Supplementary Information for

Structural basis of light-induced redox regulation in the Calvin-Benson cycle

Ciaran R. McFarlane, Nita R. Shah, Burak V. Kabasakal, Blanca Echeverria Perez
Charles A.R. Cotton, Doryen Bubeck, James W. Murray

Corresponding authors: Doryen Bubeck and James W. Murray
Email: d.bubeck@imperial.ac.uk (D.B.) and j.w.murray@imperial.ac.uk (J.W.M.)

This PDF file includes:

Supplementary Information Text

MATERIALS AND METHODS

GAPDH and CP12 expression and purification

The GAPDH (*tll1466*) and CP12 genes (*tsr1820*) were amplified by PCR from *Thermosynechococcus elongatus* BP-1 (a gift from Matthias Rögner, Rühr-Universität Bochum) genomic DNA (primers CP12-FOR TTCCGCGTGGATCCATGAGTAATCTCGAG; CP12-REV GCTGCAGATCTCGAGTTATTAGTCGTCG; GAPDH-FOR GTCCGCGTGGATCCATGGTTAGAGTCGC; GAPDH-REV GCCGGATCCTCGAGCTACTAAGCCCAGTGG), and cloned by restriction-ligation into a modified prSET-A expression vector(1) with a thrombin-cleavable hexahistidine tag.

Escherichia coli KRX (Promega) transformed with the GAPDH or CP12 plasmids were grown in Terrific Broth with 100 µg/ml ampicillin at 37°C to OD_{600} 0.6-0.8 followed by overnight induction at 18°C with 0.1% (w/v) rhamnose. Cells were harvested by centrifugation, resuspended in 50 mM Tris-HCl, 150 mM NaCl, pH 7.9 and lysed by sonication. Insoluble material was removed by centrifugation for 1 hour at 80,000 g. Protein was purified by Ni-NTA affinity chromatography. The 50 ml of lysate supernatant was incubated with 3 ml of Ni-NTA resin (Qiagen). The resin was washed three times with 15 ml of 50 mM Tris-HCl, 150 mM NaCl, 30 mM imidazole followed by elution in the same buffer with 300 mM imidazole.

The eluted proteins were cleaved overnight at 4°C with 100 units of thrombin (Sigma-Aldrich), concentrated to 1 ml (Vivaspin concentrators, Sartorius) and further purified by size-exclusion chromatography (Proteosec 6-600 column, Generon). The flow rate was 1 ml/min (Äkta Purifier) with a mobile phase of 50 mM Tris-HCl, 150 mM NaCl, pH 7.9. The GAPDH-CP12 complex was made by mixing with a molar ratio of 10 CP12 per GAPDH monomer, followed by size-exclusion chromatography.

Purification of ternary complex

PRK was purified from *T. elongatus* cell lysate. Throughout the purification, PRK-containing fractions were identified by assaying PRK activity. *T. elongatus* was grown in batches of 3 l in DTN media (2) and 20 l were used per preparation. Cultures were grown to OD_{720} 1.0 at 45°C under constant illumination at 20 µmol m^{-2} s^{-1}. Cells were pelleted and washed in 50 mM Tris-HCl, 150 mM NaCl, 10 mM $MgCl_2$, pH 7.9 and lysed with a cell disruptor at 25 kpsi (Constant systems T5) followed by centrifugation for 1 hour at 185,000 g.

The lysate supernatant (~150-200 ml) was dialyzed for 16 hours against 50 mM Tris-HCl pH 7.9. The dialyzed sample was 0.22 µm filtered and loaded onto a DEAE-s (Toyopearl) 75 ml anion exchange column using a peristaltic pump at 2 ml/min (P-1 GE Healthcare). Protein was eluted with an NaCl gradient from 0 to 150 mM over seven column volumes (Äkta Purifier). The anion exchange fractions with PRK activity were pooled and dialyzed for 16 hours against 10 mM phosphate buffer pH 6.5. Protein was then loaded onto a 35 ml hydroxyapatite column (Bio-rad) using a peristaltic pump at 2 ml/min (P-1 GE Healthcare). Protein was eluted using a gradient to 250 mM phosphate buffer pH 7.0 over five column volumes.

PRK-containing fractions from the hydroxyapatite elution were then further purified using size-exclusion chromatography (Proteosec 6-600, Generon). This was carried out with a flow rate of 0.5 ml/min and a mobile phase of 50 mM Tris-HCl, 150 mM NaCl, 10 mM $MgCl_2$, pH 7.9. The yield of partially purified PRK was ~500 µg protein, which was concentrated to about 1 mg/ml, the main contaminant was allophycocyanin, about 50% by mass. The partially purified PRK was incubated overnight with a 5-fold molar excess of GAPDH-CP12 binary complex with 1 mM NAD^+, 1 mM trans-1,2-dihydroxy-4,5-dithane (oxidized DTT) and 1 mM ADP (Sigma-Aldrich). The ternary complex was purified using size-exclusion chromatography (Proteosec 6-600 column, Generon). This was carried out with a flow rate of 0.5 ml/min (ÄKTA Purifier) and a mobile phase of 50 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, pH 7.9. The eluted complex was identified using SDS-PAGE and ESI-QUAD-TOF

mass spectroscopy(3). Protein concentration was determined using Pierce-BCA assays (ThermoFischer).

Enzyme activity assays

GAPDH activity was measured by following the reduction of $NAD(P)^+$ at 340 nm with DL-glyceraldehyde-3-phosphate and arsenate. Each 200 µl reaction contained 5 mM DL-glyceraldehyde-3-phosphate (Sigma-Aldrich), 2.5 mM sodium arsenate (Sigma-Aldrich) in 10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, pH 7.9. GAPDH activity to measure the oxidation of NAD(P)H was measured using a coupled enzyme reaction that produced the GAPDH substrate 1,3-bisphosphoglycerate. Each 200 µl reaction contained 5 mM 3-phosphoglycerate, 5 mM ATP and 1 unit of yeast phosphoglycerate kinase (Sigma) in 10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, pH 7.9. Reactions were performed at 25 °C in a quartz cuvette (Hellma) and absorbance at 340 nm was measured using a 1 nm slit-length and 1 Hz sampling rate (Shimadzu).

PRK activity was measured by coupling the production of ADP product to the reduction of NAD^+ using ADP-hexokinase and glucose-6-phosphate dehydrogenase (SI Appendix, Fig. S5A) and following increase in A_{340}. In each reaction, the coupling enzyme reagent contained 10 µg/ml ADP-hexokinase, 4 units of glucose-6-phosphate dehydrogenase (Sigma-Aldrich), 10 mM glucose and 500 µM NAD^+ in 10 mM Tris, 50 mM NaCl, 10 mM $MgCl_2$, pH 7.9. Reducing conditions for assays were achieved by pre-incubation of the sample for 15 min with 10 mM DTT.

The ADP-dependent hexokinase gene (WP_004069859) from *Thermococcus litoralis* DSM-5473 gDNA (Deutsche Sammlung von Mikroorganismen und Zellkulturen)(4) was PCR amplified and cloned by Gibson assembly into the modified pRSET-A vector and purified by affinity chromatography similarly to GAPDH and CP12.

For the purification of PRK, activity was measured at 25 °C in a 150 µl reaction containing the coupled enzyme assay mix with 5 µl of protein sample. Activity was measured with 100 µM ATP plus 100 µM ribulose-5-phosphate (Ru5P, Sigma-Aldrich) and corrected for background ATPase activity by measuring the same sample with 100 µM ATP and no Ru5P (SI Appendix, Fig. S5B).

Michaelis-Menten kinetics for both GAPDH and PRK were determined by measuring the initial reaction rates for 30 s at increasing concentrations of substrate in the presence of the equivalent 55 nmols of GAPDH monomer and 25 nmols PRK monomer. Traces were plotted and analyzed using Origin (OriginLab). Each trace was measured for 100 s with the initial rate measured by fitting the first 30 s to a linear curve using a least-squares function. Activity at each substrate concentration was measured in triplicate and plotted with standard deviation.

Native-gel electrophoresis

Native-gel PAGE was performed on GAPDH-CP12-PRK samples by incubating 20 µl of 100 µg/ml aliquots overnight in combinations of DTT, NADPH, ADP, each at 2.5 mM. Samples were loaded onto 4-20% gradient Tris-Glycine gels (NuPAGE, Invitrogen). The loading buffer contained 62.5 mM Tris-HCl, 0.05% bromophenol blue, 10% glycerol, pH 6.8. The running buffer was 25 mM Tris base, 192 mM glycine. Gels were loaded and run on an XCell SureLock Mini-Cell electrophoresis system (Thermo-Fischer) at constant-current set at 20 mA for 4 hours followed by staining with Instant Blue Coomassie (Expedion). Protein bands were identified using ESI-QUAD-TOF mass spectroscopy performed by the St. Andrews mass spectrometry service (3).

Protein crystallization

GAPDH and GAPDH-CP12 complexes at 50 mg/ml were screened for crystallization with sitting drop vapor diffusion at 17°C using commercial sparse matrix screens JCSG-plus, Wizard 3 and 4 (Molecular Dimensions), PEG ION 1 and 2 (Hampton) commercial screens with a Mosquito robot (TTPLabTech). Some hits were further optimized in manually set up hanging drop vapor diffusion experiments. For the structure of GAPDH with four CP12 bound, GAPDH-CP12 was incubated with a further 10-fold molar excess of CP12 to GAPDH monomer. The structure of GAPDH with NADPH bound was obtained by mixing GAPDH with 5 mM of glyceraldehyde 3-phosphate, 3-phosphoglycerate, $NADP^+$, NADPH and 1 unit/ml bisphosphoglycerate kinase prior to crystallization

turning over the bound NAD^+ in GAPDH, replacing it with $NADP^+$. Crystallization conditions are given in SI Appendix, Table 1. Crystals were cryo-protected in the mother liquor with 30% volume PEG 400 added, then flash-cooled in liquid nitrogen.

X-ray structure determination

X-ray diffraction data were collected at the Diamond Light Source synchrotron, and processed with the xia2 pipeline(5). GAPDH crystal structures were solved by molecular replacement in Phaser(6) with a model based on our earlier *T. elongatus* GAPDH structure (PDB 4BOY). The models were rebuilt in Coot(7) with cycles of refinement in phenix.refine(8), and validated with MolProbity(9). The CP12 models could be built into the difference density. Data collection and refinement information are in SI Appendix, Table S1. Coordinates and structure factors were deposited in the PDB with accessions 6GFP, 6GFQ, 6GFR, 6GFO, 6GG7, 6GHR, 6GHL. Structure figures were prepared using PyMol (Shrödinger) and UCSF Chimera(10).

Negative stain EM

Negative stain EM was used to assess sample quality during purification. 2.5 µl of the GAPDH-CP12-PRK complex at ~30 µg/ml was applied to glow-discharged carbon-coated copper grids (400 mesh) and stained with 2% uranyl acetate. Data were collected on a 120 keV Tecnai T12 microscope with a pixel size of 2.6 Å.

Sample preparation and cryoEM data acquisition

Immediately following size-exclusion chromatography, 2.5 µl GAPDH-CP12-PRK complex at 300 µg/ml in 50 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM NAD^+, 1 mM ADP, 1 mM trans-1,2-dihydroxy-4,5-dithane was adsorbed to a glow-discharged holey carbon grid (C-Flat 1.2/1.3 grid or Quantifoil R2/2), which was overlaid with a thin layer of amorphous carbon. Using a Vitrobot mark III (Thermo Fisher Scientific), grids were blotted for 3.5 seconds at “blot force” 3 and plunge frozen in liquid ethane cooled to liquid nitrogen temperature. For initial model generation, data were collected on a 200 keV Tecnai F20 electron microscope (Thermo Fisher Scientific) fitted with a Falcon II direct electron detector (Thermo Fisher Scientific). Images were recorded at a pixel size of 2.05 Å and with a defocus range of 4 to 6 µm underfocus. Electron micrograph movies for the high resolution reconstruction were collected on a 300 keV Titan Krios (Thermo Fisher Scientific) equipped with a Quantum K2 Summit direct electron detector (Gatan). Data were collected in counting mode, and image stacks were recorded at 6 frames per second for 9 seconds with an accumulated dose of 44 $e^-/Å^2$. Images were recorded at pixel size of 1.047 Å and with a defocus range of -1.75 to -3.5 µm. All data were acquired using image acquisition software EPU (Thermo Fisher Scientific).

CryoEM image processing

An initial model was generated using data collected at 200 keV on a Tecnai F20 electron microscope. Contrast transfer function (CTF) parameters were estimated using CTFFIND4 (11), and micrographs were curated based on figure of merit value and ice quality. 14,000 particles were manually picked from the remaining micrographs, and 2,708 particles were discarded after 2D classification and selection. The initial model generation tool within RELION (12) was used to create a model which was refined with C2 symmetry constraints to a resolution of 17.4 Å. Additional rounds of 2D classification enabled the selection of a more homogenous subset, and the remaining 5,828 particles were used to refine the particle orientations with D2 symmetry constraints. The resulting 11.2 Å resolution reconstruction was then used as an initial model for the refinement of Krios-collected data.

For the high resolution cryoEM reconstruction, 4880 micrograph movies were collected at 300 keV on the Titan Krios electron microscope. Electron micrograph movie frames were aligned by MotionCor2 (13), discarding the first and last frames. CTF parameters were estimated using CTFFIND4 (11). Any movies containing low figure of merit scores, substantial drift, low contrast, thick ice, or crystalline ice were discarded from further analysis. 2D class averages from cryoEM data were used as templates for the automated particle picking tool in RELION. 585,526 initial particles were subjected to multiple rounds of 2D classification and selection. The initial model derived from data collected at 200 keV was strongly low-pass filtered (40 Å) to prevent model bias and used as a starting model for a gold-standard 3D autorefinement of images. These orientations served as the

starting point for tracking beam-induced movement of individual particles, which was corrected using particle polishing within RELION. Additional rounds of 2D classification after particle polishing were performed to improve the sample homogeneity. 197,212 selected particles contributed to a reconstruction which was refined using D2 symmetry to 3.9 Å. To assess the extent of conformational and stoichiometric heterogeneity of the GAPDH-CP12-PRK complex, we also performed a refinement without symmetry constraints. The resulting reconstruction refined to 4.3 Å average resolution and revealed heterogeneity across the four CP12-PRK binding sites (Fig. S8). Particles were classified into 6 groups allowing orientations to be refined after each iteration. The two most populated classes with a total of 130,737 particles, which also reported the highest resolution estimates, were combined in a 3D autorefinement imposing either C1 or D2 symmetry. The C1 symmetry refined density map indicated that all four CP12-PRK interfaces were fully occupied. The symmetry of the D2-refined particle dataset was then expanded such that each particle was assigned four orientations that corresponded to its D2 symmetry redundant views. This expanded data set was 3D autorefined with no symmetry imposed and limited to local search parameters only, thus ensuring the expanded particles could not revert to another redundant orientation. The overall resolution of 4.0 Å was calculated using the mask-corrected Fourier shell correlation (FSC) with local resolution ranging from 3.6 to 6.2 Å (Fig. S7). Though the original D2-symmetrized reconstruction generated from 197,212 particles had a slightly higher average resolution and very similar local resolution estimates when compared to the symmetry expanded reconstruction; the latter of the two maps has better defined features (Fig. S12), and thus chosen for model building and deposition. The final reconstruction was deposited in the EMDB with accession code EMD-0071. Handedness of the reconstruction was determined by fitting the chiral GAPDH crystal structure into the 3.9 Å map using the 'fit in map' tool in Chimera(10). There was a clear difference between correlation coefficients when comparing the different hands (CC = 0.86 versus 0.65).

CryoEM model Building and refinement

The crystallographic GAPDH-CP12 model was placed in the B-factor sharpened, local resolution-filtered map with Chimera, and the CP12 N-terminus and linker was refitted into the density. A PRK model based on an archaeal PRK (PDB 5B3F) (14) was placed in the map and extensively rebuilt and extended with the *T. elongatus* sequence, and then updated in the light of a cyanobacterial PRK structure becoming available (PDB 6HZK) (15).

The asymmetric unit was rebuilt in COOT (7), expanded to complete the pseudo-D2 symmetry and refined with phenix real space refine (16) using rotamer restraints. The final model was validated with MolProbity (9). A locally sharpened map made with phenix.auto_sharpen (17) was also used in model building. The final model is full-length in the three proteins, however side chains for less well ordered regions were omitted in the final coordinates, with most missing from the N-terminal regions of PRK. Parameters for the cryoEM ternary complex model are given in SI Appendix, Table 2. Coordinates for the ternary complex were deposited in the PDB with accession 6GVE.

Data Availability

The atomic coordinates and structure factors for GAPDH-NAD^+, GAPDH-$NADP^+$, GAPDH-CP12-conf1-conf2, GAPDH-CP12-conf3, GAPDH-CP12-conf4, GAPDH-$CP12_2$ and GAPDH-$CP12_4$ have been deposited in the Protein Data Bank under accession codes 6GFR, 6GFP, 6GFO, 6GHR, 6GHL, 6GFQ and 6GG7 respectively. The GAPDH-CP12-PRK cryoEM map has been deposited in the Electron Microscopy Databank under EMDB-0071, and the atomic model has been deposited in the PDB, accession 6GVE.

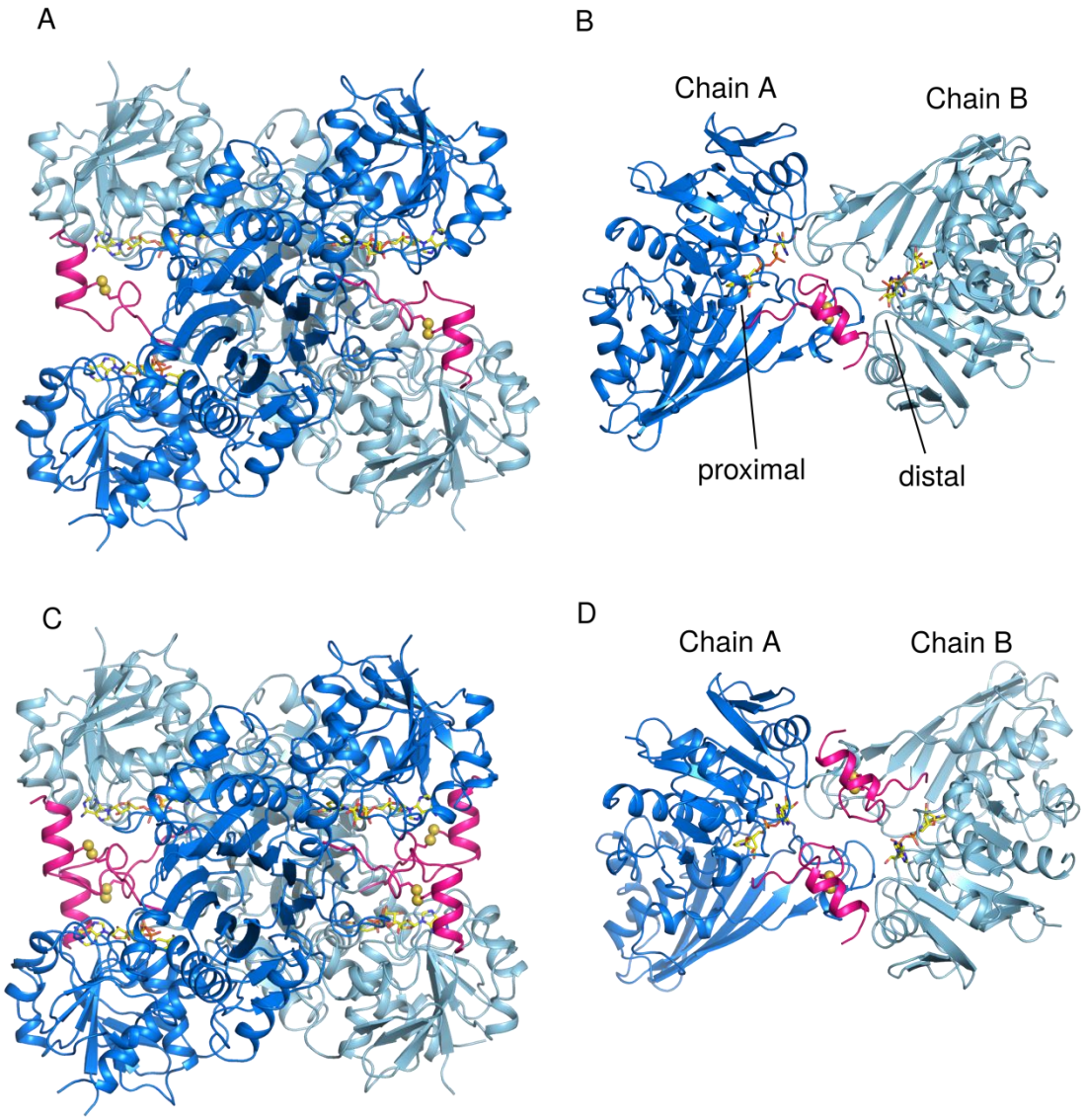


Fig. S1 GAPDH with different numbers of CP12 bound.

(A) GAPDH tetramer with two CP12 bound (drawn from 6GFQ), cartoon models colored by chain, GAPDH chain A (light blue), GAPDH chain B (marine) and CP12 (pink). (B) View of dimer from (A), showing proximal CP12 binding (chain A), and distal (chain B). (C) GAPDH tetramer with four CP12 bound, produced by incubating GAPDH-CP12 with excess CP12 (drawn from 6GG7). (D) View of dimer from (C), showing two CP12 bound.

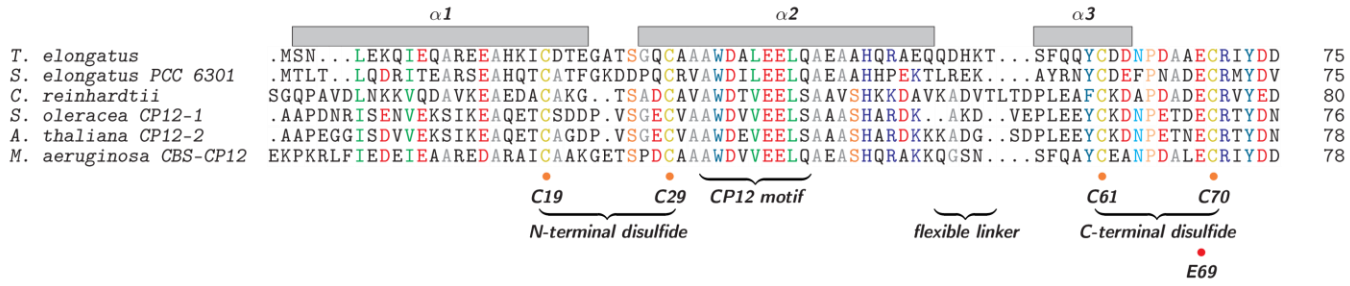


Fig. S2 Alignment of representative CP12 sequences.

CP12 sequences from *T. elongatus*, *S. elongatus* PCC 6301*, A. thaliana*, *S. oleracea*, *C. reinhardtii* and the CP12 domain of the CBS-CP12 protein from *Microcystis aeruginosa* PCC 7806 (18). The leaf-associated CP12-1 isoform has been chosen from the higher plants. CP12 secondary structure is marked above the alignment, and the two disulfides below, with the CP12 motif, flexible linker and conserved Glu69 that would clash with $NADP^+$ in GAPDH.

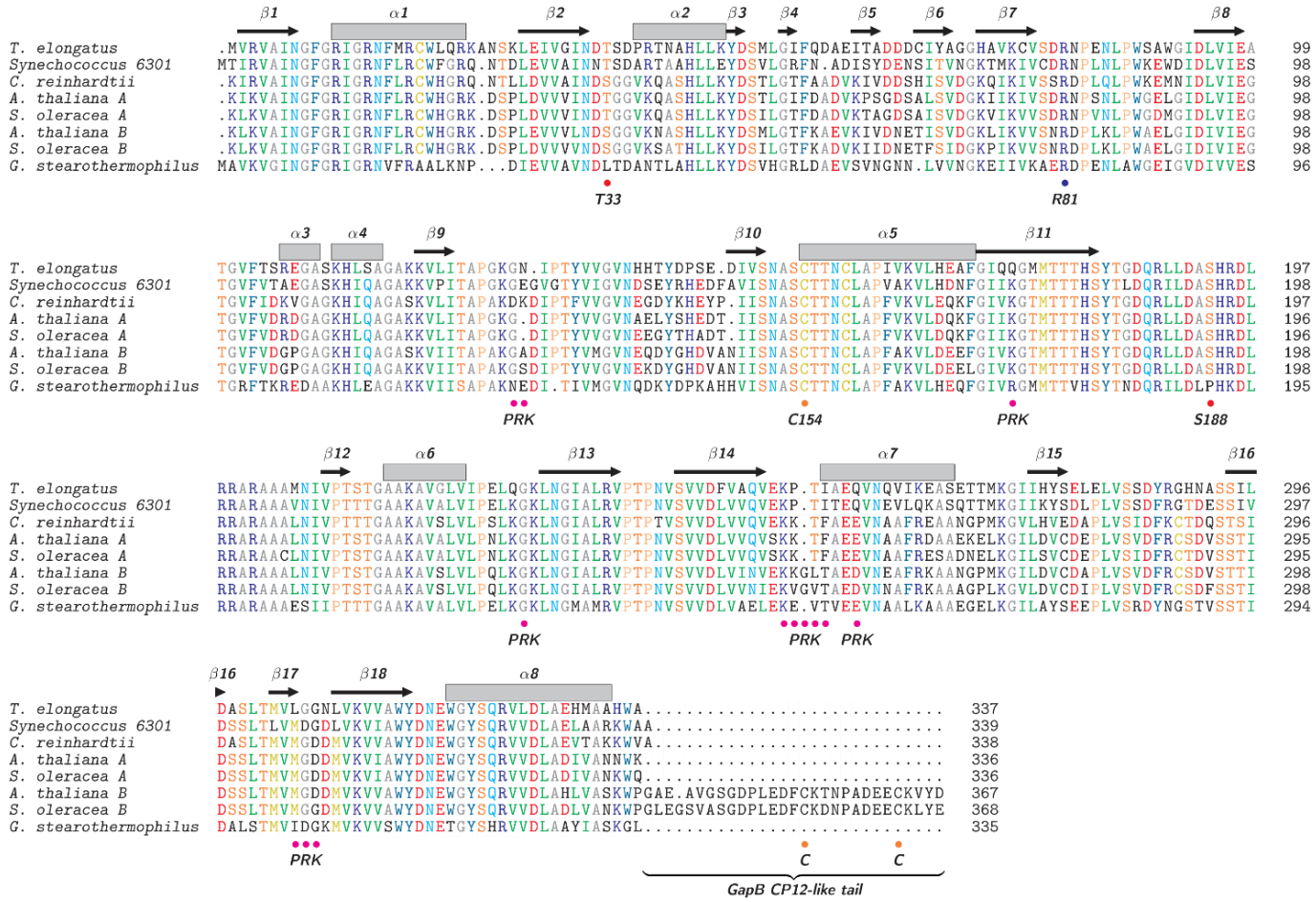


Fig. S3 Alignment of representative GAPDH sequences.

Alignment of GAPDH from cyanoboacteria, *Chlamydomonas*, higher plant GAPDH-A and GAPDH-B and non-photosynthetic GAPDH from *Geobacillus stearothermophilus.* The secondary structure of *T. elongatus* GAPDH is shown above the alignment. The active site cysteine 154 is shown, as are Thr33 and Ser188 of the spinach A4-GAPDH, which are involved in NADP binding (27). Arg81 moves depending in whether NAD^+ or $NADP^+$ are bound. Residues interacting with PRK in the ternary complex are labeled with magenta circles.

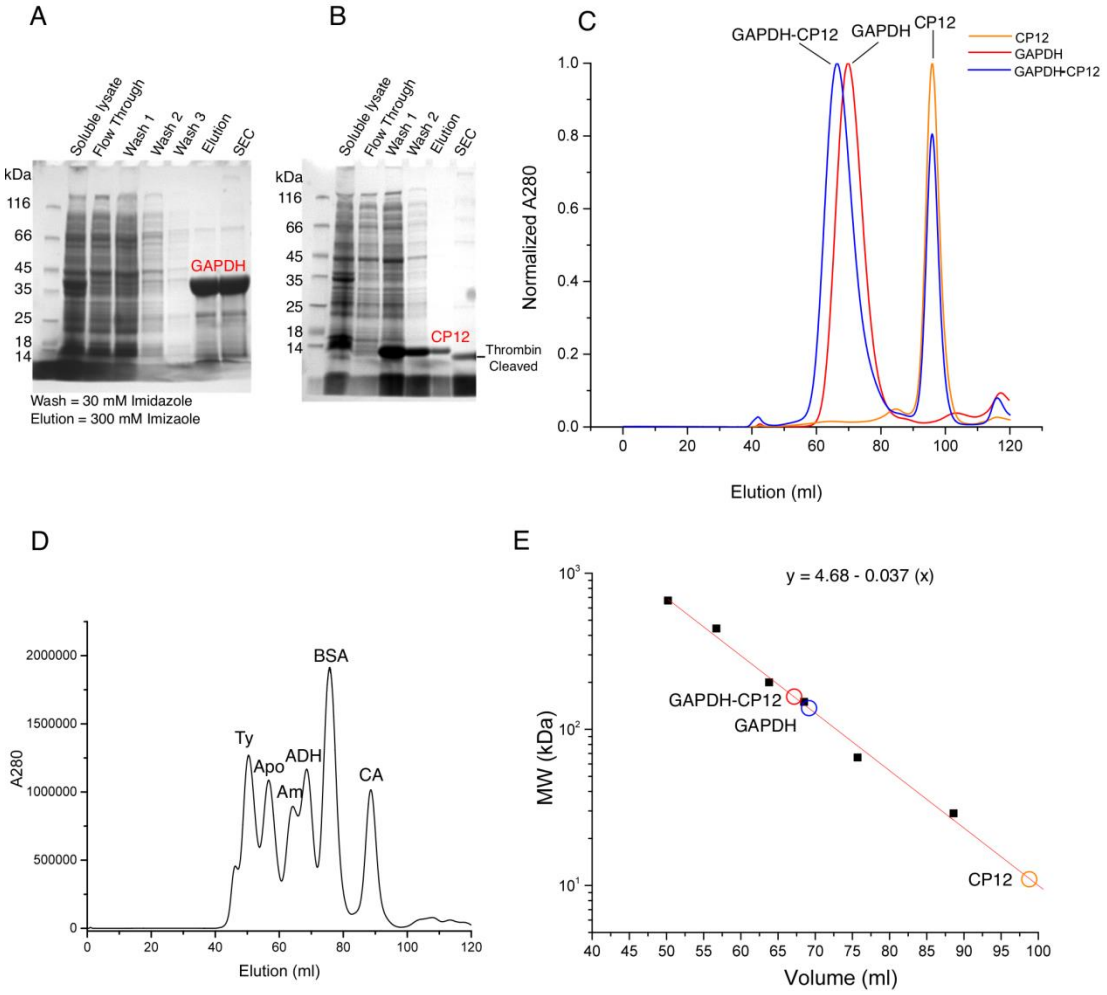


Fig. S4 Purification and reconstitution of GAPDH-CP12 complex.

(A) SDS-PAGE of Ni-NTA affinity purification for GAPDH and (B) CP12, followed by size-exclusion chromatography (SEC). (C) Size-exclusion chromatograms normalized against the highest peak to show the elution profiles of each species. The GAPDH-CP12 complex (163 kDa) peak is shifted relative to GAPDH alone (146 kDa) and CP12 alone (8.5 kDa).

(D) Standard chromatogram of proteins of known molecular weight (MW). The standards were thyroglobulin (Ty), 669 kDa; apoferritin (Apo), 443 kDa; β-amylase (Am), 200 kDa; Alcohol dehydrogenase (ADH); 150 kDa; bovine serum albumin (BSA); 66 kDa and carbonic anhydrase (CA), 29 kDa. (E) Elution volumes plotted on a semi-log graph and fitted to a straight line, y=4.68-0.037(x) where x is the elution volume. The molecular weight equals 10^y. Based on the elution profile GAPDH had an apparent mass of 133 kDa (predicted mass of 146 kDa), CP12 had an apparent mass of 12 kDa (predicted mass of 8.5 kDa) and GAPDH-CP12 had an apparent mass of 161 kDa (predicted mass of 155 kDa). The apparent 28 kDa increase in GAPDH mass corresponds to an increase of approximately two CP12 chains.

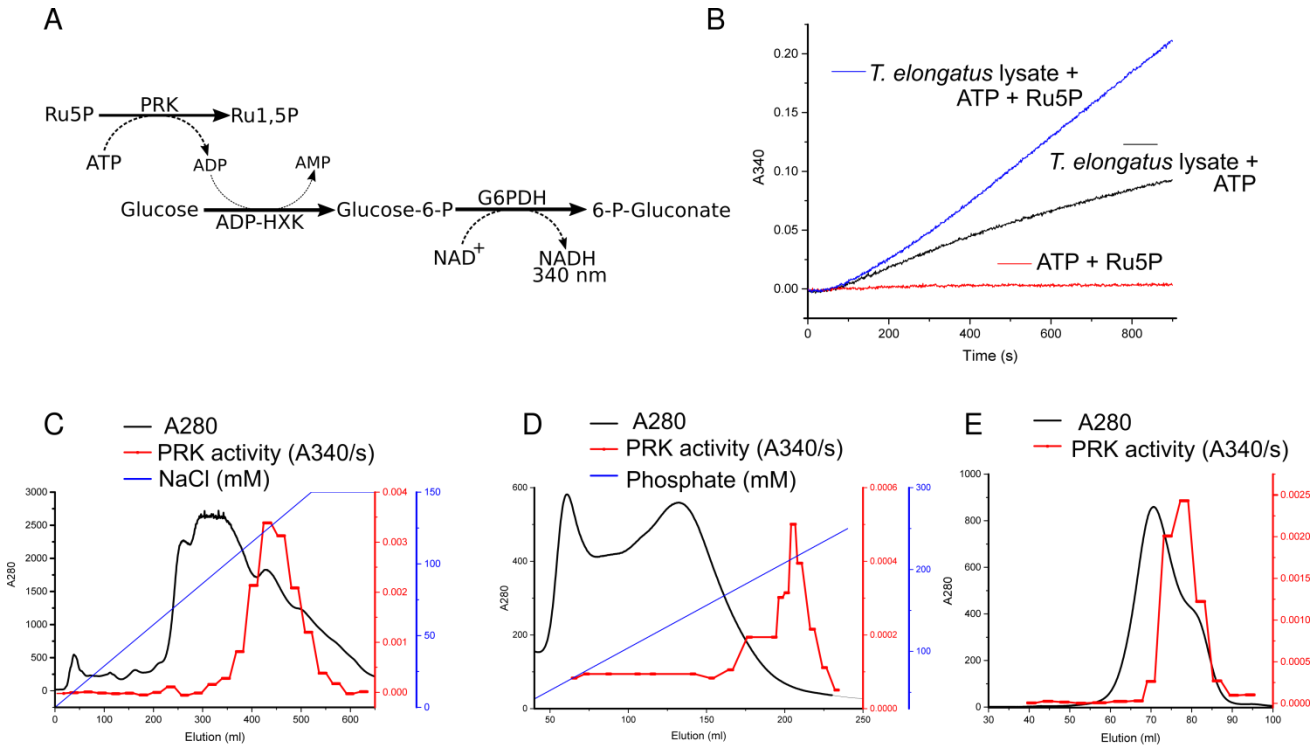
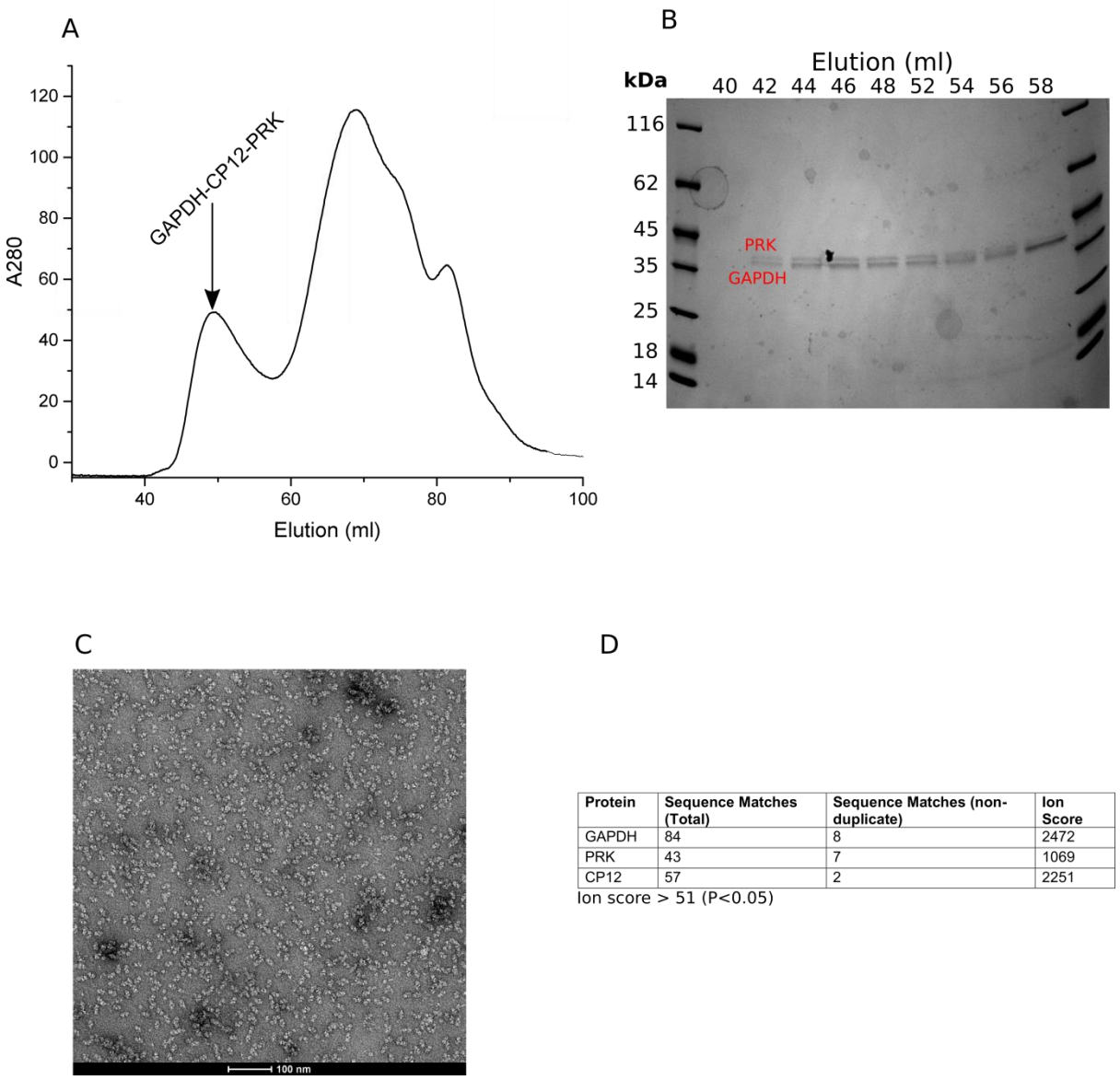


Fig. S5 Purification of PRK from *T. elongatus* by purifying ribulose-5-phosphate kinase activity.

(A) Schematic of PRK activity measurement by coupling the PRK-dependent production of ADP to the reduction of NAD^+ by ADP-hexokinase and glucose 6-phosphate dehydrogenase, which was measured at 340 nm. (B) PRK activity in lysate was the difference in activity between sample with ATP + Ru5P (blue) and with ATP (black). The assay mixture alone produced low background activity in the absence *T. elongatus* protein (red). (C) Anion exchange chromatography with A280 (mAU) measuring protein elution (black), the NaCl gradient (blue) and the specific PRK activity (red). (D) Hydroxyapatite with A280 (black), phosphate gradient (blue) and PRK activity (red). (E) Size-exclusion chromatography with A280 (black) and PRK activity (red).



Protein	Sequence Matches (Total)	Sequence Matches (non-duplicate)	Ion Score
GAPDH	84	8	2472
PRK	43	7	1069
CP12	57	2	2251

Ion score > 51 (P<0.05)

Fig. S6 Reconstitution and purification of GAPDH-CP12-PRK ternary complex

(A) Size-exclusion A280 profile showing the GAPDH-CP12-PRK complex eluting at a high molecular weight peak (arrow at 47 ml corresponding to ~500 kDa). (B) SDS-PAGE of this elution shows that the peak is composed of purified complex with a visible doublet corresponding to GAPDH (36.6 kDa) and PRK (38.0 kDa). CP12 is not visible on the gel. (C) Negatively-stained purified GAPDH-CP12-PRK complexes show that the sample is homogenous, monodisperse, and adopts a range of orientations on the carbon-coated grid. (D) Mass spectroscopy QUAD-TOF identifying GAPDH, PRK and CP12 in the complex. The ion score for an MS/MS match is -10Log(P), where P is the calculated probability that the observed match between the experimental data and the database sequence is a random event. Data from St. Andrews protein ID service (3).

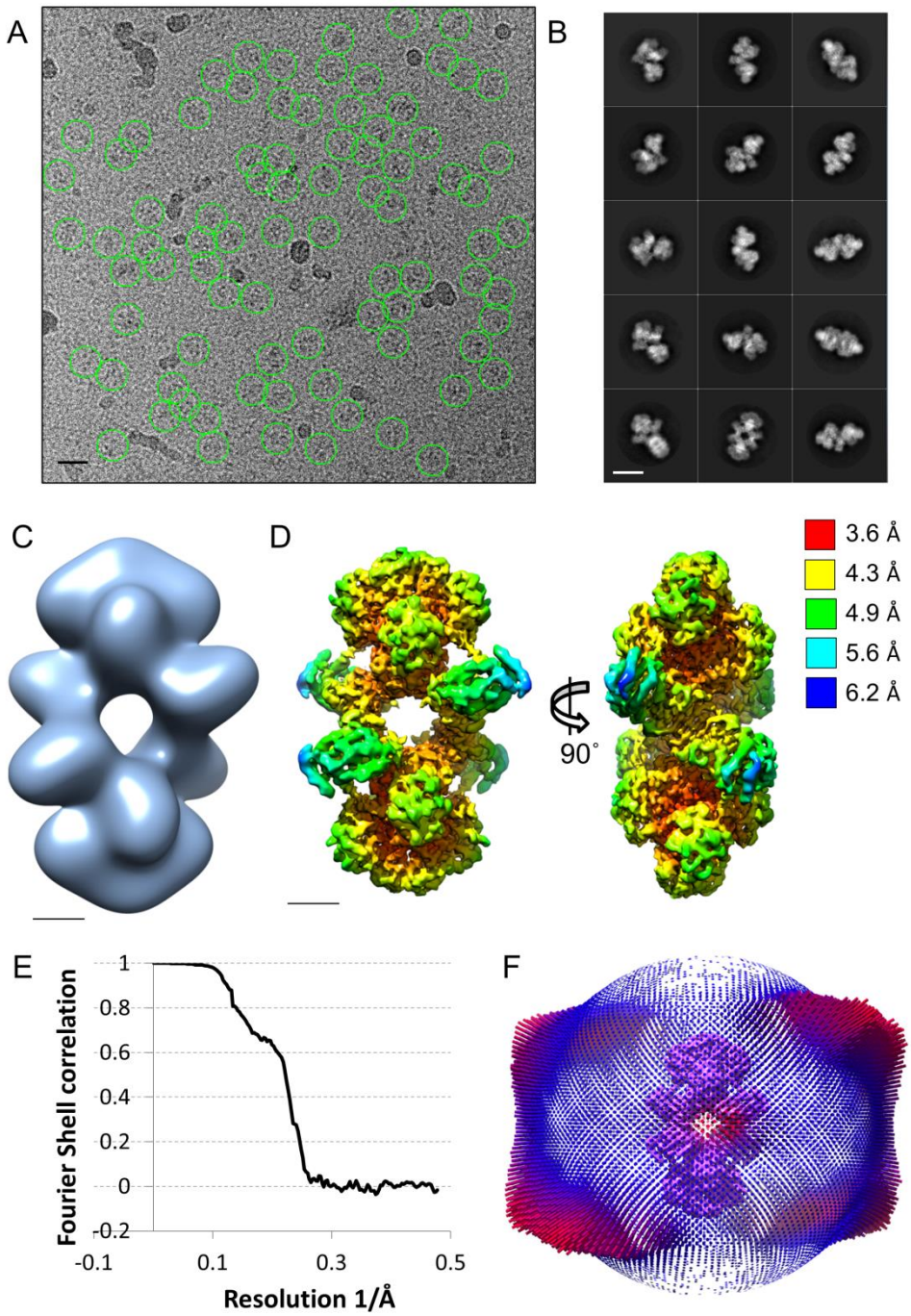


Fig. S7 Single-particle reconstruction of GAPDH-CP12-PRK. Fig. S5 Single-particle reconstruction of GAPDH-CP12-PRK.

(A) Representative electron micrograph of the GAPDH-CP12-PRK complex on a carbon-coated holey carbon grid, frozen in vitreous ice. Particles circled in green; scale bar 250 Å. (B) Selected 2D class averages derived from cryoEM images, scale bar 125 Å. (C) Model used to initiate the high-resolution refinement low-pass filtered to 40 Å, scale bar 30 Å. (D) Final 3D density map filtered based on local resolution and colored according to panel, scale bar 30 Å. (E) Mask-corrected Fourier shell correlation (FSC) curves computed from the unfiltered half-maps in RELION. (F) Angular distribution for the reconstruction. Height of the cylinder at each projection direction is proportional to the number of particle images.

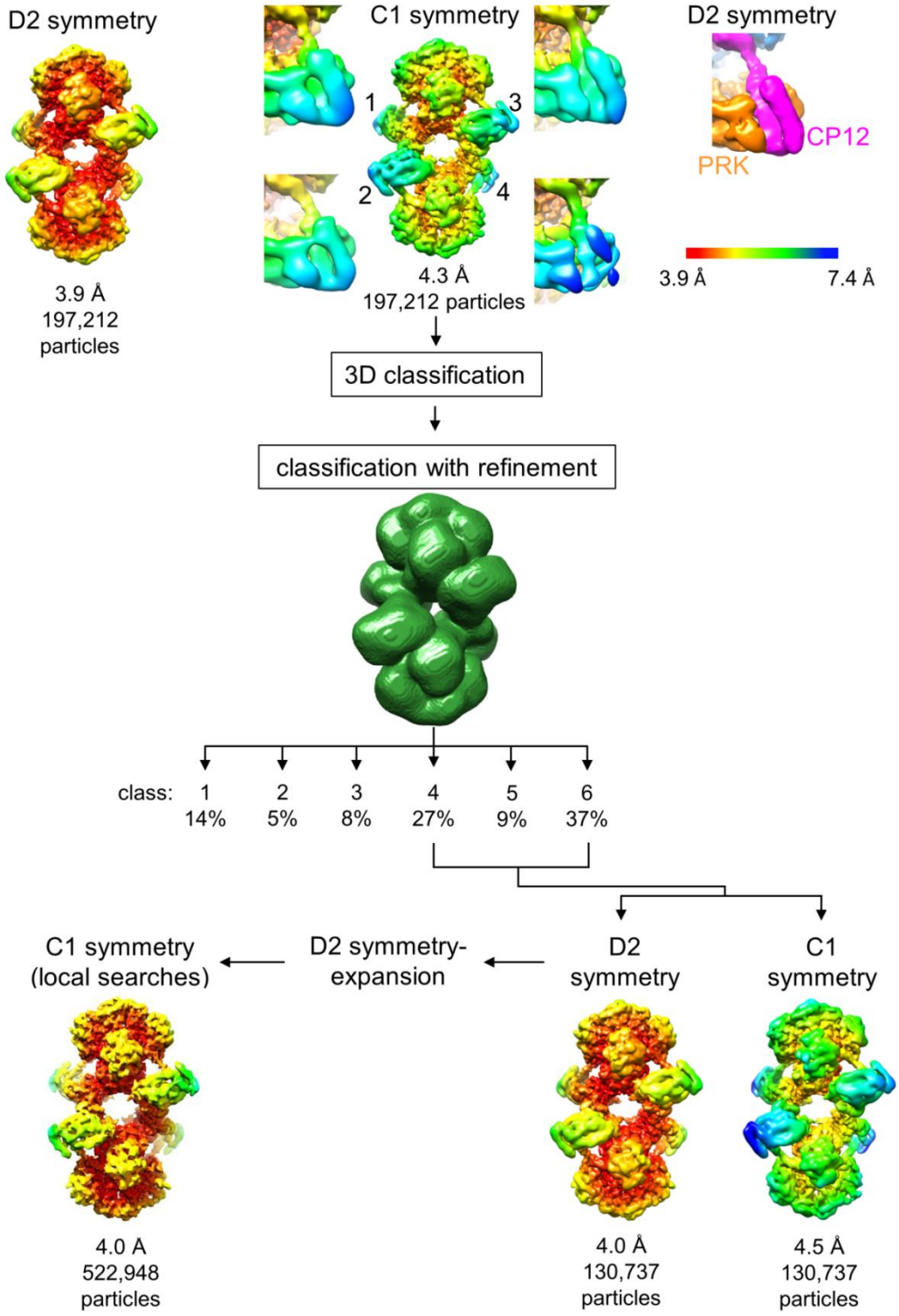


Fig. S8 3D classification and reconstruction strategy for GAPDH-CP12-PRK.
The GAPDH-CP12-PRK asymmetric reconstruction (C1 symmetry at the top) and the D2-symmetrized map of all 197,212 particles (top left) are filtered according to local resolution. The four CP12-PRK binding interfaces are indicated with numbers on the C1 reconstruction. The four insets show each interface viewed from the same orientation and colored as in the C1 map (top). The corresponding density from the D2 symmetrized reconstruction is colored according to protein components and is shown for reference on the top, right. Interface 4 has inferior local resolution and decreased occupancy compared to the other interfaces. To select for particles with full occupancy at

all four interfaces, a 3D classification with refinement was performed on the entire complex in which particle orientations were refined at each iteration; the green density represents the classification mask, which encompasses the entire complex. The classes with the best resolved density were combined and further refined in either D2 or without symmetry (C1). The C1 reconstruction showed full occupancy for all 4 interfaces. Particles refined under D2 symmetry were expanded such that each particle was assigned 4 orientations that corresponded to its D2 symmetry redundant views. These symmetry expanded particles were then refined with only local search parameters with no symmetry constrains (C1 symmetry) to obtain the optimal map (see Fig. S12). All 3D density maps are filtered based on local resolution estimates and colored according to the same scale, ranging from 3.9 Å (red) to 7.4 Å (blue). Overall resolution and particle numbers for each map are indicated.

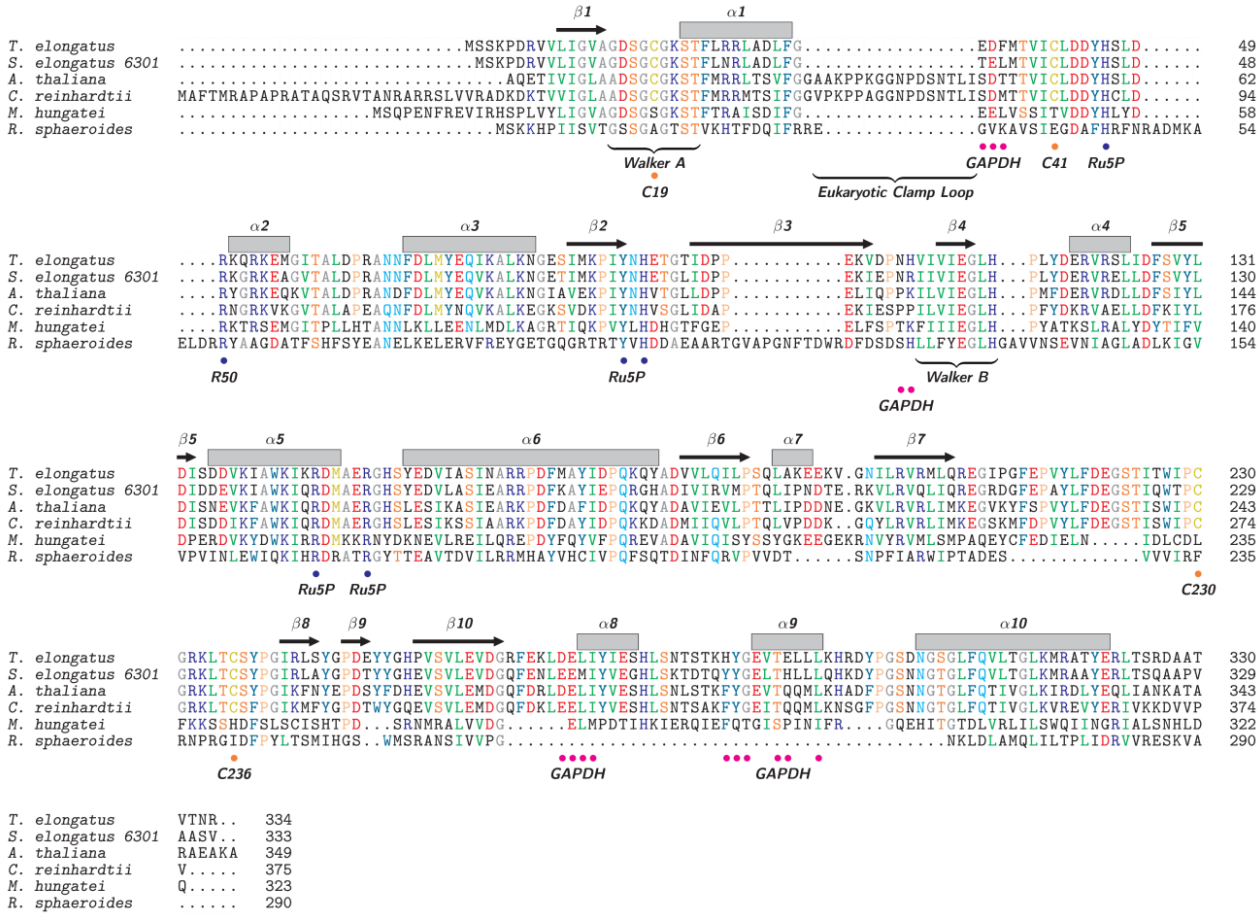


Fig. S9 Alignment of representative PRK sequences.

The secondary structure of *T. elongatus* PRK is shown above the alignment. Residues involved in binding ribulose 5-phosphate are marked with blue circles, as identified by site-directed mutagenesis of the *R. sphaeroides* enzyme (19–23). Nucleotide-binding, Walker A (P-loop), Walker B motifs and the clamp loop (24) are marked, as are the two pairs of conserved cysteine residues in green-type PRKs. Residues interacting with GAPDH in the ternary complex are labelled with magenta circles. All sequence alignments were done with ClustalX (25) and drawn with TeXShade (26).

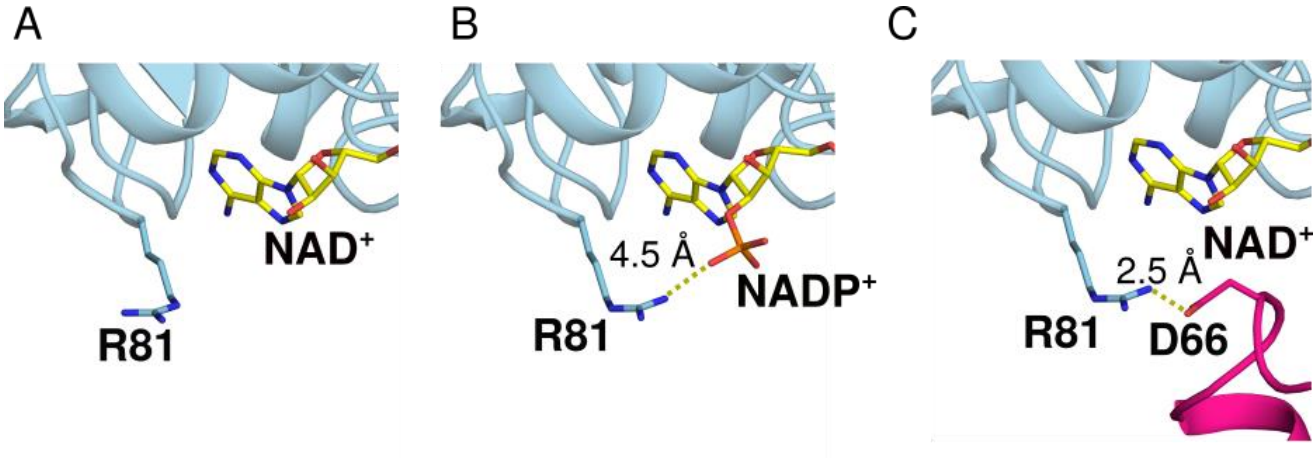


Fig. S10. GAPDH active site with different dinucleotide cofactors and CP12 binding, showing movement of Arg 81.
(A) GAPDH with NAD^+ bound (6GFR) and Arg81 guanidino group facing away from the active site. (B) GAPDH with $NADP^+$ bound (6GFP), showing salt bridge between Arg81 and the 2'- phosphate of $NADP^+$. (C) GAPDH with NAD+ bound and CP12 (6GFQ), showing a hydrogen bond between Arg81 and Asp66 of CP12, the guanidino group is in the same conformation as when binding the 2'-phosphate of $NADP^+$.

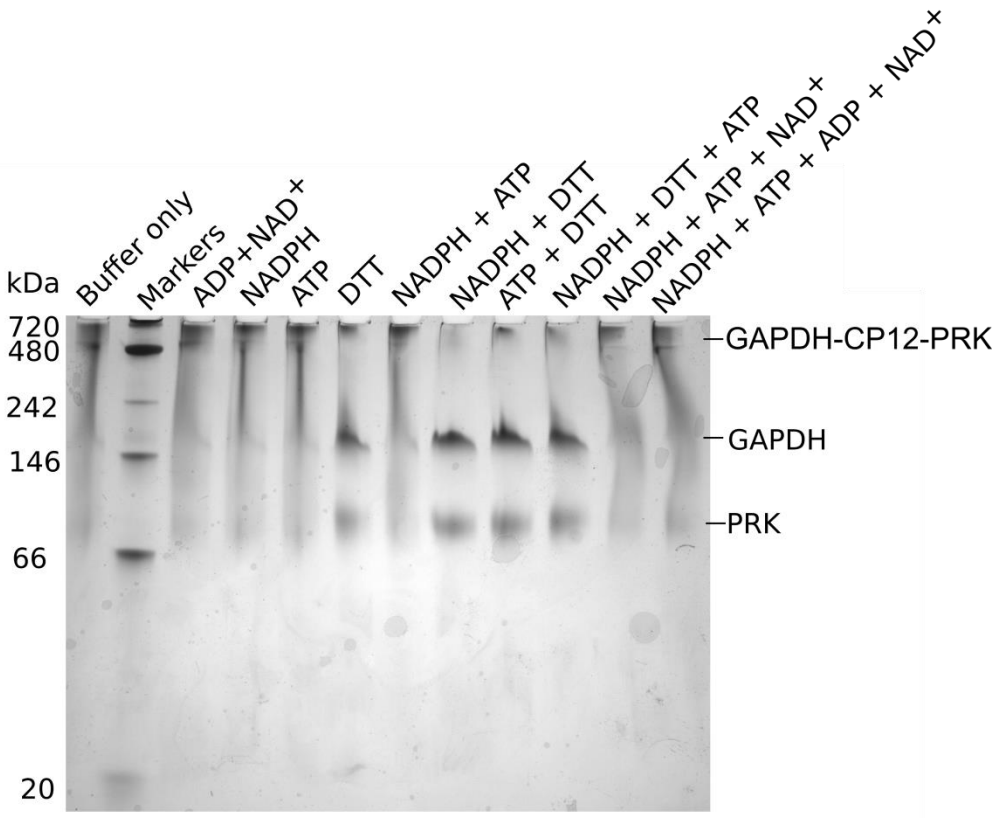


Fig. S11 Native gel electrophoresis of GAPDH-CP12-PRK complex.

Each lane contained GAPDH-CP12-PRK ternary complex and the combination of additives shown at 2.5 mM each. Only in the presence of 2.5 mM DTT is the complex dissociated into separate GAPDH and PRK bands. Combinations of ATP, ADP, NADPH and NAD^+ at 2.5 mM did not affect the stability of the complex. Bands were cut out of the gel and identified by mass spectrometry.

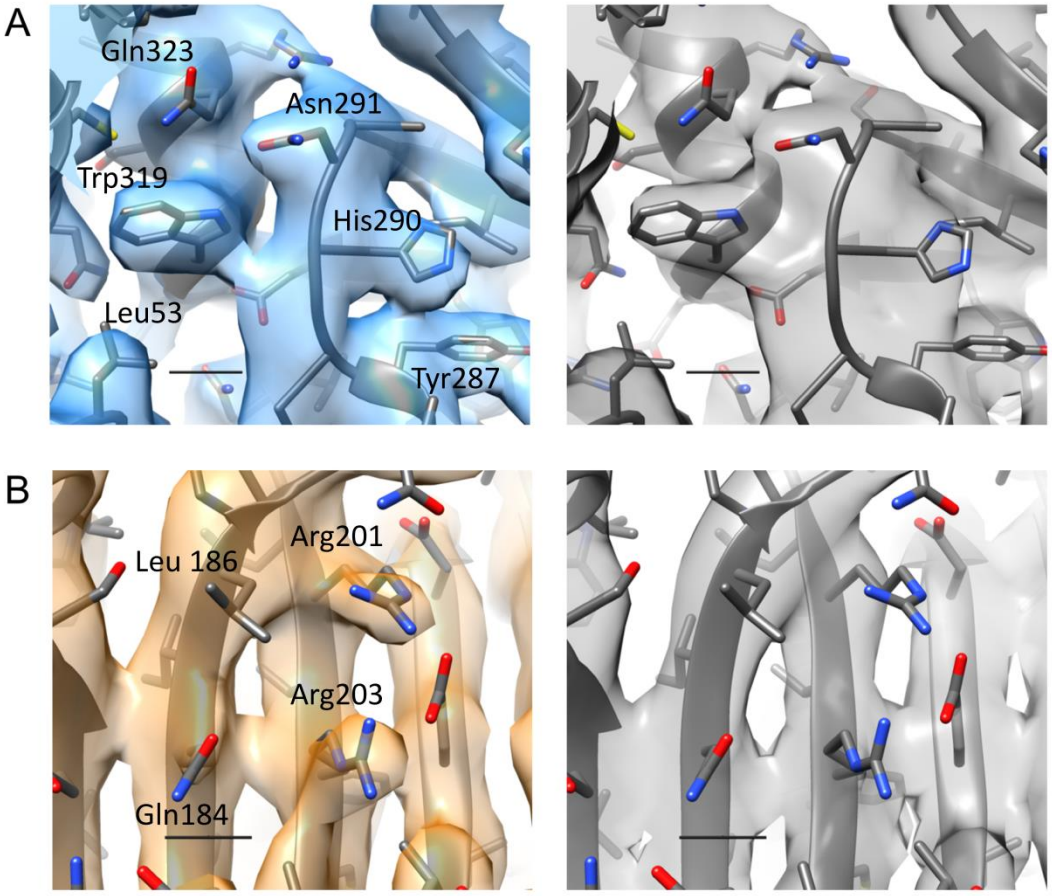


Fig. S12 Comparison of GAPDH-CP12-PRK cryoEM map features.
The D2 symmetry expanded C1-refined optimum density map (blue and orange densities, left panels) (bottom left in Fig. S8) has more distinct features compared to the D2-refined map of the original particle set (grey densities, right panels) (top left in Fig. S8), even though the local resolution estimates are very similar. Comparing a region of GAPDH density (A) and PRK density (B). Scale bars represent 3 Å.

Structure	GAPDH-CP12 $NADP^+$	GAPDH-CP12-2 C-termini visible	GAPDH-NAD^+	GAPDH-CP12 conformation 1 and conformation 2	GAPDH-CP12 4 bound	GAPDH-CP12 conformation 3	GAPDH-CP12 conformation 4
PDB entry	6GFP	6GFQ	6GFR	6GFO	6GG7	6GHR	6GHL
Crystallization condition	100 mM Hepes pH 7.0 16% PEG3350 4% Tacsimate (5 mM G3P, NADP+, BPG, NADPH)	4% Tacsimate, pH 7 8% PEG3350	12% PEG3350 2% Tacsimate, 100 mM Hepes, pH 7	1% Tryptone, 0.001 M Azide, 0.005 M Azide, 12% PEG3350	1% Tryptone, 25% PEG, 100 mM Hepes pH 8	4% Tacsimate, pH 7 12% PEG3350	100 mM Hepes pH7.0 16% PEG3350 4% Tacsimate
Beam line	I03-DLS	I03-DLS	I24-DLS	I03-DLS	I24-DLS	I04-DLS	I03-DLS
Wavelength	0.9762	0.9762	0.9686	0.9762	0.9686	0.9795	0.97625
Resolution range / Å	44.74 - 1.54 (1.595 - 1.54)	41.77 - 1.4 (1.45 - 1.4)	69.93 - 1.919 (1.988 - 1.919)	70.63 - 2.1 (2.175 - 2.1)	60.42 - 1.32 (1.367 - 1.32)	71.23 - 2.249 (2.329 - 2.249)	40.4 - 2.378 (2.463 - 2.378)
Space group	P 43 21 2	P 43 21 2	P 43 21 2	P 21 21 2	P 43 21 2	P 21 21 2	P 21 21 21
Unit cell / Å / degrees	141.469 141.469 74.4253 90 90 90	141.217 141.217 76.2382 90 90 90	139.865 139.865 80.5014 90 90 90	136.446 141.049 81.5925 90 90 90	141.052 141.052 75.9309 90 90 90	133.438 146.964 81.4279 90 90 90	80.7901 138.801 139.772 90 90 90
Total reflections	1399889 (133889)	1548856 (47025)	788139 (75513)	581446 (56619)	2276429 (222027)	267223 (22072)	406558 (39630)
Unique reflections	111354 (10960)	147904 (12946)	61333 (5951)	92479 (9129)	178279 (17624)	76310 (7510)	63755 (6186)
Multiplicity	12.6 (12.2)	10.5 (3.6)	12.9 (12.7)	6.3 (6.2)	12.8 (12.6)	3.5 (2.9)	6.4 (6.4)
Completeness (%)	99.98 (99.95)	98.12 (86.93)	99.88 (99.08)	99.94 (99.91)	99.99 (100.00)	99.35 (98.79)	99.76 (98.45)
Mean I/sigma(I)	13.92 (1.94)	19.71 (1.91)	7.94 (2.03)	6.90 (2.45)	11.43 (1.90)	7.68 (1.82)	6.82 (1.57)
Wilson B-factor	19.03	15.28	24.85	26.80	13.62	33.67	40.54
R-merge	0.09574 (0.8891)	0.05763 (0.499)	0.2205 (0.9235)	0.1387 (0.4962)	0.1266 (0.9185)	0.09854 (0.4633)	0.1687 (0.8387)
R-meas	0.09982 (0.9278)	0.06027 (0.5766)	0.2297 (0.9625)	0.1514 (0.5416)	0.132 (0.9573)	0.1163 (0.565)	0.1838 (0.9138)
R-pim	0.02803 (0.2636)	0.01732 (0.2804)	0.06388 (0.2692)	0.0601 (0.2152)	0.03692 (0.2683)	0.06067 (0.3183)	0.07219 (0.3589)
CC1/2	0.999 (0.567)	0.999 (0.724)	0.995 (0.543)	0.991 (0.905)	0.988 (0.586)	0.993 (0.709)	0.994 (0.467)

CC*	1 (0.851)	1 (0.916)	0.999 (0.839)	0.998 (0.975)	0.997 (0.86)	0.998 (0.911)	0.998 (0.798)
Reflections used in refinement	111347 (10960)	147899 (12945)	61314 (5950)	92440 (9126)	178272 (17624)	76274 (7507)	63724 (6181)
Reflections used for R-free	5516 (524)	7422 (654)	3050 (281)	4599 (455)	8785 (895)	3800 (363)	3149 (323)
R-work	0.1639 (0.2830)	0.1457 (0.2404)	0.1648 (0.2916)	0.1707 (0.2194)	0.1452 (0.2503)	0.1904 (0.2762)	0.1969 (0.3089)
R-free	0.1833 (0.2814)	0.1736 (0.2823)	0.2064 (0.3256)	0.2185 (0.2752)	0.1732 (0.2738)	0.2268 (0.3199)	0.2470 (0.3471)
CC(work)	0.973 (0.777)	0.971 (0.877)	0.970 (0.820)	0.960 (0.945)	0.973 (0.827)	0.963 (0.849)	0.963 (0.673)
CC(free)	0.972 (0.731)	0.961 (0.811)	0.960 (0.729)	0.949 (0.871)	0.962 (0.798)	0.949 (0.752)	0.930 (0.616)
Number of non-hydrogen atoms	6034	6001	5654	11974	6330	11517	11346
macromolecules	5204	5433	5162	11455	5678	11061	11034
ligands	202	90	101	181	88	181	183
solvent	628	478	391	338	564	275	129
Protein residues	677	700	678	1499	722	1450	1446
RMS(bonds)	0.007	0.006	0.007	0.008	0.006	0.004	0.005
RMS(angles)	1.19	1.22	1.18	1.18	1.22	0.96	1.05
Ramachandran favored (%)	96.73	96.24	96.28	96.15	96.49	95.81	94.89
Ramachandran allowed (%)	2.83	3.47	3.27	3.58	3.23	3.91	4.41
Ramachandran outliers (%)	0.45	0.29	0.45	0.27	0.28	0.28	0.70
Rotamer outliers (%)	1.42	0.34	0.36	0.08	0.32	0.08	0.59
Clashscore	2.90	1.64	2.96	3.61	1.75	2.65	11.22
Average B-factor / Å²	27.08	21.50	33.94	38.67	21.25	44.52	52.78
macromolecules	26.13	20.63	33.56	38.97	19.80	44.81	52.74
ligands	18.63	16.95	29.32	28.69	16.58	42.75	65.76
solvent	37.66	32.31	40.11	34.08	36.57	34.04	37.70
Number of TLS groups	8		8	19		21	16

Table S1 Data collection and refinement statistics for GAPDH crystal structures, with CP12, NAD^+ and $NADP^+$. Statistics for the highest-resolution shell are shown in parentheses.

model to map fit (CC around atoms)	0.76
RMS deviations (bonds /Å)	0.009
RMS deviations (bond angles /Å)	0.960
All-atom clashscore	7.01(87^{th} percentile all resolutions)
molprobity score raw (percentile)	1.68 (90th percentile all resolutions)
Ramachandran favoured (%)	95.75
outlier (%)	0.0
Rotamer outliers	0.11%
Favored rotamers	96.7%

Table S2 Model refinement statistics of the GAPDH-CP12-PRK cryoEM structure. Validation using MolProbity, additional details in the Materials and Methods.

sample	substrate	K_m (µM)	k_{cat} (s^{-1})	k_{cat} / K_m
GAPDH-CP12-PRK^{Ox}	ATP	-	-	-
GAPDH-CP12-PRK^{Red}	ATP	40 ± 2	13.6 ± 0.3	0.34 ± 0.02

Table S3 PRK activity kinetics data.

PRK activity with ATP. Rates were measured in triplicate, plotted with standard deviation (Fig. 1) and fitted using the Michaelis-Menten formula. Apparent K_m values were derived from the curve fitting and k_{cat} values calculated from the derived V_{max} and moles of PRK active sites. Dash indicates no activity was measurable.

sample	substrate	K_m (µM)		k_{cat} (s^{-1})		k_{cat} / K_m	
GAPDH	NAD^+	37	± 5	6.7	± 0.4	0.18	± 0.03
GAPDH	$NADP^+$	40	± 3	7.4	± 0.3	0.18	± 0.01
GAPDH-CP12^{Ox}	NAD^+	11	± 2	1.5	± 0.01	0.14	± 0.02
GAPDH-CP12^{Ox}	$NADP^+$	138	± 8	9.5	± 0.2	0.07	± 0.004
GAPDH-CP12-PRK^{Ox}	NAD^+	14	± 4	1.4	± 0.005	0.10	± 0.02
GAPDH-CP12-PRK^{Ox}	$NADP^+$	-		-		-	

Table S4 GAPDH activity kinetics data.

GAPDH activity with NAD^+ and $NADP^+$ of GAPDH, GAPDH-CP12 and GAPDH-CP12-PRK. Rates under oxidizing conditions (Ox) were measured in triplicate, plotted with standard deviation (Fig. 1) and fitted using the Michaelis-Menten formula. Apparent K_m values were derived from the curve fitting and k_{cat} values calculated from the derived V_{max} and moles of PRK active sites. Dash indicates no activity was measurable.

References

1. Michoux F, Takasaka K, Boehm M, Nixon PJPJ, Murray JWJW (2010) Structure of CyanoP at 2.8 A: implications for the evolution and function of the PsbP subunit of photosystem II . *Biochemistry* 49(35):7411–3.
2. Castenholz R (1988) Culturing methods for cyanobacteria. *Methods in Enzymology*, eds Lester P, Glazer AN (Academic Press), pp 68–93.
3. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20(18):3551–3567.
4. Belkin S, Jannasch HW (1985) A new extremely thermophilic, sulfur-reducing heterotrophic, marine bacterium. *Arch Microbiol* 141(3):181–186.
5. Winter G (2010) *xia2*: An expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* 43(1):186–190.
6. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658–674.
7. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
8. Adams P, et al. (2010) *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D* 66(2):213–221.
9. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biol Crystallogr* 66(1):12–21.
10. Pettersen E, et al. (2004) UCSF Chimera-a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
11. Rohou A, Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192(2):216–221.
12. Kimanius D, Forsberg B, Scheres SH, Lindahl E (2016) Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife* 5.
13. Zheng S, et al. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* 14(4):331–332.
14. Kono T, et al. (2017) A RuBisCO-mediated carbon metabolic pathway in methanogenic archaea. *Nat Commun* 8:14007.
15. Wilson RH, Hayer-Hartl M, Bracher A (2019) Crystal structure of phosphoribulokinase from

Synechococcus sp. strain PCC 6301. *Acta Crystallogr Sect F* 75(4):278–289.
16. Afonine P, et al. (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr Sect D, Struct Biol* 74(Pt 6):531–544.
17. Terwilliger T, Sobolev O, Afonine P, Adams P (2018) Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr Sect D, Struct Biol* 74(Pt 6):545–559.
18. Hackenberg C, et al. (2018) Structural and functional insights into the unique CBS-CP12 fusion protein family in cyanobacteria. *Proc Natl Acad Sci U S A*.
19. Runquist J, Miziorko H (2006) Functional contribution of a conserved, mobile loop histidine of phosphoribulokinase. *Protein Sci* 15(4):837–842.
20. Sandbaken MG, Runquist JA, Barbieri JT, Miziorko HM (1992) Identification of the phosphoribulokinase sugar phosphate binding domain. *Biochemistry* 31(14):3715–3719.
21. Runquist J, Koteiche H, Harrison D, Miziorko H (1998) Phosphoribulokinase: Mutagenesis of the Mobile Lid and “P-Loop.” *Photosynthesis: Mechanisms and Effects*, ed Garab G (Springer Netherlands), pp 3399–3402.
22. Kung G, Runquist JA, Miziorko HM, Harrison DH (1999) Identification of the allosteric regulatory site in bacterial phosphoribulokinase. *Biochemistry* 38(46):15157–15165.
23. Runquist JA, Ríos SE, Vinarov DA, Miziorko HM (2001) Functional evaluation of serine/threonine residues in the P-Loop of *Rhodobacter sphaeroides* phosphoribulokinase. *Biochemistry* 40(48):14530–14537.
24. Gurrieri L, et al. (2019) *Arabidopsis* and *Chlamydomonas* phosphoribulokinase crystal structures complete the redox structural proteome of the Calvin-Benson cycle. *Proc Natl Acad Sci*. doi:10.1073/pnas.1820639116.
25. Larkin M a, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–8.
26. Beitz E (2000) TeXshade: shading and labeling of multiple sequence alignments using LaTeX2e. *Bioinformatics* 16(2):135–139.
27. Sparla F, et al. (2004) Coenzyme site-directed mutants of photosynthetic A4-GAPDH show selectively reduced NADPH-dependent catalysis, similar to regulatory AB-GAPDH inhibited by oxidized thioredoxin. *J Mol Biol* 340(5):1025–1037.