

Structure and identification of the native PLP synthase complex from *Methanosarcina acetivorans* lysate

Angela Agnew,^{1,2,3} Ethan Humm,¹ Kang Zhou,² Robert P. Gunsalus,^{1,4} Z. Hong Zhou^{1,2,3}

AUTHOR AFFILIATIONS See affiliation list on p. 13.

ABSTRACT Many protein-protein interactions behave differently in biochemically purified forms as compared to their *in vivo* states. As such, determining native protein structures may elucidate structural states previously unknown for even well-characterized proteins. Here, we apply the bottom-up structural proteomics method, *cryoID*, toward a model methanogenic archaeon. While they are keystone organisms in the global carbon cycle and active members of the human microbiome, there is a general lack of characterization of methanogen enzyme structure and function. Through the *cryoID* approach, we successfully reconstructed and identified the native *Methanosarcina acetivorans* pyridoxal 5'-phosphate (PLP) synthase (PdxS) complex directly from cryogenic electron microscopy (cryo-EM) images of fractionated cellular lysate. We found that the native PdxS complex exists as a homo-dodecamer of PdxS subunits, and the previously proposed supracomplex containing both the synthase (PdxS) and glutaminase (PdxT) was not observed in cellular lysate. Our structure shows that the native PdxS monomer fashions a single 8 α /8 β TIM-barrel domain, surrounded by seven additional helices to mediate solvent and interface contacts. A density is present at the active site in the cryo-EM map and is interpreted as ribose 5-phosphate. In addition to being the first reconstruction of the PdxS enzyme from a heterogeneous cellular sample, our results reveal a departure from previously published archaeal PdxS crystal structures, lacking the 37-amino-acid insertion present in these prior cases. This study demonstrates the potential of applying the *cryoID* workflow to capture native structural states at atomic resolution for archaeal systems, for which traditional biochemical sample preparation is nontrivial.

IMPORTANCE Archaea are one of the three domains of life, classified as a phylogenetically distinct lineage. There is a paucity of known enzyme structures from organisms of this domain, and this is often exacerbated by characteristically difficult growth conditions and a lack of readily available molecular biology toolkits to study proteins in archaeal cells. As a result, there is a gap in knowledge concerning the mechanisms governing archaeal protein behavior and their impacts on both the environment and human health; case in point, the synthesis of the widely utilized cofactor pyridoxal 5'-phosphate (PLP; a vitamere of vitamin B₆, which humans cannot produce). By leveraging the power of single-particle cryo-EM and map-to-primary sequence identification, we determine the native structure of PLP synthase from cellular lysate. Our workflow allows the (i) rapid examination of new or less characterized systems with minimal sample requirements and (ii) discovery of structural states inaccessible by recombinant expression.

KEYWORDS cryo-EM, cryoID, PLP synthase, *Methanosarcina acetivorans*

While protein complexes from model eukaryotic and prokaryotic systems are highly represented in structural biology literature, there remain gaps in knowledge in

Editor Caroline S. Harwood, University of Washington School of Medicine, Seattle, Washington, USA

Address correspondence to Z. Hong Zhou, Hong.Zhou@UCLA.edu, or Robert P. Gunsalus, RobG@microbio.ucla.edu.

The authors declare no conflict of interest.

See the funding table on p. 14.

Received 9 October 2024

Accepted 4 November 2024

Published 26 November 2024

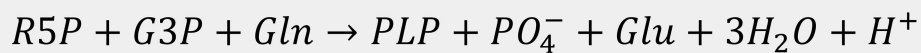
Copyright © 2024 Agnew et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

cases of dynamic cellular processes and taxonomically neglected species. As members of one of the three domains of life alongside bacteria and eukaryotes, archaeal organisms (1) have been chronically understudied due to several issues that have hindered accurate documentation of the domain. These include, but are not limited to, differences in bacterial identification strategies relative to archaeal metabolism and cell structure, the frequent difficulty of culturing archaea, and the lack of a robust body of archaeal functional genomic annotations in existing databases (2). Furthermore, many of these species can be recalcitrant to traditional molecular biology techniques due to extremes of cell culture conditions (e.g., temperature, pH, osmolarity, and anaerobiosis) plus low cell yield. Recombinant approaches are not applicable when unique archaeal cofactors are required for protein function. While clonal and CRISPR-Cas9 techniques have been recently introduced for methanogenic archaeal clades, they are neither currently commercially available nor widely distributed across the range of archaeal model species (3–5). However, often these unique cases provide crucial evolutionary context for foundational biology.

Methanogens are a subset of the kingdom Euryarchaeota of the domain Archaea and are classified by their unique ability to anoxygenically produce methane. They are the primary source of biogenic methane on earth and drive the anoxic carbon cycle. These archaea are ubiquitous in anaerobic habitats, such as wetlands, marine sediments, sewage plants, and the GI tracts of cold- and warm-blooded animals (6). Methanogenesis is an ancient metabolism, with the first methanogen likely existing not long after bacteria and archaea diverged (7). Though still responsible for producing 1 gigaton of methane annually, much of the methane produced is subsequently metabolized by methanotrophic microorganisms living in adjoining anaerobic and aerobic environments (6). The activity of methanotrophs is still insufficient to offset combined biogenic and nonbiogenic sources of methane, such that the impact of microorganisms cannot be ignored in the effort to curtail greenhouse gas emissions.

Yet another feature of several methanogenic archaea is their relationship with the human microbiome. Archaeal species have now been identified from skin, the GI tract, and respiratory tract (8, 9). The consequence of cross-feeding with fermentative bacteria in this context is promoting the overgrowth of pathogenic microbes, yet this field of research is still in its infancy. Still, there are some methanogenic archaea now tied to periodontitis, irritable bowel syndrome, and colorectal cancer (10, 11). Since H_2 is a methanogenic substrate, methanogens keep environmental concentrations of H_2 low, energetically benefiting the fermentative metabolism of other bacteria. While these effects are not known to arise from virulence factors of the archaea themselves, these outcomes do result from their metabolic behavior and adherence to sites of infection. A mechanistic understanding of their key enzymes may give rise to therapeutic targets that can ameliorate these polymicrobial diseases.

One such target which we probe in this study is pyridoxal 5'-phosphate (PLP), a vitamere of vitamin B6, which is an essential cofactor for human neurological and immune health yet is one that humans and animals cannot natively synthesize. PLP is the biologically active state of vitamin B6, participating in over 160 crucial enzymatic processes, which include the metabolism of glycogen, amino acids, and lipids (Fig. 1) (12, 13). Particularly due to its roles in the production of neurotransmitters and modulation of interleukin-2, PLP deficiencies have been correlated with neurological disorders, such as epileptic encephalopathy, schizophrenia, and Parkinson's disease, as well as immune system dysregulation and heightened inflammation (14, 15). Since humans can only obtain vitamin B6 from their diet or from endogenous microbial species, it is clear that microbiome composition and activity have an effect on vitamin B6 utilization (16). As components of the gut and oral microbiome, vitamin B6 production may represent another route through which archaea have specific relevance to human health outcomes. Furthermore, the fact that humans lack the genes for *de novo* vitamin B6 synthesis suggests that PLP synthases are a potentially druggable target for microbial disease. Some methanogenic archaea can grow mutualistically with environmental



Reaction formula for PLP synthase catalysis

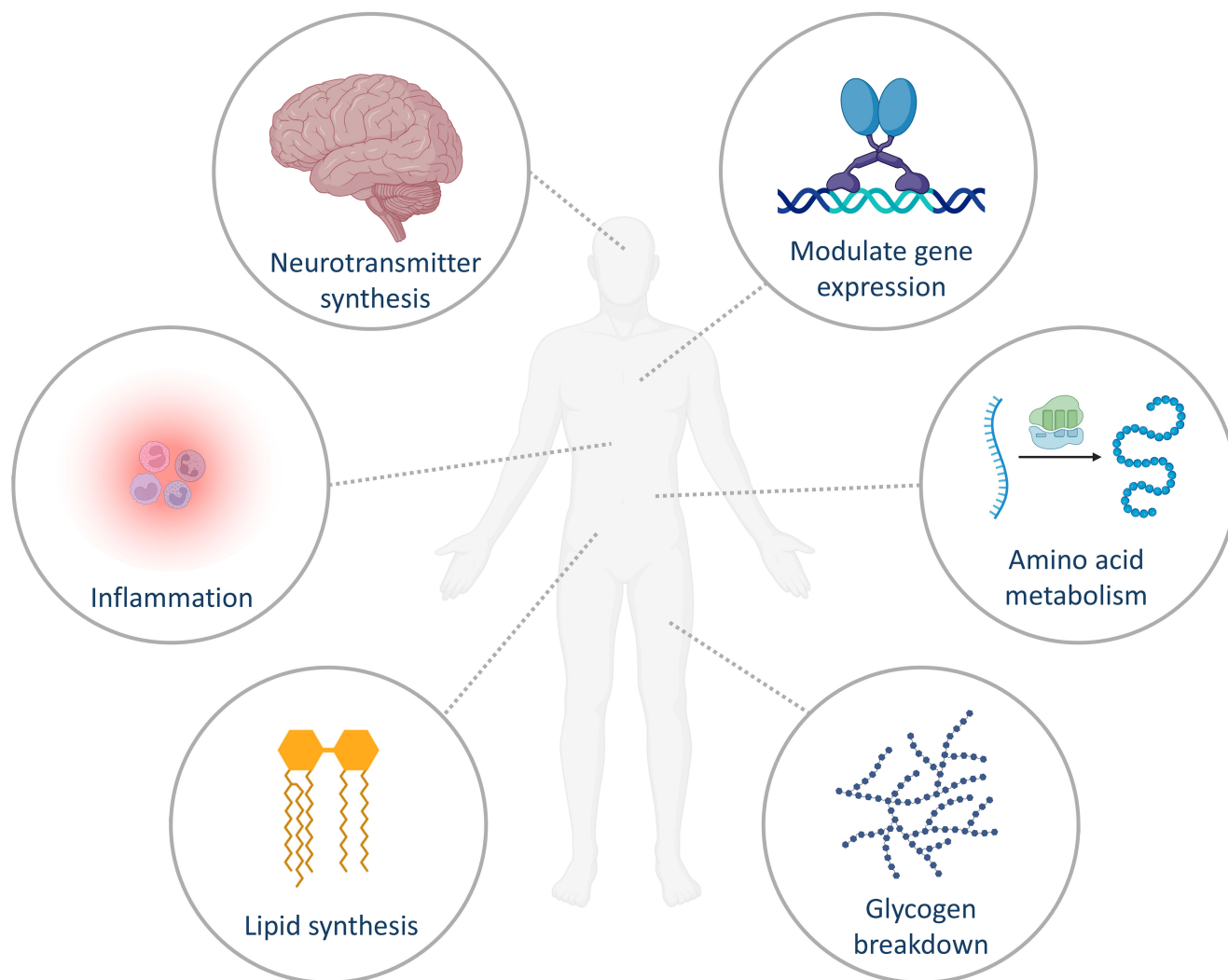


FIG 1 PLP is an essential cofactor for many basic cellular processes implicated in human health, including biosynthesis and metabolism of amino acids, carbohydrates, lipids, and nucleic acids; neurotransmitter production; modulating steroidal receptor gene expression; and regulating immune function and inflammatory responses (17–19).

bacterial colonies, marking them for examination in the development of therapeutics against pathogenic bacterial species.

Of current structural biology techniques, cryogenic electron microscopy (cryo-EM) singularly enables us to achieve high-resolution structures of large-scale, frozen-hydrated protein complexes in their native states (20). Previous publications have pioneered a pipeline that allows for autonomous reconstruction and identification of unknown proteins from images of an enriched cellular milieu, using an in-house software termed *cryoID* (21–23). By avoiding extensive recombinant modifications, proteins may maintain the conformations, complexation, and chemical modifications crucial to their activity in

living cells. To date, this method has not been applied to an archaeal system in pursuit of cryo-EM reconstruction of native proteins.

Here, we have applied the *cryoID* approach to study the model methanogen species *Methanosarcina acetivorans*, chosen for its comparatively rapid culture growth and established characterization, which is apt for exploratory studies. By directly imaging and identifying native cytosolic proteins, we have successfully determined the PdxS subunit of its pyridoxal 5'-phosphate (PLP) synthase. Our reconstruction of the dodecamer reveals departures from the only two known crystal structures from Euryarchaeota Archaea and demonstrates the promise of the *cryoID* workflow for samples from less characterized systems, particularly those from Archaea, where molecular biology approaches have been limited.

RESULTS

Cryo-EM reconstruction

In efforts to simultaneously ensure a sufficiently high population of dominant species in this sample while preserving as many native interactions as possible, per the *cryoID* approach, a glycerol density gradient alone was used to fractionate the lysate of 100 μ L of pelleted cells from a 50-mL culture. A 2-mL fraction window containing particles 8–12 nm in diameter was pooled for structural examination due to the promising 2D class averages noted in negative stain electron microscopy (EM) screening (Fig. 2B). Even after selecting a bespoke size range, resultant images of stained and frozen samples make evident the retained heterogeneity after separation (Fig. 2A through C). Similar classes were noted between negative stain and cryo-EM 2D averaging, and by employing a single-particle analysis approach with the cryoSPARC suite (24), we reconstructed a cryo-EM map to 3.38 Å with D6 symmetry (Fig. 2D and E). Even with just several views predominating, a high-resolution map could be reconstructed due to its high symmetry. This made it rather feasible to “mix and match” classes, which look extremely different from each other as putative top and side views, providing they have similar dimensions with high and compatible symmetries, and eventually discover an *ab initio* reconstruction that produced a rational and continuous density. After repicking the data set with a deep learning (Topaz) model trained on particles contributing to the best *ab initio* model, we successfully expanded our particle data set selected specifically for this bespoke species (25). Using the 3.38-Å resolution cryo-EM reconstruction, we could identify the particle as the 387-kD synthase (PdxS, UniProt accession no. [Q8TQH6](#)) subunit of the *M. acetivorans* pyridoxal 5' phosphate synthase (previously shown to appear in supramolecular complex with its glutaminase subunit, PdxT [Table 1] [26, 27]) through the *cryoID* software (21) as well as model amino acids 11–301 of the monomeric subunit, thus bypassing the need for traditional mass spectrometry identification.

Structural analysis

Like other homologous PLP synthase structures, the *M. acetivorans* PdxS dodecameric complex consists of two layers of homohexameric rings that stack cylindrically, with an outer diameter of 108 Å and an inner diameter of 42 Å (Fig. 2D). Each monomer adapts a triose phosphate isomerase (TIM) barrel fold that houses the active site for deoxyxylulose 5-phosphate-independent PLP synthesis from D-ribose 5-phosphate (R5P), D-glyceraldehyde 3-phosphate, and an ammonia moiety from glutamine. Ribulose 5-phosphate and dihydroxyacetone phosphate are also acceptable substrates, as PdxS homologs demonstrate triose and pentose isomerase activity in addition to PLP synthesis (as suggested by the domain architecture) (17). There is strong density in this binding pocket adjacent to residues annotated as binding sites for R5P, suggestive of captured particles undergoing native catalysis. The first 10 residues lack clear density for modeling, as was similarly observed with the *Arabidopsis thaliana* Pdx1.2/1.3 pseudoenzyme structures (28), indicating that this region is relatively flexible when unbound to its glutaminase, PdxT. The rest of the monomeric secondary structure elements are named

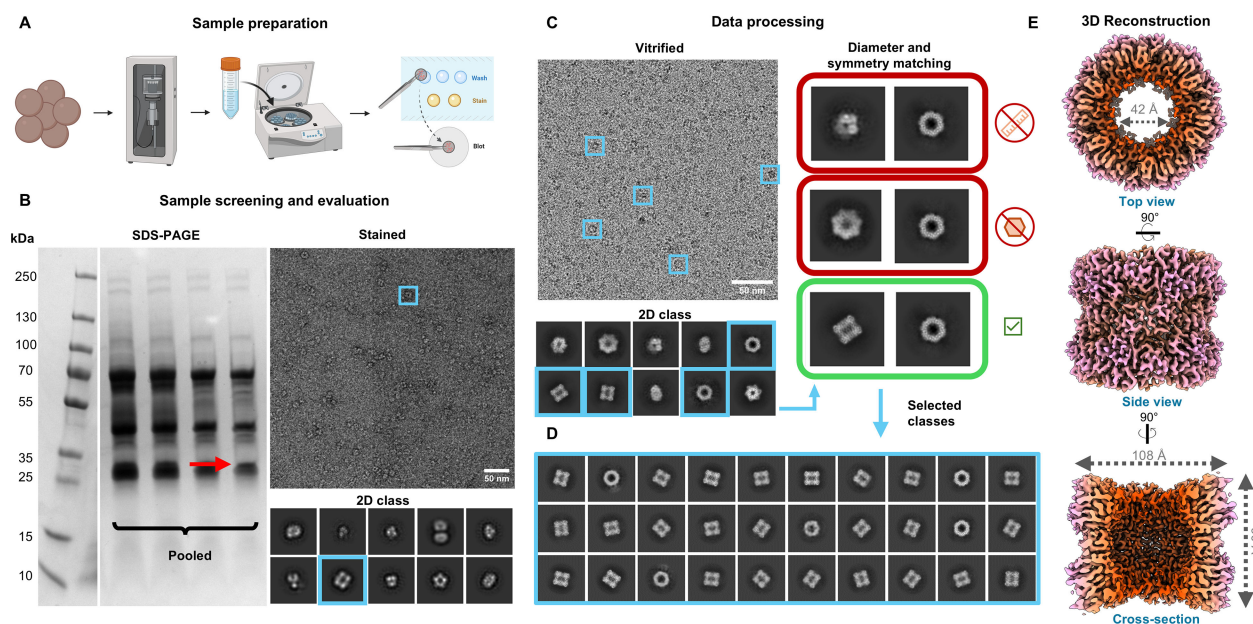


FIG 2 *CryoID* workflow for the reconstruction of PdxS. (A) *M. acetivorans* cells are lysed by sonication, fractionated by size, and applied to TEM grids. (B) Sample screening and evaluation. The left panel contains a Coomassie-stained SDS-PAGE gel with the selected fractions of interest, with the band corresponding to the PdxS molecular weight indicated by the red arrow. On the right, a representative negative stain TEM micrograph of this selected density gradient fraction and corresponding 2D class averages are displayed. An example PdxS particle and its corresponding 2D class are both boxed in blue. (C–E) Cryo data reconstruction workflow. An example cryo-EM micrograph (C) illustrates the extent of sample heterogeneity, with particles of interest corresponding to PdxS boxed in blue. Like in panel B, example PdxS particles and the corresponding 2D class are boxed in blue. These particles are also used in the panel “diameter and symmetry matching” for better illustration (C–E). Different 2D classes hypothesized to represent orthogonal views were compared based upon matching dimensions, symmetry compatibility, and ultimately reasonable *ab initio* reconstructions. In this example, the first two classes mismatch in their longest diameters, and the second two are a highly unlikely symmetry pair. (D) After selecting initial classes that represent top and side views, the data set was repicked with a Topaz model trained on the selected classes, giving rise to the second set of 2D classes exclusively showing different views of our chosen particle (D). Using D6 symmetric nonhomogenous reconstruction in cryoSPARC, we obtain the map seen in panel E. Density map is colored by cylindrical radius.

by sequence, as depicted in Fig. 3B. There are in total 15 alpha helices and 8 beta sheets, with helices $\alpha 1$ –8 (excluding prime and double-prime alpha helices) participating in the TIM-barrel fold with beta sheets $\beta 1$ – $\beta 8$.

Interactions between PdxS subunits

The atomic model of the PdxS complex shows high levels of surface and charge complementarity in its inter-subunit binding interfaces (Fig. 4). The monomer itself resembles an asymmetric, inverted triangle, demonstrating on either side a convex (A) or concave (B) facet posed for lateral binding. Monomer tips also interdigitate at the hexamer-hexamer interface, creating altogether three distinct types of lateral contacts that give rise to the dodecamer, here described as AB, AA, and BB (Fig. 4B). A network of six inter-subunit hydrogen bonds anneal the facets of interface AB together, while two such interactions comprise AA, and another four comprise BB (while each has 1 and 2 unique bonds, respectively, these are doubled due to the inherent twofold symmetry of the “self” interfaces). The residues contributing to the AB contacts are as follows. Two hydrogen bonds form between glutamate 98 of helix 3 and arginine 233 of helix 7 (Fig. 4E). Also posed on $\alpha 7$ is aspartate 227, which forms two hydrogen bonds; one with histidine 93 from loop $\beta 3$ to $\alpha 3$, and arginine 90 on $\beta 3$ (Fig. 4E). Arginine 67 from loop $\alpha 2'$ to $\alpha 2$ forms a backbone contact with methionine 282 of loop $\alpha 8''$ – $\alpha 9$. Lastly, aspartate 118 from loop $\beta 4$ to $\alpha 4$ forms a hydrogen bond with lysine 172 of $\alpha 6$. The “self” interfaces (AA and BB) are dominated by electrostatic interactions between alpha helices 6 and 6'. Comprising the AA interface, the one unique sidechain bond pair is between arginine

TABLE 1 Comparison of existing PLP structures and departing characteristics including acquisition methods and oligomeric states^a

System	Domain	Structure method(s)	Recombinant expression	Glutaminase bound	37-aa insert present	Oligomeric state
<i>M. acetivorans</i>	A	cryo-EM	No	No	No	Homo 12-mer
<i>Arabidopsis thaliana</i>	E	XRD, cryo-EM	Yes	No	No	Homo 12-mer
<i>Saccharomyces cerevisiae</i>	E	XRD	Yes	No	No	Homo 6-mer
<i>Bacillus subtilis</i>	B	XRD	Yes	Yes	No	Hetero 24-mer
<i>Staphylococcus aureus</i>	B	XRD	Yes	Yes	No	Hetero 24-mer
<i>Plasmodium berghei</i>	E	XRD	Yes	Yes	No	Hetero 24-mer
<i>Thermotoga maritima</i>	B	XRD	Yes	Yes	No	Hetero 24-mer
<i>Thermus thermophilus</i>	B	XRD	Yes	No	No	Homo 12-mer
<i>Geobacillus kaustophilus</i>	B	XRD	Yes	No	No	Homo 12-mer
<i>Geobacillus stearothermophilus</i>	B	XRD	Yes	No	No	Homo 12-mer
<i>Mycobacterium tuberculosis</i>	B	XRD	Yes	No	No	Homo 12-mer
<i>Methanocaldococcus jannaschii</i>	A	XRD	Yes	No	Yes	Homo 12-mer
<i>Pyrococcus horikoshii</i>	A	XRD	Yes	No	Yes	Homo 6-mer

^aThe described systems include *A. thaliana* (5LNR, 7LB5) (28, 29), *B. subtilis* (2NV2) (27), *S. aureus* (8U7J) (30), *S. cerevisiae* (3FEM) (31), *P. berghei* (4ADS) (32), *T. maritima* (2ISS) (26), *T. thermophilus* (2ZBT) (33), *G. kaustophilus* (4WY0) (34), *G. stearothermophilus* (1ZNN) (35), *M. tuberculosis* (2ISS) (36), *M. jannaschii* (2YZR) (37), and *P. horikoshii* (4FIQ) (38).

194 (α6') and glutamate 197 (α6'). For the contacts involved in BB, the two unique bonds are shared between arginine 179 (α6) and glutamate 188 (α6').

The overall shape of the monomer is topologically complementary to itself, given that monomers within a hexamer self-associate via the concave-convex fit of the A and B interfaces. The alpha helices external to the (β/α)₈ TIM barrel motif (αN, α2', α6', α6'', α8', α8'', and α9) are also responsible for some morphological stabilization in addition to protecting the active site from solvent exposure (αN and α2'). These stabilizing contacts include a protrusion of the T115-Y121 loop into a hydrophobic patch along the crevice between the adjacent monomer's α6 and α7 helices along the AB interface (Fig. 4F). Another such region is located at the same interface, anterior to the aforementioned loop; α7 and α8'' share valine- and (iso)leucine-mediated contacts with the adjacent monomeric unit's α3 helix (Fig. 4F).

In our maps, no instances of PdxT binding were observed, even after extensive 3D classification and examination under a low-density threshold. Furthermore, upon screening the higher molecular weight fractions of the glycerol density gradient, no PdxS-PdxT supracomplexes were detected as determined by 2D classification and examination of individual micrographs. In homologous proteins for which there are crystal structures, the αN helix of PdxS is instrumental to PdxT binding (40). The *M. acetivorans* αN helix displays several solvent-exposed hydrophobic residues (I16, G19, F20, M23), with regions of dense negative and positive charge above and below, respectively, which is highly suggestive of a putative binding interface for the glutaminase. That being said, given that this sample represents a composite of the most dominant states in our cellular lysate, our work may suggest a novel state wherein PdxT binds extremely transiently and, upon delivery of substrates, need not remain docked to the PdxS homo-dodecamer. This is supported by recent findings which demonstrate transient, sub-stoichiometric binding of Pdx1 and Pdx2 in *Staphylococcus aureus* PLP synthase by SAXS (30).

Active site chemistry and evolutionary conservation

On the interior of the TIM barrel, we observe an elongated density adjacent to residues annotated as substrate-binding sites (Fig. 5A). The PLP synthase active site chemistry as currently understood is rather complex, involving ring openings, closings,

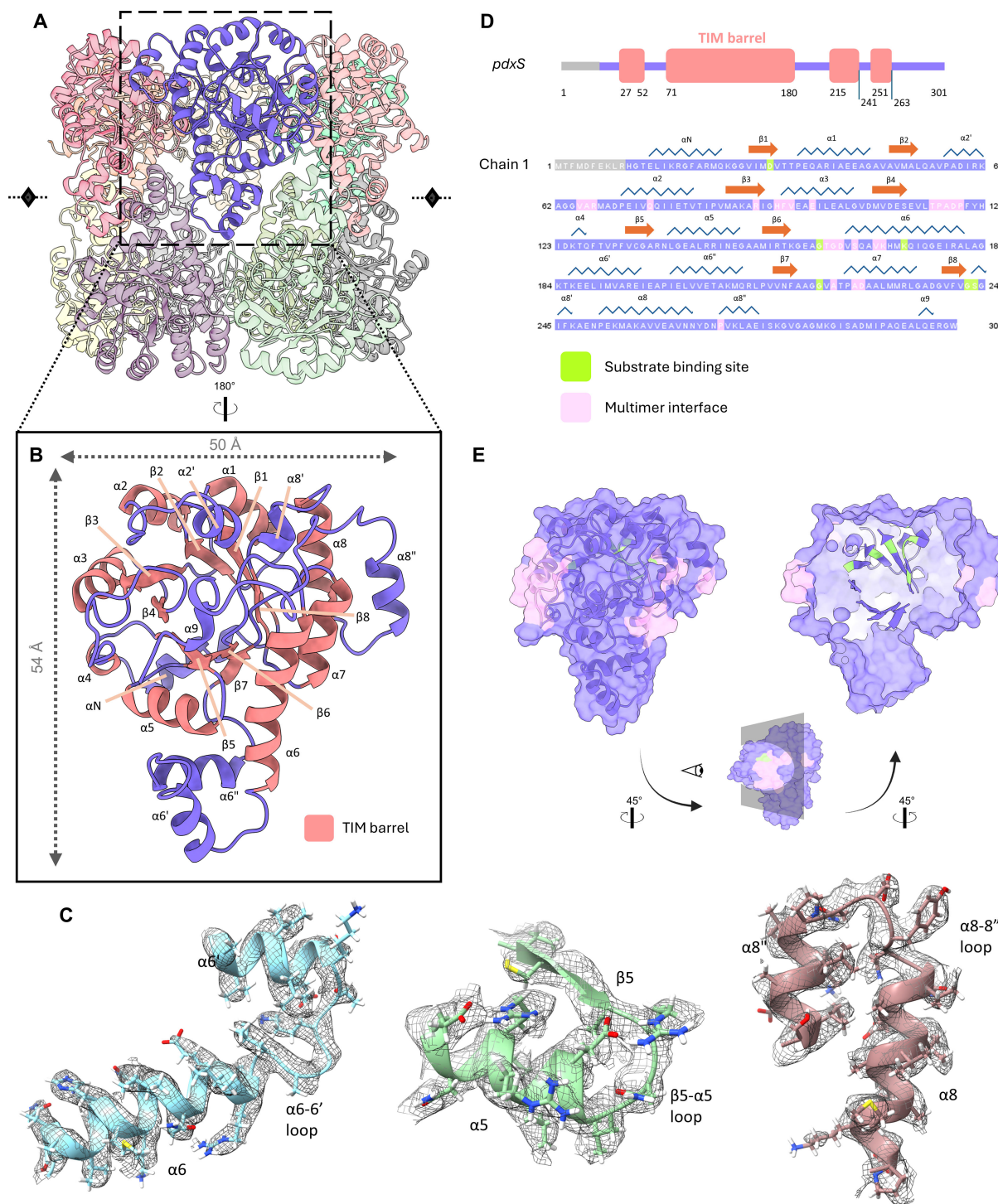


FIG 3 Monomer structure of PdxS. (A) Multimer structure of PdxS with one subunit highlighted. The monomer subunit is depicted as a ribbon representation with secondary structure elements labeled (B) and with selected sections docked into the cryo-EM density (represented as a mesh) to demonstrate sidechain fit (C). (D) Domain structure of the PdxS monomer and its complete primary sequence with secondary structure elements, binding site residues, and multimer interface interactions annotated according to UniProt entry (39). (E) Space-filling representation of the monomer model with multimer interfaces colored in pink and R5P binding site residues colored in green, along with a second cut-away view to reveal the beta barrel ribbon depiction within.

and isomerizations but generally proceeds with a linearized R5P being stabilized via formation of a Schiff base of its C1 with an active site lysine residue (40). A secondary lysine residue likewise bonds with the C5 of R5P and is eventually responsible for

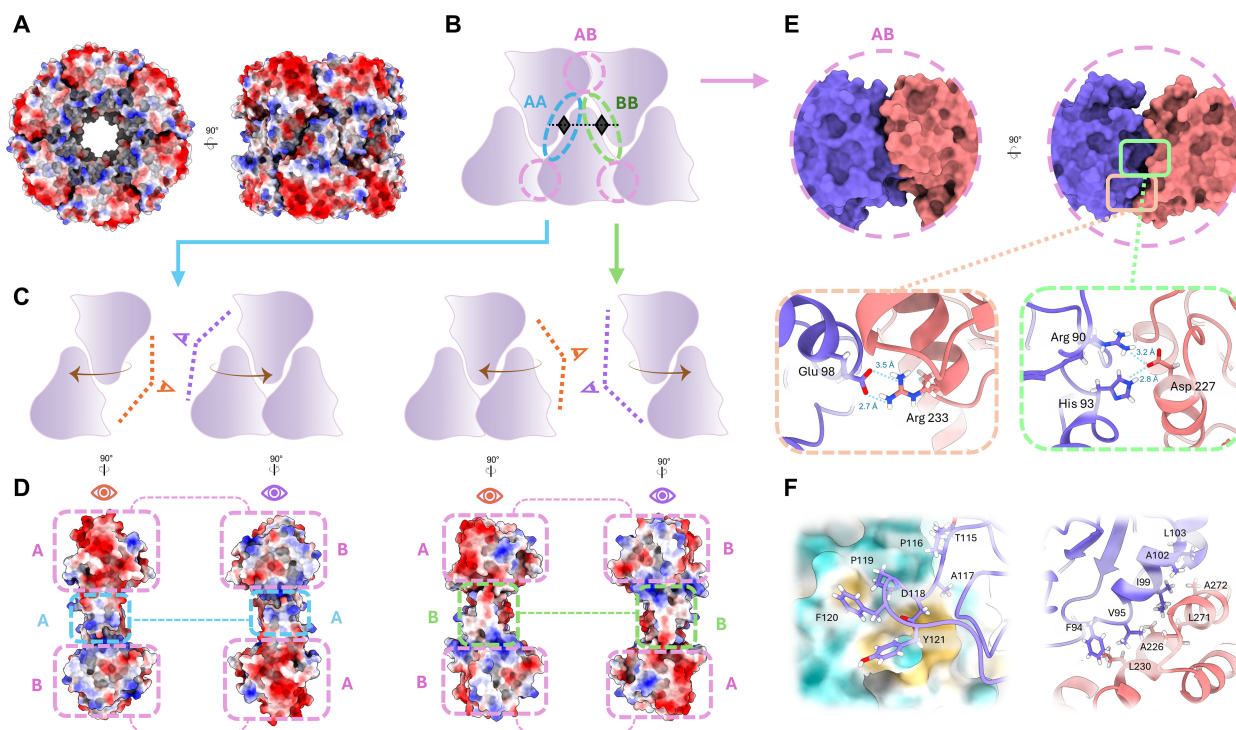


FIG 4 Charge and shape complementarity in oligomeric interfaces. (A) Top and side view of the space-filling representation of the PdxS dodecamer with surface electrostatics depicted in red (acidic) and blue (basic). (B) Cartoon schematic of intersubunit interactions defining each of three types of contacts, based on splitting the asymmetric monomer unit into two unique lateral faces, A and B. Each contact region is assigned a corresponding annotation color. (C) Illustration of views for each interface to be depicted in panel D. (D) Electrostatic surface coloring of complementary interfaces with each unique contact region annotated in its corresponding color as defined in (B). The AA and BB regions are self-contacts. Overall, acidic and basic patches along each surface are complementary. (E) Side and top views of the AB interface and select hydrogen-bonding partners contributing to its interface stabilization. (F) Representations of the hydrophobic loop (left panel) and hydrophobic helix-helix interactions (right panel) that stabilize lateral multimer binding. The left panel illustrates the B face of AB in a hydrophobicity surface coloring with yellow representing hydrophobic residues and the A face as a ribbon representation with projecting residues represented as sticks. The right panel shows the distribution of hydrophobic residues between face A's $\alpha 3$ helix and face B's $\alpha 7$ and $\alpha 8''$ helices.

swinging the intermediate to a second phosphate-binding site exterior to the TIM barrel (29). Based on comparison with previous crystal structures where PdxS was co-crystallized with R5P and various intermediate states (29), the density is most likely to belong to either bound R5P or an R5P Schiff-base intermediate. Docking an R5P moiety with the molecule oriented such that the phosphate group is positioned in the bulbous density nestled between $\beta 6$ and loops 156–164 then positions the oxygen 1 of R5P next to lysine 88, which is where the structurally equivalent lysine in other models (e.g., *Plasmodium berghei*) forms a Schiff-base intermediate with R5P (Fig. 5B through E). This also positions R5P to engage in a hydrogen bond with aspartate 31. One caveat to this assignment is that the oblong density adjacent to the putative phosphate density does, however, run in a perpendicular direction to the conserved R5P position. Repositioning the R5P to align with this perpendicular density could suggest the engagement of the second active site lysine (K156) in the formation of a second Schiff base with C5, but the density in our map for K156 is very strongly oriented in its alternate conformation facing away from the TIM barrel active site (Fig. 5C). It is clear, therefore, that multiple substrate states have likely been captured in this averaged density map.

Upon performing a phylogenetic analysis of the PLP synthase sequences from other biological systems for which crystal structures have been deposited, it is apparent that the *M. acetivorans* PdxS shares more sequence and structural similarity with the other deposited bacterial and eukaryotic PLP synthases than with the two archaeal systems so far determined (*Pyrococcus horikoshii* and *Methanococcus jannaschii*, Fig. 6) (37, 38). This

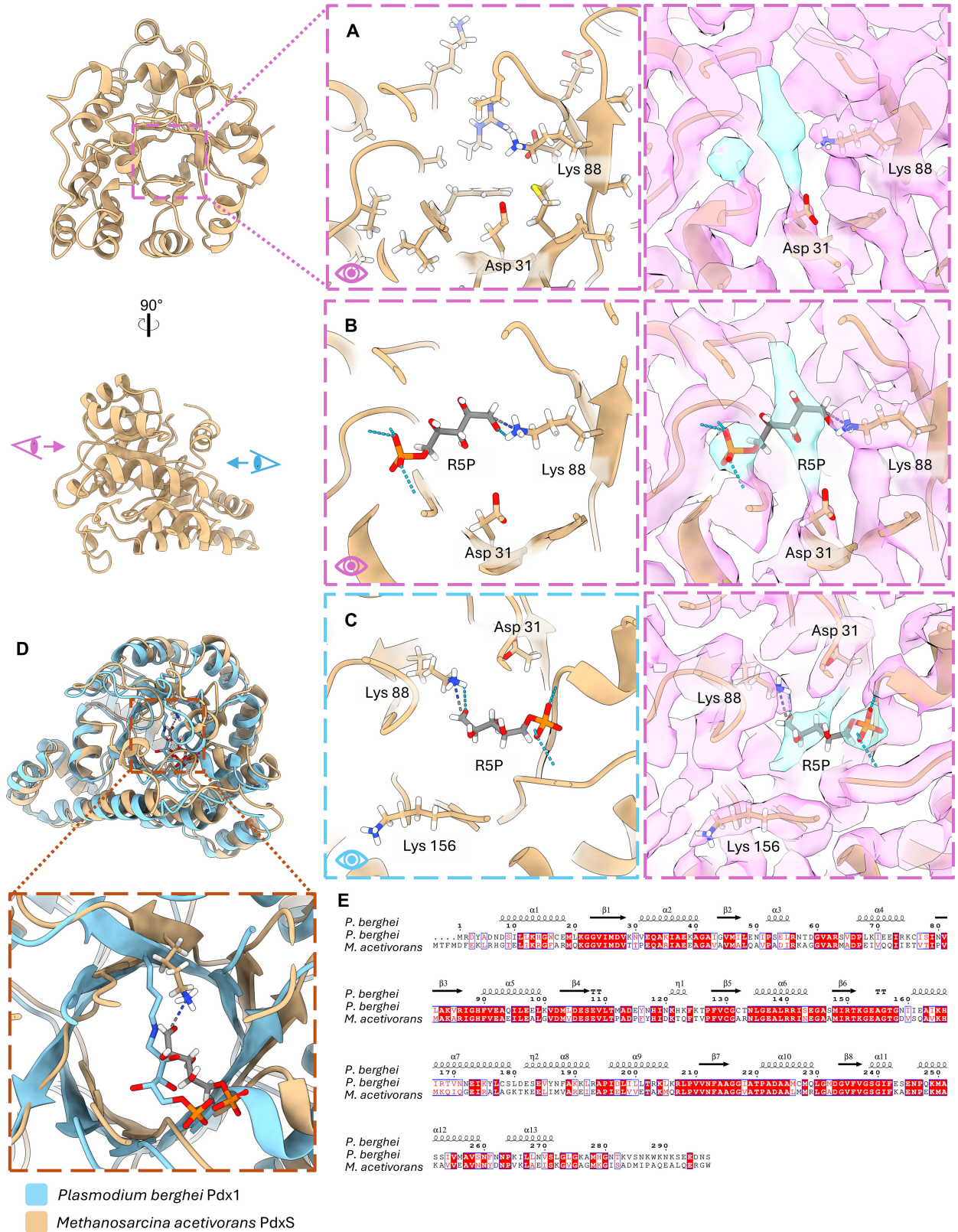


FIG 5 Active site densities and potential ligand modeling. (A) Ribbon model and corresponding cryo-EM density of Pdx in its active site. Nearby active site residues are represented as stick models. The elongated density and its neighboring spherical density colored in blue are not occupied by any nearby side chains. (B and C) Possible R5P configuration in active site density, with the phosphate group docked into the spherical density. A potential Schiff base is represented (Continued on next page)

Fig 5 (Continued)

with K88 as a dashed dark blue and gray line. Hydrogen bonds between the phosphate group and nearby backbone are depicted with a dashed light blue line. (C) Alternative view of the docked R5P, revealing the strong density for the second active site lysine (K156) in its extra-active site conformation. (D and E) Comparison of *Plasmodium berghei* and *M. acetivorans* Pdx structure (D) and primary sequence (E). The crystal structure for *P. berghei* bound with R5P is aligned with our structure, including our hypothesized R5P placement in the density.

is due to a 37-aa insertion that appears in various classes across several archaeal phyla, including Euryarchaeota (a kingdom shared with *M. acetivorans*), though notably this mainly appears within thermophilic genera (Fig. 6). The insertion gives rise to an extra alpha helix and beta sheets between equivalent helices to *M. acetivorans* α6' and α6''.

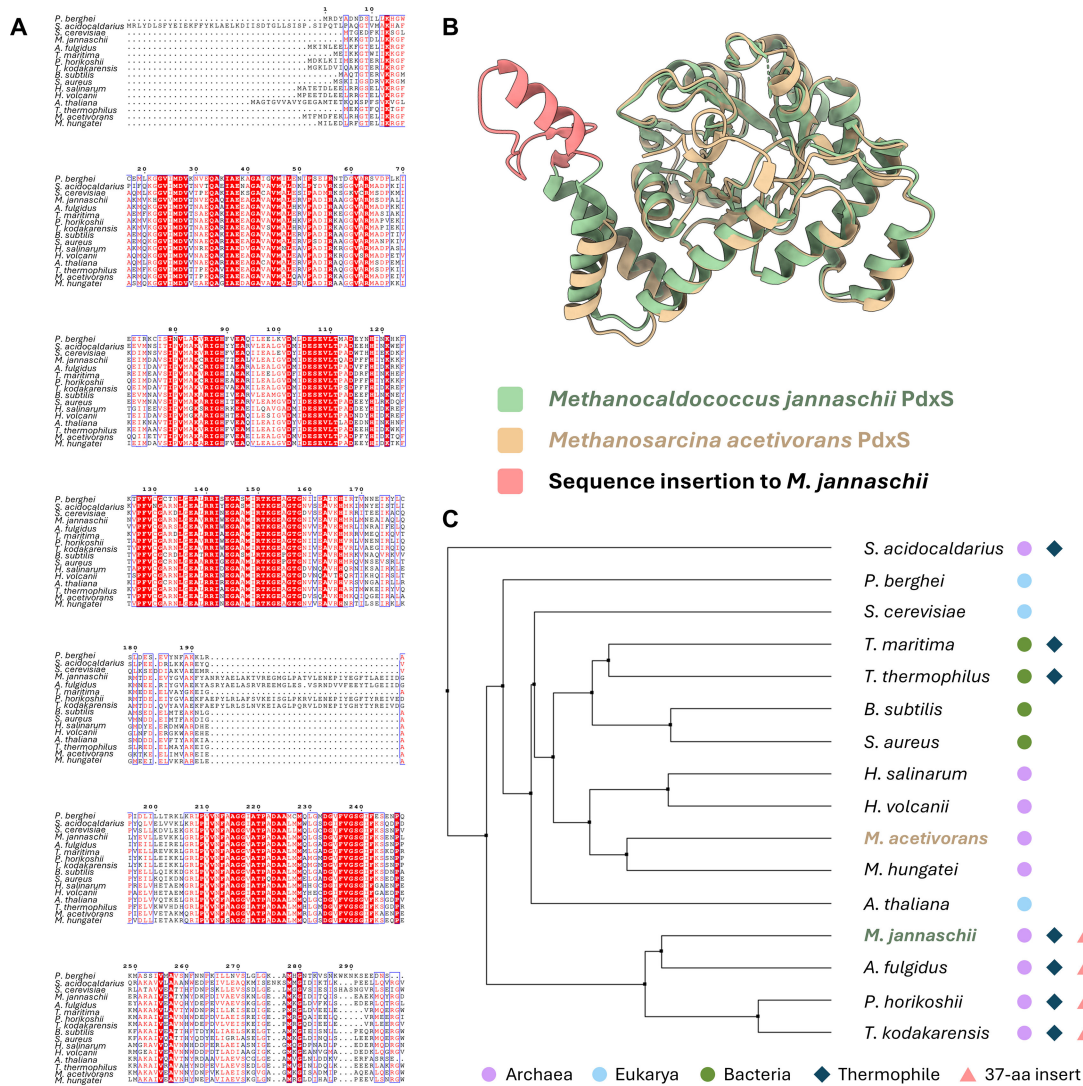


FIG 6 Sequence insertions in previously deposited archaeal PLP synthase subunits and other thermophilic archaeal species. (A) Sequence alignment between select species for which Pdx1/S structures have been deposited and additional thermophilic and mesophilic archaea. *P. horikoshii* and *M. jannaschii* both contain 37-aa inserts that are lacking in all other deposited Pdx1/S structures. (B) Structure alignment of the *M. jannaschii* and *M. acetivorans* PdxS monomers, with the sequence insertion in *M. jannaschii* highlighted in coral. (C) Average distance tree based on the ClustalW sequence alignment of primary sequences listed in (A). Domain, thermophilicity, and presence of a sequence insertion are denoted by corresponding symbols. Note that *Sulfolobus acidocaldarius* uniquely contains an N-terminal insertion instead, of also 37 amino acids.

DISCUSSION

In this study, we have determined the first native structure of the PdxS component of PLP synthase from an archaeon by cryo-EM and the *cryoID* approach. We demonstrate that the high degree of structural and sequence conservation witnessed in other crystal structures across bacteria and eukaryotes is also shared in this archaeal lineage. Intriguingly, our model shows greater divergence from the sole deposited archaeal PLP synthase structures at current than all other taxa for which the synthase structure is known. Given that those two species are extremophiles, the sequence insertion may have arisen due to their high-temperature conditions. Indeed, the additional helix and loops provide further electrostatic and hydrophobic interactions between subunits within hexamers, potentially further stabilizing the oligomerized state in the face of immense heat from these thermophilic organisms' environment (38). This is further supported by sequence alignments with other thermophilic and mesophilic archaea, which reveal that these sequence insertions are confined to PdxS sequences from archaeal thermophiles (Fig. 6A). Indeed, the top 100 hits from an NCBI BLAST search for this insertion contain only organisms from the extreme thermophilic Euryarchaeota lineage. Interestingly, the extreme thermophilic Proteoarchaeota (*Sulfolobus acidocaldarius*), the mesophilic Euryarchaeota, and bacterial extreme thermophiles lack the 37-aa insert noted in *P. horikoshii* and *M. jannaschii*, suggesting a relatively ancient evolutionary divergence from the Euryarchaeota thermophiles as well as indicating the existence of two structurally distinct classes of thermophilic PLP synthases (Fig. 6C).

Notably, density corresponding to bound PdxT was absent from our refined structure. Even after focused classification, evidence for any significant population of PdxT could not be observed. Prior crystal structures demonstrate 24-mer complexes but only under conditions of extremely high protein concentration, which are intrinsically required for crystallization. This suggests that natively, PdxT may bind the homo-dodecameric complex very weakly. It has been recently shown that PdxS can even function with direct ammonia binding, though ammonia delivery via the glutaminase promotes the efficiency of the synthase (30). Still, perhaps this indicates that the intracellular occupancy or expression of PdxT is often low. The aforementioned study also finds that in the *S. aureus* PLP synthase, the majority of particles form sub-stoichiometric complexes, with the number of homo-dodecamers outnumbering 24-mers, which supports the notion that the PdxS-PdxT interaction is generally unstable.

Our present work also demonstrates the promise of leveraging the *cryoID* workflow for native proteins harvested directly from cellular lysate. With only crude size-based fractionation, current single-particle analysis capabilities were successful in selecting, reconstructing, and identifying a native protein. Furthermore, we observe in our density some evidence of active site occupancy. Because of the multiple orientations of R5P or R5P intermediates this density accommodates, it is likely that we captured an amalgamation of many catalytic states. With further imaging and a greater particle set, these states can be classified and separately reconstructed. As such, this approach is particularly promising for studying native interactions involved in catalytic cycles of enzymes or for those systems which lack developed or established molecular biology techniques for extensive gene editing, endogenous protein purification or require unique archaeal cofactors. Structural characterization of biomolecules is crucial to (i) the development of targeted therapies in disease-causing systems and (ii) the pursuit of evolutionary descriptions of function. Therefore, it is essential to leverage the gross information retrieval capabilities of bottom-up cryo-EM toward understudied organisms, such as those in archaeal taxa.

MATERIALS AND METHODS

Sample preparation

M. acetivorans C2A (DSM 2834) was cultivated in 100 mL serum bottles with an N₂:CO₂ (80:20) headspace and 50 mL medium as previously described (41). Methanol was used as the sole carbon and energy supply (50 mM). Cells were harvested at OD ~1 by centrifugation at 10,000 × *g* for 10 min at 5°C in an IEC tabletop centrifuge. Cells were resuspended in chilled lysis buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, 5 mM MgCl, 2 mM dithiothreitol, and protease inhibitor cocktail) and lysed via 1.5–150 Watt adjustable sonication at 20–25 kHz (40% power using ten 2 s pulses at 4°C). Cell lysate was incubated with benzonase for 10 min before centrifugation (12,000 × *g* for 6 min at 4°C) to clarify nucleic acid and insoluble cellular debris. The supernatant was decanted, applied to a 10%–30% glycerol density gradient, and centrifuged at 110,000 × *g* for 20 hours at 4°C. Furthermore, 500 μL gradient fractions were evaluated by SDS-PAGE and negative stain electron microscopy, and selected fractions containing particles 8–12 nm in diameter were pooled together for subsequent grid preparation.

Electron microscopy of stained and vitrified samples

For negative stain EM screening, 3 μL aliquots were applied to glow-discharged formvar/carbon-coated grids (300 mesh, Ted Pella) and incubated for 1 min, before negative staining with 2% uranyl acetate. Samples were screened on an FEI Tecnai F20 electron microscope operated at 200 keV, and images were recorded on a TIETZ F415MP 16-megapixel CCD camera at a nominal magnification of 60,000×.

Prior to freezing, n-dodecyl-beta-maltoside was added to the sample to a final concentration of 0.0043% (one half the CMC) to alleviate protein denaturation arising from aggregation at the air-water interface. Subsequently, 3 μL aliquots were applied to glow-discharged holey carbon grids (Quantifoil R1.2/1.3 300 mesh, Ted Pella) and incubated for 10 s before automated blotting and flash freezing in liquid ethane with a Vitrobot Mark IV vitrification system (Thermo Fisher Scientific). Various freezing conditions—including chamber temperature, humidity, blotting time, blotting force, and drain time after blotting—were screened on the aforementioned instrument used during negative stain evaluation. Optimal conditions were obtained with a 90-s glow discharge, a chamber temperature of 4°C and 100% humidity, 5-s blotting time, blot force of 0, and 0-s drain time. Optimized cryo-EM grids were stored in liquid nitrogen until cryo-EM data collection.

Imaging for downstream processing was performed on a Titan Krios 300 kV electron microscope (Thermo Fisher Scientific) equipped with a Gatan Imaging Filter (GIF) Quantum LS and a Gatan K3 direct electron detector. A total of 14,890 movies were recorded with SerialEM in super-resolution mode at a nominal magnification of 81,000×, yielding a calibrated pixel size of 0.55 Å/pixel at the specimen level (42). The GIF slit width was set to 20 eV. Each movie contained 44 frames with an exposure time of 2.2 s per frame, giving a total estimated electron dose of 50 e⁻/Å².

Structure determination

Movie frames were aligned with patch motion correction in the cryoSPARC v4 suite, after which the calibrated pixel size was 1.1 Å/pixel (24). Defocus values were determined with the cryoSPARC patch CTF estimation job. Autonomous, reference-free particle picking was first performed with the Blob Picker tool, specifically selecting for particles in the size range of 80–120 Å diameter. Initially 14,627,499 particles were picked from 14,890 micrographs and boxed out by 300 × 300 pixels. After several iterations of 2D classification to remove junk classes, three distinct classes demonstrating alignment to 6 Å resolution emerged. These classes were used as templates to repick the data set, which after iterative 2D classification yielded 17 well-aligned classes of 614,949 particles. This particle subset was subsequently used to train the deep learning-based particle picker Topaz, after which the entire data set was again repicked with the trained model, yielding

3,724,083 particles (25). Another round of iterative 2D classification gave rise to many well-aligned, distinct 2D classes. Of these, several classes appeared to represent a 100 Å D6-symmetric structure. These were used for the training of a separate Topaz model, and the data set was picked a third time with a model trained on just these views. From the resultant 2D classes, a total of 99,971 particles were selected and subjected to *ab initio* reconstruction in cryoSPARC.

This initial model was used as a reference for nonuniform refinement, and with D6 symmetry enforced, the map was reconstructed to 3.59 Å. After 3D classification and iterative CTF refinement and 3D refinement, the final resolution was 3.38 Å based on the gold-standard FSC 0.143 criterion. B-factor, local resolution, and FSC curves were all calculated in cryoSPARC.

Structure identification with cryoID

After feeding *cryoID* our sharpened map, the program generated two query sequences built into two helical domains. The model was extended on both termini as permitted by the density, resulting in the following query sequences: (i) GGKYLKPLGLGKGLGGKGLL and (ii) LPGKLGKGLG. The top ranked output was UniProt Q8TQH6, or PdxS, which was further validated by docking the AlphaFold3-predicted structure into our density.

Atomic modeling, model refinement, and graphics visualization

Modeling was performed by docking the AlphaFold3-predicted structure of the PdxS monomer into the cryo-EM density and confirming that the output identification was rational. The model was then iteratively real space refined in the ISOLDE package of UCSF ChimeraX and validated in PHENIX, until outlier scores became negligible (43–45). A last round of real space refinement was performed in PHENIX and evaluated using the PDB validation server (46). Visualization of all maps and models was performed with UCSF ChimeraX. All sequence alignments were performed with ClustalW and visualized with ESPript 3 (47, 48). The average distance tree was constructed with JalView (49).

ACKNOWLEDGMENTS

This project is supported partly by grants from the US NIH (R01GM071940 to Z.H.Z.) and the National Science Foundation (1911781 to R.P.G.) and U.S. Department of Energy (DOE) Office of Science (BER) (contract DE-FC-02-02ER63421 to R.P.G.). We acknowledge the use of instruments at the Electron Imaging Center for Nanomachines supported by UCLA and grants from the NIH (1510OD018111) and the National Science Foundation (DBI-1338135 and DMR-1548924). Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number T32GM145388 and by the National Institute of Dental & Craniofacial Research of the National Institutes of Health under Award Number T90DE030860. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

We thank David Strugatsky for his assistance in cryo-EM data collection and Roman Aguirre and Shervin Nia for reading the manuscript.

Z.H.Z. and R.P.G. conceived the project; R.P.G., E.H., and A.A. prepared samples; A.A. and K.Z. recorded cryo-EM images, A.A. processed the data and built the atomic model; A.A., Z.H.Z., and R.P.G. interpreted results and wrote the paper; all authors edited and approved the paper.

AUTHOR AFFILIATIONS

¹Department of Microbiology, Immunology & Molecular Genetics, University of California Los Angeles, Los Angeles, California, USA

²California NanoSystems Institute, University of California Los Angeles, Los Angeles, California, USA

³Department of Chemistry and Biochemistry, University of California Los Angeles, Los Angeles, California, USA

⁴UCLA-DOE Institute, Los Angeles, California, USA

AUTHOR ORCID*s*

Angela Agnew  <http://orcid.org/0000-0002-4926-9217>

Kang Zhou  <http://orcid.org/0000-0002-3034-2808>

Robert P. Gunsalus  <http://orcid.org/0000-0002-1937-8412>

Z. Hong Zhou  <http://orcid.org/0000-0002-8373-4717>

FUNDING

Funder	Grant(s)	Author(s)
HHS NIH National Institute of General Medical Sciences (NIGMS)	R01GM071940, T32GM145388	Angela Agnew Z. Hong Zhou
National Science Foundation (NSF)	1911781	Robert P. Gunsalus
U.S. Department of Energy (DOE)	DE-FC-02-372 02ER63421	Robert P. Gunsalus
HHS NIH National Institute of Dental and Craniofacial Research (NIDCR)	T90DE030860	Angela Agnew

AUTHOR CONTRIBUTIONS

Angela Agnew, Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Ethan Humm, Methodology | Kang Zhou, Data curation, Methodology | Robert P. Gunsalus, Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review and editing | Z. Hong Zhou, Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Writing – review and editing

DIRECT CONTRIBUTION

This article is a direct contribution from Robert P. Gunsalus, a Fellow of the American Academy of Microbiology, who arranged for and secured reviews by Michael Adams, The University of Georgia, and Bo Hu, University of Texas Medical School at Houston.

DATA AVAILABILITY

The PdxS cryo-EM map has been deposited in the Electron Microscopy Data Bank under accession number [EMD-47202](#). The coordinates of the PdxS model have been deposited in the Protein Data Bank under accession number [9DVF](#). 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.

REFERENCES

- Woese CR, Kandler O, Wheelis ML. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* 87:4576–4579. <https://doi.org/10.1073/pnas.87.12.4576>
- Koskinen K, Pausan MR, Perras AK, Beck M, Bang C, Mora M, Schilhabel A, Schmitz R, Moissl-Eichinger C. 2017. First insights into the diverse human archaeome: specific detection of archaea in the gastrointestinal tract, lung, and nose and on skin. *MBio* 8:e00824-17. <https://doi.org/10.1128/mBio.00824-17>
- Metcalf WW, Zhang JK, Apolinario E, Sowers KR, Wolfe RS. 1997. A genetic system for Archaea of the genus *Methanosarcina*: liposome-mediated transformation and construction of shuttle vectors. *Proc Natl Acad Sci U S A* 94:2626–2631. <https://doi.org/10.1073/pnas.94.6.2626>
- Nayak DD, Metcalf WW. 2017. Cas9-mediated genome editing in the methanogenic archaeon *Methanosarcina acetivorans*. *Proc Natl Acad Sci U S A* 114:2976–2981. <https://doi.org/10.1073/pnas.1618596114>
- Kohler PRA, Metcalf WW. 2012. Genetic manipulation of *Methanosarcina* spp. *Front Microbiol* 3:259. <https://doi.org/10.3389/fmicb.2012.00259>
- Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R. 2008. Methanogenic archaea: ecologically relevant differences in energy

- conservation. *Nat Rev Microbiol* 6:579–591. <https://doi.org/10.1038/nrmicro1931>
7. Wang Y, Wegener G, Williams TA, Xie R, Hou J, Tian C, Zhang Y, Wang F, Xiao X. 2021. A methylotrophic origin of methanogenesis and early divergence of anaerobic multicarbon alkane metabolism. *Sci Adv* 7:eabj1453. <https://doi.org/10.1126/sciadv.abj1453>
 8. Borrel G, Brügère J-F, Gribaldo S, Schmitz RA, Moissl-Eichinger C. 2020. The host-associated archaeome. *Nat Rev Microbiol* 18:622–636. <https://doi.org/10.1038/s41579-020-0407-y>
 9. Candelieri F, Sola L, Raimondi S, Rossi M, Amaretti A. 2024. Good and bad dispositions between archaea and bacteria in the human gut: new insights from metagenomic survey and co-occurrence analysis. *Synth Syst Biotechnol* 9:88–98. <https://doi.org/10.1016/j.synbio.2023.12.007>
 10. Cai M, Tang X. 2022. Human archaea and associated metabolites in health and disease. *Biochemistry* 61:2835–2840. <https://doi.org/10.1021/acs.biochem.2c00232>
 11. Coker OO, Wu WKK, Wong SH, Sung JJY, Yu J. 2020. Altered gut archaea composition and interaction with bacteria are associated with colorectal cancer. *Gastroenterology* 159:1459–1470. <https://doi.org/10.1053/j.gastro.2020.06.042>
 12. Stach K, Stach W, Augoff K. 2021. Vitamin B6 in health and disease. *Nutrients* 13:3229. <https://doi.org/10.3390/nu13093229>
 13. Abad AND, Seshadri K, Ohashi M, Delgadillo DA, de Moraes LS, Nagasawa KK, Liu M, Johnson S, Nelson HM, Tang Y. 2024. Discovery and characterization of pyridoxal 5'-phosphate-dependent cycloleucine synthases. *J Am Chem Soc* 146:14672–14684. <https://doi.org/10.1021/jacs.4c02142>
 14. di Salvo ML, Safo MK, Contestabile R. 2012. Biomedical aspects of pyridoxal 5'-phosphate availability. *Front Biosci Elite Ed* 4:897–913. <https://doi.org/10.2741/E428>
 15. Zhang P, Tsuchiya K, Kinoshita T, Kushiya H, Suidasari S, Hatakeyama M, Imura H, Kato N, Suda T. 2016. Vitamin B6 prevents IL-1 β protein production by inhibiting NLRP3 inflammasome activation. *J Biol Chem* 291:24517–24527. <https://doi.org/10.1074/jbc.M116.743815>
 16. Wan Z, Zheng J, Zhu Z, Sang L, Zhu J, Luo S, Zhao Y, Wang R, Zhang Y, Hao K, Chen L, Du J, Kan J, He H. 2022. Intermediate role of gut microbiota in vitamin B nutrition and its influences on human health. *Front Nutr* 9:1031502. <https://doi.org/10.3389/fnut.2022.1031502>
 17. Rosenberg J, Ischebeck T, Commichau FM. 2017. Vitamin B6 metabolism in microbes and approaches for fermentative production. *Biotechnol Adv* 35:31–40. <https://doi.org/10.1016/j.biotechadv.2016.11.004>
 18. Mooney S, Leuendorf JE, Hendrickson C, Hellmann H. 2009. Vitamin B6: a long known compound of surprising complexity. *Molecules* 14:329–351. <https://doi.org/10.3390/molecules14010329>
 19. Oka T. 2001. Modulation of gene expression by vitamin B6. *Nutr Res Rev* 14:257–266. <https://doi.org/10.1079/NRR200125>
 20. Saibil HR. 2022. Cryo-EM in molecular and cellular biology. *Mol Cell* 82:274–284. <https://doi.org/10.1016/j.molcel.2021.12.016>
 21. Ho C-M, Li X, Lai M, Terwilliger TC, Beck JR, Wohlschlegel J, Goldberg DE, Fitzpatrick AWP, Zhou ZH. 2020. Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. *Nat Methods* 17:79–85. <https://doi.org/10.1038/s41592-019-0637-y>
 22. Xia X, Liu S, Zhou ZH. 2022. Structure, dynamics and assembly of the ankyrin complex on human red blood cell membrane. *Nat Struct Mol Biol* 29:698–705. <https://doi.org/10.1038/s41594-022-00779-7>
 23. Liu S, Xia X, Calvo E, Zhou ZH. 2023. Native structure of mosquito salivary protein uncovers domains relevant to pathogen transmission. *Nat Commun* 14:899. <https://doi.org/10.1038/s41467-023-36577-y>
 24. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14:290–296. <https://doi.org/10.1038/nmeth.4169>
 25. Bepler T, Morin A, Rapp M, Brasch J, Shapiro L, Noble AJ, Berger B. 2019. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* 16:1153–1160. <https://doi.org/10.1038/s41592-019-0575-8>
 26. Zein F, Zhang Y, Kang YN, Burns K, Begley TP, Ealick SE. 2006. Structural insights into the mechanism of the PLP synthase holoenzyme from *Thermotoga maritima*. *Biochemistry* 45:14609–14620. <https://doi.org/10.1021/bi061464y>
 27. Strohmeier M, Raschle T, Mazurkiewicz J, Rippe K, Sinning I, Fitzpatrick TB, Tews I. 2006. Structure of a bacterial pyridoxal 5'-phosphate synthase complex. *Proc Natl Acad Sci U S A* 103:19284–19289. <https://doi.org/10.1073/pnas.0604950103>
 28. Novikova IV, Zhou M, Du C, Parra M, Kim DN, VanAernum ZL, Shaw JB, Hellmann H, Wysocki VH, Evans JE. 2021. Tunable heteroassembly of a plant pseudoenzyme-enzyme complex. *ACS Chem Biol* 16:2315–2325. <https://doi.org/10.1021/acscchembio.1c00475>
 29. Rodrigues MJ, Windeisen V, Zhang Y, Guédez G, Weber S, Strohmeier M, Hanes JW, Royant A, Evans G, Sinning I, Ealick SE, Begley TP, Tews I. 2017. Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis. *Nat Chem Biol* 13:290–294. <https://doi.org/10.1038/nchembio.2273>
 30. Barra A.L.C, Ullah N, Brognaro H, Gutierrez RF, Wrenger C, Betzel C, Nascimento AS. 2024. Structure and dynamics of the staphylococcal pyridoxal 5-phosphate synthase complex reveal transient interactions at the enzyme interface. *J Biol Chem* 300:107404. <https://doi.org/10.1016/j.jbc.2024.107404>
 31. Neuwirth M, Strohmeier M, Windeisen V, Wallner S, Deller S, Rippe K, Sinning I, Macheroux P, Tews I. 2009. X-ray crystal structure of *Saccharomyces cerevisiae* Pdx1 provides insights into the oligomeric nature of PLP synthases. *FEBS Lett* 583:2179–2186. <https://doi.org/10.1016/j.febslet.2009.06.009>
 32. Guédez G, Hipp K, Windeisen V, Derrer B, Gengenbacher M, Böttcher B, Sinning I, Kappes B, Tews I. 2012. Assembly of the eukaryotic PLP-synthase complex from *Plasmodium* and activation of the Pdx1 enzyme. *Structure* 20:172–184. <https://doi.org/10.1016/j.str.2011.11.015>
 33. Manzoku M, Ebihara A, Yokoyama S, Kuramitsu S, RIKEN Structural Genomics/Proteomics Initiative (RSGI). 2007. Crystal structure of pyridoxine biosynthesis protein from *Thermus thermophilus* HB8. <https://doi.org/10.2210/pdb2ZBT/pdb>
 34. Smith AM, Brown WC, Harms E, Smith JL. 2015. Crystal structures capture three states in the catalytic cycle of a pyridoxal phosphate (PLP) synthase. *J Biol Chem* 290:5226–5239. <https://doi.org/10.1074/jbc.M114.626382>
 35. Zhu J, Burgner JW, Harms E, Belitsky BR, Smith JL. 2005. A new arrangement of (β / α) $_8$ barrels in the synthase subunit of PLP synthase. *J Biol Chem* 280:27914–27923. <https://doi.org/10.1074/jbc.M503642200>
 36. Kim S, Kim KJ. 2013. Crystal structure of *Mycobacterium tuberculosis* Rv2606c: a pyridoxal biosynthesis lyase. *Biochem Biophys Res Commun* 435:255–259. <https://doi.org/10.1016/j.bbrc.2013.04.068>
 37. Manzoku M, Ebihara A, Yokoyama S, Kuramitsu S, RIKEN Structural Genomics/Proteomics Initiative (RSGI). 2007. Crystal structure of pyridoxine biosynthesis protein from *Methanocaldococcus jannaschii*. <https://doi.org/10.2210/pdb2y2r/pdb>
 38. Matsuura A, Yoon JY, Yoon HJ, Lee HH, Suh SW. 2012. Crystal structure of pyridoxal biosynthesis lyase PdxS from *Pyrococcus horikoshii*. *Mol Cells* 34:407–412. <https://doi.org/10.1007/s10059-012-0198-8>
 39. UniProt Consortium. 2023. UniProt: the universal protein knowledge-base in 2023. *Nucleic Acids Res* 51:D523–D531. <https://doi.org/10.1093/nar/gkac1052>
 40. Barra ALC, Ullah N, Morão LG, Wrenger C, Betzel C, Nascimento AS. 2021. Structural dynamics and perspectives of vitamin B6 biosynthesis enzymes in *Plasmodium*: advances and open questions. *Front Cell Infect Microbiol* 11:688380. <https://doi.org/10.3389/fcimb.2021.688380>
 41. Sowers KR, Boone JE, Gunsalus RP. 1993. Disaggregation of *Methanosarcina* spp. and growth as single cells at elevated osmolarity. *Appl Environ Microbiol* 59:3832–3839. <https://doi.org/10.1128/aem.59.11.3832-3839.1993>
 42. Mastronarde DN. 2003. SerialEM: a program for automated tilt series acquisition on tecnai microscopes using prediction of specimen position. *Microsc Microanal* 9:1182–1183. <https://doi.org/10.1017/S1431927603445911>
 43. Croll TI. 2018. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* 74:519–530. <https://doi.org/10.1107/S2059798318002425>
 44. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. 2023. UCSF ChimeraX: tools for structure building and analysis. *Protein Sci* 32:e4792. <https://doi.org/10.1002/pro.4792>
 45. Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant

- MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, Terwilliger TC, Urzhumtsev AG, Videau LL, Williams CJ, Adams PD. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* 75:861–877. <https://doi.org/10.1107/S2059798319011471>
46. Berman H, Henrick K, Nakamura H. 2003. Announcing the worldwide Protein Data Bank. *Nat Struct Biol* 10:980. <https://doi.org/10.1038/nsb1203-980>
47. Madeira F, Madhusoodanan N, Lee J, Eusebi A, Niewielska A, Tivey ARN, Lopez R, Butcher S. 2024. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res* 52:W521–W525. <https://doi.org/10.1093/nar/gkae241>
48. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42:W320–W324. <https://doi.org/10.1093/nar/gku316>
49. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>