Published in final edited form as:

Cell Rep. 2023 June 27; 42(6): 112609. doi:10.1016/j.celrep.2023.112609.

High-resolution structural-omics of human liver enzymes

Chih-Chia Su[1,2], **Meinan Lyu**[1,2], **Zhemin Zhang**[1,2], **Masaru Miyagi**[1], **Wei Huang**[1], **Derek J. Taylor**[1], **Edward W. Yu**[1,3,*]

[1]Department of Pharmacology, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

[2]These authors contributed equally

[3]Lead contact

SUMMARY

We applied raw human liver microsome lysate to a holey carbon grid and used cryo-electron microscopy (cryo-EM) to define its composition. From this sample we identified and simultaneously determined high-resolution structural information for ten unique human liver enzymes involved in diverse cellular processes. Notably, we determined the structure of the endoplasmic bifunctional protein H6PD, where the N- and C-terminal domains independently possess glucose-6-phosphate dehydrogenase and 6-phosphogluconolactonase enzymatic activity, respectively. We also obtained the structure of heterodimeric human GANAB, an ER glycoprotein quality-control machinery that contains a catalytic α subunit and a noncatalytic β subunit. In addition, we observed a decameric peroxidase, PRDX4, which directly contacts a disulfide isomerase-related protein, ERp46. Structural data suggest that several glycosylations, bound endogenous compounds, and ions associate with these human liver enzymes. These results highlight the importance of cryo-EM in facilitating the elucidation of human organ proteomics at the atomic level.

In brief

Su et al. use cryo-electron microscopy to simultaneously identify and solve structures of ten different enzymes or complexes from a raw human liver lysate. Their work highlights the potential of this methodology in facilitating structural-omics of a biological system at the atomic level.

Graphical Abstract:

---

[*]Correspondence: edward.w.yu@case.edu.

AUTHOR CONTRIBUTIONS

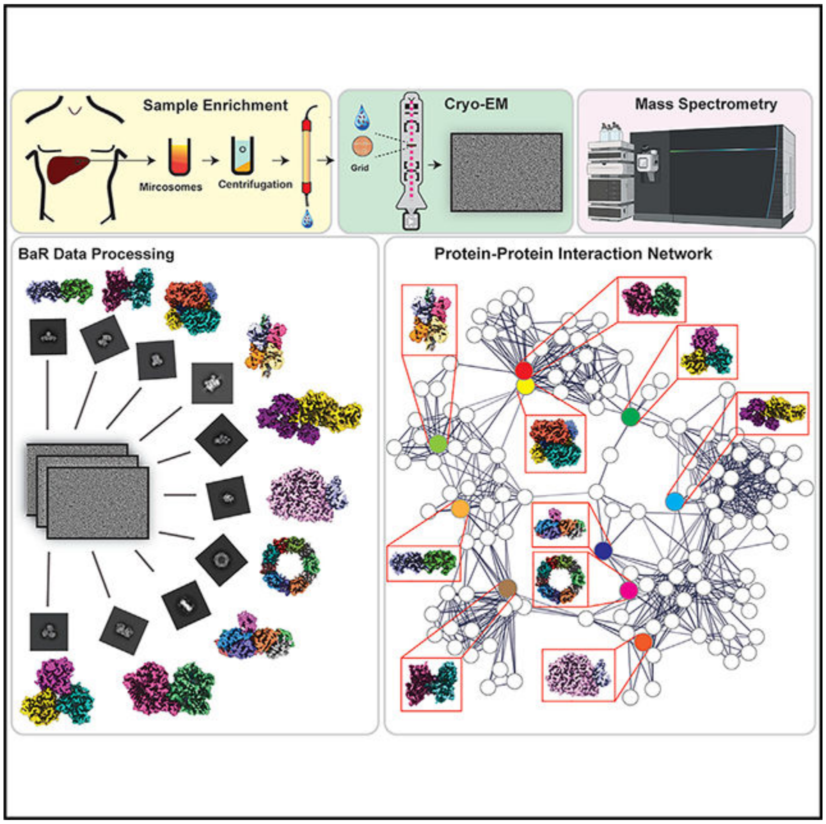
Conceptualization, C.-C.S., M.L., Z.Z., and E.W.Y.; methodology, C.-C.S., M.L., Z.Z., M.M., W.H., D.J.T., and E.W.Y.; investigation, C.-C.S., M.L., Z.Z., M.M., and E.W.Y.; writing – original draft, E.W.Y.; writing – review & editing, C.-C.S., M.L., Z.Z., M.M., W.H., D.J.T., and E.W.Y.; supervision, E.W.Y.; funding acquisition, E.W.Y.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112609.



INTRODUCTION

Systems proteomics is a powerful discipline in biomedical research. Recently, the advancements of mass spectrometry (MS) have opened up an avenue for the field to resolve molecular details of proteome variation in different human tissues and organs. This approach has resulted in the large-scale Human Proteome Project,[1] which has greatly increased our understanding of the localization, expression, and interactions of different proteins within our body. A complement to this effort is the Human Protein Atlas program,[2] which explores the human proteome using genome-wide antibody-based profiling on tissue microarrays. Recently a map of the human tissue proteome,[3] based on an integrated omics approach that involves quantitative transcriptomics at the tissue and organ scale, has been elucidated to spatially localize proteins down to the single-cell level.

This integrated systems approach for studying tissues and organs has been thought to be inaccessible in the fields of X-ray crystallography and cryo-electron microscopy (cryo-EM), where these techniques historically focus on elucidating structures of a particular target biomacromolecule at the atomic level. These structural biology techniques typically require homogeneous and pure samples to obtain useful structural information. Consequently, structural biology tools have had difficulties extending to omics approaches, where specimens possess tens of thousands of proteins and biomacromolecules in a complex, heterogeneous environment.

To address this challenge, we recently developed a "Build and Retrieve" (BaR) methodology.[4] This is an iterative methodology capable of performing *in silico* purification and sorting of images of several different classes of biomacromolecules within a large heterogeneous dataset. BaR is powerful in that it allows us to deconvolute images of a mixture of proteins and produce high-resolution cryo-EM maps for individual proteins from a heterogeneous, multiprotein system. We rationalized that we can adapt BaR for the study of systems structural proteomics, which would allow for the identification and structural determination of proteins and possibly the elucidation of protein-protein interaction networks at atomic resolution.

To develop an approach to use cryo-EM to study human tissues and organs in the context of systems proteomics, we chose to analyze human liver microsomes. The liver is an organ only found in vertebrates. In humans, it is both the largest internal organ and the largest gland. It is responsible for many vital physiological functions, including storage of glycogens and vitamins, removal of toxic compounds, synthesis of plasma proteins, lipid metabolism, and the production of bile, hormones, vitamins, and other biomolecules necessary for digestion and growth.[5] It also helps regulate plasma components. Transcriptome analysis suggests that 72% of all human proteins are produced in the liver, and over 900 of these genes produce a significantly elevated level of protein expression.[3] In addition, genome-scale metabolic model analysis indicates that the human liver carries tissues with the most metabolic activity.[3]

To highlight the potential of the BaR platform in handling human organ samples, we enriched proteins from raw lysate of human liver microsomes using size-exclusion chromatography. We obtained two major peaks with sizes corresponding to 100–250 kDa and 300–650 kDa, respectively. We then individually collected single-particle images of these two human liver enzyme enriched peaks and processed the cryo-EM data using the BaR methodology. From this approach we were able to simultaneously identify and solve cryo-EM structures of ten different human liver enzymes with links to processes involved in carbohydrate storage, vitamin storage, lipid metabolism, protein synthesis, and detoxification.

RESULTS

We separately loaded the two enriched human liver microsome lysate peaks onto holey carbon grids. We then performed cryo-EM data collection for each enzyme peak. Excitingly, our BaR platform, which includes preliminary 2D and 3D classifications, building initial maps, retrieving full particle sets, and final structural refinements, allowed us to simultaneously identify and solve structures of a variety of enzymes found in human livers (Figures S1–S4 and Table S1). A similar approach has been used successfully to simultaneously elucidate structural information for individual proteins and protein-nucleic acid complexes from a heterogeneous mixture of five protein complexes.[6]

Based on the BaR protocol (Figures S2–S4), we were able to solve three new human enzyme or enzyme complex structures. These new structures are the hexose-6-phosphate dehydrogenase (H6PD) enzyme, the heterodimeric enzyme glucosidase II (GANAB) complex, and the peroxiredoxin 4-endoplasmic reticulum protein 46 (PRDX4-ERp46) complex. In addition, we resolved the cryo-EM structure of the microsomal triglyceride transfer protein complex (MTP) and structures of six other human liver enzymes, including aldehyde oxidase 1 (AOX1), carboxylesterase 1 (CES1), glycogen phosphorylase (GP), peroxiredoxin 4 (PRDX4), retinaldehyde dehydrogenase 1 (ALDH1A1), and glutamate dehydrogenase 1 (GLUD1). Liquid chromatography-mass spectrometry (LC-MS) was employed to ascertain the presence of each enzyme in the samples (Table S2). These enzymes and complexes are all involved in very important biological processes such as lipid metabolism and transport, detoxification, carbohydrate metabolism, vitamin storage, and glutamate metabolism and energy homeostasis (Figures S1, S3, and S4).

To envision whether these ten human liver enzymes interact or incorporate with each other, we built an interconnected network based on the STRING database.[7] We observed that all these ten identified enzymes interconnect with each other in the network (Figure S1).

Novel structures of human liver enzymes and complexes

BaR is an untargeted structural biology approach to elucidate structural information of a raw sample. Using this approach, we identified the GDH/6PGL endoplasmic bifunctional enzyme H6PD, which plays a critical role in maintaining blood glucose levels and carbohydrate storage. We also found two human enzyme complexes, GANAB and PRDX4-ERp46, which are engaged in protein synthesis and detoxification, respectively. Here we report the first high-resolution structures of these three enzymes and complexes.

Structure of the hexose-6-phosphate dehydrogenase enzyme—We obtained a total of 194,411 projections of this protein class from our microsomal sample. Using the BaR methodology, we were able to identify this protein as the H6PD enzyme, an enzyme that plays a key role in carbohydrate metabolism. The first structure of H6PD was resolved to a resolution of 3.02 Å (Figures 1A, 1B, and S5; Table S1).

The microsomal H6PD enzyme, also known as the GDH/6PGL endoplasmic bifunctional protein, is a 789-amino-acid protein encoded by the human gene *H6PD*.[5] The N-terminal region (~500 residues) and C-terminal region (~250 residues) of H6PD are homologous to human glucose-6-phosphate dehydrogenase (GDH) and 6-phosphogluconolactonase (6PGL), respectively. It is a luminal bifunctional enzyme of the endoplasmic reticulum (ER), possessing both activities of the cytosolic enzymes GDH and 6PGL.[8,9] H6PD is mostly populated in the liver, although lesser amounts of this enzyme exist in most tissues. H6PD is capable of catalyzing both the first and second steps of the pentose phosphate pathway.[10,11] Despite its importance in the pentose phosphate pathway,[10,11] to date there has been no structural information obtained from this liver enzyme. Our high-resolution data allowed us to build a *de novo* model of this enzyme comprising 500 amino acids (residues 24–55 and 58–525).

This dehydrogenase is dimeric in form (Figures 1A and 1B). Each monomer of H6PD can be divided into an αβ domain (six α helices and six β strands), an all-α-helical domain (six α helices), and an all-β-stranded domain (12 β strands). The extensive dimer interface is created by the β domain, where an extended flexible loop from each subunit of the β domain protrudes into the β domain of the other subunit, interconnecting the two molecules of H6PD and securing dimerization. In particular, residues R236, F337, M357, D363, R437, R509, D512, F513, and E514 of one subunit of H6PD and residues Q375′, Q376′, V380′, Q381′, S382′, E383′, and Q392′ of the adjacent subunit, as well as their counterparts, are engaged in forming the dimer interface.

Post-translational oligosaccharide modifications have been found to play a role in localization, solubility, and stability of many eukaryotic proteins and enzymes. In the H6PD structure, an N-linked glycosylation site was found at residue N157 of each subunit of H6PD (Figure 1B). It is observed that each asparagine is connected to an N-acetylglucosamine moiety (NAG). This structural feature is in good agreement with previous glycoproteomics analysis which showed that residue N157 of H6PD is glycosylated.[12] Our cryo-EM data indeed confirms that this residue is glycosylated and directly linked to NAG.

Structure of the heterodimeric enzyme glucosidase II complex—We collected a total of 129,601 single-particle projections for this class of images. The BaR protocol allowed us to identify this protein as the human GANAB enzyme. We then determined the first structural information of this human enzyme to a resolution of 2.92 Å (Figures 2A, 2B, and S6; Table S1).

Human GANAB is a heterodimeric protein that plays an important role in protein folding and quality control. It is an ER-resident enzyme that contains a catalytic α subunit and a noncatalytic accessory β subunit. GANAB catalyzes the hydrolysis of glucose residues of peptide-bound oligosaccharides and triggers quality-control assessment of glycoprotein folding.[13]

The full-length α subunit of GANAB contains 944 amino acids. Our final cryo-EM structure includes 859 amino acids (residues 33–185, 222–328, and 346–944) for the α subunit of this enzyme. The structure indicates that the α subunit of GANAB can be divided into an N-terminal domain, a catalytic domain, and a C-terminal domain. The N-terminal domain is mostly β-stranded, except residues 45–52, which form an a helix at the N-terminal end. The catalytic domain adopts a fold of the GH31 family[14] that contains a mixture of the α-helical and β-sheet secondary structural elements. The C-terminal domain can be subdivided into proximal and distal C-terminal domains, where these two subdomains create seven-stranded and ten-stranded β barrels, respectively.

The β subunit of GANAB is quite disordered. Of the 528-amino-acid β subunit, we could only include 93 residues (residues 25–117) in our final model, where the secondary structure of these residues mainly presents as an unstructured random loop. Two spherical-shaped extra densities are found to coordinate within the β subunit (Figures 2C and 2D). Notably, the X-ray structure of the mouse homologous enzyme α-glucosidase II[15] indicates that there are two bound Ca^{2+} ions located at the β subunit of this mouse enzyme. We, therefore, assume that these two extra spherical densities arise from two Ca^{2+} ions that are hexacoordinated within this human enzyme (Figure 2D).

Interestingly, *N*-linked glycosylation was observed at residue N97. Based on our cryo-EM density (Figure 2B), this glycan chain is likely in the form of NAG-NAG-BMA-MAN with NAG, BMA, and MAN coded for *N*-acetylglucosamine, β-D-mannose, and α-D-mannose, respectively. This observation is indeed in good agreement with a biochemical study reporting that human GANAB is glycosylated.[16] In addition, a crystal structure of mouse α-glucosidase II (homologous to human GANAB) depicts that this mouse enzyme is glycosylated with a chain NAG-NAG-BMA-MAN at residue N97.[15]

Structure of the peroxiredoxin 4-endoplasmic reticulum protein 46 complex—

We obtained 20,035 single-particle projections for this class of images. Surprisingly, the BaR methodology ascertained that these images originated from the PRDX4-ERp46 (ER protein 46) complex. We then solved the first structure of this complex to a resolution of 3.52 Å (Figures 3A, 3B, and S7; Table S1).

Full-length human PRDX4 and ERp46 contain 271 and 432 amino acids, respectively. Our final cryo-EM structure of this complex includes 186 residues (residues 75–260) for PRDX4 and 106 residues (residues 190–295) for ERp46. Human PRDX4 is a thio-specific peroxidase that catalyzes the reduction of hydrogen peroxide to water or organic hydrogen peroxides to alcohols. It is an antioxidant enzyme belonging to the PRDX family and is the only known secretory form in this family that scavenges reactive oxygen species in both the intracellular and extracellular spaces.[17] As such, PRDX4 is one of the most important defense mechanisms in the human body. The ER protein ERp46 plays a major role in regulating insulin content and may also contribute to glucose toxicity by affecting insulin production.[18]

The cryo-EM structure indicates that the PRDX4 enzyme is decameric in oligomerization (Figures 3A and 3B). This decamer consists of five pairs of PRDX4 dimer, where the five pairs assemble as a donut-shaped structure with the N-terminal and C-terminal ends pointing toward the inside and outside of the ring, respectively. Each monomer of the peroxiredoxin is made up of seven β strands that form a central β-sheet core. The architecture is such that the central β sheet is surrounded by four α helices to constitute the monomer.

Within the PRDX4-ERp46 structure, an ERp46 molecule is found to bind by two protomers of the PRDX4 decamer. ERp46 is positioned above the interface between two PRDX4 protomers, such that the C-terminal tail of the PRDX4 protomers intimately contact the ERp46 protein (Figures 3A and 3B). Specifically, E254 of an individual PRDX4 protomer forms a hydrogen bond with R260 of ERp46 to anchor this protein. The backbone oxygen of E243 also forms a hydrogen bond with the NE2 atom of H219 of the bound ERp46 protein. In addition, A247 and F98′ (where the prime designates residues from the next subunit) of the two PRDX4 protomers contribute to perform hydrophobic interactions with Y262 and W216 of ERp46 to enhance its attachment. Additionally, C245 of the PRDX4 protomer contacts C217 of ERp46 to form a disulfide bond to secure binding (Figure 3C).

Structures of other human liver enzymes and complexes

In addition to the aforementioned novel structures, the BaR methodology enabled us to identify and solve the cryo-EM structures to high resolution of seven other enzymes and complexes. These structures include the MTP complex and the AOX1, CES1, GP, PRDX4, ALDH1A1, and GLUD1 enzymes. As the structures and functions of these enzymes are well documented, we here only briefly describe their cryo-EM structural features. Our work indeed indicates that the BaR methodology is powerful in that it allows us to simultaneously review identities and solve cryo-EM structures of various enzymes from a raw sample. Notably, the BaR platform also enables us to identify bound ions, small molecules, and endogenous compounds as well as post-translational modifications in these human liver enzymes.

Structure of the microsomal triglyceride transfer protein complex—We obtained 249,877 single-particle counts for this class of enzyme in our liver microsome sample. BaR allowed us to identify that these particles correspond to the human MTP enzyme complex. The density for reconstruction of the MTP complex is somewhat anisotropic, which can be attributed to certain preferential orientations of the single-molecule particles in the sample (Figure S8). Nevertheless, we managed to solve the cryo-EM structure of MTP to a resolution of 3.07 Å using the BaR platform (Figures 4 and S8; Table S1).

MTP resides at the ER and is responsible for shuttling triglycerides, cholesteryl esters, and phospholipids to nascent apoB in the lumen of the ER. MTP is a heterodimeric protein that contains an 894-residue MTP α subunit (MTPα) and a 492-residue protein disulfide isomerase β subunit (PDI). Our final cryo-EM structural model contains 781 and 438 amino acids for the MTPα and PDI subunits, respectively.

The structure depicts that MTPα comprises an N-terminal β-stranded domain (residues 18–297), an α-helical domain (residues 298–603), and a C-terminal β-stranded domain (residues 604–894). The PDI subunit within the complex also contains four domains (a: residues 26–131; b: residues 137–232; b′: residues 235–349; a′: residues 369–479), and each domain possesses a mixture of α and β secondary structural elements (Figures 4A and 4B). These domain structures are in good agreement with those of the previously determined X-ray structure of MTP.[19] Superimposition of our cryo-EM structure onto that of the X-ray structure of human MTP[19] (PDB: 6I7S) gives rise to a root-mean-square deviation (RMSD) of 1.20 Å (for 581 Cα atoms).

The interaction between MTPα and PDI is a prerequisite for the function of this protein complex. It is observed that the a domain of PDI intimately coordinates with the α-helical domain of MTPα (Figure 4C), whereas the a′ domain of PDI directly contacts the C-terminal β-stranded domain of MTPα (Figure 4D). In addition, the b′ domain of PDI is found to associate with the α-helical and C-terminal β-stranded domains of MTPα to strengthen the binding between these two subunits of the complex (Figure 4E).

Structure of aldehyde oxidase 1—Within the cryo-EM grid of our microsome sample, we imaged 185,973 single-particle projections for this protein class. We identified that this protein is the human AOX1 detoxification enzyme. We resolved the cryo-EM structure of this enzyme to a resolution of 2.91 Å (Figures 5 and S9; Table S1).

AOX1 is a complex molybdoflavoprotein that belongs a family of the molybdenum-containing enzymes.[20] Each molecule of AOX1 contains 1,338 amino acids. Our final cryo-EM structure includes 1,215 residues per subunit. In this homodimer, each monomer of AOX1 can be divided into three distinct domains consistent with the X-ray structure of the AOX1 enzyme[21] (Figures 5A and 5B). Superimposition of our cryo-EM structure onto the X-ray structure of human AOX1[21] (PDB: 4UHW) leads to an RMSD of 0.79 (for 1,215 Cα atoms).

The X-ray structure of AOX1[21] contains several cofactors, including two [2Fe-2S] iron-sulfur clusters, a flavin adenine dinucleotide (FAD), and a molybdenum cofactor (Moco). Interestingly, our cryo-EM map shows four extra densities within this enzyme. Two of these extra densities are compatible with two [2Fe-2S] clusters. The shapes of the other two densities are similar to those of FAD and Moco. We, therefore, assume that these four cofactors are bound in our enzyme and include them in our final cryo-EM model.

The small N-terminal domain (domain I: residues 4–168) forms the iron-sulfur binding domain, where it houses the two bound [2Fe-2S] clusters (Figure 5C). One of these [2Fe-2S] clusters is anchored by four cysteines (C114, C117, C149, and C151) to secure the binding. Similarly, C44, C49, C52, and C74 are the four cysteines responsible for binding the second [2Fe-2S] cluster (Figure 5D). It appears that domain II (residues 220–537) harbors an FAD. The binding of FAD is extensive. Residues P263, V264, I265, N268, T269, S270, V271, L312, L344, S354, H358, I360, H363, D365, D367, L368, L405, I410, L411, and R429 are involved in stabilizing the binding (Figure 5E). The large C-terminal domain (domain III: residues 562–1314) creates the Moco binding site and accommodates the molybdenum catalytic center, where residues Q113, Q776, F807, R921, M1047, Q1049, V1091, Q1203, and L1268 are engaged in anchoring Moco (Figure 5F). The locations of the two [2Fe-2S], one FAD, and one Moco binding sites are clearly identified based on the cryo-EM densities. These binding site locations are also in good agreement with those observed from the X-ray structure.[21]

Structure of carboxylesterase 1—We obtained 210,285 single-particle counts for this class of protein from our cryo-EM grid. Based on BaR, we were able to identify this enzyme as CES1, also known as monocyte esterase and cholesterol ester hydrolase. We solved the cryo-EM structure of this enzyme to a resolution of 2.67 Å (Figures 6A–6D and S10; Table S1).

CES1 is a trimeric enzyme consisting of 567 amino acids per subunit (Figures 6A and 6B). Our final cryo-EM structure includes 532 amino acids (residues 22–553) per subunit. Each subunit of CES1 is composed of a catalytic domain, an oligomerization domain, and a regulatory domain (Figure 6B). Both the oligomerization and regulatory domains consist entirely of α helices; however, the catalytic domain contains a mixture of eight α-helical and 13 β-stranded secondary structural elements. The overall cryo-EM structure of CES1 is consistent with crystal structures of both rabbit liver and human liver carboxylesterases.[22,23] Superimposition of our cryo-EM structure onto the X-ray structure of human CES1[24] (PDB: 1YAJ) gives rise to an RMSD of 0.39 (for 529 Cα atoms).

Interestingly, an *N*-linked post-translational oligosaccharide modification is observed at residue N79 of CES1 in the cryo-EM structure, where this residue is connected with an NAG-NAG-BMA carbohydrate chain with NAG and BMA corresponding to the *N*-acetylglucosamine and β-D-mannose moieties, respectively (Figure 6C). This observation is in good agreement with a previous mutagenesis study reporting that this human enzyme is glycosylated at position N79.[25]

As expected, the catalytic domain of human liver CES1 possesses the serine hydrolase hydrogen-bonded triad, consisting of residues S221, E354, and H468 (Figure 6D). Interestingly, an extra cryo-EM density with a shape compatible with an ethyl acetate (EA) molecule, a known substrate of CES1, is found within this catalytic domain. The location of this extra density coincides with that of bound EA in the X-ray structure of human CES1.[24] We, therefore, include a bound EA molecule within the catalytic site of our CES1 cryo-EM structure. There are at least 13 residues, including L97, S221, V254, L255, L359, L363, M425, involved in binding this EA substrate (Figure 6D).

Structure of glycogen phosphorylase—We obtained 463,522 single-particle counts for this protein class in our cryo-EM dataset. Using BaR, we were able to identify that this protein is the human liver GP enzyme. We then solved the cryo-EM structure of this enzyme to a resolution of 2.65 Å (Figures 6E–6G and S11; Table S1). Seven hundred and eighty out of the 847 amino acid residues per subunit of this dimeric enzyme were included in our final structural model.

Human liver GP directly influences the blood glucose level via the process of glycogenolysis.[26] This enzyme exists as a homodimer (Figures 6E and 6F). Overall, the cryo-EM structure of human liver GP is in good agreement with the X-ray structures of other GP enzymes.[26,27] Superimposition of our cryo-EM structure onto the X-ray structure of GP[28] (PDB: 2QLL) provides an RMSD of 1.22 (for 767 Cα atoms). A pyridoxal phosphate (PLP) molecule was observed to bind in the substrate binding site of the X-ray structure of GP.[28] Interestingly, our cryo-EM map also indicates an extra density with a shape compatible with PLP located at the PLP binding site of the enzyme. Thus, we include an endogenous PLP ligand in our final cryo-EM structural model.

Each protomer consists of 34 α helices and 16 β strands that can be separated into N-terminal and C-terminal domains. The endogenous PLP molecule is found to anchor at the PLP binding site. There are at least 11 residues, including Y90, G134, G135, R138, W491, K568, K574, Y648, V650, T676, and K680, responsible for binding PLP (Figure 6G).

Structure of peroxiredoxin 4—We collected 28,380 single-particle counts for this class of protein images. The BaR protocol allowed us to identify that this protein is the PRDX4 enzyme. We determined the structure of this enzyme to a resolution of 2.84 Å (Figures 3D, 3E, and S7; Table S1).

Of the 271 amino acids that make up the full-length enzyme, residues 72–245 of PRDX4 were included in the final structural model. In agreement with the X-ray structure of PRDX4,[29] the cryo-EM structure of PRDX4 indicates that this enzyme is decameric in oligomerization. Superimposition of this cryo-EM structure onto the X-ray structure of human PRDX4[29] (PDB: 3TJB) leads to an RMSD of 0.45 Å (for 170 Cα atoms). Our cryo-EM structure indicates that the donut-shaped structure of the PRDX4 decamer is almost identical to that of the PRDX4-ERp46 complex. However, there is a drastic conformational difference between the C-terminal residues of PRDX4 and those of PRDX4-ERp46. In the apo form of PRDX4, residues 246–265 are disordered and not included in the final structural model. However, in the PRDX4-ERp46 structure, we observe that the C-terminal tail of the PRDX4 protomer shifts toward the bound ERp46 and directly interacts with this protein.

Structure of retinaldehyde dehydrogenase 1—We collected a total of 138,661 single-particle cryo-EM projections for this class of protein images. The BaR methodology allowed us to construct a high-resolution cryo-EM map from these projections, and subsequently we were able to identify this protein as the ALDH1A1 enzyme. We then resolved the structure of this enzyme to a resolution of 2.64 Å (Figures 7A–7D and S12; Table S1).

The ALDH1A1 cytosolic enzyme plays a key role in the biosynthesis of retinoic acid. Each molecule of ALDH1A1 contains 501 amino acids and assembles as a tetramer (Figures 7A and 7B). We built a *de novo* model of each subunit of this enzyme consisting of residues 9–501.

Like other crystal structures of ALDHs,[30,31] each monomer of ALDH1A1 comprises a coenzyme binding domain, a catalytic domain, and an oligomerization domain. Superimposition of our cryo-EM structure onto the X-ray structure of human ALDH1A1[20] (PDB: 4WB9) allows us to obtain an RMSD of 0.53 Å (for 487 Cα atoms). A bound nicotinamide adenine dinucleotide (NAD^+) cofactor was found within the X-ray structure of each subunit of the ALDH1A1 enzyme.[20] The high-quality density of our cryo-EM map also unambiguously depicts an extra density located at the coenzyme binding site of each subunit of ALDH1A1, where the shape of this extra density fits well with a NAD^+ cofactor (Figure 7C). We therefore include this cofactor in our final cryo-EM structure. Within the NAD^+ binding site there are at least 17 residues, including I166, P168, W169, K193, E196, Q197, P227, F244, A231, S234, V250, S247, C303, Q350, K353, E400, and F402, involved in binding NAD^+ (Figure 7D).

Structure of glutamate dehydrogenase 1—We obtained 10,295 cryo-EM single particles for this class of protein. The BaR methodology led us to identify that this protein is GLUD1. We then solved its cryo-EM structure to a resolution of 3.24 Å (Figures 7E–7G and S13; Table S1).

GLUD1 plays a key role in nitrogen and glutamate metabolism and energy homeostasis. Human liver GLUD1 is a 558-aminoacid enzyme, where residues 63–556 were included in our final cryo-EM structure. Like the existing X-ray structures of bovine and human GLUD1,[32,33] the cryo-EM structure of this liver GLUD1 enzyme is hexameric in form (Figures 7E and 7F). The six GLUD1 molecules are packed to form a dimer of trimers. Each subunit consists of the N-terminal glutamate binding domain (residues 63–266), the NAD^+ binding domain (residues 267–488 and 528–556), and the antenna domain (residues

449–501) (Figure 7G). In addition, the C-terminal end of this enzyme forms a pivot helix (residues 502–527). Superimposition of the cryo-EM structure onto the X-ray structure of human apo-GLUD1[33] (PDB: 1L1F) gives an RMSD of 1.00 Å (for 444 Cα atoms), indicating that these two structures are nearly identical. As with the X-ray structure of human apo-GLUD1[33] we did not observe any bound ligands, suggesting that our cryo-EM structure should represent the apo form of this enzyme.

DISCUSSION

Cellular function is a very complex phenomenon that is orchestrated by a network of interactions between different proteins and biomacromolecules. These biomacromolecules often further coordinate with ions, small molecules, metabolites, and endogenous compounds *in situ*. Therefore, a deeper understanding of cellular function and organization requires relatively thorough models of interactome networks, particularly at the proteome level. To appreciate the functional complexity of living systems, such as tissues and organs, the approach of systems biology is preferable because it can empower us to gain a better comprehensive understanding of complex biological processes.

The rapid technological advancements of single-particle cryo-EM have opened up a new avenue for 3D structural determination of biomacromolecules at high resolution.[34] It is now possible to obtain structural information of samples below 100 kDa using this technique.[35] A few recent studies have shown that it is feasible to apply cryo-EM to elucidate biomacromolecular structural information from native cell extracts to high resolution.[4,36,37] However, solving multiple protein structures directly from a raw, heterogeneous sample remains a difficult task.

We previously developed a bottom-up systems structural proteomic BaR approach to identify and obtain high-resolution structural models of multiple proteins from a raw sample.[4] We demonstrated that the BaR methodology makes it possible to construct high-resolution cryo-EM maps, which in turn allow us to solve structures of a number of relatively small (<100 kDa) and less abundant (<10%) unidentified proteins within a single heterogeneous sample. Importantly, these results highlight the potential of cryo-EM for systems structural proteomics when coupling with the complementary technique of MS. In the present work, we apply BaR to study human liver microsomes in order to further establish an approach of using BaR to elucidate the proteome of human tissues and organs to high resolution. Our data indeed show that the BaR methodology is capable of enabling identification and high-resolution structural determination of ten different proteins and complexes simultaneously from a raw human organ sample to resolutions between 2.64 Å and 3.52 Å. Of the ten solved structures, seven have been previously determined from recombinant proteins using X-ray crystallography. All seven crystal structures are in good agreement with their corresponding cryo-EM structures resolved from this work. Additionally, it is observed that these ten enzymes are interconnected and participate in a range of interactions based on the interaction network (Figure S1).

Our work provides strong corroboration that BaR can empower us to solve unknown structures of enzymes and complexes. We were able to identify the presence of human

H6PD, GANAB, and PRDX-ERp46 and solve the first structures of these enzymes and complexes from a raw sample based on the BaR protocol.

Notably, the BaR methodology can also be used to identify endogenous ligands and/or native ions. This is evidenced by the cases of GP, ALDH1A1 and CES1, where PLP, NAD^+, and EA were found to anchor in these enzymes, respectively. In addition, it is worth mentioning that four endogenous ligands, including two [2Fe-2S], one FAD, and one Moco, were observed to simultaneously bind in the AOX1 protein. We were also able to identify a bound native Ca^{2+} ion in the GANAB protein.

Besides the identification of bound ligands and ions, BaR can be used to elucidate structures of post-translational modifications and glycosylations of enzymes based on the cryo-EM maps. Interestingly, we noticed from our work that the human liver enzymes H6PD, GANAB, and CES1 are glycosylated.

Limitations of the study

Even though BaR is capable of simultaneously solving structures of different proteins from a raw sample, presently there are several limitations in this methodology. It is difficult to identify a particular protein if the population of the protein is low (<5%) in the heterogeneous sample. In addition, it is not easy to unambiguously identify a protein if the resolution of the map is not high enough (better than 4-Å resolution). Furthermore, the preferential orientation of protein images presents a potential problem, as BaR may not be able to offer a solution for a protein class with insufficient orientation views. This problem may limit the quality of the initial model and create difficulties during initial retrieval, although we may be able to overcome this shortcoming through multiple iterations of the BaR protocol. In addition, it has been observed that we could overcome the orientation bias problem by introducing additives to modify grid properties during cryo-EM sample preparation.[38] We may also be able to resolve the problem of preferential orientation simply by collecting tilted cryo-EM datasets.[39]

Based on our experimental results, it appears that there is no strong correlation between the population ranking of proteins via MS and single-particle counts of these proteins based on the BaR protocol. This is quite surprising, as both BaR and LC-tandem MS focus on analyzing the population of proteins. This discrepancy could be attributed to additional steps needed to prepare the cryo-EM grid for sample imaging. For instance, the procedures for preparing a cryo-EM grid include a blotting process, whereby the filter paper may have different absorptivity to different proteins. A large population of proteins may be lost from the grid due to this blotting process, which may significantly alter the composition of protein components and their relative populations. In addition, denaturation and degradation of protein particles during grid freezing may partially account for the discrepancy. To narrow the difference between the experimental results of BaR and proteomics, it is worthwhile to explore the possibility of using the method of blotting-free vitrification[40–42] for cryo-EM grid preparation. This approach may eventually help improve the agreement between BaR and proteomics analysis.

Although the BaR platform is designed for using cryo-EM to study tissue and organ samples in the context of systems proteomics, it still has a long distance to go in order to allow for simultaneous identification and determination of structures of hundreds of proteins in a single sample. Regardless, our work strongly indicates that we can use BaR effectively in many cases of sample impurity and heterogeneity and utilize the cryo-EM structural approach to simultaneously solve structures of a variety of enzymes from a human organ sample at high resolution. BaR can complement affinity purification coupled with MS[43] and proximity-dependent biotinylation identification (BioID)[44] to unravel protein-protein interaction networks that mediate a variety of essential biological processes. Coupled with proteomics and interactomics, including those techniques based on affinity purification and co-fractionation, the BaR methodology should help illuminate the details of biological networks at high resolution. It is expected that in the future, cryo-EM will enable a new perspective on the elucidation of the human proteome and interactome at the atomic level.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Edward W. Yu (edward.w.yu@case.edu).

Materials availability—This study did not generate unique reagents.

Data and code availability

- Coordinates and EM maps can be found at the following PDB and EMCB accession numbers: 8EM2 (PDB) and EMD-28232 (EMBD) for H6PD; 8EMR (PDB) and EMD-28262 (EMDB) for GANAB; 8EKY (PDB) and EMD-28217 (EMDB) for PRDX4-ERp46; 8EOJ (PDB) and EMD-28377 (EMDB) for MTP; 8EOR(PDB) and EMD-28465 (EMDB) for CES1; 8EMT(PDB) and EMD-28264 (EMDB) for AOX1; 8EMS (PDB) and EMD-28263 (EMDB) for GP; 8ENE (PDB) and EMD-28271 (EMDB) for ALDH1A1; 8EKW (PDB) and EMD-28214 (EMDB) for PRDX4 and 7UZM (PDB) and EMD-26915 (EMID) for GLUD1. The raw cryo-EM data have been deposited in EMPIAR (https://www.ebi.ac.uk/empiar/). The raw mass spectrometry proteomics data has been deposited in ProteomeXchange via the PRIDE database (https://www.ebi.ac.uk/pride/). Accession numbers, codes and links to access the structural data, raw cryo-EM images and raw mass spectrometry proteomics data are also listed in the key resources table and are publicly available as of the date of publication. All data reported in this paper are available upon request from the lead contact.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human liver microsomes were purchased from SEKISUI XenoTech (Kansas City, KS) (catalog number H0610). There were 50 donors in the sample group, consisting of 30 males and 20 females with ages ranging between 5 and 83 years old. The ethnic profile of this group was Caucasian (44), African American (1), Asian (1) and Hispanic (4).

METHOD DETAILS

Human liver microsome lysate—The human liver microsomes were resuspended in buffer containing 20 mM Tris-HCl (pH 7.5), 100 mM NaCl and 5 mM Na-cholate. Insoluble material from the microsome lysate was removed by ultracentrifugation at 20,000 × g. The extracted lysate, containing a variety of unknown enzymes, was enriched using a Superdex 200 column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 7.5, and 100 mM NaCl.

Electron microscopy sample preparation and data collection—The human liver microsome lysate sample was concentrated to 0.5 mg/mL. This sample was applied to glow-discharged holey carbon grids (Quantifoil Cu R1.2/1.3, 300 mesh), blotted for 7 s and then plunge-frozen in liquid ethane using a Vitrobot (Thermo Fisher). The grids were then transferred into cartridges. A Titan Krios cryo-electron transmission microscope (Thermo Fisher) was used to collect cryo-EM images. The images were recorded at 1–2.5 μm defocus on a K3 direct electron detector (Gatan) using super-resolution and correlated-double sampling (CDS) modes at nominal 81K magnification, corresponding to a sampling interval of 1.08 Å/pixel (super resolution, 0.54 Å/pixel). Each micrograph was exposed for 4.5 s with 11.71 e−/A2/sec dose rate (total specimen dose, 41.25 e−/A2). 45 frames were captured per specimen area using SerialEM.[49]

Data processing—The human microsome raw lysate were processed using a similar protocol as described previously.[4] Super-resolution image stacks were aligned and binned by 2 using cryoSPARC[46] to give a final pixel size of 1.08 Å/pixel. CTFs were estimated using patch CTF in cryoSPARC.[46] After manual inspection to discard poor images and estimate particle size, the blob picker in cryoSPARC[46] was used to select particles from subsets of micrographs. These particles were classified with one round of 2D classification and clear templates were selected for template picking in cryoSPARC.[46] Template picker was used to select initial particle sets. Several iterative rounds of 2D classifications were used to clean these particle sets with different circular masks to account for different particle sizes. Featureless classes were removed from each step, resulting in cleaned heterogeneous particle stacks for further processing.

From these cleaned heterogeneous particle sets, particles were classified and final maps were solved using an iterative method termed “Build and Retrieve” (BaR) processing (which is summarized in Figures S2–S4). Briefly, *ab initio* methods were used to “build” initial 3D maps from the complete dataset and particles were then “retrieved” based on the maps. To build the initial maps, particles were separated using 2D classification paired with 3D *ab initio* and heterogeneous classifications. The combination of these techniques was used to divide the dataset into subsets with similar structural features. Each subset was then treated as an individual particle set, where several rounds of 2D classification followed by

3D *ab initio* reconstruction were used to improve the quality of each subset. Non-uniform refinement was applied to "build" the preliminary maps.

To increase particle counts of each classified class, these initial maps were used to "retrieve" particles from the raw particle stack without 2D classification. 3D heterogeneous refinement using the *ab initio* maps, determined from the "build" phase of BaR, was applied to the cleaned heterogeneous particle sets. The new particle subsets were then cleaned using multiple rounds of 2D and 3D *ab initio* classifications. In the case of GLUD1, we eliminated poorly resolved 3D images during the process of 3D classification, resulting in only retaining 10,295 particles for final structural determination.

To determine the correct operating symmetry, we first used the C1 symmetry to conduct the *ab initio* reconstruction. We then performed refinement by introducing symmetry operators and evaluated the resulting maps to assess the influence of these symmetries based on the map quality. In the cases of GP and AOX1, we only identified pseudo symmetry components in these proteins. We therefore eliminated the symmetry constraint and performed final 3D refinement using C1 symmetry (Table S1).

Non-uniform refinement using cryoSPARC[46] with symmetry imposed was used to refine all symmetrical protein structures, while non-uniform refinement followed by local refinement with non-uniform sampling was used to refine those protein structures without a distinct symmetry. The maps were further sharped using cryoSPARC[46] and RELION.[50]

The high quality of the maps enabled us to trace most of the Cα atoms in these protein molecules. The main-chain (Cα) of these human liver enzymes were manually traced using the program Coot.[45] The phenix.sequence_from_map program in PHENIX[47] was then used to generate the best-fitting sequences. The identities of these enzymes were revealed based on the protein sequences, and their presence was further confirmed by proteomics.

Model building and refinement—Model buildings of H6PD, GANAB, PRDX4-ERp46, MTP, AOX1, CES1, GP, PRDX4, ALDH1A1 and GLUD1 were based on the cryo-EM maps generated from the BaR methodology. The subsequent model rebuilding processes were performed using Coot.[45] Structure refinements were done using the phenix.real_space_refine program[51] from the PHENIX suite.[47] The final atomic models were evaluated using MolProbity.[52] The statistics associated with data collection, 3D reconstruction and model refinement are included in Tables S1. The final structures (Figures 1–7) were drawn using Chimera X [128].

Proteomic analysis—The protein of human liver microsome homogenate in the two peaks from the size-exclusion chromatography enrichment (200 μL) was digested by trypsin. The resulting peptides were desalted by a C18 Microspin column (Nest Group, Ipswich, MA) per the manufacturer's instruction and analyzed by LC-MS/MS using a ThermoScientific Fusion Lumos mass spectrometry system.[53] Proteins were identified by comparing all of the experimental peptide MS/MS spectra against the UniProt human database using the Andromeda search engine integrated into the MaxQuant version 1.6.3.3.[54,55] Carbamidomethylation of cysteine was set as a fixed modification, whereas variable modifications included oxidation of methionine to methionine sulfoxide and acetylation of N-terminal amino groups. For peptide/protein identification, strict trypsin specificity was applied, the minimum peptide length was set to 7, the maximum missed cleavage was set to 2, and the cutoff false discovery rate was set to 0.01. Match between runs (match time window: 0.7 min; alignment time window: 20 min) and label-free quantitation (LFQ) options were enabled. The LFQ minimum ratio count was set to 2. The remaining parameters were kept as default.

Interaction network—The interaction network connecting the 10 identified enzymes was predicted using the STRING database.[7] The enzymes were input as a multiprotein search. Line thickness was used to depict interaction confidence.

QUANTIFICATION AND STATISTICAL ANALYSIS

Standard GS-FSC (Gold Standard-Fourier Shell Correlation) curves at a threshold of 0.143 were computed using cryoSPARC[46] to obtain final resolutions of protein models. The final atomic models were evaluated using MolProbity.[52]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Dr. Philip A. Klenotic for proofreading the manuscript. This work was supported by an NIH grant R01AI145069 (E.W.Y.). The mass spectrometer used was purchased with an NIH Shared Instrument grant (S10 RR031537). We thank Belinda Willard and Ling Li for the acquisition of MS data. We are grateful to the Cryo-Electron Microscopy Core at the CWRU School of Medicine and Dr. Kunpeng Li for access to the sample preparation and cryo-EM instrumentation.

REFERENCES

1. Paik Y-K, Jeong S-K, Omenn GS, Uhlen M, Hanash S, Cho SY, Lee H-J, Na K, Choi E-Y, Yan F, et al. (2012). The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome. Nat. Biotechnol. 30, 221–223. 10.1038/nbt.2152. [PubMed: 22398612]
2. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, Zwahlen M, Kampf C, Wester K, Hober S, et al. (2010). Towards a knowledge-based human protein Atlas. Nat. Biotechnol. 28, 1248–1250. 10.1038/nbt1210-1248. [PubMed: 21139605]
3. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. (2015). Tissue-based map of the human proteome. Science 347, 1260419. 10.1126/science.1260419. [PubMed: 25613900]
4. Su C-C, Lyu M, Morgan CE, Bolla JR, Robinson CV, and Yu EW(2021). A ‘Build and Retrieve’ methodology to simultaneously solve cryoEM structures of membrane proteins. Nat. Methods 18, 69–75. 10.1038/s41592-020-01021-2. [PubMed: 33408407]
5. Abdel-Misih SRZ, and Bloomston M (2010). Liver anatomy. Surg. Clin. North Am. 90, 643–653. [PubMed: 20637938]
6. Verbeke EJ, Zhou Y, Horton AP, Mallam AL, Taylor DW, and Marcotte EM (2020). Separating distinct structures of multiple macromolecular assemblies from cryo-EM projections. J. Struct. Biol. 209, 107416. [PubMed: 31726096]
7. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, et al. (2021). The STRING database in 2021: customizable protein-protein networks,

and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 49, D605–D612. [PubMed: 33237311]

8. Mason PJ, Stevens D, Diez A, Knight SW, Scopes DA, and Vulliamy TJ (1999). Human hexose-6-phosphate dehydrogenase (Glucose 1-dehydrogenase) encoded at 1p36: coding sequence and expression. Blood Cells Mol. Dis. 25, 30–37. 10.1006/bcmd.1999.0224. [PubMed: 10349511]
9. Collard F, Collet JF, Gerin I, Veiga-da-Cunha M, and Van Schaftingen E (1999). Identification of the cDNA en coding human 6-phosphogluconolactonase, the enzyme catalyzing the second step of the pentose phosphate pathway. FEBS Lett. 459, 223–226. 10.1016/s0014-5793(99)01247-8. [PubMed: 10518023]
10. Draper N, Walker EA, Bujalska IJ, Tomlinson JW, Chalder SM, Arlt W, Lavery GG, Bedendo O, Ray DW, Laing I, et al. (2003). Mutations in the genes encoding 11 beta-hydroxysteroid dehydrogenase type 1 and hexose-6-phosphate dehydrogenase interact to cause cortisone reductase deficiency. Nat. Genet. 34, 434–439. 10.1038/ng1214. [PubMed: 12858176]
11. Lavery GG, Walker EA, Tiganescu A, Ride JP, Shackleton CHL, Tomlinson JW, Connell JMC, Ray DW, Biason-Lauber A, Malunowicz EM, et al. (2008). Steroid biomarkers and genetic studies reveal in-activating mutations in hexose-6-phosphate dehydrogenase in patients with cortisone reductase deficiency. J. Clin. Endocrinol. Metab. 93, 3827–3832. 10.1210/jc.2008-0743. [PubMed: 18628520]
12. Chen R, Jiang X, Sun D, Han G, Wang F, Ye M, Wang L, and Zou H (2009). Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. J. Proteome Res. 8, 651–661. [PubMed: 19159218]
13. Porath B, Gainullin VG, Cornec-Le Gall E, Dillinger EK, Heyer CM,Hopp K, Edwards ME, Madsen CD, Mauritz SR, Banks CJ, et al. (2016). Mutations in GANAB, encoding the glucosidase IIa subunit, cause autosomal-dominant polycystic kidney and liver disease. Am. J. Hum. Genet. 98, 1193–1207. 10.1016/j.ajhg.2016.05.004. [PubMed: 27259053]
14. Henrissat B (1991). A classification of glycosyl hydrolases based on amino-acid-sequence similarities. Biochem. J. 280, 309–316. 10.1042/bj2800309. [PubMed: 1747104]
15. Caputo AT, Alonzi DS, Marti L, Reca IB, Kiappes JL, Struwe WB, Cross A, Basu S, Lowe ED, Darlot B, et al. (2016). Structures of mammalian ER alpha-glucosidase II capture the binding modes of broad-spectrum iminosugar antivirals. Proc. Natl. Acad. Sci. USA 113, E4630–E4638. 10.1073/pnas.1604463113. [PubMed: 27462106]
16. Martiniuk F, Ellenbogen A, and Hirschhorn R (1985). Identity of neutral alpha-glucosidase AB and the glycoprotein processing enzyme glucosidase II. Biochemical and genetic studies. J. Biol. Chem. 260, 1238–1242. [PubMed: 3881423]
17. Moore KJ, Kunjathoor VV, Koehn SL, Manning JJ, Tseng AA, Silver JM, McKee M, and Freeman MW (2005). Loss of receptor-mediated lipid uptake via scavenger receptor A or CD36 pathways does not ameliorate atherosclerosis in hyperlipidemic mice. J. Clin. Invest. 115, 2192–2201. 10.1172/jci24061. [PubMed: 16075060]
18. Alberti A, Karamessinis P, Peroulis M, Kypreou K, Kavvadas P, Pagakis S, Politis PK, and Charonis A (2009). ERp46 is reduced by high glucose and regulates insulin content in pancreatic beta-cells. Am. J. Physiol. Endocrinol. Metab. 297, E812–E821. 10.1152/ajpendo.00053.2009. [PubMed: 19622788]
19. Biterova EI, Isupov MN, Keegan RM, Lebedev AA, Sohail AA, Liaqat I, Alanen HI, and Ruddock LW (2019). The crystal structure of human microsomal triglyceride transfer protein. Proc. Natl. Acad. Sci. USA 116, 17251–17260. 10.1073/pnas.1903029116. [PubMed: 31395737]
20. Coelho C, Mahro M, Trincão J, Carvalho ATP, Ramos MJ, Terao M, Garattini E, Leimkühler S, and Romão MJ (2012). The first mammalian aldehyde oxidase crystal structure insights into substrate specificity. J. Biol. Chem. 287, 40690–40702. 10.1074/jbc.M112.390419. [PubMed: 23019336]
21. Coelho C, Foti A, Hartmann T, Santos-Silva T, Leimkühler S, and Romão MJ (2015). Structural insights into xenobiotic and inhibitor binding to human aldehyde oxidase. Nat. Chem. Biol. 11, 779–783. 10.1038/nchembio.1895. [PubMed: 26322824]
22. Bencharit S, Morton CL, Howard-Williams EL, Danks MK, Potter PM, and Redinbo MR (2002). Structural insights into CPT-11 activation by mammalian carboxylesterases. Nat. Struct. Biol. 9, 337–342. 10.1038/nsb790. [PubMed: 11967565]

23. Bencharit S, Morton CL, Xue Y, Potter PM, and Redinbo MR(2003). Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme. Nat. Struct. Biol. 10, 349–356. 10.1038/nsb919. [PubMed: 12679808]
24. Fleming CD, Bencharit S, Edwards CC, Hyatt JL, Tsurkan L, Bai F, Fraga C, Morton CL, Howard-Williams EL, Potter PM, and Redinbo MR (2005). Structural insights into drug processing by human carboxylesterase 1: tamoxifen, mevastatin, and inhibition by benzil. J. Mol. Biol. 352, 165–177. [PubMed: 16081098]
25. Alam M, Vance DE, and Lehner R (2002). Structure-function analysis of human triacylglycerol hydrolase by site-directed mutagenesis: identification of the catalytic triad and a glycosylation site. Biochemistry 41, 6679–6687. [PubMed: 12022871]
26. Rath VL, Ammirati M, LeMotte PK, Fennell KF, Mansour MN, Danley DE, Hynes TR, Schulte GK, Wasilko DJ, and Pandit J (2000). Activation of human liver glycogen phosphorylase by alteration of the secondary structure and packing of the catalytic core. Mol. Cell 6, 139–148. 10.1016/s1097-2765(00)00015-0. [PubMed: 10949035]
27. Barford D, and Johnson LN (1989). The allosteric transition of glycogen-phosphorylase. Nature 340, 609–616. 10.1038/340609a0. [PubMed: 2770867]
28. Pautsch A, Stadler N, Wissdorf O, Langkopf E, Moreth W, and Streicher R (2008). Molecular recognition of the protein phosphatase 1 glycogen targeting subunit by glycogen phosphorylase. J. Biol. Chem. 283, 8913–8918. [PubMed: 18198182]
29. Cao Z, Tavender TJ, Roszak AW, Cogdell RJ, and Bulleid NJ(2011). Crystal structure of reduced and of oxidized peroxiredoxin IV enzyme reveals a stable oxidized decamer and a non-disulfide-bonded intermediate in the catalytic cycle. J. Biol. Chem. 286, 42257–42266. 10.1074/jbc.M111.298810. [PubMed: 21994946]
30. Morgan CA, and Hurley TD (2015). Development of a high-throughput in vitro assay to identify selective inhibitors for human ALDH1A1. Chem. Biol. Interact. 234, 29–37. 10.1016/j.cbi.2014.10.028. [PubMed: 25450233]
31. Moore SA, Baker HM, Blythe TJ, Kitson KE, Kitson TM, and Baker EN (1998). Sheep liver cytosolic aldehyde dehydrogenase: the structure reveals the basis for the retinal specificity of class 1 aldehyde de-hydrogenases. Structure 6, 1541–1551. 10.1016/s0969-2126(98)00152-x. [PubMed: 9862807]
32. Peterson PE, and Smith TJ (1999). The structure of bovine glutamate dehydrogenase provides insights into the mechanism of allostery. Structure 7, 769–782. [PubMed: 10425679]
33. Smith TJ, Schmidt T, Fang J, Wu J, Siuzdak G, and Stanley CA (2002). The structure of apo human glutamate dehydrogenase details subunit communication and allostery. J. Mol. Biol. 318, 765–777. [PubMed: 12054821]
34. Vinothkumar KR, and Henderson R (2016). Single particle electron cry-omicroscopy: trends, issues and future perspective. Q. Rev. Biophys. 49, e13–e25. 10.1017/s0033583516000068. [PubMed: 27658821]
35. Herzik MA Jr., Wu M, and Lander GC (2019). High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. Nat. Commun. 10, 1032. [PubMed: 30833564]
36. Kastritis PL, O’Reilly FJ, Bock T, Li Y, Rogon MZ, Buczak K, Romanov N, Betts MJ, Bui KH, Hagen WJ, et al. (2017). Capturing protein communities by structural proteomics in a thermophilic eukaryote. Mol. Syst. Biol. 13, 936. 10.15252/msb.20167412. [PubMed: 28743795]
37. Ho CM, Li X, Lai M, Terwilliger TC, Beck JR, Wohlschlegel J, Goldberg DE, Fitzpatrick AWP, and Zhou ZH (2020). Bottom-up structural proteomics: cryo-EM of protein complexes enriched from the cellular milieu. Nat. Methods 17, 79–85. 10.1038/s41592-019-0637-y. [PubMed: 31768063]
38. Drulyte I, Johnson RM, Hesketh EL, Hurdiss DL, Scarff CA, Porav SA, Ranson NA, Muench SP, and Thompson RF (2018). Approaches to altering particle distributions in cryo-electron microscopy sample preparation. Acta Crystallogr. D Struct. Biol. 74, 560–571. [PubMed: 29872006]
39. Tan YZ, Baldwin PR, Davis JH, Williamson JR, Potter CS, Carragher B, and Lyumkis D (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat. Methods 14, 793–796. [PubMed: 28671674]

40. Arnold SA, Albiez S, Bieri A, Syntychaki A, Adaixo R, McLeod RA, Goldie KN, Stahlberg H, and Braun T (2017). Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts. J. Struct. Biol. 197, 220–226. [PubMed: 27864160]
41. Levitz TS, Weckener M, Fong I, Naismith JH, Drennan CL, Brignole EJ, Clare DK, and Darrow MC (2022). Approaches to using the chameleon: robust, automated, fast-plunge cryoEM specimen preparation. Front. Mol. Biosci. 9, 903148. [PubMed: 35813832]
42. Ravelli RBG, Nijpels FJT, Henderikx RJM, Weissenberger G, Thewessem S, Gijsbers A, Beulen BWAMM, López-Iglesias C, and Peters PJ (2020). Cryo-EM structures from sub-nl volumes using pin-printing and jet vitrification. Nat. Commun. 11, 2563. [PubMed: 32444637]
43. Varjosalo M, Sacco R, Stukalov A, van Drogen A, Planyavsky M, Hauri S, Aebersold R, Bennett KL, Colinge J, Gstaiger M, and Superti-Furga G (2013). Interlaboratory reproducibility of large-scale human protein-complex analysis by standardized AP-MSMS. Nat. Methods 10, 307–314. 10.1038/nmeth.2400. [PubMed: 23455922]
44. Roux KJ, Kim DI, Raida M, and Burke B (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. 196, 801–810. 10.1083/jcb.201112098. [PubMed: 22412018]
45. Emsley P, and Cowtan K (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132. [PubMed: 15572765]
46. Punjani A, Rubinstein JL, Fleet DJ, and Brubaker MA (2017). Cryo-SPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296. [PubMed: 28165473]
47. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, and Terwilliger TC (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. 58, 1948–1954. [PubMed: 12393927]
48. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, and Ferrin TE (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. 30, 70–82. [PubMed: 32881101]
49. Mastronarde DN (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. 152, 36–51. [PubMed: 16182563]
50. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, and Scheres SH (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife 7, e42166. [PubMed: 30412051]
51. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, and Adams PD (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. D Struct. Biol. 74, 531–544. [PubMed: 29872004]
52. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, and Richardson DC (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21. [PubMed: 20057044]
53. Robinson NJ, Miyagi M, Scarborough JA, Scott JG, Taylor DJ, and Schiemann WP (2021). SLX4IP promotes RAP1 SUMOylation by PIAS1 to coordinate telomere maintenance through NF-kB and Notch signaling. Sci. Signal. 14, eabe9613. [PubMed: 34187905]
54. Cox J, and Mann M (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372. [PubMed: 19029910]
55. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, and Mann M (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. 10, 1794–1805. [PubMed: 21254760]

Highlights

- Using cryo-EM to define the composition of a raw human liver lysate
- Simultaneously identifying and solving structures of ten enzymes or complexes
- Noting the potential of studying high-resolution structural-omics of human enzymes

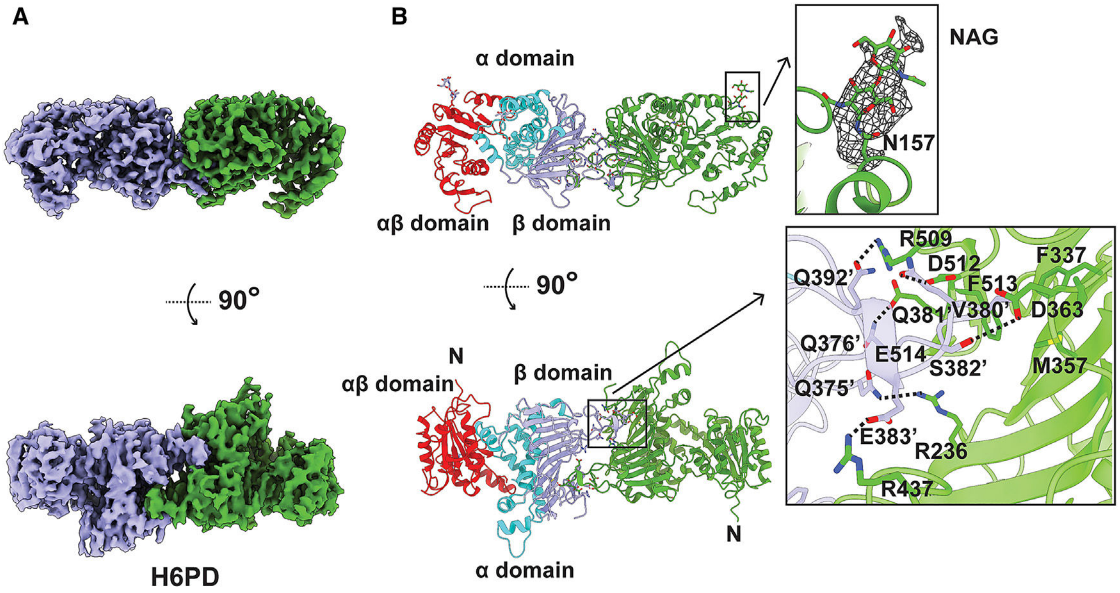


Figure 1. Cryo-EM structure of human H6PD

(A) Cryo-EM density map of dimeric H6PD. The two protomers of H6PD are colored slate and green, respectively.

(B) Ribbon diagram of the 3.02-Å resolution structure of dimeric H6PD. The αβ domain, α domain, and β domain of one subunit of H6PD are colored red, cyan, and slate, respectively. The other subunit of H6PD is colored green. The cryo-EM densities of the two NAG moieties in the green subunit of H6PD are in gray meshes. Residues that are involved in dimerization of H6PD at the dimer interface are highlighted with sticks. The dotted lines indicate hydrogen bonds.

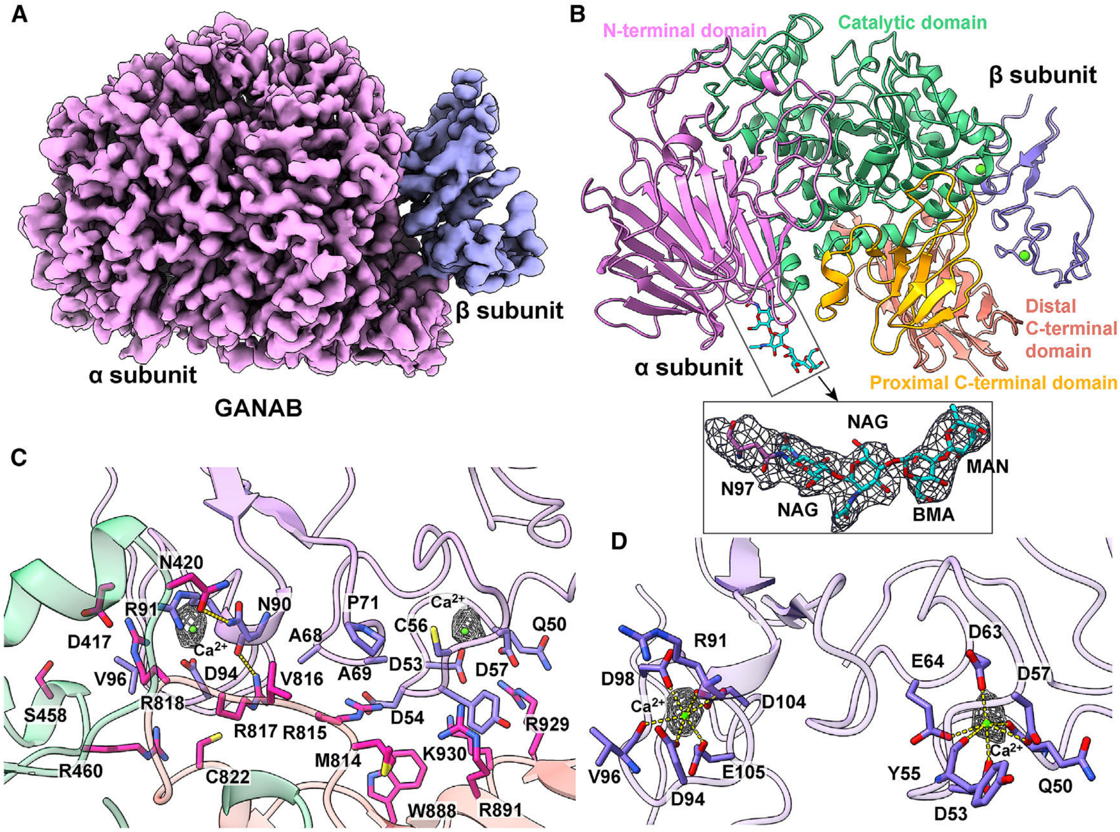


Figure 2. Cryo-EM structure of human GANAB

(A) Cryo-EM density map of the heterodimeric GANAB enzyme, consisting of the α and β subunits. The a and b subunits are colored magenta and slate, respectively.

(B) Ribbon diagram of the 2.92-Å resolution structure of GANAB. The N-terminal, catalytic, proximal C-terminal, and distal C-terminal domains of the α subunit are colored magenta, green, yellow-orange, and salmon, respectively. The β subunit is colored slate. The two bound Ca^{2+} ion at the β subunit are shown as green spheres. An N-linked glycosylation was observed at residue N97. The density of the NAG-NAG-BMA-MAN glycan chain is in gray mesh.

(C) Protein-protein interface between the α and β subunits of GANAB. The α- and β-subunit residues that are important for heterodimerization are highlighted with magenta and slate sticks. The two bound Ca^{2+} ions of the β subunit are shown as green spheres. The dotted lines indicate hydrogen bonds.

(D) The Ca^{2+} ion binding sites of the β subunit. The two bound Ca^{2+} ions are shown as green spheres. Residues that are involved in Ca^{2+} binding are shown as slate sticks.

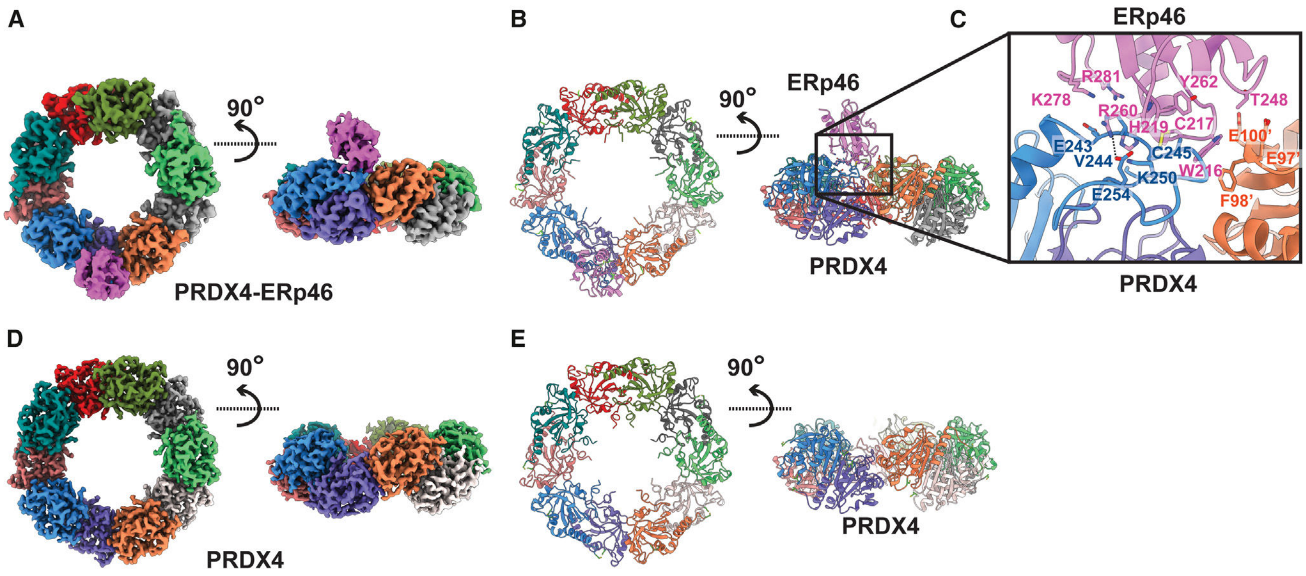


Figure 3. Cryo-EM structures of human PRDX4 and PRDX4-ERp46

(A) Cryo-EM density map of the PRDX4-ERp46 complex. The ten protomers are in different colors. The bound ERp46 protein is colored magenta.

(B) Ribbon diagram of the 3.52-Å resolution structure of PRDX4-ERp46. The ten protomers are in different colors. The bound ERp46 protein is colored magenta.

(C) Interaction between PRDX4 and ERp46. The two protomers of PRDX4 are colored blue and orange. The bound ERp46 protein is colored magenta. Residues that are important in PRDX4-ERp46 interactions are shown as sticks.

(D) Cryo-EM density map of the decameric PRDX4 enzyme. The ten protomers are in different colors.

(E) Ribbon diagram of the 2.84-Å resolution structure of PRDX4. The ten protomers are in different colors.

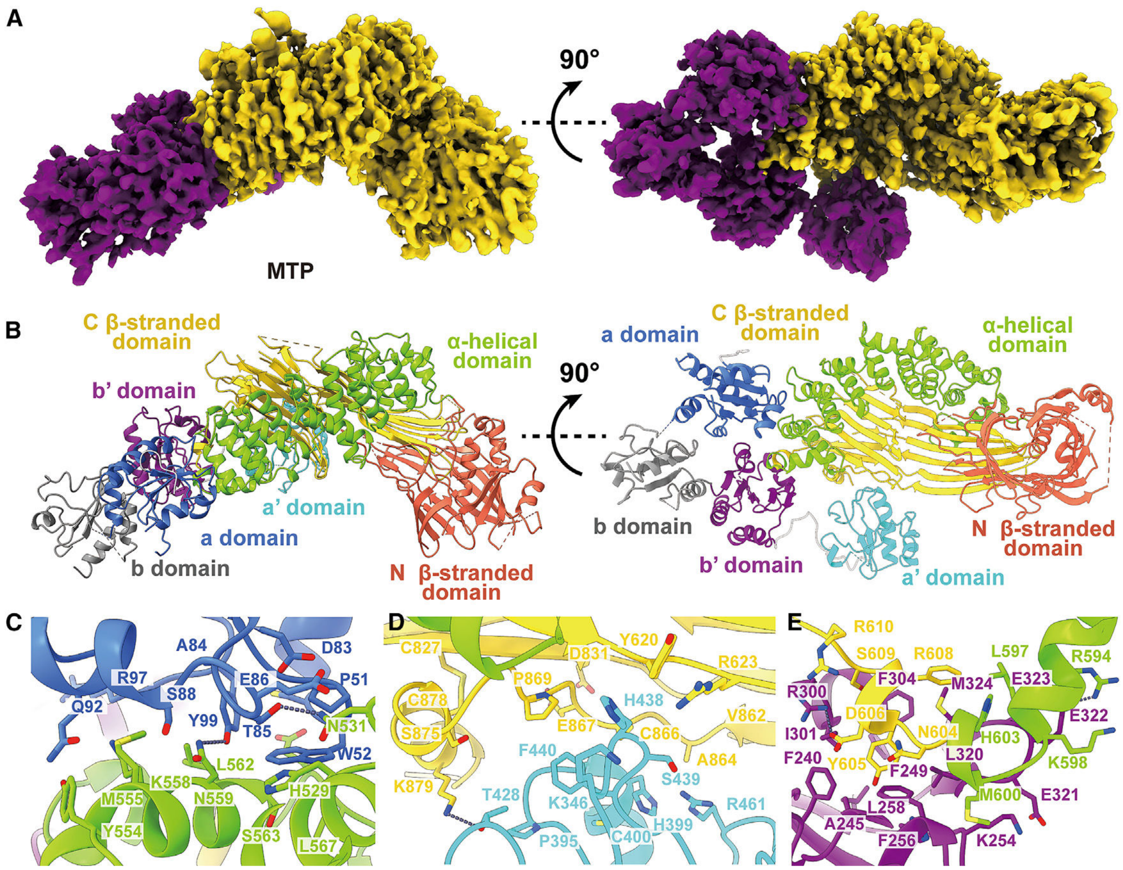


Figure 4. Cryo-EM structure of human MTP

(A) Cryo-EM density map of the heterodimeric MTP enzyme, consisting of the MTPα and PDI subunits. The MTPα and PDI subunits are colored yellow and purple, respectively.

(B) Ribbon diagram of the 3.07-Å resolution structure of MTP. The N-terminal β-stranded, α-helical, and C-terminal β-stranded domains of MTPα are colored orange, green, and yellow, respectively. The a, b, b′, and a′ domains of PDI are colored blue, gray, purple, and cyan, respectively. The α-helical and C-terminal β-stranded domains of MTPα interact with the a, b′, and a′ domains of PDI at the heterodimeric interface.

(C–E) Specific interactions of MTPα and PDI at the protein-protein interface. The α-helical and C-terminal β-stranded domains of MTPα are colored green and yellow, respectively. The a, b′, and a′ domains of PDI are colored blue, purple, and cyan, respectively. The dotted lines indicate hydrogen bonds.

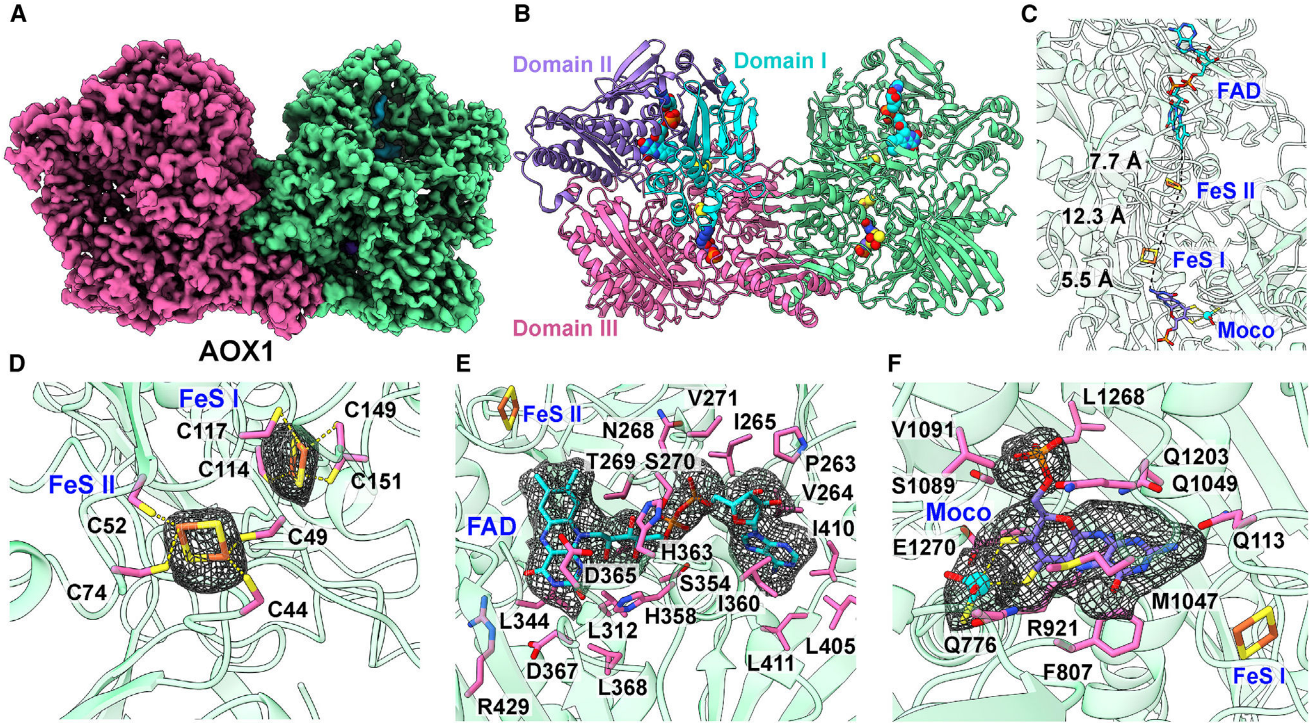


Figure 5. Cryo-EM structure of human AOX1

(A) Cryo-EM density map of dimeric AOX1. The two protomers of AOX1 are colored magenta and green, respectively.

(B) Ribbon diagram of the 2.91-Å resolution structure of dimeric AOX1. Domains I, II, and III of the AOX1 protomers are colored cyan, slate, and magenta, respectively. The other AOX1 protomer is colored green. The bound FAD, Moco, and two [2Fe-2S] cofactors are shown as spheres.

(C) Locations of the bound FAD, Moco, and two [2Fe-2S] cofactors. The FAD and Moco cofactors as well as two [2Fe-2S] clusters are shown as sticks. The dottedlines indicate the distances between these cofactors (12.3 Å between the two [2Fe-2S] clusters, FeS I and FeS II; 5.5 Å between Moco and FeS I; 7.7 Å between FeS II and FAD).

(D) The [2Fe-2S] binding sites. The densities of both FeS I and FeS II are in gray meshes. Residues that are involved in [2Fe-2S] binding are shown as magenta sticks.

(E) The FAD binding site. The density of bound FAD is in gray mesh. Residues that are involved in FAD binding are shown as magenta sticks.

(F) The Moco binding site. The density of bound Moco is in gray mesh. Residues that are involved in Moco binding are shown as magenta sticks.

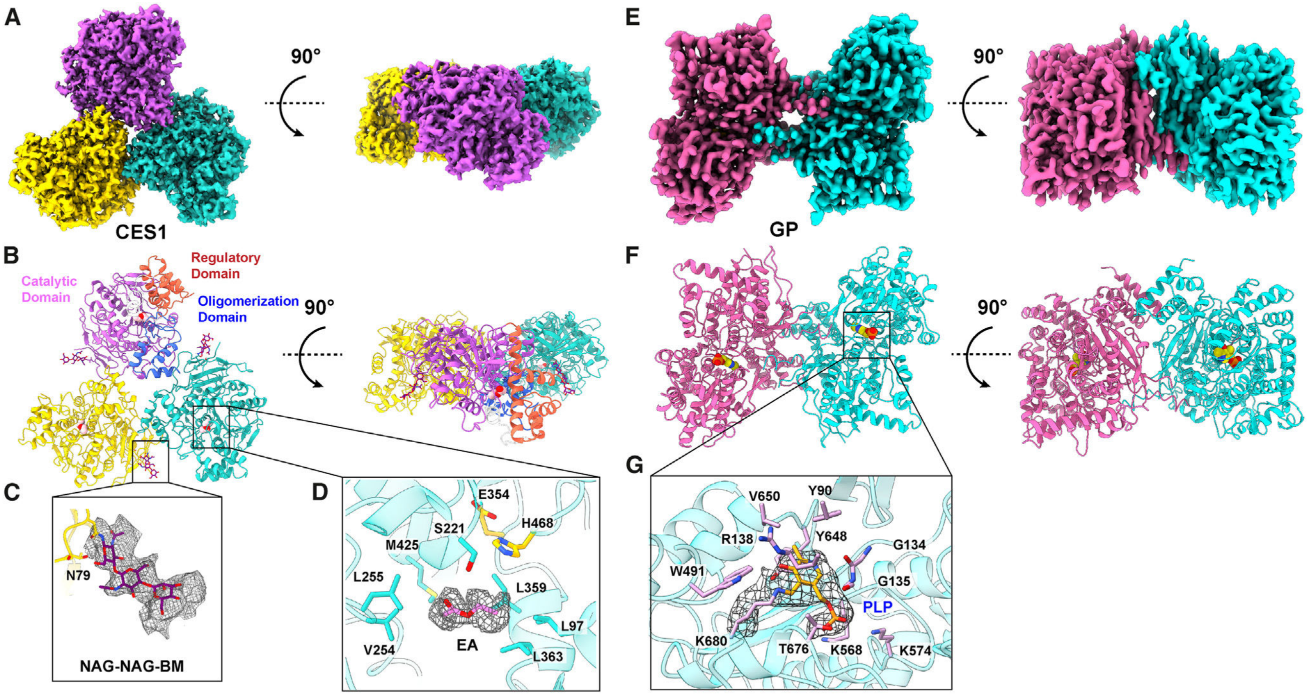


Figure 6. Cryo-EM structures of human CES1 and GP

(A) Cryo-EM density map of trimeric CES1. The three protomers of CES1 are colored magenta, yellow, and cyan, respectively.

(B) Ribbon diagram of the 2.67-Å resolution structure of trimeric CES1. The catalytic, oligomerization, and regulatory domains of one of the CES1 protomers are colored magenta, blue, and orange, respectively. The two other protomers are colored yellow and cyan, respectively.

(C) Cryo-EM density of the NAG-NAG-BMA carbohydrate chain at residue N79 of CES1. The density of NAG-NAG-BMA is in gray mesh. The NAG-NAG-BMA moiety is colored magenta. Residue N79 is shown as yellow sticks.

(D) The ethyl acetate (EA) binding site. Density of the bound EA is in gray mesh. Residues involved in forming the inhibitor binding site are colored cyan. ResiduesE354 and H468, which form a hydrogen-bonded triad with S221, are colored yellow.

(E) Cryo-EM density map of dimeric GP. The two protomers of GP are colored magenta and cyan, respectively.

(F) Ribbon diagram of the 2.65-Å resolution structure of dimeric GP. The two bound PLP molecules within the dimer are shown as yellow spheres.

(G) The PLP binding site. Each GP protomer forms a PLP binding site. Residues Y90, G134, G135, R138, W491, K568, K574, Y648, V650, T676, and K680 are responsible for the binding. The cryo-EM density of bound PLP is in gray mesh. The bound PLP molecule is shown as yellow sticks. Residues involved in PLP binding are colored pink.

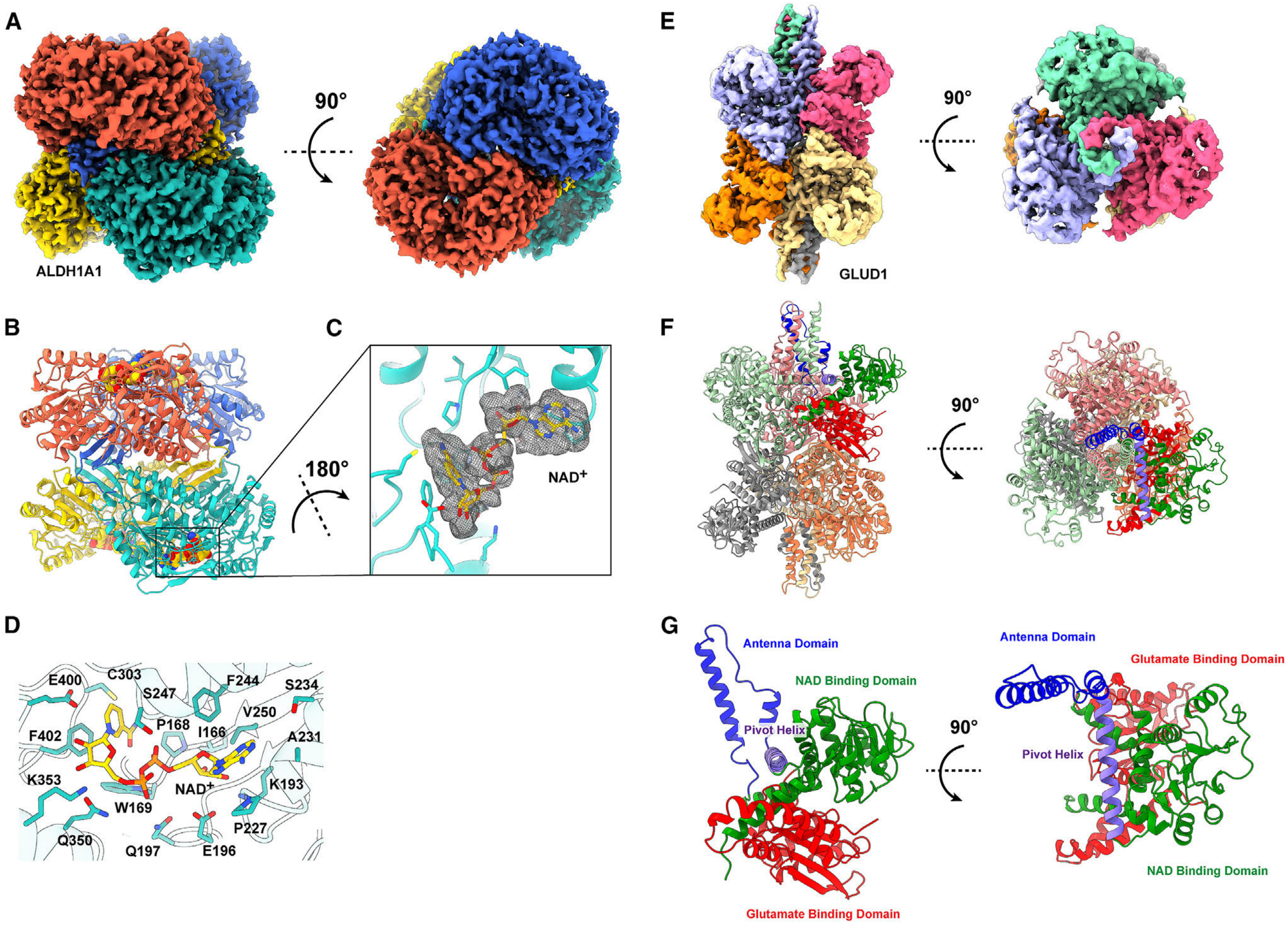


Figure 7. Cryo-EM structures of human ALDH1A1 and GLUD1

(A) Cryo-EM density map of tetrameric ALDH1A1. The four protomers of ALDH1A1 are colored orange, blue, yellow, and green-cyan, respectively.

(B) Ribbon diagram of the 2.64-Å resolution structure of tetrameric ALDH1A1. The four protomers are colored orange, blue, yellow, and green-cyan, respectively.

(C) Cryo-EM density of bound nicotinamide adenine dinucleotide (NAD^+). The density of NAD^+ is in gray meshes.

(D) The NAD^+ binding site. The bound NAD^+ molecule is shown as yellow sticks. Residues involved in NAD^+ binding are highlighted with green-cyan sticks.

(E) Cryo-EM density map of hexameric GLUD1. The six protomers of GLUD1 are colored purple, pink, orange, sand brown, gray, and green-cyan, respectively.

(F) Ribbon diagram of the 3.24-Å resolution structure of tetrameric GLUD1. The six protomers are colored purple, pink, orange, sand brown, gray, and green-cyan, respectively.

(G) The antenna, pivot, NAD^+ binding, and glutamate binding domains are colored blue, purple, green, and red, respectively.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
human liver microsomes	SEKISUI XenoTech	Cat# H0610
Chemicals, peptides, and recombinant proteins		
Na-cholate	Sigma Life Science	Cat# C6445
Acetonitrile	Millipore Sigma	Cat# 1000291000
Ammonium bicarbonate	Sigma-Aldrich	Cat# 5330050050
Iodoacetamide	GE Healthcare	Cat# RPN6302
Trypsin/Lys-C Mix	Promega	Cat# V5073
Deposited data		
Raw cryo-EM data: Peak 1 (300–650 kDa)	This paper	EMPIAR-11249 DOI: https://doi.org/10.6019/EMPIAR-11249
Raw cryo-EM data: Peak 2 (100–250 kDa)	This paper	EMPIAR-11250 DOI: https://doi.org/10.6019/EMPIAR-11250
Cryo-EM structure and map of H6PD	This paper	PDB ID: 8EM2 EMDB ID: EMD-28232 DOI: https://doi.org/10.2210/pdb8EM2/pdb
Cryo-EM structure and map of GANAB	This paper	PDB ID: 8EMR EMDB ID: EMD-28262 DOI: https://doi.org/10.2210/pdb8EMR/pdb
Cryo-EM structure and map of PRDX4-ERp46	This paper	PDB ID: 8EKY EMDB ID: EMD-28217 DOI: https://doi.org/10.2210/pdb8EKY/pdb
Cryo-EM structure and map of MTP	This paper	PDB ID: 8EOJ EMDB ID: EMD28377 DOI: https://doi.org/10.2210/pdb8EOJ/pdb
Cryo-EM structure and map of CES1	This paper	PDB ID: 8EOR EMDB ID: EMD-28465 DOI: https://doi.org/10.2210/pdb8EOR/pdb
Cryo-EM structure and map of AOX1	This paper	PDB ID: 8EMT EMDB ID: EMD-28264 DOI: https://doi.org/10.2210/pdb8EMT/pdb
Cryo-EM structure and map of GP	This paper	PDB ID: 8EMS EMDB ID: EMD-28263 DOI: https://doi.org/10.2210/pdb8EMS/pdb

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cryo-EM structure and map of ALDH1A1	This paper	PDB ID: 8ENE EMDB ID: EMD-28271 DOI: https://doi.org/10.2210/pdb8ENE/pdb
Cryo-EM structure and map of PRDX4	This paper	PDB ID: 8EKW EMDB ID: EMD-28214 DOI: https://doi.org/10.2210/pdb8EKW/pdb
Cryo-EM structure and map of GLUD1	This paper	PDB ID: 7UZM EMDB ID: EMD-26915 DOI: https://doi.org/10.2210/pdb7UZM/pdb
Raw mass spectrometry proteomics data	This paper	PXD040733
Software and algorithms		
Coot	Emsley et al.[45]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
cryoSPARC v3	Punjani et al.[46]	https://cryosparc.com/
MaxQuant	Cox and Mann	https://www.maxquant.org/
PHENIX	Adams et al.[47]	https://phenix-online.org/
STRING database	STRING	https://string-db.org/
UCSF-Chimera X	Pettersen et al.[48]	https://www.cgl.ucsf.edu/chimerax/
Other		
Superdex 200 column	GE Healthcare	Cat# GE28-9909-44
C18 Microspin column	Nest Group	Cat# SEM SS18V
Quantifoil R 1.2/1.3 Cu 300 mesh, copper	Quantifoil	Cat# Q33336