydrogenase factory"

CellPress

Article

Integrative structure of a 10-megadalton eukaryotic pyruvate dehydrogenase complex from native cell extracts

Fotis L. Kyrilis,[1,2,6] Dmitry A. Semchonok,[1,6] Ioannis Skalidis,[1,2] Christian Tüting,[1] Farzad Hamdi,[1] Francis J. O'Reilly,[3] Juri Rappsilber,[3,4] and Panagiotis L. Kastritis[1,2,5,7,*]

[1]Interdisciplinary Research Center HALOmem, Charles Tanford Protein Center, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Straße 3a, Halle/Saale, Germany
[2]Institute of Biochemistry and Biotechnology, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Straße 3, Halle/Saale, Germany
[3]Bioanalytics, Institute of Biotechnology, Technische Universität Berlin, 13355 Berlin, Germany
[4]Wellcome Centre for Cell Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, Scotland, United Kingdom
[5]Biozentrum, Martin Luther University Halle-Wittenberg, Weinbergweg 22, Halle/Saale, Germany
[6]These authors contributed equally
[7]Lead contact
*Correspondence: panagiotis.kastritis@bct.uni-halle.de
https://doi.org/10.1016/j.celrep.2021.108727

SUMMARY

The pyruvate dehydrogenase complex (PDHc) is a giant enzymatic assembly involved in pyruvate oxidation. PDHc components have been characterized in isolation, but the complex's quaternary structure has remained elusive due to sheer size, heterogeneity, and plasticity. Here, we identify fully assembled *Chaetomium thermophilum* α-keto acid dehydrogenase complexes in native cell extracts and characterize their domain arrangements utilizing mass spectrometry, activity assays, crosslinking, electron microscopy (EM), and computational modeling. We report the cryo-EM structure of the PDHc core and observe unique features of the previously unknown native state. The asymmetric reconstruction of the 10-MDa PDHc resolves spatial proximity of its components, agrees with stoichiometric data (60 E2p:12 E3BP:∼20 E1p: ≤ 12 E3), and proposes a minimum reaction path among component enzymes. The PDHc shows the presence of a dynamic pyruvate oxidation compartment, organized by core and peripheral protein species. Our data provide a framework for further understanding PDHc and α-keto acid dehydrogenase complex structure and function.

INTRODUCTION

The oxidative decarboxylation of pyruvate is a key metabolic reaction during pyruvate catabolism. This vital step in aerobic metabolism is performed by a giant enzymatic complex, the pyruvate dehydrogenase complex (PDHc), which forms acetyl-coenzyme A (CoA), CO_2, and NADH (H+) (Bisswanger, 1981; Patel et al., 2014; Perham, 1991; Reed, 2001; Roche and Reed, 1972). The PDHc belongs to the α-keto acid/2-oxo-acid dehydrogenase complex (OADH complexes) family, together with the 2-oxoglutarate dehydrogenase complex (OGDHc, also known as α-ketoglutarate dehydrogenase complex) and branched-chain keto-acid dehydrogenase complexes (BCKDHc). In eukaryotes, they have been localized to the mitochondrial inner-membrane-matrix interface (Patel et al., 1996), being major metabolic checkpoints that finely tune sugar and amino acid degradation (Patel et al., 1996) (Figure 1A). These complexes are also interesting due to their additional roles as mitochondrial autoantigens (Bassendine et al., 1997) and their implication in several pathologies such as viral infection (Liu et al., 2019), neurodegeneration (Jak-

kamsetti et al., 2019; Mkrtchyan et al., 2015), and inflammation (Palmieri et al., 2020). OADH complexes range from 4 to 10 million Da (MDa), making them some of the largest and most sophisticated enzymatic systems of the cell (Reed, 2001), and their overall native architecture is currently elusive.

Three major components assemble OADH complexes and are present in multiple copy numbers (Perham, 1991): the E1 (pyruvate/2-oxoglutarate/BCKADH; EC 1.2.4.1/1.2.4.2/1.2.4.4, E1p, E1o, and E1b, respectively); E2 (dihydrolipoyl [or dihydrolipoamide] transacetylase/transsuccinylase [or acetyl/succinyltransferase]; EC 2.3.1.12/2.3.1.61/2.3.1.168, E2p, E2o, and E2b, respectively); and E3 (dihydrolipoyl [or (dihydro)lipoamide] dehydrogenase; EC 1.8.1.4). Five coenzymes are shared among all the family members: TPP, NAD+, FAD, lipoic acid, and CoA (CoASH). E3 is common to all OADH complexes and shared with the glycine cleavage system. The members of the OADH family perform five sequential reactions that decarboxylate α-keto acids and covalently attach the remaining acyl group to CoA (Figure 1B). The PDHc converts pyruvate to acetyl-CoA, which then enters the citric acid cycle (tricarboxylic acid cycle
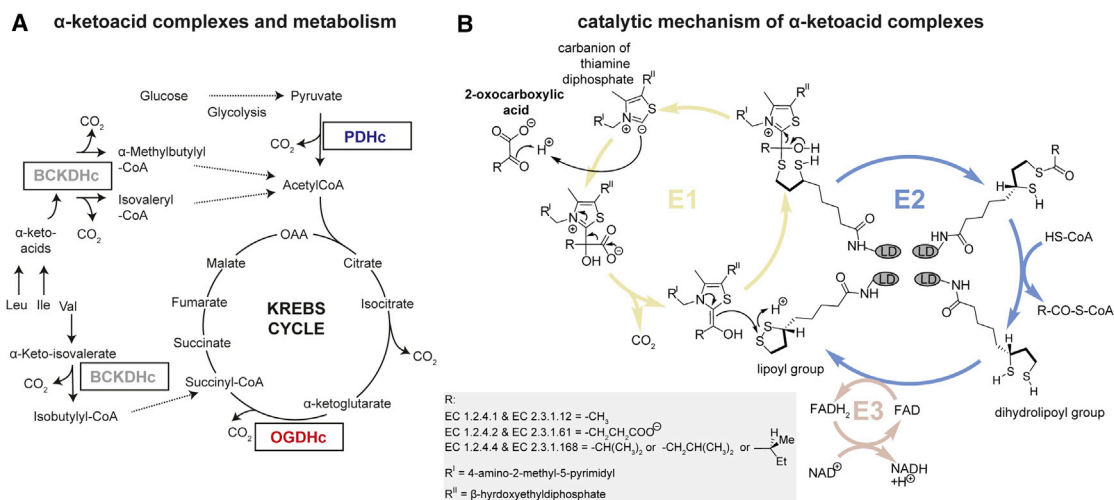
Figure 1. Pathways involving α-keto acid dehydrogenase complexes
(A) Biochemical pathway with PDHc, OGDHc, and BCKDHc as major participants.
(B) Reaction mechanism performed by the E1, E2, and E3 enzymes of the α-keto acid dehydrogenase complexes.

(TCA), Krebs cycle). The OGDHc acts within TCA, converting 2-oxoglutarate to succinyl-CoA. BCKDHc provides various acyl-CoA precursors for TCA by converting with broad specificity branched, short-chain α-keto acids. During the reaction cycle, an E1 captures and decarboxylates an α-keto acid and then transfers it to a mobile lipoyl domain (LD) of an E2. The activated E2 LD transports the acyl group (the nature of the acyl group depends on the particular E2) to the E2's catalytic domain, which subsequently transfers it to CoA, forming the acyl-CoA. The acyl-CoA is then released, and the LD of the E2 is regenerated (reoxidized) for another cycle by the E3, which transfers protons to NAD+, forming NADH. There are notable differences between prokaryotic and eukaryotic enzymatic complexes and among the PDHc, OGDHc, and BCKDHc. They differ in additional components, the number of LDs, subunit stoichiometry, higher-order architecture, and overall abundance in the cell (Ciszak et al., 2006; Klyuyeva et al., 2019; Patel et al., 2014; White et al., 2018).

The PDHc has been the most thoroughly studied of the OADH complexes. The ordered domains of E1p, E2p, and E3 proteins and their respective protein complexes have been structurally characterized in isolation (e.g., in human, the heterotetramer of E1p [Kato et al., 2008], the core module of E2p [Jiang et al., 2018], and the homodimer of E3 [Szabo et al., 2019]). Electron microscopy (EM) studies revealed the architecture of the eukaryotic PDHc E2p core and provided insights into its function (Forsberg et al., 2020; Prajapati et al., 2019; Yu et al., 2008; Zhou et al., 2001). The PDHc includes the E3-binding protein (E3BP) that tethers the E3. The E3BP localizes at the PDHc core, replacing some E2p in mammals (Hiromasa et al., 2004), or inside the core as four separate complexes in fungi (Forsberg et al., 2020). The number of E3BPs present remains debatable (Brautigam et al., 2006) due to the underlying structural intricacy (Zhou et al., 2001) and the absence of high-resolution information. If the stoichiometry and structure of the E2p and E3BP are unknown, the E1p and E3 subunits cannot be accurately placed because the E1p and E3 attach to the flexible peripheral-subunit binding

domains (PSBDs) of the E2p and E3BP, respectively. Another complication for structural studies is that the catalytic mechanism depends on the highly flexible LD and on flexible linkers that tether the E1p and E3 subunits proximal to the core (Lengyel et al., 2008). Because of these complexities, overall 3D architectures of PDHc and other OADH complexes remain unknown.

Here, we present a workflow to analyze OADH complexes from native cell extracts of *C. thermophilum*, a thermophilic eukaryote. Proteins from *C. thermophilum* are stable due to their thermal adaptation and, thus, are advantageous for structural studies (Figure S1A) (Cheng et al., 2019). Using native cell extracts with limited purification can help maintain the native structure of protein complexes (Ho et al., 2020; Kyrilis et al., 2019; McCafferty et al., 2020; Verbeke et al., 2018, 2020). We combine cryo-EM with kinetic assays; mass spectrometry (MS); chemical crosslinking; and computational, biophysical modeling to unveil stoichiometries and proximity principles for the E2p, E3BP, E1p, and E3 localized in the PDHc, illuminating aspects of PDHc structure and function.

RESULTS

Enrichment of active OADH complexes with a single fractionation step

We set up optimal growth conditions of *C. thermophilum* to 45°C and propagated mycelium (Figure S1B) as in Kellner et al. (2016). To isolate complexes, we fractionated lysate by size-exclusion chromatography (SEC) (Figures 2A, S1C, and S1D). We vitrified cell extracts including the highest molecular weight complexes, which were previously reproducibly identified by MS (Kastritis et al., 2017), and recognized prominent, structurally unknown molecular signatures in the cryo-EM micrographs (Figure 2B). We recognized a plethora of structures, including the fatty acid synthase (FAS) metabolon (Kastritis et al., 2017), double- and single-membrane structures, liposomes with encapsulated biomolecules, filaments, and other higher-order complexes of
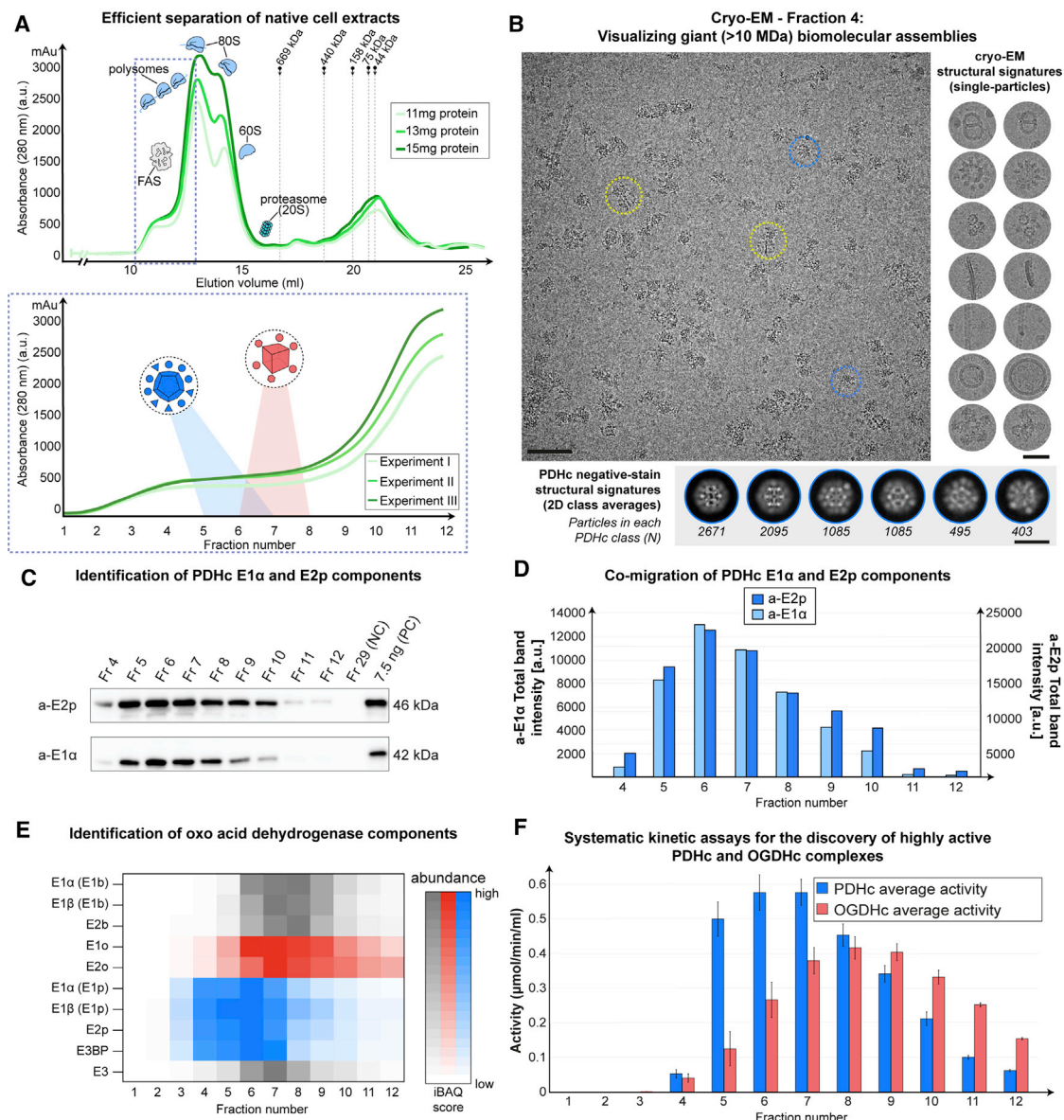
Figure 2. Biochemical characterization of native cell extracts aiming to capture α-keto acid dehydrogenase complexes

(A) SEC profile of *C. thermophilum* native extract; inset shows recovered protein material at high-molecular-weight fractions.

(B) Cryo-EM of fraction 4 (>10-MDa complexes). Insets show the metabolon of FAS with the bound carboxylase (Kastritis et al., 2017) and PDHc particles in yellow and blue circles, respectively. On the right, single particles described in STAR methods. On the bottom, class averages of PDHc derived from NS. Scale bars: inset, 100 nm; extracted single particles and PDHc class averages, 60 nm.

(C) Immunodetection of E2p and E1α in cell extracts; NC and PC stand for negative and positive controls (see STAR methods protocols).

(D) Calculation of total band intensity showing co-migration of identified components across fractions.

(E) MS data show abundance of α-keto acid dehydrogenase subunits.

(F) Activity assays for PDHc (blue) and OGDHc (red) E1. Standard deviation is shown after averaging three biological replicates that are reported in Figure S3.

intriguing architecture (Figures 2B and S2A). After negative staining (NS), we identified particles that match the particle diameter and subunit architecture of the PDHc in these early cell extract fractions (Figure 2B). We also produced antibodies (Abs) for *C. thermophilum* E1α and E2p protein constructs, which showed co-migration of E1α and E2p (Figures 2C and 2D). This agrees with quantitative protein identifications from MS experiments

(Kastritis et al., 2017). We identified a significant amount of co-eluting OADH complex components (Figure 2E) and performed sequence- and complex-specific abundance analysis (Figures S2B–S2J; Table S1). MS data show a high abundance of PDHc components with a peak at fraction 6, whereas for OGDHc and BCKDHc components, abundance peaked at fractions 7 and 8, respectively (Figures 2E and S2). Identification of sequences

and domain architecture of *C. thermophilum* OADH complexes revealed that they are comparable to their human counterparts, sharing 38%–69% of sequence similarity (Figures S2B–S2D). A significant difference is that the E2p of *C. thermophilum* has one LD, similar to yeast, in comparison to the human counterpart, which has two (Figure S2B).

To assess the activity of fractions containing the PDHc and OGDHc, we performed a coupled continuous spectrophotometric assay, which targets the E1 activity. Starting from cell lysis, followed by ultracentrifugation, concentration (Figure S3A), and the collected fractions after SEC (Figures S3A–S3D), both complexes retained activities and showed a peak in activity at the corresponding fractions that were abundant (Figures 2F and S3) as determined by MS (Figures 2E and S2G–S2J). The highest activity for the PDHc is seen in fraction 6 (Figure S3C). An active E1 tetramer corresponds to a calculated molecular weight of only ~160 kDa (Kresze and Ronft, 1981). In our fractions, E1 was (1) identified in the high MDa fractions; (2) co-eluted with all other proteins of PDHc and OGDHc; and (3) measured to be catalytically active, suggested it is bound in the respective higher-order complexes.

Assignment of structural signatures by large-scale EM imaging

We looked into the fractions with EM to identify structural features of fully assembled OADH complexes. We collected micrographs from NS samples for fractions 5–11, acquiring a total of 5,666 images (Figure 3A; Table S2), and performed subsequent image processing (Figures 3A–3C; Table S2). These fractions were chosen because (1) OADH complexes are enriched in those; (2) protein concentration was higher than in the previous fractions (Figure S1C), so we expected more particles; and (3) minor populations of amorphous aggregates present in the earlier fractions are avoided. The overall presence of aggregates was minor, as monitored by EM (Figures 2B and 3A) and following fractions at 320 nm among the replicates (Figure S1D). In the fractions, only a few OADH complexes (OGDHc/BCKDHc) can be readily observed in acquired micrographs (Figure 3A). This is in agreement with MS, where the highest protein abundance of summed OADH complexes does not pass the 40% threshold of overall fraction abundance (Figures S2G–S2J). After 2D classification of 2,043,294 particles in total, class averages emerged of the two prominent cores (Figure 3B). Single particles classified to possess different cores were grouped; subsequently, 3D reconstructed in each fraction, they were recovered without any symmetry information at low resolution (Table S2), but cores were retrieved with prominent structural features that reflect their symmetry (correct number of faces and vertices). A superposition of *C. thermophilum* models based on human E2p, E2o, and E2b cores with the overall density calculated from every fraction presented high cross-correlation (CC) values and high overlap of models to volumes, therefore pointing to an explanation of the architecture of the distinct cores (Figures S4B–S4F). After superimposing the cores, a fraction of the density did not belong to core structures (Figures 3C and S4B–S4F). Cores were surrounded by densities that localized asymmetrically and not in a diffuse pattern; this localization of external densities of the two structural species was reproducible in the independent 3D reconstructions retrieved from the various fractions (Figure 3C).

We produced homology models of structurally resolved regions related to all protein components to validate structural relationships within the OADH complex family. We analyzed crosslinking MS data from cell extracts (Kastritis et al., 2017) and mapped a total of 118 crosslinks (Figures S5A–S5C). In total, 77% of the crosslinks (n = 91) map to flexible regions of the OADH complexes, further corroborating the underlying plasticity of those assemblies (Figure S5A). The rest (n = 27) map to ordered domains and recapitulate the log-normal distribution of crosslink distances (16.1 ± 6.5 Å), with only one crosslink showing minor violation of the 30.0-Å cut-off distance (E1o, Lys542-Lys171, 33.2 Å) (Figures S5B and S5C).

Reconstructed particles with cubic cores cannot be distinguished as being either OGDHc or BCKDHc because they have similar architectures (Figures S4B–S4E) and elute with similar abundance (Figures S2H–S2J). However, the reconstructed dodecahedral particles can be assigned as being PDHc because single-particle abundance (Figure 3B) further agrees with quantitative MS (Figures S2G–S2J), immunodetection (Figures 2C and 2D), and activity data (Figure 2F). The 3D reconstructions from the same fraction strongly resemble the known eukaryotic E2p PDHc core (Figure 3C). To our knowledge, there is no other complex at comparable abundance to the PDHc in cells that shares the same core architecture, in contrast to the other OADH complexes. Application of known symmetry to resolve the PDHc further corroborates the size and architecture of the respective E2p core (Figures S5D–S5I). Yet, energy calculations and comparison to its mesophilic counterpart did not reveal any significant contribution of calculated non-covalent forces to the observed stabilization of thermophilic complexes (Table S3). Finally, it is also prominent from EM (Figure 3) and derived core models (Figures S4 and S5) that the native size of PDHc is larger than the rest of the OADH complexes. This agrees with both their retention times (Figure 2C) and the distinct activity peaks of PDHc and OGDHc E1 enzymes (Figure 3F).

Reconstructions of the PDHc and other OADH complexes reveal localized clustering of external subunits

To characterize the localization of the peripheral densities for annotated PDHc and remaining OADH complexes, we collected a considerable number of single particles from fraction 6 (Figures 3A and 3B) and retrieved a gallery of class averages, including ~54% of single particles (Figure S4A). We finally retrieved 55 and 36 reference-free class averages for the PDHc and the remaining OADH complexes, including totals of 282,698 and 17,649 particles, respectively. Distinct spatial and molecular features, manifested as localized densities, were observed surrounding the E2 cores and exhibited high-contrast densities (Figure 4A). For the PDHc, quantification of this signal recapitulated 2–6 localized densities per class average (Figure 4A). By extrapolating class average statistics to the total number of averaged particles therein, we calculate that each PDHc particle should at least have 3.10 ± 1.27 localized densities, following stringent outer-density cluster classification. Correspondingly, for the remaining OADH complexes, derived statistics are lower, possibly
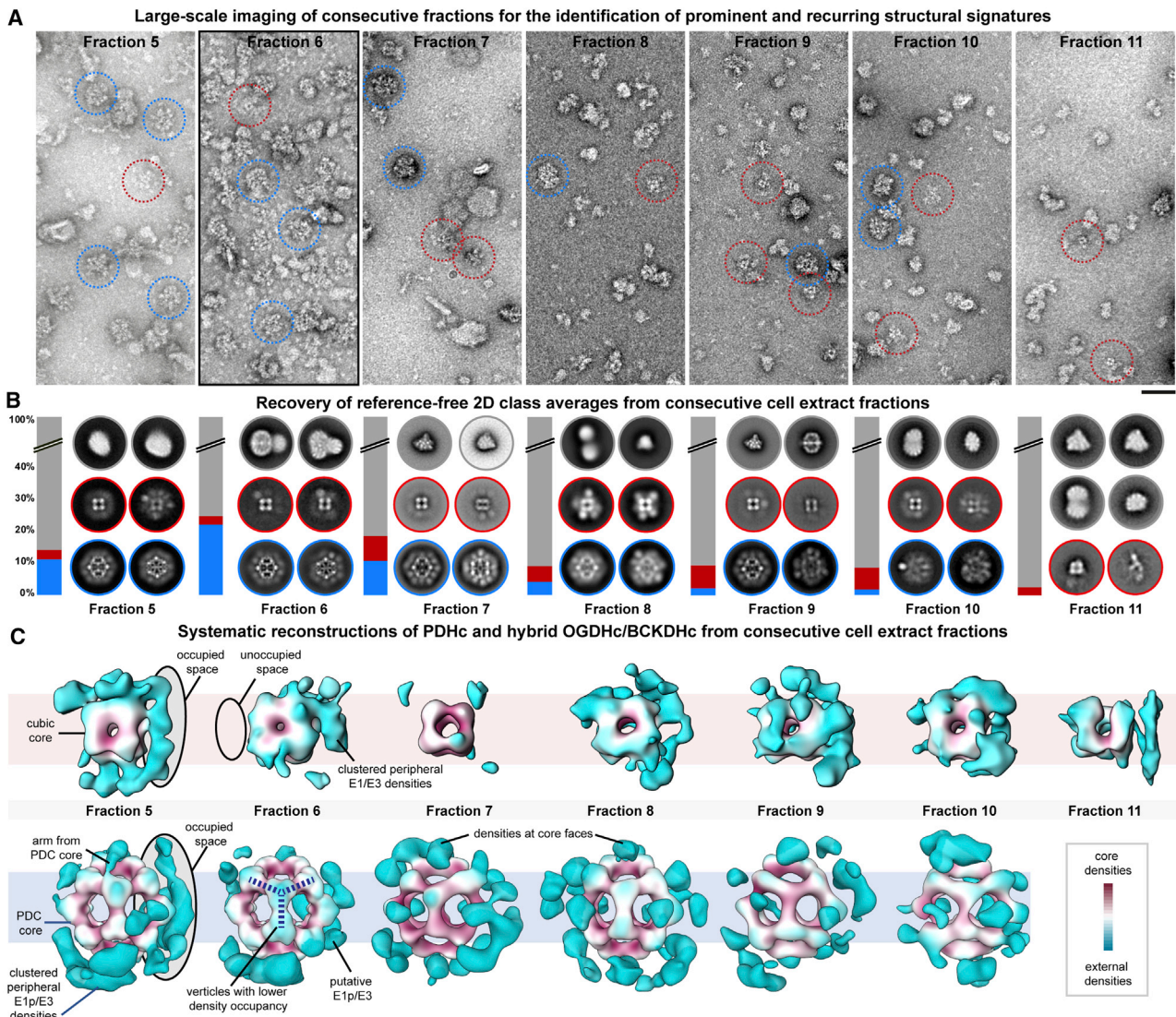
A Large-scale imaging of consecutive fractions for the identification of prominent and recurring structural signatures

B Recovery of reference-free 2D class averages from consecutive cell extract fractions

C Systematic reconstructions of PDHc and hybrid OGDHc/BCKDHc from consecutive cell extract fractions

Figure 3. Identification of OADH complexes in native cell extracts by EM
(A) Typical micrographs showing distinct single particles for fractions 5–11. Blue and red circles correspond to prominent signatures of particles with dodecahedral and octahedral cores, respectively. Scale bar: 60 nm.
(B) Quantification of particles resulting in class averages showing characteristics of dodecahedral (blue), octahedral (red), and other types of structures (gray). Inset class averages show a typical class average from each category of molecules. For fraction 6, all class averages are shown in Figure S4A.
(C) 3D reconstructions with C1 symmetry of complexes with cubic (top) and dodecahedral cores (bottom).

indicating higher flexibility, mixed signal, and/or "diffusability" of external densities. Still, we calculate that out of 17,649 single particles, at least 1.25 ± 0.57 localized densities are observed per single particle (Table S4).

To localize those manifested densities in 3D, we performed asymmetric reconstruction of the PDHc and OGDHc/BCKDHcs from derived class averages, each including 267,547 and 9,000 particles, respectively, and reaching ~26.6 Å and ~44.6 Å resolution (Fourier shell correlation [[FSC] = 0.5; Figure S5I). Despite limited resolution, we observe asymmetric clustering (Figures 3C and 4B) and densities that frequently stem from the vertices of the cores. We hypothesize that these densities of the PDHc

and the hybrid OGDHc/BCKDHc stemming from the core correspond to N-ter extensions of the E2 core proteins that tether the E1 (Figure 4B). At this resolution, we cannot exclude that these densities also correspond to E3BP flexible regions in the case of the PDHc, which are known to tether the E3. Most of the densities that are resolved were not previously seen (shown in purple; Figures 4B and 4C), mainly because both molecules were reconstructed symmetrically when studied; they were also rarely studied from the native source. Our reconstructions suggest the presence of external densities, which include 50% and 25% of the overall volume of the PDHc and the other OADH complexes, respectively (Figure 4C).
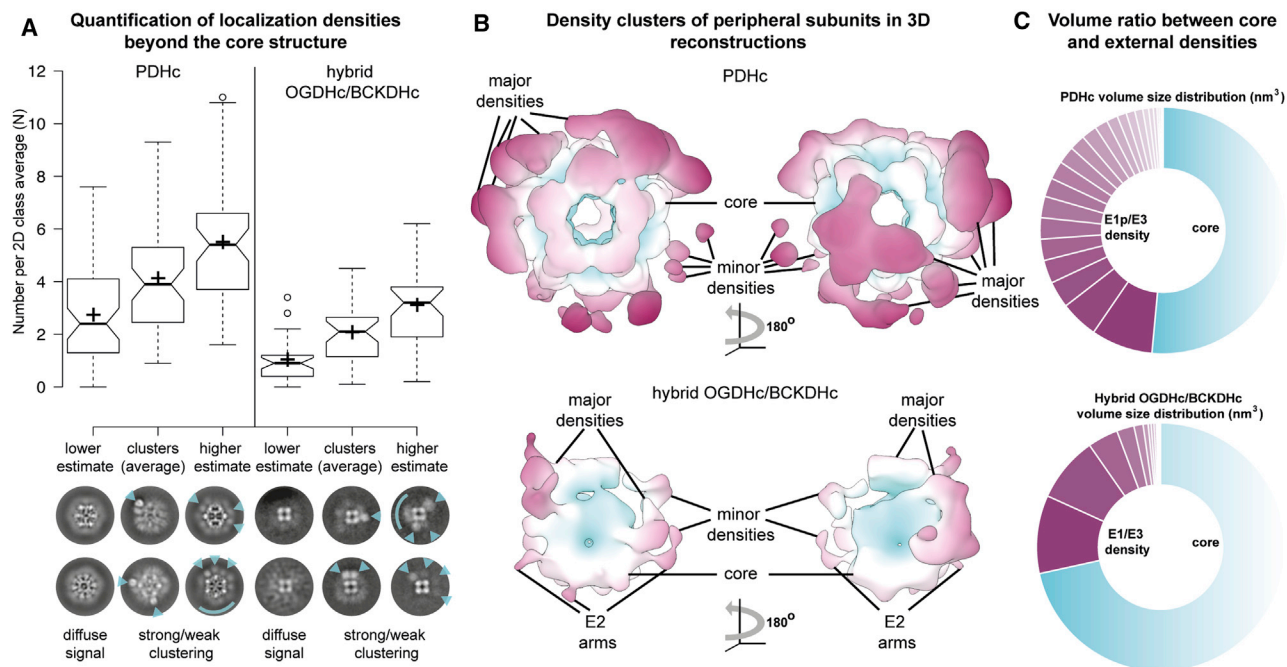
Figure 4. Analysis of external densities of α-keto acid dehydrogenase complexes

(A) Class averages exhibit strong external densities (clusters). Upper part shows boxplots of number of clusters per reference-free 2D class average. Lower part shows class averages with diffuse, weak, and strong signals (all shown in Table S4).

(B) A large concentration of external densities stem from the core structure as arms. Smaller densities are also present but constitute a minority of the overall structure.

(C) Ratio of volumetric data corresponding to the core (blue) and the outer part of the α-keto acid dehydrogenase complexes (purple). In total, 50% and 25% of the overall volume is covered by external densities, where E1p and E3 dimers are supposedly localized, respectively, together with associated kinases and phosphatases.

The cryo-EM structure of the native core of the PDHc reveals flexibility of folded domains and a basis for thermophilic interface formation

Our models of OADH E2 cores, modeled from the EM and cross-linking data, did not show significant energetic contributions to their interfaces as compared to their mesophilic counterparts (Table S3). Especially for PDHc, a higher-resolution reconstruction was derived (14.2 Å, FSC = 0.5; Figure S5D) as compared to the OGDHc (32.7 Å, FSC = 0.5; Figures S4C and S5I). In addition, compared to the human counterpart, although the fit of the derived E2p core model is in good agreement (Figure S5E), we observed that a few structural features were not well covered by the density: (1) part of the *N-ter* region Q227-N251 (T417 to I443 in human) of the E2p (Figures S5F and S5G) that forms a helix-turn-sheet-loop fold (Figures S2E and S2F) (Jiang et al., 2018; Mattevi et al., 1992), covering the top, external part of the trimeric asymmetric unit of the core (Figures S5G and S5H); and (2) inter-trimeric contacts mediated by helix G341–G359 of the homodimer (G531–G551 in human) (Figures S5G and S5H).

To verify the abovementioned differences, we vitrified fraction 6 and resolved the PDHc core at 6.9 Å by cryo-EM (FSC = 0.143; Figures 5 and S6; Table S5). Our previously resolved NS E2p model fit well (Figures S5F and S5H), and we were able to refine it because densities are uniform and α helices are distinguishable (Figures 5A and 5B). The region G341–G359 of the E2p homodimer, present at the inter-trimeric interface, is now resolved

(Figures 5C and S5H), showing that low resolution or structural unfolding artifacts due to staining may have limited localization of this structural region. By repeating energy calculations of the intra- and inter-trimeric interfaces, we derive species- and interface-specific differences. The intra-trimeric interface in the *C. thermophilum* E2p core is larger than in human (Figure 5D) and is also governed by large desolvation energy. Desolvation energy also governs the *C. thermophilum* inter-trimeric interface, which is almost 2 times higher than its counterpart. For the inter-trimeric interface in particular, weaker electrostatics are calculated (Figure 5D), indicating possible compensation effects for higher stability of the thermophilic interface. This higher contribution of solvation reflects optimization of the hydrophobic interactions in both the E2p core interfaces, and this biophysical adaptation was reported for other thermophilic proteins (Sawle and Ghosh, 2011) and complexes (Ma et al., 2010).

Another striking difference between native PDHc core and other E2p core structures is the absence of an equivalent density for the human *N-ter* region T417–I443, which is highly conserved in *C. thermophilum* and spans residues Q227–N251 (Figures 5E and S5H). This region was shown before to not be an active part of the catalytic domain (Mattevi et al., 1992), but constructs used to date include that *N-ter* extension. This region interacts with the active site, the latter also being highly conserved between human and *C. thermophilum* (Figure 5F). We superimposed the crystallographically determined molecules of oxidized CoA and
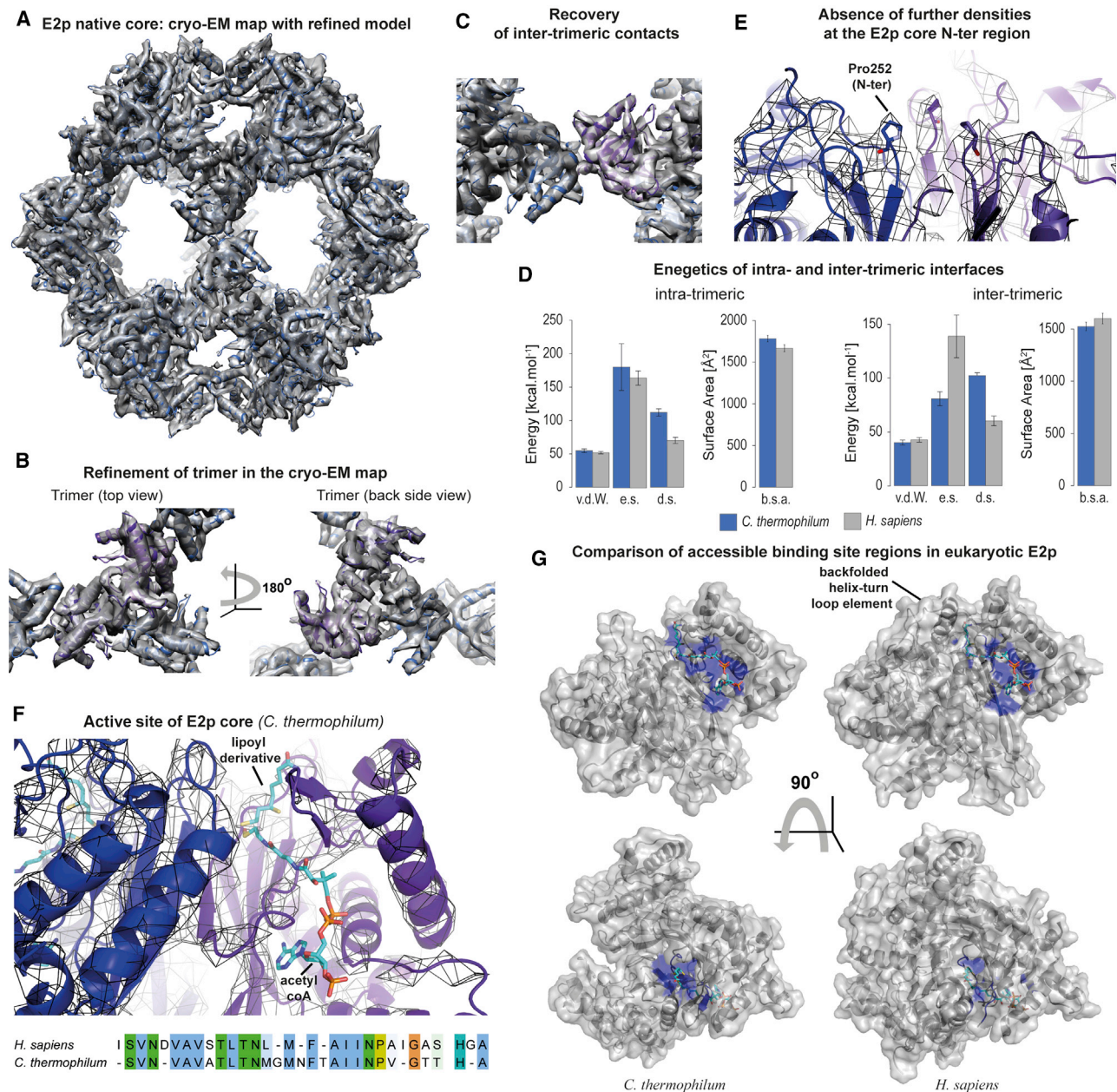
Figure 5. The cryo-EM PDHc core structure from native cell extracts

(A) Architecture of the *C. thermophilum* E2p core.

(B) View of the asymmetric unit of the E2p trimer.

(C) Recovery of inter-trimeric contacts and side helices that were not identified in the NS reconstruction (Figure S5).

(D) Comparative energy calculations for thermophilic and mesophilic interfaces. Average and standard deviations shown correspond to energetics for the top four calculated structures. v.d.w., van der Waals; e.s., electrostatics score; D.s., desolvation score; b.s.a, buried surface area.

(E) The core *N-ter* density does not continue beyond Pro252 (see also Figure S5).

(F) The *C. thermophilum* E2p active site. Fitting of a lipoyl derivative (lpm from PDB: 1EAE) and CoA (CoA from PDB: 1EAD) is possible, without clashes. Below, conserved active site residues are shown.

(G) Comparison of active site accessible surface areas of human and *C. thermophilum* E2p cores.

the lipoyl derivative (6,8-dimercapto-octanoic acid amide) from the *Azotobacter vinelandii* PDHc (Mattevi et al., 1993). Although side-chain resolution or cofactor Coulomb potential maps are not resolvable, the active site can accommodate both without

violation (Figure 5F). The two substrates that participate in the trans-acetylation reaction enter the active site from opposite directions; CoA enters from the interior of the core, and the acetylated lipoamide moiety enters from the outside. At this exact
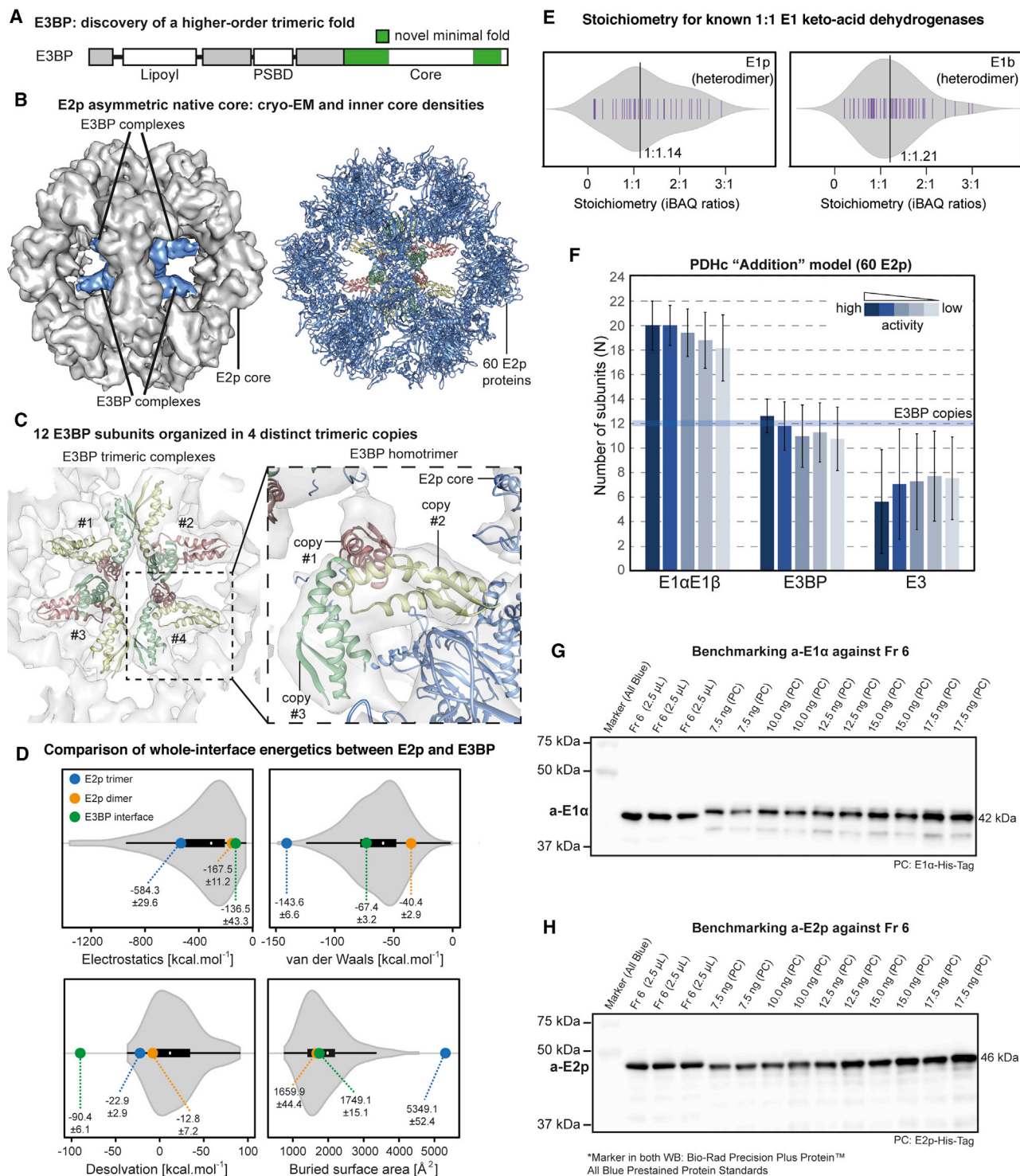
Figure 6. Localization and modeling of E3BP in the E2p scaffold and subsequent stoichiometry estimates

(A) Identified structural homology of *C. thermophilum* E3BP at the *C-ter* (in green).

(B) On the left, the C1 PDHc core reconstruction exhibits densities inside the E2p (blue density and red/yellow/green molecules).

(C) Trimeric E3BP models show significant curvature, recapitulated by the density.

(D) Energy calculations for various E2 interfaces and E3BP.

(E) Intensity-based absolute quantification (iBAQ) ratios of E1 complexes in various fractions across the replicates.

(legend continued on next page)

outer region, due to the structural absence of the conserved additional back-folded structural element of the *N-ter* region in *C. thermophilum*, accessible active site regions span a considerably larger surface as compared to the human counterpart (Figure 5G). Capturing the complex in the fraction circumvents possible determination of artificial states, as shown by the *N-ter* E2p region flexibility.

Four E3BP trimers are embedded within the PDHc with previously undetected oligomerization interface

We identified structural homologs for *ct*E3BP *C-ter* regions (residues 270–442; Figure 6A) and constructed a minimal fold (Figure S6A) based on structural similarity with the *Bos taurus C-ter* domain of the E2b (Kato et al., 2006). Looking inside the PDHc core (Figure 5), we observe strong averaged densities (Figure S6B) that were previously assumed to belong to E3BP in native yeast PDHc (Stoops et al., 1997). To gain further insights, we reconstructed the native core without symmetry at 8.7 Å (FSC = 0.143; Figure S6; Table S5) and retrieved four robust densities within the core, each resembling trimeric complexes of C3 symmetry (Figures 6B, 6C, and S6C). These trimeric complexes were also found in the *Neurospora crassa* E2p/E3BP core and shown to belong to E3BP (Forsberg et al., 2020). Here, we expand this observation to a PDHc from a thermophile.

Our extended E3BP core model recapitulates the shape of one arm of the inside trimeric density (Figure S6D). By considering C3 symmetry, we resolved an E3BP trimer (Figure 6C) with a clear, clash-free interface after fitting (Figures S6D and S6E) and refinement (Figure S6F). The *ct*E3BP model not only validates the few previously modeled helical regions of its homolog in *N. crassa*, but also unveils the interaction interface and its energetics as compared to the E2p trimer and E2p dimer, respectively (Figure 6D). The E2p/E3BP core interfaces recapitulate energetics found in 144 common protein-protein interfaces (Kastritis et al., 2014), with a few exceptions: (1) the E2p trimer has a large interface that exhibits similarities to permanent interfaces (extensive surface area and lower electrostatics; Bahadur et al., 2004); and (2) the E3BP trimer shows favorable desolvation, pointing to the involvement of additional protein-protein interactions as previously shown for other complexes with favorable desolvation (Casasoli et al., 2009). This is not surprising, since we resolved the core part of E3BP but not its additional flexible regions.

The PDHc is composed of 1 E2p 60-mer, 4 E3BP trimers, ~20 E1p tetramers, and a maximum of 12 E3 dimers

PDHc stoichiometry exhibits variations for core component proteins (E2p, E3BP), recently reported for human (Prajapati et al., 2019) and *N. crassa* (Forsberg et al., 2020) PDHc, observing the substitution (Hiromasa et al., 2004) or the addition model (Sanderson et al., 1996), respectively. E1α, E1β, and E3 decorate the core with diverse stoichiometries (Prajapati et al., 2019; Stoops et al., 1997). Here, we analyzed MS data and retrieved

abundance ratios for PDHc components, capturing a distribution with a median of a correct 1:1 stoichiometry for E1p and E1b across fractions (Figure 6E). PDHc component ratios in fractions with E1p activity are consistent; therefore, activity is not directly related to the extensive absence/presence of PDHc components but mostly to PDHc concentration (Figure 6F). Cryo-EM models of *C. thermophilum* E2p/E3BP core and ratios from MS data support a stoichiometry of a 60-meric E2p core and four E3BP trimers ("addition model"), in agreement with other reports in fungal PDHc (Forsberg et al., 2020; Stoops et al., 1997).

The maximum number of bound complexes is predicted to be 12 E3 and 60 E1p, but flexible arms of E3BP and E2p are not always saturated (Smolle et al., 2006). MS data predicted 2–12 E3 monomers (6 E3s) and ~20 E1α E1β dimers (~10 E2p), possibly underestimating their copy numbers because (1) E3 is found in all OADH complexes and (2) stoichiometry from iBAQ data is inherently limited (Wohlgemuth et al., 2015). We also calculated copy numbers of E1p by densitometry analysis, semi-quantifying E1α and E2p within fraction 6 (Figures 6G, 6H, and S7A). No bands were detected when E2p-His-Tag or E1α-His-Tag was immunoblotted using Abs against a-E1α and a-E2p, respectively (Figure S7A). By comparing to pure amounts (Figures 6G, 6H, and S7A), ~3 µg of E2p and ~1.7 µg of E1α are present in fraction 6, corresponding to monomeric molar concentrations of 166.2 ± 20.6 nmol·L^{-1} E2p and 107.2 ± 29.3 nmol·L^{-1} E1α. This points to ~19 copies of E1p being present per PDHc, ranging from ~16 to ~22 copies. Cryo-EM, kinetic assays, MS, and western blot (WB) data reveal that PDHc is composed of 1 E2p 60-mer, 4 E3BP trimers, ~20 E1p tetramers and a maximum of 12 E3 dimers, summing up to a complex composed of ~176 polypeptide chains.

Localized E1p and E3 subunits in the outer densities of the PDHc core

Utilizing the same micrographs and cryo-EM single particles that were used to reconstruct the E2p/E3BP cores (Figures 5, 6, and S6G–S6J), we also determined the structure of the asymmetric, native PDHc at a resolution of ~22 Å (FSC = 0.5; Table S5; Figures S6K and S6L). The total volume is 11,070–13,800 nm^3 and is recapitulated by calculating the summed volume of 20-Å-low-pass-filtered PDHc components (60 E2p, E2p core, 12 E3BP, 20 E1p, and 2–12 E3 dimers), totaling 10,608–13,930 nm^3. This shows that PDHc components can be accommodated in the reconstruction. It is possible that densities above the faces represent E3 due to the resolved localization of the E3BP *C-ter* and the distance constraints imposed by the 50-residue-long linker region. E1p components are expected to localize closer to the intersection of the dodecahedral edges, formed only by the E2p.

Derived class averages showed a systematic higher-order density outside the core (Figure 7A). This density manifested in pairs, reflecting a sort of communication, and was present in class

---

(F) Assuming 60 E2p, iBAQ ratios predict 12 E3BP subunits, but E1p and E3 dimer copy numbers are underestimated. Predictions are consistent across fractions with variable PDHc activity. Average and standard deviations shown correspond to stoichiometric calculations for fractions 5–11. Bar to the left includes measurements for fractions 5–7 (n = 3); and bars to the right include measurements for fractions 5–8 (n = 4), 5–9 (n = 5), 5–10 (n = 6), and 5–11 (n = 7), respectively.

(G and H) Benchmarking of Abs against E1α-His-Tag and E2p-His-Tag protein products with a technical triplicate of fraction 6 that includes native *C. thermophilum* E1α and E2p allows semi-quantification by densitometry.
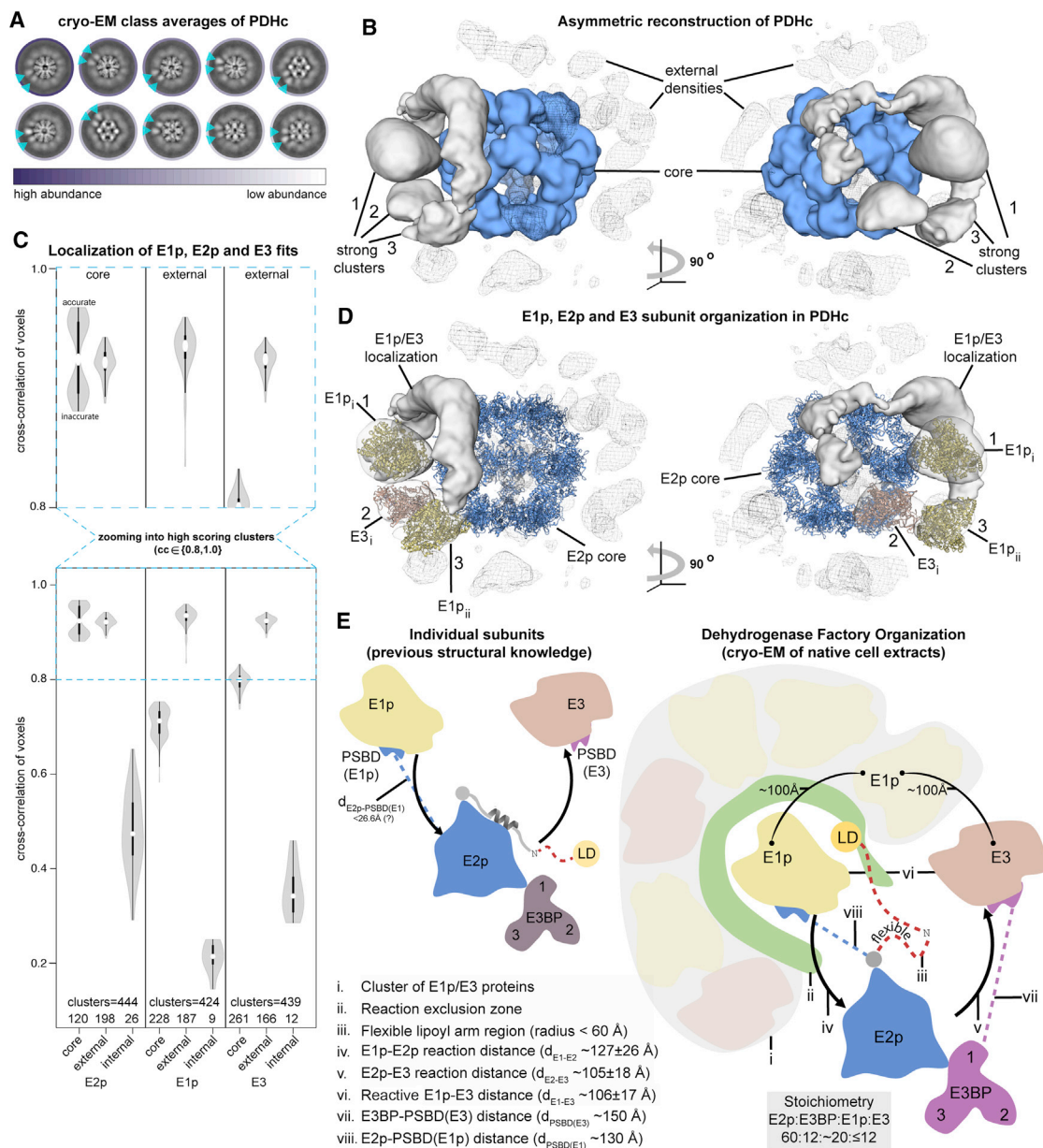
Figure 7. Asymmetric reconstruction of the full PDHc from native cell extracts

(A) Class averages showing localized binary densities (inset triangles). With purple to white, color-coded class-averages according to abundance.

(B) The PDHc map was segmented with Segger (Pintilie and Chiu, 2012); localized densities (wire), strong clusters of densities (silver), and E2p core (blue) are shown. Outer-density clusters are denoted as 1, 2, and 3.

(C) Systematic fitting of low-pass-filtered E1p, E2p, and E3 enzymes in the reconstruction (Figures S7C and S7D). E1p$_i$, E1p$_{ii}$, and E3$_i$ correspond to the best-fit conformation.

(D) Top fits are shown for the E1p, E2p, and E3. The rest of the densities are not fit due to ambiguity in localizing the enzymes.

(E) On the left, the model of PDHc is shown where the LD *C-ter* is stably attached to the E2p core, covering the binding site of the lipoyl. The arm transfers intermediates from the E1p to the E2p and then to the E3. E3BP is described inside the E2p, but its intra-molecular interface is unknown. On the right, the PDH factory model observed with our integrative methodology.

averages (Figure 7A). The asymmetric reconstruction clearly resolves E2p/E3BP densities previously modeled (Figures 5, 6, and 7B) but also recovers surrounding densities, previously reported at lower resolution in the NS C1 reconstruction (Figures 3C and 4). These outer densities are captured in two independent biological replicates and in the NS reconstructions of particles from the same fraction (Figure S7B), showing a prominent conical shape surrounded by other conical, elongated, and tubular densities (Figure 7B). This density corresponds to a MDa ultrastructure and may include external clusters of E1p/E3 subunits.

We derived molecular models for the E1p and E3 subunits of *C. thermophilum*, which share sufficient sequence homology to their human homologs. Together with the *C. thermophilum* E2p trimer (Figure 5), we filtered all models of enzyme domains to 20 Å and fit each to the asymmetric PDHc density, calculating 10,000 solutions for each (Figure 7C). Note that one filtered copy per each protein is fit 10,000 times in the asymmetric PDHc density. Then, each fit for each PDHc component was classified as fitting on the PDHc core, inside the PDHc core, or on the external densities, corresponding to the categories in Figure 7C ("core," "internal," and "external"). We performed the systematic fitting because we realized that at high CC values and density overlap, we can volumetrically distinguish E1p from E3s at 22 Å resolution (Figures S7C–S7F) and derive benchmarked thresholds for accurate fitting (Figures S7D and S7E). Importantly, after resolving the E2p/E3BP core, the flexible linker distances to their respective PSBDs are now known. This biochemical constraint will discard solutions if they clash with or violate these distances.

E2p localized with high CC values in the core, as expected. It also fit with lower CC to core densities (but those were at the inter-trimeric interface), at the external densities, or completely inside the core. For the E1p and E3s, higher CC was observed for fits at the non-core densities (Figure 7C). The E1p had higher CC values than the E3, possibly indicating its abundance on non-core densities. Out of these fit groups, those with high-volume overlap with the map are displayed in Figure 7D and ranked with CC values above threshold (Figures S7D–S7F). The E1p, having an overall conical shape, gives top-ranking fits inside the conical density of the reconstruction (location 1, E1p$_i$ position) and also provides another high-ranking fit in a less-resolved density, but of similar conical shape (location 3, E1p$_{ii}$ position). In both positions, E3 fits do not yield CC values above threshold. Both densities are closer to the vertices than to the faces of the native core. Equivalent top-ranking fits above threshold for the E3 (CC > 0.935) localize close to the E1p fits in a density that is rather elongated and on top of the face of the dodecahedral core (location 2, E3$_i$ position; Figure 7D). In the E3$_i$ position, no significant solution for the E1p is retrieved (Figure S7F). Due to the C2 symmetry of the E1p and E3, both solutions were recovered with high CC values for all three positions, validating that the map holds information that can discriminate symmetric solutions. A similar symmetry analysis also correctly localized the hNup107 subcomplex in the nuclear pore complex structure (Bui et al., 2013). In our reconstruction, we do not see direct interactions of E1p or E3 subunits with the E2p/E3BP core, pointing to those being highly dynamic, rare, or a minority set of events during the overall reaction. This is not at all surprising since E1 or E3 LD interactions with the core, by their very nature, must be highly transient, rapidly transferring the substrates acyl group to CoA and then releasing. In any event, the core would be expected to favor binding of acetylated LDs, likely in the presence of CoA, and neither of them was present.

Dehydrogenase factories: a discovered structural organization of the PDHc revealed by integrative structural biology

Here, we propose a model of a "dehydrogenase factory" (Figure 7E). We consider that this could be a full reaction pathway for the PDHc reaction, forming a temporal reaction chamber. The flexible parts of the dehydrogenase factory are, however, notably restricted by a density cloud surrounding single E1p and E3 subunits, as visualized by cryo-EM (Figures 7B, 7D, and 7E). This density cloud presumably is composed of E1p and E3, but we cannot exclude the possible presence of PDH-specific kinases and/or phosphatases or other binders. However, in the MS data (Table S1), PDH kinases and phosphatases were of very low abundance, and to our knowledge, no other known binders of PDHc were identified. The E1p and E3 proteins are not observed to interact directly with the E2p core but are spatially confined in relative proximity. However, this E1p/E3-density cloud confines the flexibility of the E2p and E3BP unresolved *N-ter* regions, which tether the E1p and E3 subunit to the core. The E1p-E2p active site minimum distance (d$_{minAS}$) of ~65 Å and the E2p-E3 d$_{minAS}$ of ~75 Å are in the range of both flexible linker regions (core-PSBD; PSBD-LD; Figure S2B). The E1p-E3 d$_{minAS}$ of 80 Å is in the linker range as well (PSBD-LD). In addition, the emerging distance between the last residue of the unresolved region of PSBD and the first residue of the E3BP core (Leu215-Ala271) is in a realistic range (138.8 Å or 151.1 Å), corresponding to a compacted state of the 55 residues structuring this region. None of the Euclidean distances are sterically hindered; especially for E3BP, the distance measurement transverses the face of the dodecahedron. In between the E1p, E2p, and E3 enzymes, a continuous exclusion zone is present with no apparent densities. This exclusion zone may accommodate the flexible LDs and define a trail for the delivery of intermediates among the E1p, E2p, and E3 active sites, forming the catalytically active reaction chamber of the PDH factory.

DISCUSSION

Advances and directions in the integrative structural studies of native cell extracts

Higher-resolution structural studies of native cell extracts combine proteomics, (cryo-)EM, and computational biology to assemble meaningful native architectures of large assemblies (Kyrilis et al., 2019; McCafferty et al., 2020), and our work connected activity profiles in this approach. Activity must be probed because purification procedures may inactivate complexes and offer a view of what may constitute a functional assembly. For PDHc, subcomplexes are those directly involved in the partial reactions of pyruvate oxidation (E1p, E3BP, E2p, E3), whereas the complex is the entity performing the full catalytic reaction (PDHc).

The study of cell extracts is limited by relative enrichment of molecules that may differ across organisms, confining options for comparative studies. Importantly, the addition of results from highly pure material must be considered because unambiguous data obtained will complement the study of native extracts. Here, we relied on results from pure material to retrieve (1) in-fraction activities of OADH complexes; (2) *C. thermophilum* models of OADH complexes to build, enrich, validate, and support the PDH factory model; (3) approximate copy numbers of E1p in the PDHc; and (4) PDHc domain component analysis to report distances of PSBDs to the E2p/E3BP core and among E1p, E2p, and E3 active sites.

Lysing cells to recover material will affect the native state; electron cryo-tomography (cryo-ET) of lamellae of mitochondria

from focused ion beam milling (cryo-FIB) could be an option to probe the PDHc structure, but extreme mitochondrial protein concentration (up to ~560 mg mL^{-1}; Srere, 1980) prohibits in-cell identification. Conclusively, it is unknown if the retrieved "native state" of the PDHc from extracts resembles the in-cell structure and function, but considering the limitations of the methods available to date, our integrative methodology, with its advantages and disadvantages (Kyrilis et al., 2019), provides a closer view off the native biomolecular structure.

Comparison of eukaryotic PDHcs uncovers adaptations of the native state

The PDHc or its core structures are produced using icosahedral symmetry (Jiang et al., 2018; Prajapati et al., 2019; Stoops et al., 1997; Yu et al., 2008; Zhou et al., 2001), but non-icosahedrally symmetric parts will be averaged out or eventually manifest as artificial densities, especially for the PDHc, as recently suggested (Forsberg et al., 2020). For the mammalian core, a "substitution model" is favored with 40:20 (Brautigam et al., 2009) or 48:12 E2p:E3BP proteins (Hiromasa et al., 2004). In fungi, an "addition model" is favored with 60:12 E2p:E3BP proteins, and the location of E3BP has been observed in a lower-symmetry reconstruction of the E2p/E3BP core (Forsberg et al., 2020). Additionally, all E2p cores reported to date resolve an *N-ter* back-folded conformation of E2p, limiting accessibility of the LD to the E2p active site. It is still unknown how the LD interacts with the E2p (or any PDHc component) active sites because no interface data are available, and only model compounds have been solved within the E2p active site or bound to LD. Consequences of applied symmetry are substantial for recovery of the positions of the E1p and E3 relative to the core. The distance to the E1p is only 26.6 Å (seven residues) if the flexible E2p core *N-ter* is folded. This distance contradicts previous hypotheses on E1p localization (e.g., up to 95 Å from the E2p; Zhou et al., 2001). The E1p was also hypothesized to be ~50 Å away from the E2p core, close to the E3, almost at "interacting" distance (Yu et al., 2008). Even if E2p-E1p vicinity could be feasible—agreeing, for example, with the *in vitro* E2p/E3BP/E1p complex shown in NS micrographs (Prajapati et al., 2019)—the external densities in the native *C. thermophilum* PDHc are observed ~150 Å from the PDHc core (Figure 7). This implies that the folded state of the linker is not favoring an active, native PDHc. Placement of the E1p on an asymmetrically resolved density settles the proximity of its active site relative to the E2p. The location of the E3 in mammalian PDHc is speculative because E3BP was never really localized. Even when the E3BP position was postulated in yeast (Stoops et al., 1997) and shown in *N. crassa* with a partial model (Forsberg et al., 2020), the E3 is hypothesized to be close to the pentagonal openings of the core. We do not observe any density in this location but localize the E3 farther above the pentagonal face. The E3BP *C-ter* is also captured in our reconstruction, and the distance to its E3 PSBD is 138.1 Å. Considering the linker length (Leu215-Ala271), the E3 is reasonably placed, and the linker is compacted.

Dehydrogenase factories: a plausible model for localized pyruvate oxidation reactions

Our 3D model suggests an architecture, components, and approximate distances that overall define a low-resolution struc-

ture of a major structural element of native PDHc, the PDH factory. It manifests as a high-molecular-weight density on the outer part of the PDHc core and provides an attractive model that may explain a catalytic mechanism for the complex where there is restricted movement of involved enzymes by the very same molecules that would otherwise participate in the pyruvate oxidation reaction. They additionally define the relative proximity of binding sites, and, although it is observed that a greater movement of the LD is possible, this movement is confined in the exclusion zone. Our data ultimately propose the minimum distance of LD trafficking to perform a full PDHc reaction to be ~140 Å, as measured by the summation of minimum distances among the three active sites.

Various studies have postulated two different mechanisms for the higher-order interaction of the components of OADH complexes: the "multiple random coupling" (MRC) (Hackert et al., 1983) and the "division-of-labor" (DOL) (Milne et al., 2006; Prajapati et al., 2019). The observed PDH factory provides evidence for a modified DOL mechanism, where a non-homogeneous distribution of different PDHc components is assumed but does not necessarily agree with a mechanism for spatially organized, distinct (E2p-E1p and E3BP-E3) clusters. It rather points to a clustering behavior of increased sophistication of the E1p and E3 in relation to the PDHc core, where the E1p and E3 interdependently separate for some to be involved in the enzymatic reaction and others to spatially confine it. We can only speculate if structures similar to dehydrogenase factories exist, perhaps, for mammalian PDHc or the remaining OADH complexes. Interestingly, another molecular reaction chamber has been observed recently with lower-resolution methods responsible for chromatin modification (Gilbert and van Leeuwen, 2020); it is intriguing to hypothesize that transient MDa reaction chambers could be additionally responsible for pathway/reaction regulation.

Overall, we provided insights into the spatial relationships of the E1p, E2p, E3BP, and E3. Open questions remain. For example, is localization and determination of the dynamic lipoyl arms and a higher number of E1p and E3 proteins possible? Are there additional components of these MDa machines (as shown in Figure S7G for the PDHc)? How does the complex outer architecture change in response to (de-)phosphorylation and interaction with its various kinases and phosphatases? And, ultimately, are high-resolution insights into native architectural and functional relationships of the OADH complex family feasible, considering the immense structural complexity of those metabolons?

Answers could be provided by cryo-EM in combination with finer fractionation of cell extracts joined with high-density crosslinking MS, utilizing stable biomolecular assemblies, kinetic assays for all subunits and/or pure components, large-scale image acquisition efforts, and efficient asymmetric reconstruction. Ultimately, integrative biology of native cell extracts can shed light onto the immense complexity of the giant machines behind cellular respiration, a pathway that has fascinated biochemists since Hans Krebs (Krebs and Johnson, 1937). It continues to be of vital interest due to the involvement of those α-keto acid dehydrogenase complexes in numerous functions and disease.

Cell Reports
Article

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Model organism
- METHOD DETAILS
  - Mycelium harvesting, culture lysis and sample preparation
  - Size exclusion chromatography
  - Activity assays
  - Transmission electron microscopy of negatively stained samples and data collection
  - Cryo-EM sample preparation and data collection
  - Image processing of negative stain data and volumetric calculations
  - Image processing of cryo-EM data
  - Sequence alignments, homology modeling, interface energy statistics and model refinement
  - Localization, fitting and refinement of E3BP in the asymmetric C1 core reconstruction of PDHc
  - Antibody production
  - Densitometric protein quantification
  - Fit of E1p, E2p and E3 proteins in the PDHc C1 reconstruction and measured active site distances
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2021.108727.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

F.L.K. designed and carried out the experimental work, assisted by F.J.O. in mass spectrometry and D.A.S. in negative-staining EM. F.H. acquired the micrographs. D.A.S. calculated the negative-staining EM 3D reconstructions and I.S. the cryo-EM reconstructions. C.T. performed computational modeling. P.L.K. and J.R. secured funding. P.L.K. conceived the project and wrote the manuscript with contributions from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. D Struct. Biol. 74, 531–544.

Bahadur, R.P., Chakrabarti, P., Rodier, F., and Janin, J. (2004). A dissection of specific and non-specific protein-protein interfaces. J. Mol. Biol. 336, 943–955.

Bassendine, M.F., Jones, D.E., and Yeaman, S.J. (1997). Biochemistry and autoimmune response to the 2-oxoacid dehydrogenase complexes in primary biliary cirrhosis. Semin. Liver Dis. 17, 49–60.

Bisswanger, H. (1981). Substrate specificity of the pyruvate dehydrogenase complex from Escherichia coli. J. Biol. Chem. 256, 815–822.

Brautigam, C.A., Wynn, R.M., Chuang, J.L., Machius, M., Tomchick, D.R., and Chuang, D.T. (2006). Structural insight into interactions between dihydrolipoamide dehydrogenase (E3) and E3 binding protein of human pyruvate dehydrogenase complex. Structure 14, 611–621.

Brautigam, C.A., Wynn, R.M., Chuang, J.L., and Chuang, D.T. (2009). Subunit and catalytic component stoichiometries of an in vitro reconstituted human pyruvate dehydrogenase complex. J. Biol. Chem. 284, 13086–13098.

Bui, K.H., von Appen, A., DiGuilio, A.L., Ori, A., Sparks, L., Mackmull, M.T., Bock, T., Hagen, W., Andrés-Pons, A., Glavy, J.S., and Beck, M. (2013). Integrated structural analysis of the human nuclear pore complex scaffold. Cell 155, 1233–1243.

Casasoli, M., Federici, L., Spinelli, F., Di Matteo, A., Vella, N., Scaloni, F., Fernandez-Recio, J., Cervone, F., and De Lorenzo, G. (2009). Integration of evolutionary and desolvation energy analysis identifies functional sites in a plant immunity protein. Proc. Natl. Acad. Sci. USA 106, 7666–7671.

Cheng, J., Baßler, J., Fischer, P., Lau, B., Kellner, N., Kunze, R., Griesel, S., Kallas, M., Berninghausen, O., Strauss, D., et al. (2019). Thermophile 90S Pre-ribosome Structures Reveal the Reverse Order of Co-transcriptional 18S rRNA Subdomain Integration. Mol. Cell 75, 1256–1269.e7.

Ciszak, E.M., Makal, A., Hong, Y.S., Vettaikkorumakankauv, A.K., Korotchkina, L.G., and Patel, M.S. (2006). How dihydrolipoamide dehydrogenase-binding protein binds dihydrolipoamide dehydrogenase in the human pyruvate dehydrogenase complex. J. Biol. Chem. 281, 648–655.

de la Rosa-Trevín, J.M., Quintana, A., Del Cano, L., Zaldívar, A., Foche, I., Gutiérrez, J., Gómez-Blanco, J., Burguet-Castell, J., Cuenca-Alba, J., Abrishami, V., et al. (2016). Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. J. Struct. Biol. 195, 93–99.

Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U., and Sali, A. (2006). Comparative protein structure modeling using Modeller. Curr. Protoc. Bioinformatics 54, 5.6.1–5.6.37.

Forsberg, B.O., Aibara, S., Howard, R.J., Mortezaei, N., and Lindahl, E. (2020). Arrangement and symmetry of the fungal E3BP-containing core of the pyruvate dehydrogenase complex. Nat. Commun. 11, 4667.

Gilbert, N., and van Leeuwen, F. (2020). Chromatin modified in a molecular reaction chamber. Nature 579, 503–504.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 27, 14–25.

Hackert, M.L., Oliver, R.M., and Reed, L.J. (1983). Evidence for a multiple random coupling mechanism in the alpha-ketoglutarate dehydrogenase multi-enzyme complex of Escherichia coli: a computer model analysis. Proc. Natl. Acad. Sci. USA 80, 2226–2230.

Hiromasa, Y., Fujisawa, T., Aso, Y., and Roche, T.E. (2004). 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.

Ho, C.M., Li, X., Lai, M., Terwilliger, T.C., Beck, J.R., Wohlschlegel, J., Goldberg, D.E., Fitzpatrick, A.W.P., and Zhou, Z.H. (2020). Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. Nat. Methods 17, 79–85.

Hunter, S., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., et al. (2009). InterPro: the integrative protein signature database. Nucleic Acids Res. 37, D211–D215.

Jakkamsetti, V., Marin-Valencia, I., Ma, Q., Good, L.B., Terrill, T., Rajasekaran, K., Pichumani, K., Khemtong, C., Hooshyar, M.A., Sundarrajan, C., et al. (2019). Brain metabolism modulates neuronal excitability in a mouse model of pyruvate dehydrogenase deficiency. Sci. Transl. Med. 11, eaan0457.

Jiang, J., Baiesc, F.L., Hiromasa, Y., Yu, X., Hui, W.H., Dai, X., Roche, T.E., and Zhou, Z.H. (2018). Atomic Structure of the E2 Inner Core of Human Pyruvate Dehydrogenase Complex. Biochemistry 57, 2325–2334.

Kastritis, P.L., Rodrigues, J.P., Folkers, G.E., Boelens, R., and Bonvin, A.M. (2014). Proteins feel more than they see: fine-tuning of binding affinity by properties of the non-interacting surface. J. Mol. Biol. 426, 2632–2652.

Kastritis, P.L., O'Reilly, F.J., Bock, T., Li, Y., Rogon, M.Z., Buczak, K., Romanov, N., Betts, M.J., Bui, K.H., Hagen, W.J., et al. (2017). Capturing protein communities by structural proteomics in a thermophilic eukaryote. Mol. Syst. Biol. 13, 936.

Kato, M., Wynn, R.M., Chuang, J.L., Brautigam, C.A., Custorio, M., and Chuang, D.T. (2006). A synchronized substrate-gating mechanism revealed by cubic-core structure of the bovine branched-chain alpha-ketoacid dehydrogenase complex. EMBO J. 25, 5983–5994.

Kato, M., Wynn, R.M., Chuang, J.L., Tso, S.C., Machius, M., Li, J., and Chuang, D.T. (2008). Structural basis for inactivation of the human pyruvate dehydrogenase complex by phosphorylation: role of disordered phosphorylation loops. Structure 16, 1849–1859.

Kellner, N., Schwarz, J., Sturm, M., Fernandez-Martinez, J., Griesel, S., Zhang, W., Chait, B.T., Rout, M.P., Kück, U., and Hurt, E. (2016). Developing genetic tools to exploit Chaetomium thermophilum for biochemical analyses of eukaryotic macromolecular assemblies. Sci. Rep. 6, 20937.

Klyuyeva, A., Tuganova, A., Kedishvili, N., and Popov, K.M. (2019). Tissue-specific kinase expression and activity regulate flux through the pyruvate dehydrogenase complex. J. Biol. Chem. 294, 838–851.

Krebs, H.A., and Johnson, W.A. (1937). The role of citric acid in intermediate metabolism in animal tissues. Enzymologia 4, 148–156.

Kresze, G.B., and Ronft, H. (1981). Pyruvate dehydrogenase complex from baker's yeast. 2. Molecular structure, dissociation, and implications for the origin of mitochondria. Eur. J. Biochem. 119, 581–587.

Kyrilis, F.L., Meister, A., and Kastritis, P.L. (2019). Integrative biology of native cell extracts: a new era for structural characterization of life processes. Biol. Chem. 400, 831–846.

Lengyel, J.S., Stott, K.M., Wu, X., Brooks, B.R., Balbo, A., Schuck, P., Perham, R.N., Subramaniam, S., and Milne, J.L. (2008). Extended polypeptide linkers establish the spatial architecture of a pyruvate dehydrogenase multienzyme complex. Structure 16, 93–103.

Liu, Y., You, Y., Lu, Z., Yang, J., Li, P., Liu, L., Xu, H., Niu, Y., and Cao, X. (2019). N^6-methyladenosine RNA modification-mediated cellular metabolism rewiring inhibits viral replication. Science 365, 1171–1176.

Ma, B.G., Goncearenco, A., and Berezovsky, I.N. (2010). Thermophilic adaptation of protein complexes inferred from proteomic homology modeling. Structure 18, 819–828.

Mattevi, A., Obmolova, G., Schulze, E., Kalk, K.H., Westphal, A.H., de Kok, A., and Hol, W.G. (1992). Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. Science 255, 1544–1550.

Mattevi, A., Obmolova, G., Kalk, K.H., Teplyakov, A., and Hol, W.G. (1993). Crystallographic analysis of substrate binding and catalysis in dihydrolipoyl transacetylase (E2p). Biochemistry 32, 3887–3901.

McCafferty, C.L., Verbeke, E.J., Marcotte, E.M., and Taylor, D.W. (2020). Structural Biology in the Multi-Omics Era. J. Chem. Inf. Model. 60, 2424–2429.

Milne, J.L., Wu, X., Borgnia, M.J., Lengyel, J.S., Brooks, B.R., Shi, D., Perham, R.N., and Subramaniam, S. (2006). Molecular structure of a 9-MDa icosahedral pyruvate dehydrogenase subcomplex containing the E2 and E3 enzymes using cryoelectron microscopy. J. Biol. Chem. 281, 4364–4370.

Mkrtchyan, G., Aleshin, V., Parkhomenko, Y., Kaehne, T., Di Salvo, M.L., Parroni, A., Contestabile, R., Vovk, A., Bettendorff, L., and Bunik, V. (2015). Molecular mechanisms of the non-coenzyme action of thiamin in brain: biochemical, structural and pathway analysis. Sci. Rep. 5, 12583.

Needleman, S.B., and Wunsch, C.D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48, 443–453.

Palmieri, E.M., Gonzalez-Cotto, M., Baseler, W.A., Davies, L.C., Ghesquière, B., Maio, N., Rice, C.M., Rouault, T.A., Cassel, T., Higashi, R.M., et al. (2020). Nitric oxide orchestrates metabolic rewiring in M1 macrophages by targeting aconitase 2 and pyruvate dehydrogenase. Nat. Commun. 11, 698.

Patel, M.S., Roche, T.E., and Harris, R. (1996). Alpha-Keto Acid Dehydrogenase Complexes (Springer-Verlag).

Patel, M.S., Nemeria, N.S., Furey, W., and Jordan, F. (2014). The pyruvate dehydrogenase complexes: structure-based function and regulation. J. Biol. Chem. 289, 16615–16623.

Perham, R.N. (1991). Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. Biochemistry 30, 8501–8512.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Pintilie, G., and Chiu, W. (2012). Comparison of Segger and other methods for segmentation and rigid-body docking of molecular components in cryo-EM density maps. Biopolymers 97, 742–760.

Prajapati, S., Haselbach, D., Wittig, S., Patel, M.S., Chari, A., Schmidt, C., Stark, H., and Tittmann, K. (2019). 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.

Reed, L.J. (2001). A trail of research from lipoic acid to alpha-keto acid dehydrogenase complexes. J. Biol. Chem. 276, 38329–38336.

Roche, T.E., and Reed, L.J. (1972). Function of the nonidentical subunits of mammalian pyruvate dehydrogenase. Biochem. Biophys. Res. Commun. 48, 840–846.

Sanderson, S.J., Khan, S.S., McCartney, R.G., Miller, C., and Lindsay, J.G. (1996). Reconstitution of mammalian pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase complexes: analysis of protein X involvement and interaction of homologous and heterologous dihydrolipoamide dehydrogenases. Biochem. J. 319, 109–116.

Sawle, L., and Ghosh, K. (2011). How do thermophilic proteins and proteomes withstand high temperature? Biophys. J. 101, 217–227.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682.

Seifert, F., Ciszak, E., Korotchkina, L., Golbik, R., Spinka, M., Dominiak, P., Sidhu, S., Brauer, J., Patel, M.S., and Tittmann, K. (2007). Phosphorylation of serine 264 impedes active site accessibility in the E1 component of the human pyruvate dehydrogenase multienzyme complex. Biochemistry 46, 6277–6287.

Smolle, M., Prior, A.E., Brown, A.E., Cooper, A., Byron, O., and Lindsay, J.G. (2006). A new level of architectural complexity in the human pyruvate dehydrogenase complex. J. Biol. Chem. *281*, 19772–19780.

Sorzano, C.O., Marabini, R., Velázquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H., Carazo, J.M., and Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. J. Struct. Biol. *148*, 194–204.

Srere, P.A. (1980). The Infrastructure of the Mitochondrial Matrix. Trends Biochem. Sci. *5*, 120–121.

Stoops, J.K., Cheng, R.H., Yazdi, M.A., Maeng, C.Y., Schroeter, J.P., Klueppelberg, U., Kolodziej, S.J., Baker, T.S., and Reed, L.J. (1997). On the unique structural organization of the Saccharomyces cerevisiae pyruvate dehydrogenase complex. J. Biol. Chem. *272*, 5757–5764.

Szabo, E., Wilk, P., Nagy, B., Zambo, Z., Bui, D., Weichsel, A., Arjunan, P., Torocsik, B., Hubert, A., Furey, W., et al. (2019). Underlying molecular alterations in human dihydrolipoamide dehydrogenase deficiency revealed by structural analyses of disease-causing enzyme variants. Hum. Mol. Genet. *28*, 3339–3354.

Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., and Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat. Methods *14*, 793–796.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. *157*, 38–46.

van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kastritis, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J., and Bonvin, A.M.J.J. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. J. Mol. Biol. *428*, 720–725.

Verbeke, E.J., Mallam, A.L., Drew, K., Marcotte, E.M., and Taylor, D.W. (2018). Classification of Single Particles from Human Cell Extract Reveals Distinct Structures. Cell Rep. *24*, 259–268.e3.

Verbeke, E.J., Zhou, Y., Horton, A.P., Mallam, A.L., Taylor, D.W., and Marcotte, E.M. (2020). Separating distinct structures of multiple macromolecular assemblies from cryo-EM projections. J. Struct. Biol. *209*, 107416.

Wagner, T., and Raunser, S. (2020). The evolution of SPHIRE-crYOLO particle picking and its application in automated cryo-EM processing workflows. Commun. Biol. *3*, 61.

White, P.J., McGarrah, R.W., Grimsrud, P.A., Tso, S.C., Yang, W.H., Haldeman, J.M., Grenier-Larouche, T., An, J., Lapworth, A.L., Astapova, I., et al. (2018). The BCKDH Kinase and Phosphatase Integrate BCAA and Lipid Metabolism via Regulation of ATP-Citrate Lyase. Cell Metab. *27*, 1281–1293.e7.

Wohlgemuth, I., Lenz, C., and Urlaub, H. (2015). Studying macromolecular complex stoichiometries by peptide-based mass spectrometry. Proteomics *15*, 862–879.

Yu, X., Hiromasa, Y., Tsen, H., Stoops, J.K., Roche, T.E., and Zhou, Z.H. (2008). Structures of the human pyruvate dehydrogenase complex cores: a highly conserved catalytic center with flexible N-terminal domains. Structure *16*, 104–114.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zhou, Z.H., McCarthy, D.B., O'Connor, C.M., Reed, L.J., and Stoops, J.K. (2001). The remarkable structural and functional organization of the eukaryotic pyruvate dehydrogenase complexes. Proc. Natl. Acad. Sci. USA *98*, 14802–14807.

Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.N., and Alva, V. (2018). A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J. Mol. Biol. *430*, 2237–2243.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*, e42166.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal antibody a-E1α against *C. thermophilum* E1α-His-Tag (37-411)	This paper (custom-made, Genscript)	RRID: AB_2888984
Rabbit polyclonal antibody a-E2p against *C. thermophilum* E2p-His-Tag (29-459)	This paper (custom-made, Genscript)	RRID: AB_2888985
Chemicals, peptides, and recombinant proteins		
C. thermophilum E1α-His-Tag (37-411)	This paper (custom-made, Genscript)	N/A
C. thermophilum E2p-His-Tag (29-459)	This paper (custom-made, Genscript)	N/A
1,4-Dithiothreit, min. 99%, p.a.	Carl Roth	6908.4
Acrylamide/Bis solution, 37.5:1	Serva	10688.01
Agar-Agar, bacteriological highly pure	Carl Roth	2266.3
Ammonium acetate, $\geq$ 97%, p.a., ACS	Carl Roth	7869.2
Ammonium persulfate	Serva	13376.02
Aprotinin from bovine lung	Sigma-Aldrich	A1153-1MG
Bestatin, 10MG	Sigma-Aldrich	10874515001
Clarity Western ECL substrate	BIO-RAD	170-5060
D-Sucrose, $\geq$ 99,5%, p.a.	Carl Roth	4621.1
D(+)-Glucose p. a., ACS, anhydrous	Carl Roth	X997.2
Dextrin for microbiology (from potato starch)	Carl Roth	3488.1
di-Potassium hydrogen phosphate trihydrate	Carl Roth	6878.1
di-Potassium hydrogen phosphate, $\geq$ 99%, p.a., anhydrous	Carl Roth	P749.1
DNase I	Sigma-Aldrich	10104159001
E-64	Sigma-Aldrich	E3132-1MG
EDTA disodium salt dihydrate, min. 99%, p.a., ACS	Carl Roth	8043.2
Glycine	Serva	23391.02
HEPES PUFFERAN®, min. 99.5%, p.-1 kg	Carl Roth	9105.3
Iron (III) sulfate hydrate, 80%, pure	Carl Roth	0492.1
Isopropanol	Carl Roth	CP41.1
Leupeptin	Sigma-Aldrich	L2884-1MG
Magnesium chloride hexahydrate, min. 99%, p.a., ACS	Carl Roth	2189.1
Magnesium sulfate heptahydrate, $\geq$ 99%, p.a., ACS	Carl Roth	P027.1
Methanol	Carl Roth	4627.6
Milk powder	Carl Roth	T145.3
Pefabloc	Sigma-Aldrich	11585916001
Pepstatin A	Sigma-Aldrich	77170-5MG
Peptone ex casein	Carl Roth	8986.1
Phosphate buffered saline tablets (PBS)	Sigma-Aldrich	P4417
Potassium chloride min. 99.5%, −1 kg	Carl Roth	6781.1

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Potassium dihydrogen phosphate, ≥ 99%, p.a., ACS	Carl Roth	3904.2
Precision plus protein all blue standards (marker)	BIO-RAD	161-0373
Sodium chloride 99,5%, p.a., ACS, ISO	Carl Roth	3957.2
Sodium dodecyl sulfate (SDS)	Carl Roth	0183.2
Sodium nitrate, ≥ 99%, p.a., ACS, ISO	Carl Roth	A136.1
TEMED	Carl Roth	2367.3
Tris	Carl Roth	AE15.2
Tris Hydrochloride	Carl Roth	9090.2
Tween 20	Carl Roth	9127.1
Tryptone	Sigma-Aldrich	T7293
Yeast Extract, micro-granulated	Carl Roth	2904.3
Critical commercial assays		
A-KG Dehydrogenase activity assay kit	Sigma-Aldrich	MAK189-1KT
PDH Activity assay kit	Sigma-Aldrich	MAK183-1KT
Deposited data		
Cryo-EM structure of human E2p core	Jiang et al., 2018	PDB: 6CT0
Proteomics data	Kastritis et al., 2017	PXD006660
Crosslinking data	Kastritis et al., 2017	PXD006626
Crystallographic structure of human E1p	Seifert et al., 2007	PDB: 2OZL
Cryo-EM structure of human E2o core	10.2210/pdb6H05/pdb	PDB: 6H05
Crystallographic structure of cattle E2b core	Kato et al., 2006)	PDB: 2II3
Crystallographic structure of E3	Szabo et al., 2019	PDB: 6I4R
Crystallographic structure of *A. vinelandii* cubic E2p core	Mattevi et al., 1993	PDB: 1EAA
Model of *C. thermophilum* PDHc E2p core from cryo-EM	This study	PDB: 7BGJ
Cryo-EM structure of native PDHc E2p core (icosahedrally-averaged and asymmetric) and native, asymmetric PDHc	This study	EMDB: EMD-12181
Negative stain reconstructions of PDHc core, Fraction 6, and asymmetric reconstructions from Fractions 5-10	This study	EMDB: EMD-12234
Negative stain reconstructions of hybrid BCKDc/OGDHc core, Fraction 6, and asymmetric reconstructions from Fractions 5-11	This study	EMDB: EMD-12233
Model of *C. thermophilum* PDHc E2p core from negative stain	This study	https://data.sbgrid.org/817
Model of *C. thermophilum* E1p PDHc	This study	https://data.sbgrid.org/817
Model of *C. thermophilum* E2o OGDHc from negative stain	This study	https://data.sbgrid.org/817
Model of *C. thermophilum* E2b BCKDHc from negative stain	This study	https://data.sbgrid.org/817
Model of *C. thermophilum* E3BP trimer from cryo-EM	This study	https://data.sbgrid.org/817) "SBGRID: 817"
Model of *C. thermophilum* E3	This study	https://data.sbgrid.org/817

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Raw cryo- and negative stain micrographs of a native extract from *C. thermophilum* used to reconstruct oxo acid dehydrogenase complexes	This study	EMPIAR: 10625
Experimental models: organisms/strains		
Chaetomium thermophilum var. thermophilum (Freeze-dried)	DSMZ	DSM 1495
Software and algorithms		
3DFSC	Tan et al., 2017	https://3dfsc.salk.edu/
Eman 2.0	Tang et al., 2007	https://blake.bcm.edu/emanwiki/EMAN2
Fiji	Schindelin et al., 2012	https://imagej.net/Fiji
Gctf	Zhang, 2016	https://www.mrc-lmb.cam.ac.uk/kzhang/
Gen5	BioTek Instruments	https://www.biotek.com/products/software-robotics-software/gen5-microplate-reader-and-imager-software/
Haddock	van Zundert et al., 2016	https://alcazar.science.uu.nl/services/HADDOCK2.2/
Image Lab Software 6.1	BIO-RAD	https://www.bio-rad.com/en-us/product/image-lab-software?ID=KRE6P5E8Z
Modeler	Eswar et al., 2006	https://salilab.org/modeller/
MotionCor2	Zheng et al., 2017	https://emcore.ucsf.edu/ucsf-software
Needle	Needleman and Wunsch, 1970	https://www.ebi.ac.uk/Tools/psa/emboss_needle/
Phenix	Afonine et al., 2018	https://www.phenix-online.org
Pymol	Schrödinger, inc	https://pymol.org/2/
Relion 3.0	Zivanov et al., 2018	https://github.com/3dem/relion
Scipion 2.0	de la Rosa-Trevín et al., 2016	https://scipion-em.github.io/docs/
Segger	Pintilie and Chiu, 2012	https://www.cgl.ucsf.edu/chimera/
Sphire-cryolo	Wagner and Raunser, 2020	http://sphire.mpg.de/
UCSF Chimera	Pettersen et al., 2004	https://www.cgl.ucsf.edu/chimera/
UCSF ChimeraX	Goddard et al., 2018	https://www.rbvi.ucsf.edu/chimerax/
UNICORN 7 Workstation for ÄKTA pure, pilot, process, Ready To Process WAVE 25	GE Healthcare Europe GmbH	https://www.cytivalifesciences.com/en/us/shop/chromatography/software/unicorn-7-p-05649
Xmipp	Sorzano et al., 2004	https://github.com/I2PC/xmipp

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jun. Prof. Dr. Panagiotis L. Kastritis (panagiotis.kastritis@bct.uni-halle.de).

Materials availability

All unique/stable reagents generated in this study will be made available on request but we may require a payment and/or a completed Materials Transfer Agreement if there is potential for commercial application.

Data and code availability

Maps derived from cryo-EM and negative stain have been deposited in the Electron Microscopy Data Bank (EMDB: EMD-12811 (for the cryo-EM maps), EMDB: EMD-12234 (for the negative stain maps of PDHc) and EMDB: EMD-12233 (for the negative stain maps of the hybrid BCKDHc/OGDHc)). The coordinates for the determined structure have been deposited in the Protein Data Bank (PDB: 7BGJ). Theoretical models and models from negative stain maps validated by crosslinking have been deposited in the SBGrid

Data Bank (SBGID: 817). Collected movies from negative staining and cryo-EM have been deposited in EMPIAR (EMPIAR: 10625).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Model organism

Chaetomium thermophilum var. *thermophilum* La Touche 1950 (DSM No.: 1495, Type strain) was obtained from DSMZ (Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH) in freeze-dried ampoule and initially cultivated at the Conditions and Media proposed by the company (Medium 188 DSMZ Media List, 45°C). To further optimize the growth a titration of the different conditions was performed including two temperatures (45°C and 50°C), three different growth media for solid cultures (DSM 188, LB and CCM) and two different media for liquid cultures (LB and CCM) (Figure S1). Mycelium from the initial culture was propagated in the following three media, all made up with 1000,00 mL of deionized water and a pH adjusted to 7.1. (a) DSM 188 was prepared with 2.00 g of $NaNO_3$, 0.50 g of $MgSO_4$ x 7 H_2O, 0.01 g of $Fe_2(SO_4)3$ X H_2O, 0.14 g KH_2PO_4, 1.2g K_2HPO_4, 15.00 g agar and 0.02 g yeast extract. (b) LB was prepared with 5.00 g yeast extract, 10.00 g tryptone, 10.00 g NaCl and 10.00 g agar. (c) CCM was prepared with 3.00 g sucrose, 0.50 g NaCl, 0.65 g K_2HPO_4 x 3 H_2O, 0.50 g $MgSO_4$ x 7 H_2O, 0.01 g $Fe_2(SO_4)_3$ x H_2O, 5.00 g tryptone, 1.00 g peptone, 1.00 g yeast extract, 15.00 g dextrin and 15.00 agar. Freshly prepared plates of the three different media were inoculated with mycelium and grown in two different temperature conditions (45°C and 50°C) for eight days (Figure S1). For liquid media of LB and CCM, Agar was excluded from the above recipes. 500 mL Erlenmeyer flasks containing 200 mL of each media were inoculated using small pieces scrapped off from freshly grown mycelium and grown for five days in two different temperature conditions (Figure S1) under shaking at 100 rpm. After optimal conditions for growth were determined, *Chaetomium thermophilum* was grown routinely at 45°C in a rotary shaker at 100 rpm and 10% CO_2. For large scale growth of mycelium, 20 mL of a pre-culture were used for the inoculation of a flask containing 800 mL of CCM media and incubated for 20 hours at 45°C in a rotary shaker at speed 100 rpm and 10% CO_2 supply. Material was either used fresh for subsequent experiments or flash frozen after harvesting in liquid nitrogen and stored at −80°C until further use.

METHOD DETAILS

Mycelium harvesting, culture lysis and sample preparation

Grown mycelium was strained through a metal sieve of 180 μm pore size and cells were washed with PBS at 2200 g 5 min for 3 times at 4°C. Residual moisture was removed and mycelium was freeze-ground using a pre-chilled mortar and stored at −80°C. Approximately 8 g of the mycelium were lysed in 20 mL Lysis buffer (100 mM HEPES pH 7.4, 95 mM NaCl, 5mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 5% Glycerol, 1 mM DTT, 10 μg/mL DNase, Pefabloc 2.5 mM, E-64 40 μM, Bestatin 130 μM, Aprotinin 0.5 μM, Leupeptin 1 μM, Pepstatin A 60 μM) using a Fastprep at 4°C (3 × 6.5 mps shaking speed for 25 s with 3 mins rest on ice). Cell lysate was centrifuged for 5 min at 4000*g* to precipitate the large cell debris and subsequently ultracentrifuged at 100000*g* for 45 mins. Supernatant was filtered through a 0,22 μm syringe filter and further concentrated by spin filtration using a 100000 Da cut-off Amicon Ultra centrifugal filter unit (approx. 20 times). Bradford measurement was performed to determine optimal concentration for further experimental steps. Small proportion (approx. 40 μl) of each step was kept for later kinetic validation and enrichment tracking of the molecules of interest as shown in Figure S3.

Size exclusion chromatography

Three different total protein concentrations of the lysate (22 mg mL^{-1}, 26 mg mL^{-1} and 30 mg mL^{-1} shown in Figure 2A) in a volume of 500 μL were injected in a 500 μL loop and separated by a Biosep SEC-S4000 (7.8 × 600) size exclusion column on an ÄKTA Pure 25M (FPLC) system in three independent experiments (Biological Triplicate). Column was pre-equilibrated using the 200 mM $CH_3COO^-NH^4$ (pH 7,4) Running Buffer, Flow rate was set at 0.15 mL.min^{-1} and fractions of 250 μl were collected (Figure S1D).

Activity assays

Kinetic activity assay kits for both Pyruvate Dehydrogenase (Catalog Number MAK183) and α-ketoglutarate Dehydrogenase (Catalog Number MAK189) were obtained from Sigma-Aldrich® and performed in different steps of the experiment (Figure S3) as well as in the fractions of interest (Figures 2F and S3). Both experiments were performed under the instructions of the provider where the activity was determined using a coupled enzyme reaction, which as described in the protocol results in a colorimetric (450nm) product proportional to the enzymatic activity present.

In detail for both Pyruvate Dehydrogenase and α-ketoglutarate Dehydrogenase activity assay of the different steps of the experiment (Figure S3). 2 μL of each step were used and tested for the activity. The pellets of step (c) and (e) of Figure S3 were resuspended in 800 μL of Lysis buffer and 2 μL out of this were used per well. 0, 2, 4, 6, 8 and 10 μL of the 1,25 mM (1,25 nmole.μL^{-1}) NADH Standard Solution (provided with the kit) were added (in duplicate) in a 96 well plate in order to generate 0 (blank), 2.5, 5, 7.5, 10 and 12.5 nM per well standards. For the fractions 1-12 the activity per fraction was tested in technical duplicate for each of the experiments of the biological triplicate (I, II and III) shown in Figures 2F and S3. For each of the fraction 2 μL (Sample Volume) out of the 250 μL of the fraction were used. As Negative control one of the Low Molecular Weight Fractions (Fraction 29 mentioned as Fr29) was used in both assays. PC I and PC II contained a Sample Volume of 5 μL and 2 μL of the Positive Control provided within the

kit respectively. Calculations were performed following the instructions of the protocol and only values obtained within the linear range of the NADH Standard curves (Figure S3, bottom right) were used for the overall activity calculations. Highest averaged obtained value of NADH Standard duplicates (both for PDHc and OGDHc) were used to plot a Standard Curve and out of this curve the equation for the NADH generated according to the absorbance was excluded. Correction for the background by subtracting the final measurement [(A$_{450}$)$_{final}$] of the standards and samples was performed according to the instructions of the assays. The change in absorbance measurement from T$_{initial}$ to T$_{final}$ for the samples was calculated (ΔA_{450} = (A$_{450}$)$_{final}$ - (A$_{450}$)$_{initial}$). Using the corrected measurements, the generated NADH (nmole/min/mL) was calculated by the following equation:

$$nM \, min^{-1} mL^{-1} \left(milliunits \, mL^{-1} \right) = \frac{S_a}{(T_{final} - T_{initial}) \, SV}$$

, where S$_a$ is the difference of the amount of NADH generated according to the equation between the initial and final time point per well, and SV is sample volume in mL.

Transmission electron microscopy of negatively stained samples and data collection

Monolithic carbon film on 300 mesh copper grids were used for sample preparation. Grids were glow discharged using a PELCO easiGlow, 15mA, grid negative, at 0.4mbar and 25 s glowing time. The samples were negatively stained on the glow discharged grids using 2% (w/w) uranyl acetate (UO$_2$(CH$_3$COO)$_2$·2H$_2$O) with the following procedure: A 3.5 μL volume from each of the fractions of interest (in concentrations of approx. 0.02 or 0.04 mg mL^{-1}) was applied on the grid for approx. 45 s followed by two times distilled water rinsing (washing) step with 15 μL of water. The grids were blotted after each step by ash-free filter papers from same side of the grid. Afterward, the uranyl acetate solution was applied for 1 min and then blotted out using the same method and then let to air-dry slowly under a cover overnight. After negatively stained samples were prepared, the grids were clipped and mounted onto a Thermo Fisher Scientific Glacios 200 kV cryo-transmission electron microscope with an objective lens focal length of 3.4 mm and spherical aberration of 2.7 mm. The beam was aligned to be parallel and perpendicular to the sample having 2.5 μm diameter. Also, a 100 μm objective aperture was used to provide a numerical aperture of about 14.7 mRad. The images were recorded using a Falcon 3EC direct electron detector in linear mode and a total dose of 30 e-/Å^2. Acquisition parameters are shown in Table S2.

Cryo-EM sample preparation and data collection

For the samples that were subjected to cryo-EM analysis, carbon-coated holey support film type R2/1 on 200 mesh copper grids from Quantifoil® were used. The grids were glow discharged in a similar manner for negative stained grids as described above. A 3.5 μL of total protein concentration of 0.3 mg mL^{-1} was applied on each grid and then each grid was plunge-frozen using Vitrobot® Mark IV System (Thermo Fisher Scientific) and standard Vitrobot Filter Paper (Grade 595 ash-free filter paper ø55/20 mm). Conditions in the chamber during the whole process were kept stable at 4°C temperature and 95% humidity. For plunging, blot force 2 and blotting time of 6 s were applied. The vitrified grids were clipped and mounted onto a Thermo Fisher Scientific Glacios 200 kV Cryo-transmission electron microscope under cryo and low humidity conditions. Images were acquired using Falcon 3EC direct electron detector in linear mode and a total dose of 30 e-/Å^2. Similar to the negative stain data acquisitions, the beam was aligned to be parallel and perpendicular to the sample having 2.5 μm diameter and the objective angle was restricted by a 100 μm objective aperture. The acquisition parameters are shown in Table S5.

Image processing of negative stain data and volumetric calculations

The collected micrographs of all negatively stained fractions were processed in Scipion 2.0 image processing framework (de la Rosa-Trevín et al., 2016). The micrographs were collected as single images for each fraction in low dose mode. CTF parameters were estimated with Gctf (Zhang, 2016) and micrographs with low CTF cross-correlation scores were discarded. Particle picking was performed using the combination of XMIPP-particle picking (Sorzano et al., 2004) and SPHIRE-crYOLO protocols (Wagner and Raunser, 2020). The 2D, 3D classifications, and 3D auto-refinement analysis were performed by RELION 3.0 (Zivanov et al., 2018) in Scipion. The picked micrographs were manually inspected to remove obvious errors and picked particles per each fraction were then classified by iterative 2D classification. Initial particle numbers are included in Table S2. Subset of particles per each fraction exhibiting cubic or icosahedral cores were then extracted and used for subsequent 3D classification without imposing symmetry. For 3D classification, 6CT0 and 1EEA models were low pass filtered with EMAN (Tang et al., 2007) and were used as references for the asymmetric reconstruction of PDHc and hybrid OGDHc/BCKDHc, respectively. Resulting subsets of particles for PDHc and hybrid OGDHc/BCKDHc per each fraction from 3D classes exhibiting distinguishable core characteristics were finally subjected to 3D refinement. All runs were performed without imposing symmetry, a spherical mask of 600 Å for PDHc and 450 Å for hybrid OGDHc/BCKDHc. For fraction 5, additional runs for 3D refinement included imposing of I or O symmetry for PDHc or hybrid OGDHc/BCKDHc, correspondingly, applying spherical masks of 300 Å and 250 Å. Resulted unmasked, unprocessed half-maps of PDHc/OGDHc per fraction were further uploaded to the EBI FSC server (https://www.ebi.ac.uk/pdbe/emdb/validation/fsc/) for unambiguous resolution estimation. Reported resolutions were based on an FSC threshold of 0.5 (Figure S5; Table S2).

For volume presentation in Figure 3C, applied density thresholds were: PDHc F5-F10: 0.115, 0.0815, 0.0489, 0.107, 0.0759, 0.0909; OGDHc, F5-F11:0.114, 0.071, 0.02, 0.0957, 0.0201, 0.0932, 0.0472. For volumetric calculations, negative stain reconstructions of PDHc and hybrid OGDHc/BCKDHc from Fraction 6 were thresholded at 0.07 and 0.066 σ values, respectively. Then, the

Segger software (Pintilie and Chiu, 2012) in Chimera was used by default and merging of the segmented volumes of the core was performed. Then, the volume of each segment was measured using Chimera and plotted in a 100% plot, as shown in Figure 4C.

Image processing of cryo-EM data

All image processing was performed in Scipion 2.0. Two datasets were collected from 2 biological replicates (Dataset 1 and Dataset 2) for fraction 6 and another dataset was collected for fraction 4 (Dataset 3). Collected movies of all datasets were corrected for beam-induced motion and drift with the use of MotionCor2 (Zheng et al., 2017), resulting in a dose-weighted, drift-corrected set of summed images. CTF estimation was performed with gctf, using the dose-weighted set, which was then used for the rest of the analysis steps. All other analysis steps were performed using Xmipp3 and RELION 3.0. Global resolutions mentioned for Datasets 1 and 2 were calculated in RELION 3.0 and sphericity was assessed by 3DFSC webserver (Tan et al., 2017).

For Dataset 1, 513 particles were manually picked with Xmipp3, extracted and used for the generation of initial 2D classes that were then used as templates for automated picking with RELION 3.0. The resulting 126,739 particles were subjected to reference-free 2D classification and the best resulting classes were used as templates for a new round of refined auto-picking, leading to a final set of 99,434 particles, that were then again extracted and underwent reference-free 2D classification. All classes that included falsely picked or junk particles were manually inspected and discarded. The remaining 89,263 particles were subjected to reference-based 3D classification into 5 classes without imposing symmetry. PDB ID: 6ct0 (Jiang et al., 2018) was used as a reference after being lowpass-filtered to 40 Å. Once again, particles belonging to junk classes were discarded and the 12,103 particles remaining were again subjected to reference-free 2D classification and selected for quality, ending up with a final set of 9,795 particles that was used for all subsequent 3D refinement steps. Three different refinement procedures were performed. The first reference-based, by imposing Icosahedral symmetry and a mask isolating the PDHc core, which resulted in a final core model of an overall resolution of 7.56 Å (FSC = 0.143). The second refined this data without imposing symmetry, resulting in a core model of 12.03 Å (FSC = 0.143). For the third, an initial model of the complete complex was generated with RELION 3.0 and was used as a template for 3D refinement of the complete PDH complex after low pass filtering, leading to a model with an overall resolution of 34.52 Å (FSC = 0.5).

For Dataset 2, the final particle set of 9,795 particles was reference-free 2D classified and the class averages were then used as templates for automated picking, after low pass filtering, resulting in 105,817 particles that were sorted for miss-picks and junk via 2D reference-free classification as well as 3D template-based classification, leading to a final number of 19,721 particles. These were used for all subsequent 3D refinement steps with the exact same parameters as those of Dataset 1. This time the three refinements lead to a symmetric core model of an overall resolution of 7.10 Å (FSC = 0.143), an asymmetric core model of 8.82 Å and a complex model of 22.69 Å (FSC = 0.5). The highest quality particles of both datasets were finally combined to a set of 29,516 particles and used for 3D refinement steps, once again with the exact same parameters as previously mentioned. This resulted in a final symmetric core model with an overall resolution of 6.90 Å (FSC = 0.143), an asymmetric core model of 8.73 Å (FSC = 0.143) and a complex model of 22.06 Å (FSC = 0.5).

For Dataset 3, 3,098 images were collected, and 41,843 particles were picked from the micrographs. Particles were 2D classified in 256 classes using RELION 3.0. 2D classes with prominent structural features and corresponding number of particles ending up in each of those classes are shown in Figure S2A. In Figure 2B, examples of single particles are shown present in fraction 4, including (from top to bottom): a double membrane structure, possibly stemming from nuclear or mitochondrial membrane extract; higher-order assembly of complexes on a membrane; two circular MDa complexes frequently found next each other; membrane decorated with protein material; filamentous structures; liposomes with encapsulated biological material; higher-order complexes with striking homopentameric architecture.

Sequence alignments, homology modeling, interface energy statistics and model refinement

For identification of sequence identity and similarity, the Needleman-Wunsch algorithm was used (Needleman and Wunsch, 1970). Full sequences, including potential signal peptides, for *H. sapiens* and *C. thermophilum*, deposited in UniProt were used. For the ctBCKDHc E1β protein (UniProt-ID G0SA92) an annotation discrepancy was found based on an erroneous genomic mapping. This error was reported and will be updated in the next UniProt update. For sequence alignment of this protein, amino acids 1 to 413 were used. For the generation of atomic models for the E2 cores of PDHc, OGDHc and BCKDHc, and also for the PDHc E1p and the E3 protein complexes, MODELER (Eswar et al., 2006) was used. VTFM and MD optimization were used as described in the manual (chapter 2.2.2). Template structures used for modeling were 6CT0 for PDHc E2p core (Jiang et al., 2018), 6H05 for OGDHc E2o core (10.2210/pdb6H05/pdb), 2II3 for BCKDH E2b core (Kato et al., 2006), 2OZL for PDHc E1 heterotetramer (Seifert et al., 2007) and 6I4R for the E3 homodimer (Szabo et al., 2019). To ensure structural integrity and optimize interface interactions, the E2 core proteins were modeled as homotrimer, the E1 proteins as heterotetramer and the E3 proteins as homodimer. The dodecahedral or cubic structure of the cores were generated by superimposing the generated trimeric models to the biological assembly of the used template structure.

For calculating the interface properties, the HADDOCK refinement server was used (van Zundert et al., 2016). For the E2 core proteins, two monomers of one trimer (m1, m2) and a third monomer of a neighboring trimer (m3) were isolated. For the intra-trimeric interface calculation, m1 was defined as chain A whereas both others were defined as chain B. For inter-trimeric interface calculation, m1 and m2 were defined as chain A and m3 as chain B. For dimeric interface (α/β) of the E1 heterotetramer, a single α and a β subunit were isolated. For the heteromeric interface ($\alpha\beta/\alpha\beta$), each one α and a β subunit were combined to a single chain. For the

homodimeric E3 complex, no modification was needed. For interface comparisons to E3BP trimer, unique interfaces were extracted from the E2p/E3BP core and subsequently refined in HADDOCK.

For the refinement of the initial model for the *ct*PDHc E2p core in the cryoEM density, the complete core model and the map were superimposed using ChimeraX (Goddard et al., 2018) and refined in real space with PHENIX (Afonine et al., 2018) using standard settings. Regions of the model that were not covered by densities were removed, i.e., the *N-ter* extension after Pro252. Prior to deposition, the core structure was condensed to a single chain using the PHENIX tool "*Find NCS operators.*"

Localization, fitting and refinement of E3BP in the asymmetric C1 core reconstruction of PDHc

To identify the folded region of E3BP, the protein sequence was analyzed using InterPro (Hunter et al., 2009). The overall domain architecture of E3BP is highly similar to E2p: A single *N-ter* LD, a peripheral subunit binding domain, and an ordered C-terminal region, each connected by a disordered region. The ordered C-terminal region (residues 270–442) was considered the core region and subsequently analyzed by HHpred (Zimmermann et al., 2018). One suitable template structure (PDB-ID 2II3) (Kato et al., 2006) was identified that covers parts of the *C-ter* and *N-ter* core region, and the *C. thermophilum* structure was produced, using MODELER by default (Eswar et al., 2006). To generate the trimeric structure, one out of the four trimeric densities observed inside the PDHc core was isolated, both from the C1 and the I reconstruction. The C1 density was then rotated by 120° to capture the underlying C3 symmetry of the inside trimeric protein complex and aligned to the symmetrized icosahedral density, followed by an unsupervised fitting of the initial model. All these geometric operations were done using UCSF Chimera (Pettersen et al., 2004). The derived C3 symmetrized trimeric structure was truncated at the C-terminal by 7 residues due to clashes. Then, structure was real-space refined in the C1 trimeric density with PHENIX (version 1.17) with additional morphing and simulated annealing by default.

Antibody production

Production of specific antibodies against PDHc E1α-His-Tag (aa 37-411) and E2p-His-Tag (aa 29-459) were commissioned by GenScript (GenScript USA Inc., NJ). In short, the codon optimized sequence of the protein-region was cloned into a pET-30a(+) vector in frame with a His-Tag. Protein expression was done in 1 L TB media and proteins were obtained from supernatant of cell lysate by a two-step purification (Ni-NTA followed by Superdex 200 size-exclusion chromatography). Proteins were stored in PBS, 10% Glycerol pH 7.4 (E2p-His-Tag) or PBS, 10% Glycerol, 0.2% SDS, pH 7.4 (E1α-His-Tag). Each rabbit (New Zealand rabbit) was immunized and affinity-purified antibody was obtained from serum. Antibodies and recombinant proteins were shipped.

Densitometric protein quantification

Recombinant protein was quantified by Bradford. For protein quantification, same volume of sample and defined amounts of recombinant protein were analyzed by Western-Blotting. Specifically, purified *C. thermophilum* E2p-His-Tag (aa 29-459) and E1α-His-Tag (aa 37-411) were used as positive control (PC). Signal intensity was quantified using Image Lab 6.1 (Bio-Rad Laboratories, Inc.). For direct comparison, derived concentration (ng/μl) was normalized by molecular weight. Marker used in immunodetection assays was Bio-Rad Precision Plus Protein All Blue Prestained Protein Standards.

Fit of E1p, E2p and E3 proteins in the PDHc C1 reconstruction and measured active site distances

To unambiguously localize the E1p, E2p and E3 proteins within the PDHc, we generated an ensemble of 10,000 low-pass filtered maps of each at 20 Å and placed them inside the core of the PDHc reconstruction. Then, we used the Chimera Fitmap function (Pettersen et al., 2004) and the PDHc asymmetric map at 0.02 sigma with a Gaussian filter of 2.5 σ. Using the Chimera metric "correlation" as readout, distribution of the hits classified as being part of the core, localized inside the core, or at external densities was performed. Among the top hits with high volume coverage, were C2-symmetricically oriented positions of the E1p and the E3 at the large density cluster showed in Figure 7. We manually examined all other hits and found that the best competing hits had considerably lower scores, caused obvious clashes and were non-plausible, e.g., mapped on, or completely inside the E2p core, or were fitted in regions of external density without sufficient density coverage.

For the calculation of active site distances, the active sites were defined by residues Histidine432 of each E2p, Glutamate81 of each E1β and Cystein78 of each E3 monomer. The distances between these active site residues are the Euclidian distance of their Cα atoms.

QUANTIFICATION AND STATISTICAL ANALYSIS

For the kinetic experiments of PDHc and OGDHc, values for the enzyme kinetics measurements in this paper were measured in technical duplicates and performed in biological triplicate. For counting the localized densities in the negative stain class averages of PDHc and hybrid OGDHc/BCKDHc, class averages were provided to 5 lab members. These class averages (55 for PDHc and 36 for the OGDHc/BCKDHc hybrid) were generated by Scipion 2.0 using the default threshold and shown in Table S4. Each lab member was independently asked to count the external densities, with a conservative and a non-conservative threshold. All 5 lab members performed the task independently. Results were collected and statistically treated to derive distribution properties of

the blind measurements (Figure 4A; Table S4). For validating the cryo-EM reconstructions of the core and the complex of PDHc, the structure calculations were performed in data coming from two independent biological replicates and then merged to report the final maps. Mass spectrometry abundance data were also performed in biological triplicates (Kastritis et al., 2017). For calculating distances of active sites in the final E1p-E2p-E3BP-E3 model, all distances were measures among all subunit active sites and their average and standard deviation was considered.

Supplemental Information

Integrative structure of a 10-megadalton

eukaryotic pyruvate dehydrogenase complex

from native cell extracts

Fotis L. Kyrilis, Dmitry A. Semchonok, Ioannis Skalidis, Christian Tüting, Farzad Hamdi, Francis J. O'Reilly, Juri Rappsilber, and Panagiotis L. Kastritis

Figure S1 (related to **Fig 2** and **Fig 3**). *C. thermophilum* as a model system for cell extract fractionation.

A *C. thermophilum* **as a model organism for eukaryotic structural studies**

C Bradford protein concentration per fraction in the biological replicates

B *C. thermophilum* **eastablishment of cell culture: CCM medium and 45 °Celcius**

Solid

Liquid

D

Experiment I — Total injected protein: 11mg — UV 280 / UV 320 / UV 260

Experiment II — Total injected protein: 13mg — UV 280 / UV 320 / UV 260

Experiment III — Total injected protein: 15mg — UV 280 / UV 320 / UV 260

zoom in the HMW Fractions
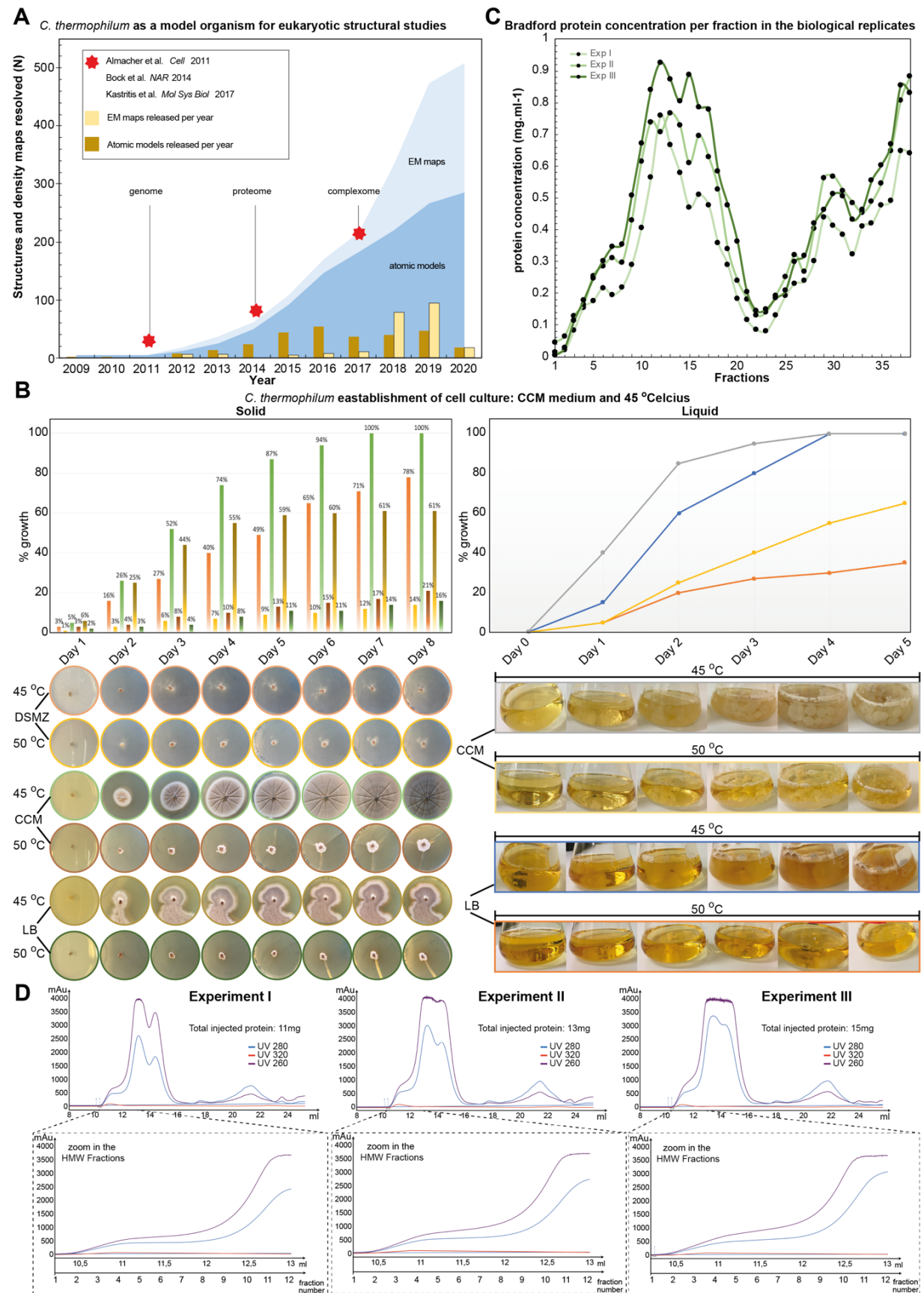
Figure S1 (related to **Fig 2** and **Fig 3**). *C. thermophilum* as a model system for cell extract fractionation. (A) Yearly cumulative increase in deposited EM electron optical density maps (light

blue) and atomic models deposited in PDB (dark blue) from *C. thermophilum*. Bar plots indicate EM maps published per year (light brown) and atomic structures published per year (dark brown). Stars represent publications communicating the genome, proteome and interactome of *C. thermophilum*. (B) Establishment of *C. thermophilum* culture at 45°C. On the left, solid culture is shown; on the right, liquid culture is shown; Plots show relative growth as function of day. Below the plots, Images acquired per day of *C. thermophilum* in different media is shown and a visual comparison of their growth. (C) Determination of protein concentration in the fractions after size exclusion chromatography; The 3 different biological replicates are shown, also reflecting the relative increase in protein concentration as function of injected protein amount of native extract. (D) Chromatography profiles of the 3 experiments and their UV absorbance at 260, 280 and 320 nm wavelengths; Below, zoom ins into the high molecular weight fractions and the amount of aggregation, as monitored by 320 nm.

Figure S2 (related to **Fig 2** and **Fig 3**). Fraction 4 cryo-EM class averages and α-keto acid dehydrogenase complex proteins, their sequence characteristics and mass spectrometry elution profiles of component proteins.

component proteins. (A) 2D class averages retrieved from cryo-EM data collected from fraction 4, where large biomolecular complexes are observed. 3 Groups of class averages are shown (Groups I-III) along with the number of particles in each class (*N*). Group I includes membrane-containing particles, Group II other complexes of striking morphology and Group III possible PDHc class averages. (B) Comparison of human vs *C. thermophilum* E2 polypeptide sequence, highlighting the different domains; Lipo1 and Lipo2: Lipoyl domains; PSBD: peripheral subunit binding domain (binding the E1); Mitochondrion is the signal peptide for mitochondrial localization. (C) Sequence alignment of different domains of the E2 protein from human and *C. thermophilum*. (D) Percentage of sequence similarity, identity and gaps among calculated alignments of *C. thermophilum* and human α-keto acid dehydrogenase complex proteins. (E) Sequence alignment of the peripheral sequence of the E2 catalytic domain and conservation of the sequence between human and *C. thermophilum*. (F) Organization of the trimeric catalytic domain of the E2 protein, where the subunit catalytic core is peripherally surrounded by the structural element described in (E), that of a loop-helix-turn-loop domain. (G-J) Co-elution of protein species identified by mass spectrometry (Kastritis et al. 2017) for all subunits of PDHc (G), OGDHc (H), BCKDHc (I) and their complexes (J), after summation of their respective protein abundances. Relative abundance is reported per fraction, which is the abundance of each of the proteins measured by their iBAQ score, over all identified proteins in each of the fractions (**Table S1** for details).

Figure S3 (related to **Fig 2**). Kinetic assays for PDHc E1p and OGDHc E1o subunits.
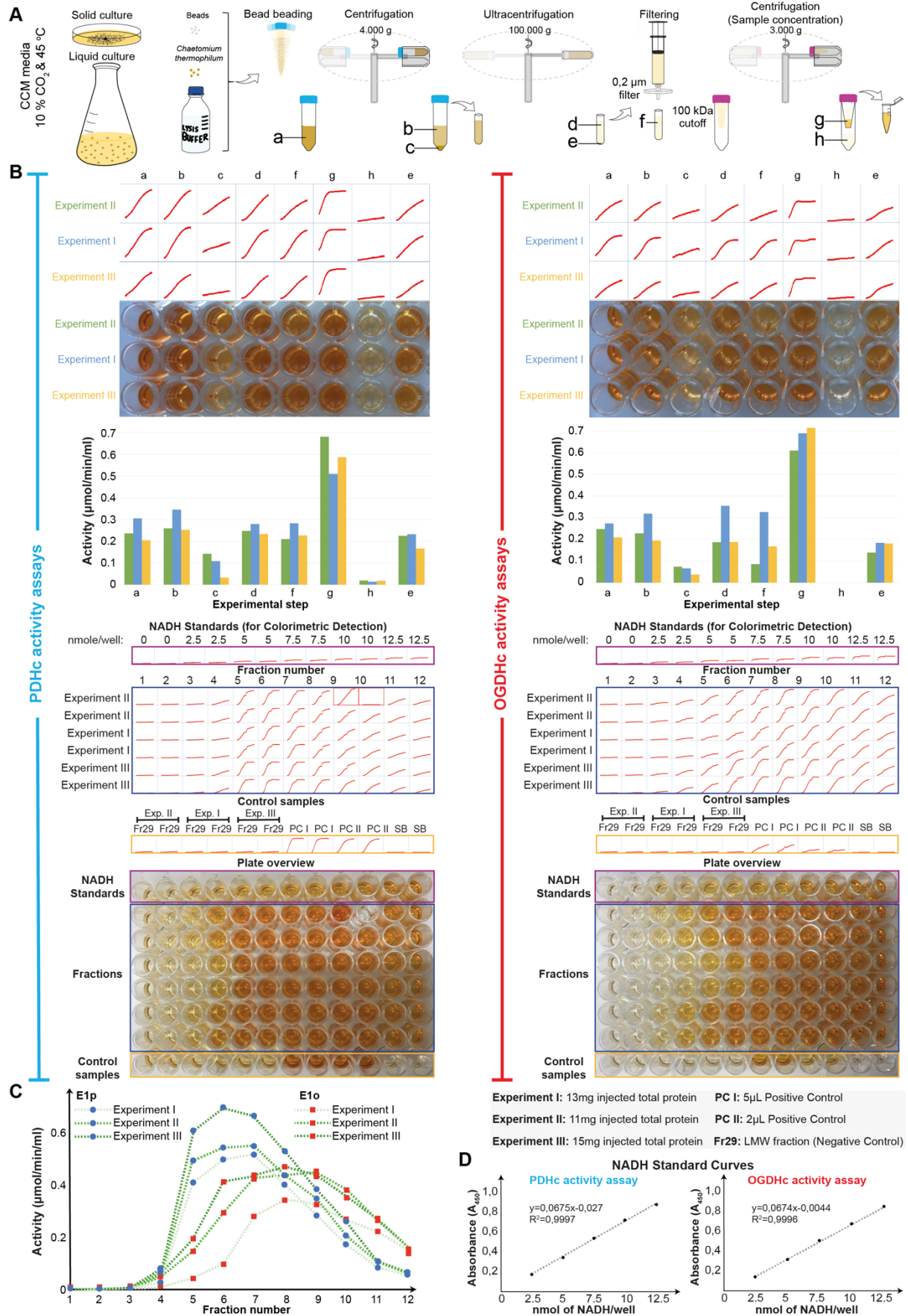
Figure S3 (related to **Fig 2**). Kinetic assays for PDHc E1p and OGDHc E1o subunits. (A) The experiment until the injection of the cell extract is shown with different steps of purification. (a) Lysate after bead homogenization; (b) and (c), lysate supernatant and pellet, respectively after the 4000 g centrifugation; (d) and (e), lysate supernatant and pellet, respectively, after the ultracentrifugation step; (f) sample after filtration and (g) and (h), sample after final concentration and filtering, (g) being the material to be injected for size exclusion chromatography and (h) complexes of <100 kDa. (B) On the left, for the PDHc activity assays, original spectrophotometric curves are shown for all steps (a-e) for all biological replicates, and the coloration of the wells, showing successful reactions; Below the image of the wells, quantification and analysis of the signal corresponds to various activities of E1p at different steps of the biochemical preparation is shown. Below the bar plots, the Standards are shown for the triplicate assays within the fractions (fractions 1-12). Control samples are also shown, and below controls, a picture of the plate with coloration of the corresponding wells is shown. Similar illustration is depicted on the right, for the kinetic assays performed in the fractions for the E1o. SB stands for Sample Blank. (C) On the bottom, the activity of PDHc and OGDHc is shown, calculated in milliunits/ml for every replicate. (D) The linear NADH standard curves are shown, used in the analysis for PDHc and OGDHc activity (see Materials and Methods for details).

Figure S4 (related to **Fig 3** and **Fig 4**). 2D class averages from negative staining data present in Fraction 6 and fitting of derived *C. thermophilum* E2 models of α-keto acid dehydrogenase complexes into 3D negative staining reconstructions.

A

| BCKDHc/ OGDHc (1.2%) | PDHc (22.5%) | PDHc- or/and OGDHc-/BCKDHc-like (4.6%) | Larger biomolecules (7.5%) | FAS (+ binder) (2.6%) | Trimers (2.4%) | Dimers & membranes (7.2%) | Smaller biomolecules (6.0%) | cannot be averaged (46.0%) |

4-step 2D classification

- iteration 1
- iteration 2
- iteration 3
- iteration 4
- cannot be averaged

B Fit of E2b core to C1 NS reconstruction

E2b core — external densities

C Reconstruction of hybrid OGDHc/BCKDHc core from NS data

octahedral hybrid core — 45°

D Fit of the E2o core to C1 NS reconstruction

E2o core — external densities

F Fit of the E2p core in C1 NS reconstruction

external densities — E2p core — 45° — E2p core — minor densities — external densities

E Fit of cubic cores to the hybrid octahedral reconstruction

E2b core — E2o core — 45° — unclear fit — 45°
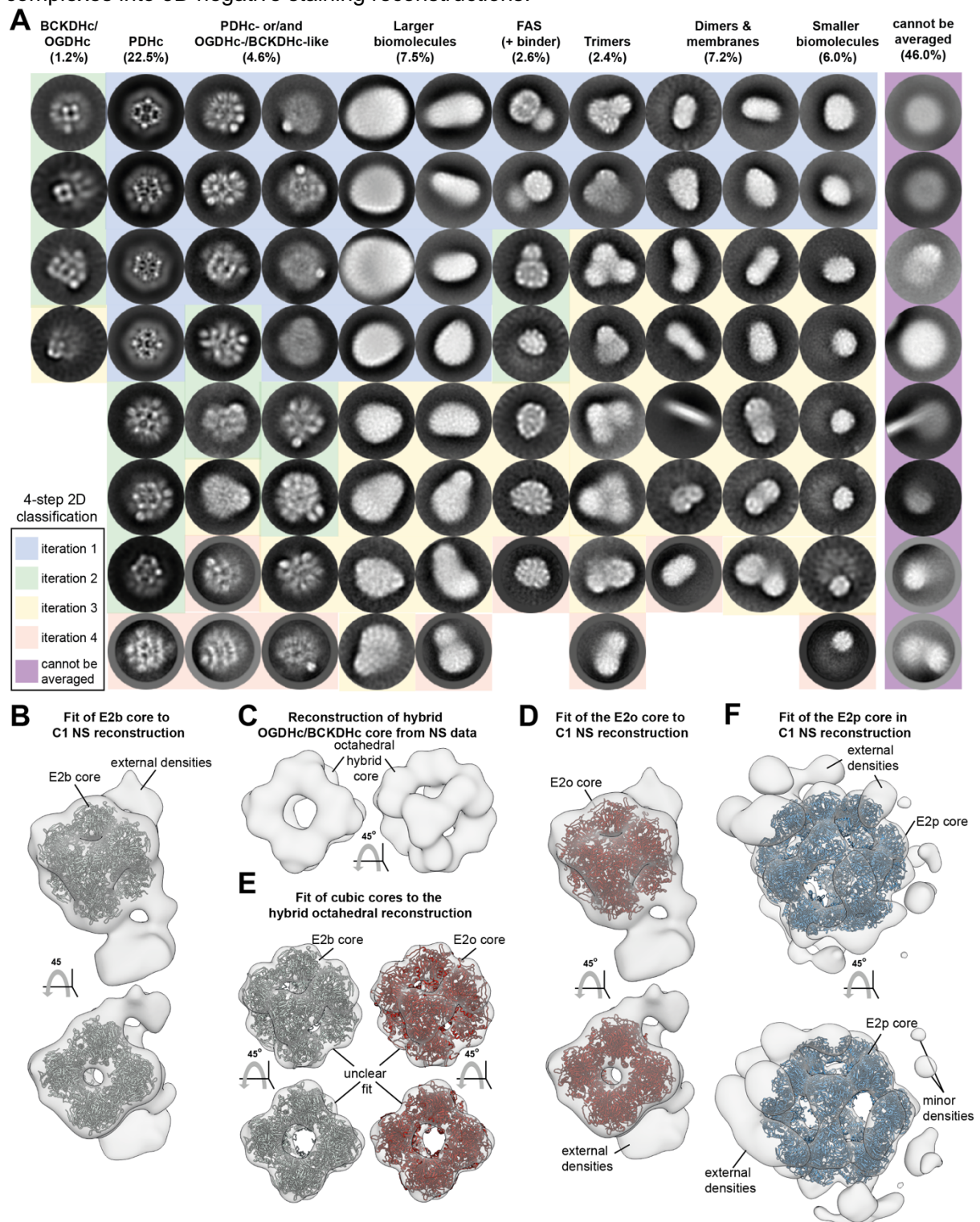
Figure S4 (related to **Fig 3** and **Fig 4**). 2D class averages from negative staining data present in Fraction 6 and fitting of derived *C. thermophilum* E2 models of α-keto acid dehydrogenase complexes into 3D negative staining reconstructions. (A) 2D class averages from the large-scale acquisition of fraction 6. Four iterations of 2D classification are shown and clear structural signatures are presented in each iteration; 46.0% of single-particles could not be averaged after the iterative classification. (B) Fit of BCKDHc E2 core in the reconstructed C1 density of fraction 6; (C) the cubic density of the hybrid OGDHc/BCKDHc reconstruction; (D) Fit of OGDHc E2 core

in the reconstructed C1 density of fraction 6; (E) Fit of cubic cores to the hybrid octahedral reconstruction and satisfaction of both higher-order assemblies; (F) Fit of PDHc E2 core in the reconstructed symmetrized density of fraction 6; Apparent densities for external subunits is highlighted.

Figure S5 (related to **Fig 4** and **Fig 5**). Analysis of crosslinks from previously communicated data (Kastritis et al. 2017) for OADH complexes and reconstructed core density of native PDHc core with the fit of derived molecular models.
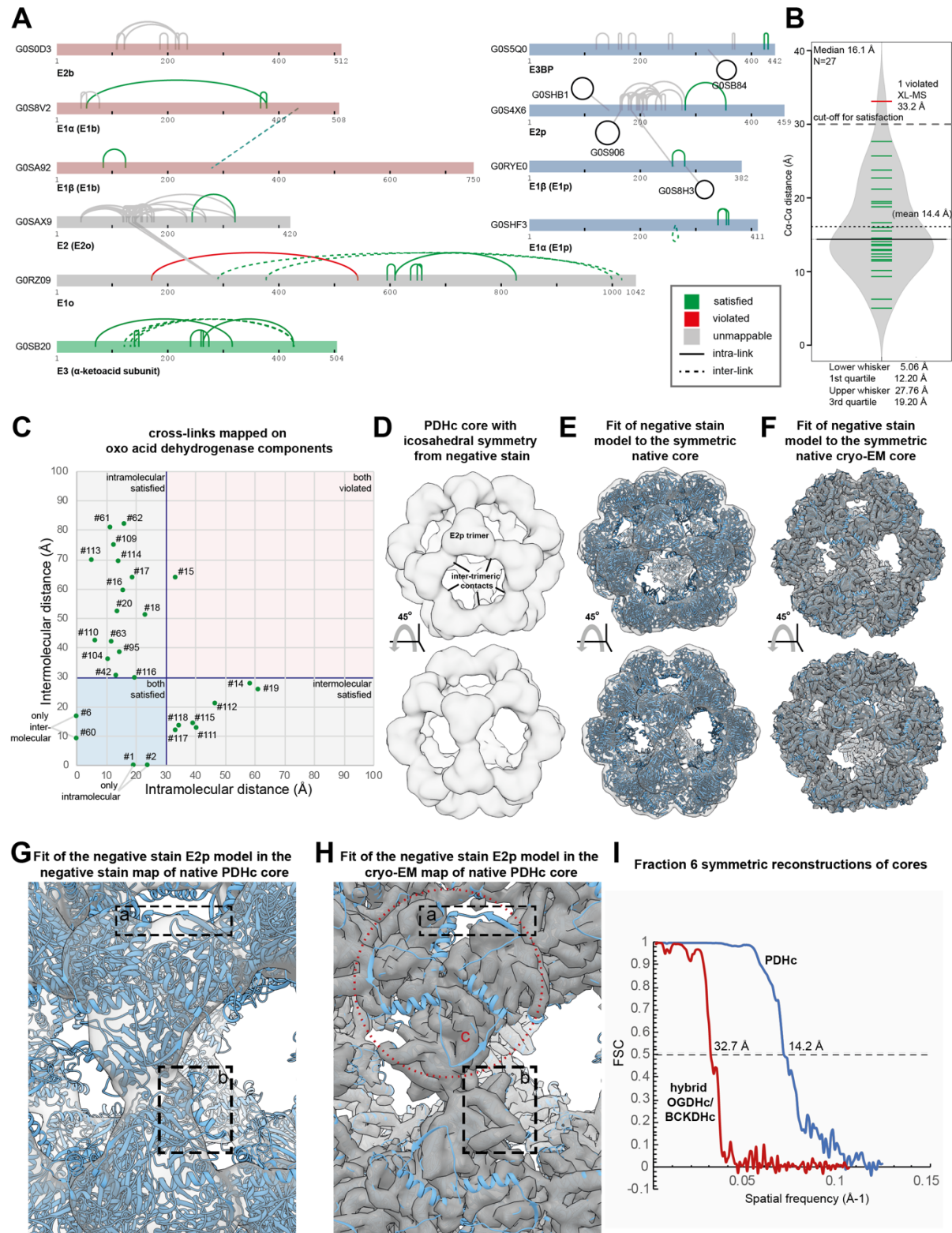
Figure S5 (related to **Fig 4** and **Fig 5**). Analysis of crosslinks from previously communicated data (Kastritis et al. 2017) for OADH complexes and reconstructed core density of native PDHc core with the fit of derived molecular models. (A) mapping of crosslinks on top of α-keto acid dehydrogenase complex protein sequences; (grey: unmappable crosslinks; green: mappable crosslinks, satisfied; red: mappable crosslink, violated). (B) Density distribution plot of all crosslinks mapping on top of the predicted of α-keto acid dehydrogenase complex protein models. (C)

Mapping of inter- and intra- molecular distances of derived crosslinks and satisfaction in either but not both of the states. Satisfied crosslinks can be either intra-molecular or inter-molecular, but results show that none is satisfied in both states: 9 correspond to satisfied inter-XL (18.4±7.3 Å), while 17 to intra-XL (14.5±5.2 Å). (D) Density from negative stain showing features of vertices and faces, as well as the trimer shape, along with distinct densities for the inter-trimeric contacts of the E2 proteins. (E) Predicted E2 model based on the negative stain density and (F) fit of the negative stain E2 atomic model in the cryo-EM map of the E2 core. (G) Zoom into (E), where the fit of the negative stain atomic model in the negative stain map of the E2 core is shown. Box (a) shows missing coverage of densities for the corresponding helix-loop-strand element (see text) and box (b) shows corresponding missing density for the helices connecting the two trimers. (H) Zoom into (F), where, clearly, densities of (a) are still missing, while densities for the helices present in the inter-trimer interface are recovered (b); (c) The N-terminal sequence of the E2 trimer is not covered by the native cryo-EM map of the PDHc core, and therefore, this region is natively flexible. (I) FSC plots for the symmetrized cores of PDHc and hybrid OGDHc/BCKDHc.

Figure S6 (related to **Fig 5, Fig 6 and Fig 7**). Method for modeling, fitting and reconstructing the E3BP trimer in the cryo-EM map, and presentation of associated validation metrics for cryo-EM reconstructions reported in this work.

A Initial core model

B Icosahedral density

C C1 density

E C3 reconstructed model

F Real-space refined model

D 120° rotation series of C1 density

G

Dataset:	1
Nr. of Images:	564
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.143
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.965 (of 1)
Global Resolution:	7.64 Å

Dataset:	2
Nr. of Images:	2029
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.143
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.966 (of 1)
Global Resolution:	6.85 Å

Dataset:	Combined
Nr. of Images:	2593
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.143
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.952 (of 1)
Global Resolution:	6.85 Å

H FSC Plots of Unprocessed Final Map PDHc Icosahedral Core

7.86 Å @ 0.5
7.49 Å @ 0.333
6.90 Å @ 0.143

I

Dataset:	1
Nr. of Images:	564
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.5
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.713 (of 1)
Global Resolution:	10.64 Å

Dataset:	2
Nr. of Images:	2029
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.5
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.705 (of 1)
Global Resolution:	8.64 Å

Dataset:	Combined
Nr. of Images:	2593
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.5
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.791 (of 1)
Global Resolution:	8.20 Å

J FSC Plots of Unprocessed Final Map PDHc C1 Core

11.02 Å @ 0.5
9.56 Å @ 0.333
8.72 Å @ 0.143

K

Dataset:	1
Nr. of Images:	564
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.5
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.869 (of 1)
Global Resolution:	21.49 Å

Dataset:	2
Nr. of Images:	2029
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.5
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.793 (of 1)
Global Resolution:	18.49 Å

Dataset:	Combined
Nr. of Images:	2593
Pixel Size:	3.18 Å/px
Cone Angle:	20.0
FSC Cutoff:	0.5
Sphericity Threshold:	0.5
High-Pass Filter:	150 Å
Sphericity:	0.751 (of 1)
Global Resolution:	15.29 Å

L FSC Plots of Unprocessed Final Map PDHc Complex

22.05 Å @ 0.5
19.36 Å @ 0.333
13.45 Å @ 0.143

Figure S6 (related to **Fig 5, Fig 6 and Fig 7**). Method for modeling, fitting and reconstructing the E3BP trimer in the cryo-EM map, and presentation of associated validation metrics for cryo-EM

reconstructions reported in this work. (A) The derived molecular model of the C-ter of ctE3BP based on structural homology. (B) The densities inside the symmetrized native ctPDHc core are symmetrized; (C) In the PDHc core asymmetric reconstruction (8.7 Å, FSC-0.143), a clear density appears exhibiting C3 symmetry. (D-E) Top fit of the E3BP model in the extracted inner density; C3 rotations of E3BP model superimpose well in the densities, recapitulating the experimentally-derived curvature (F) Real-space refinement of the model in the map generates a clash-free C3 symmetric dimer with a stable, novel interface (G) 3DFSC statistics of the symmetrized PDHc cores, reported sphericity values and derived FSC plots (see materials and methods). (H) FSC plot calculated for the symmetrized PDHc core, corresponding to the different datasets. (I) Same as panel (G), but for the asymmetric reconstruction of the PDHc core. (J) Same as (H), but for the asymmetric reconstruction of the PDHc core. (K) same as panel (G), but for the asymmetric reconstruction of the full PDHc. (L) same as (H), but for the asymmetric reconstruction of the full PDHc, and reported FSC is 0.5 (22.05 Å). Reported resolutions through the manuscript are those retrieved from RELION 3.0 (Zivanov et al., 2018).

Figure S7 (related to **Fig 7**). Structure-based analysis for the dehydrogenase factory model and single-particles of PDHc with higher-order binders, forming protein communities.

A E1p and E2p in cell extract fractions

B
Negative stain reconstruction | cryo-EM reconstruction: Replicate 1 | cryo-EM reconstruction: Replicate 2 | cryo-EM reconstruction: Combined data

⬭ recapitulation of clustered density

C Volumetric calculations for E1p and E3

R² = 0.93

D Fits of E1p EM map (20 Å) in simulated densities

Threshold

- correct fit (E1p in E1p density)
- correct localization (E1p in E1p density)
- wrong localization (E1p in E3 density)

E Fits of E3 EM map (20 Å) in simulated densities

Threshold

- correct fit (E3 in E3 density)
- correct localization (E3 in E3 density)
- wrong localization (E3 in E1p density)

F Fitting analysis of Top 25 E3 and E1p clusters

ambiguous E3 fits | ambiguous E1p fits | E3 fits | E1p fits

- density overlap > 90%
- density overlap > 50%
- density overlap < 50%
- variable density overlap

G
PDHc community with 1 binder (direct interaction) | PDHc community with 2 binders (direct interaction) | PDHc community with 3 binders (direct interaction)

25 nm

PDHc | Binder A | Binder B | Binder C

Figure S7 (related to **Fig 7**). Structure-based analysis for the dehydrogenase factory model and single-particles of P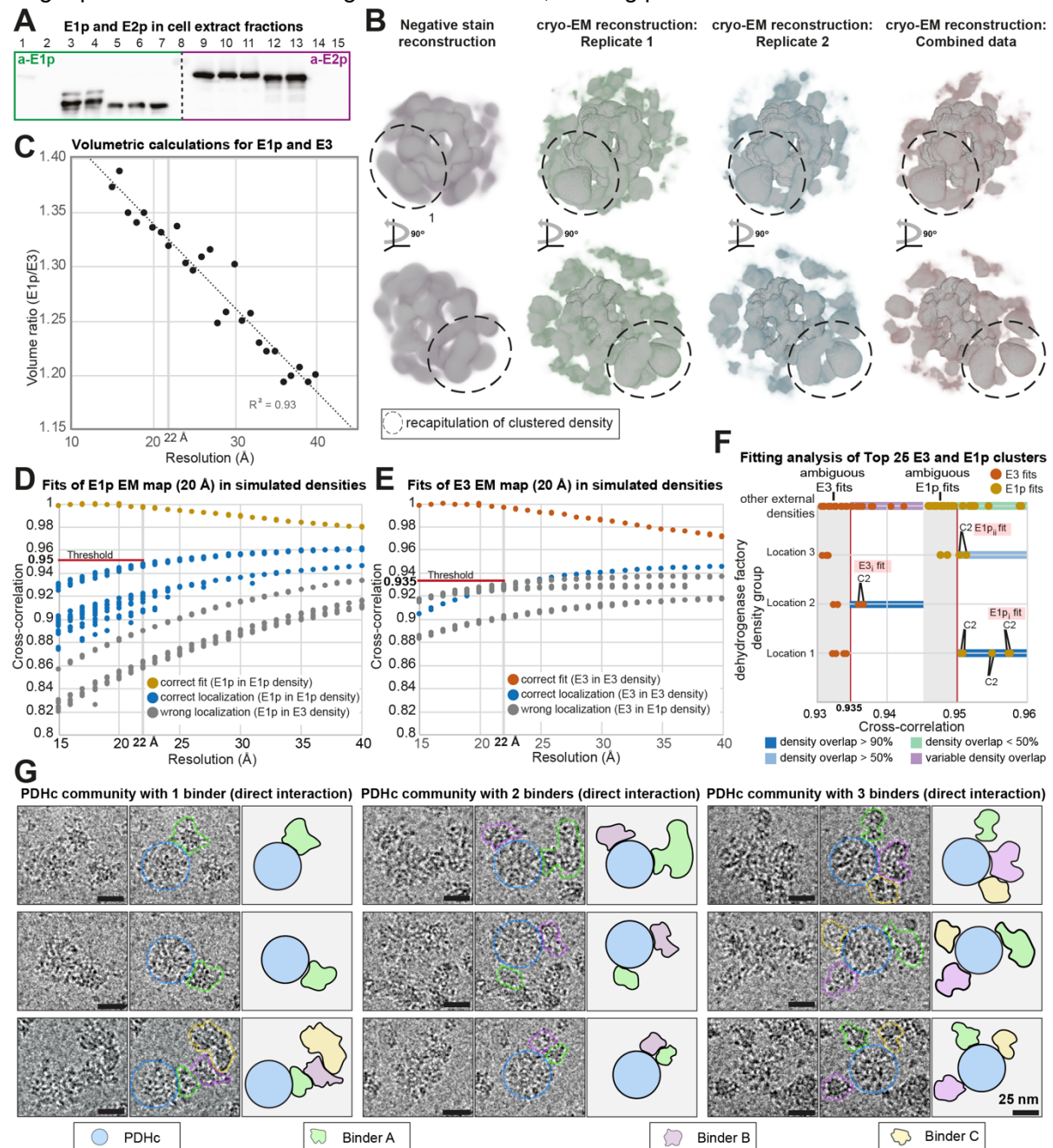DHc with higher-order binders, forming protein communities. (A) Cross-reactivity of Abs against E1p and E2 proteins and constructs is not detected, allowing efficient semi-quantification of band intensity for stoichiometry analysis. Analysis of the immunodetection experiment estimates ~20 copies of E1p for each PDHc complex (see text for details). Lanes: 1: 20.0 ng E2p-His-Tag, 2: Fr 29 (NC), 3-4: 20.0 ng and 15.0 ng E1α-His-Tag, respectively, 5-7: Fr 6 (2 µL), 8: Marker, 9-11: Fr 6 (2 µL), 12-13: 15.0 ng and 20.0 ng E2p-His-Tag, respectively, 14: Fr 29 (Negative control, NC (2 µL), E2p and E1α are not present as shown by MS data), 15: 20.0 ng E1α-His-Tag. (B) Validation using different reconstructions of the asymmetric complex of PDHc; from left to right: reconstruction from negative stain data, the first and the second cryo-dataset from different biological replicates and the final merged dataset; The density of the prominent cluster is recapitulated in all of those and is highlighted. (C) Volumetric calculations of derived densities for E1p and E3 after low-pass filtering at different resolution thresholds (15-40 Å). E1p is consistently larger than E3 in any resolution and is more than 1.3 times larger than E3 at 22 Å resolution

indicating that the 2 molecules can be distinguished. (D) Fits of E1p low-pass filtered EM map (20 Å) used for unsupervised fitting in the PDHc reconstruction shown in **Fig 7**. When systematically low-pass filtering the E1p and the E3 densities at different resolution (15-40 Å), correct fits are still recapitulated at 22 Å, with a CC threshold of 0.95. correct localization denotes that E1p fits in its density but in an incorrect orientation. (E) Same as (D), but applied to E3, and derived threshold at 22 Å for a correct fit is CC=0.935. (F) Analysis of Top 25 E3 and E1p clusters shown in Fig. 7B. Considering the density thresholds for both E1p and E3 fits (CC=0.935 and CC=0.950, respectively), only one of the molecules can be fitted above thresholds in each Location. In addition, density overlaps for each fit above threshold are extensive and C2 symmetric fits are always recapitulated, indicating resolution information is sufficient for discrimination of fits. (G) PDHc is occasionally found in protein communities with one, two, three or more MDa protein complex members when visualizing single particles that ended up in the PDHc reconstructions directly from the micrographs. PDHc may participate in a "pearl-string-like" arrangement, localizing at the edges of the community (bottom left panel). This architecture is reminiscent of other communities that were previously reported, *e.g.* the communities of FAS with its binders (Kastritis et al., 2017). Interactions of other MDa complexes with PDHc provides evidence for another layer of complexity regulating metabolons beyond the identified dehydrogenase factory.

Supplementary Tables

Table S2 (related to **Fig 3).** Data collection and analysis of negatively-stained micrographs from fractions 5-11 along with validation of derived 3D reconstructions.

	Fraction 5	Fraction 6	Fraction 7	Fraction 8	Fraction 9	Fraction 10	Fraction 5	Fraction 6	Fraction 7	Fraction 8	Fraction 9	Fraction 10	Fraction 11
Data collection and processing	PDHc	PDHc	PDHc	PDHc	PDHc	PDHc	PDHc	Hybrid OGDHc/ BCKDHc	Hybrid OGDHc/ BCKDHc	Hybrid OGDHc/ BCKDHc	Hybrid OGDHc/ BCKDHc	Hybrid OGDHc/ BCKDHc	Hybrid OGDHc/ BCKDHc
Magnification	45000X	36000X	45000X	45000X	45000X	45000X	45000X	36000X	45000X	45000X	45000X	45000X	45000X
Voltage (kV)	200	200	200	200	200	200	200	200	200	200	200	200	200
Electron exposure (e–/Å^2)	30	30	30	30	30	30	30	30	30	30	30	30	30
Applied defocus range (µm)	-0.8 to -1.6	-1.2 to -2.0	-0.6 to -1.6	-0.8 to -2.0	-1.0 to -2.0	[-2.0]	-0.8 to -1.6	-1.2 to -2.0	-0.6 to -1.6	-0.8 to -2.0	-1.0 to -2.0	[-2.0]	-1.0 to -2.0
Calculated defocus range (µm)	-0.2 to -4.0	-0.4 to -4.2	-0.2 to -3.5	-0.2 to -2.5	-0.3 to -4.6	-0.3 to -3.7	-0.2 to -4.0	-0.4 to -4.2	-0.2 to -3.5	-0.2 to -2.5	-0.3 to -4.6	-0.3 to -3.7	-0.2 to -4.0
Pixel size (Å)	3.18	4.03	3.18	3.18	3.18	3.18	3.18	4.03	3.18	3.18	3.18	3.18	3.18
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
Initial particle images (no.)	155,306	1,407,677	68,587	136,353	148,979	60,069	155,306	1,407,677	68,587	136,353	148,979	60,069	66,323
Final particle images (no.)	9,270	267,547	6,797	4,360	3,095	1,580	4,620	9,000	3,758	3,293	10,819	6,527	1,632
Map resolution (Å)	30.6	26.6	38.8	39.0	37.9	39.7	40.2	44.6	39.9	41.8	39.2	52.3	58.1
FSC threshold	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

Table S3 (related to **Fig 5**). Haddock energy statistics for derived interfaces of various models of related protein complexes of α-keto acid dehydrogenase complexes.

Enzymatic assembly	Interface	Organism	Structure	van der Waals energy (kcal·mol⁻¹)	electrostatic energy (kcal·mol⁻¹)	desolvation energy (kcal·mol⁻¹)	buried surface area (Å²)
			Paired t-test, p-value	*0.675 (n.s)*	*0.422 (n.s.)*	*0.9516 (n.s.)*	*0.8070 (n.s)*
E1p	α/β-Interface	*C. thermophilum*	Homology model	-521.9 +/-3.8	-1524.7 +/-61.2	-134.4 +/-9.4	12846.6 +/-141.3
		Homo sapiens	2OZL	-573.8 +/-11.0	-1446.2 +/-50.3	-147.4 +/-9.3	13593.7 +/-37.7
	αβ/αβ-Interface	*C. thermophilum*	Homology model	-135.0 +/-1.5	-115.6 +/-23.4	-56.8 +/-3.7	3572.3 +/-40.3
		Homo sapiens	2OZL	-144.5 +/-2.4	-296.8 +/-25.1	-36.2 +/-4.1	3439.6 +/-37.3
E2p core PDHc (negative stain)	Intra-trimeric	*C. thermophilum*	Homology model	-117.7 +/- 2.8	-364.1 +/- 21.9	-40.3 +/- 5.3	3234.6 +/- 34.3
		Homo sapiens	6CT0	-114.7 +/- 3.4	-246.5 +/- 21.9	-50.7 +/- 0.5	3046.8 +/- 54.4
	Inter-trimeric	*C. thermophilum*	Homology model	-48.7 +/- 3.4	-102.5 +/- 12.5	-29.4 +/- 3.8	1575.5 +/- 45.9
		Homo sapiens	6CT0	-43.6 +/- 1.2	-237.3 +/- 18.9	-31.6 +/- 4.9	1627.1 +/- 36.7
E2o core OGDHc (negative stain)	Intra-trimeric	*C. thermophilum*	Homology model	-64.8 +/-3.9	-32.2 +/-12.8	-15.7 +/-6.6	2051.6 +/-56.4
		Homo sapiens	6H05	-53.2 +/-1.8	-84.9 +/-22.4	-16.5 +/-6.4	1869.1 +/-72.9
	Inter-trimeric	*C. thermophilum*	Homology model	-31.3 +/-1.8	-66.1 +/-23.6	-14.9 +/-1.0	1121.7 +/-55.2
		Homo sapiens	6H05	-27.9 +/-1.7	-124.6 +/-16.0	-16.2 +/-2.7	1127.1 +/-59.0
E2b core BCKDHc (negative stain)	Intra-trimeric	*C. thermophilum*	Homology model	-124.9 +/- 1.8	-332.8 +/- 26.6	-42.5 +/- 4.4	3072.2 +/- 51.0
		Bos taurus	2II3	-115.4 +/- 3.6	-358.0 +/- 8.4	-11.3 +/- 1.9	2907.1 +/- 39.2
	Inter-trimeric	*C. thermophilum*	Homology model	-86.3 +/- 9.1	-224.8 +/- 52.4	4.4 +/- 6.5	2340.2 +/- 48.1
		Bos taurus	2II3	-74.8 +/- 2.5	-415.7 +/- 24.9	-6.3 +/- 5.6	2263.8 +/- 47.6
E3	Dimeric	*C. thermophilum*	Homology model	-280.8 +/-5.2	-873.4 +/-26.1	-1.0 +/-4.2	7614.9 +/-51.9
		Homo sapiens	6I4R	-289.6 +/-7.5	-739.0 +/-39.1	-11.5 +/-6.1	7779.7 +/-74.9
E2p core PDHc (cryo-EM)	Intra-trimeric	*C. thermophilum*	This publication	-62.4 +/- 1.1	-204.8 +/- 8.7	-103.5 +/- 3.9	1908.0 +/- 26.0
		Homo sapiens	6CT0*	-54.4 +/- 2.2	-133.3 +/- 4.5	-41.3 +/- 2.2	1633.8 +/- 38.7
	Inter-trimeric	*C. thermophilum*	This publication	-54.1 +/- 3.9	-70.5 +/- 11.3	-94.9 +/- 6.8	1781.5 +/- 49.6
		Homo sapiens	6CT0*	-45.5 +/- 1.4	-256.1 +/- 16.2	-32.7 +/- 4.4	1686.7 +/- 37.8

Table S5 (related to **Fig 5-7).** Cryo-EM data collection for fraction 6. Analysis of cryo-EM data for icosahedral PDHc core, C1 PDHc core and C1 PDHc complex reconstructions and derived model statistics.

	Combined			Dataset 1			Dataset 2		
Data collection and processing									
Magnification				45000X			45000X		
Voltage (kV)				200			200		
Electron exposure (e–/Å^2)				30			30		
Defocus range (µm)				-0.8 to -2.0			-0.8 to -2.0		
Pixel size (Å)				3.18			3.18		
Symmetry imposed (model)	C1 (complex)	I (core)	C1 (core)	C1 (complex)	I (core)	C1 (core)	C1 (complex)	I (core)	C1 (core)
Initial particle images (no.)	-	-	-	99,434	99,434	99,434	105,817	105,817	105,817
Final particle images (no.)	-	-	-	9,795	9,765	9,795	19,721	19,721	19,721
Map resolution (Å)	22.06	6.90	8.73	34.53	7.56	12.03	24.06	7.10	8.82
FSC threshold	0.5	0.143	0.143	0.5	0.143	0.143	0.5	0.143	0.143

Refinement	
Initial model used (PDB code)	6ct0
Model resolution (Å)	6.90
FSC threshold	0.143
Model resolution range (Å)	794.1 – 6.4
Map sharpening *B* factor (Å^2)	-
Model composition Non-hydrogen atoms	94500 (monomer 1575)
Protein residues	12480 (monomer 208)
Ligands	0
B factors (Å^2)	
Protein	103.25
Ligand	-
R.m.s. deviations	
Bond lengths (Å)	0.40
Bond angles (°)	0.62
Clashscore	50.0
Ramachandran favoured (%)	95.1
allowed (%)	4.9
outliers (%)	0.00