Architecture of the bundle-shaped phycobilisome from *Gloeobacter violaceus*

Anna D. Burtseva[1,2a], Yury B. Slonimskiy[1a], Timur N. Baymukhametov[3], Maria A. Sinetova[4], Eugene G. Maksimov[5], Vladimir O. Popov[1], Konstantin M. Boyko[1], Nikolai N. Sluchanko[1,*]

[1] A.N. Bach Institute of Biochemistry, Federal Research Center of Biotechnology of the Russian Academy of Sciences, Moscow 119071, Russia
[2] Moscow Center for Advanced Studies, Kulakova Str. 20, Moscow 123592, Russia
[3] Kurchatov Complex of NBICS Nature-Like Technologies, Structural Biology Department, National Research Center "Kurchatov Institute", Akademika Kurchatova pl., 1, Moscow 123182, Russia
[4] K.A. Timiryazev Institute of Plant Physiology Russian Academy of Sciences, 127276 Moscow, Russia
[5] M.V. Lomonosov Moscow State University, Faculty of Biology, 119991 Moscow, Russia

a – co-first authors

*Corresponding author:
Nikolai N. Sluchanko
A.N. Bach Institute of biochemistry
Federal Research Center of Biotechnology of the Russian Academy of Sciences
Leninsky prospect 33, building 1
119071 Moscow, Russian Federation
Tel.: +74956603430
Email: nikolai.sluchanko@mail.ru

Abstract

Cyanobacteria optimize light harvesting for photosynthesis using large, soluble antenna systems called phycobilisomes (PBS). These megacomplexes, composed of spectrally distinct phycobiliproteins and linker proteins, allow efficient energy transfer to reaction centers of the photosystems. While PBS structures from many cyanobacteria have been resolved in recent years, the PBS of *Gloeobacter violaceus*—a phylogenetically ancient, thylakoid-less cyanobacterium—remains poorly characterized. Here, we present the cryo-EM structure of the *G. violaceus* PBS, a >10-MDa complex with a unique architecture consisting in 516 polypeptide chains totaling nearly 90,000 residues and harboring over 860 bilin chromophores. Unlike canonical PBS with radiating rods, the *Gloeobacter* PBS features diverging, conformationally flexible rod bundles formed by stacked phycoerythrin (PE) and phycocyanin (PC) hexamers. These rods emerge from a pentacylindrical allophycocyanin (APC) core, which is encircled by additional PC hexamers—a structural feature atypical for other cyanobacteria. We identify the role of the two *Gloeobacter*-specific multidomain linker proteins, Glr1262 and Glr2806, as critical stabilizers of this bundle-shaped architecture. Structural analysis suggests these linkers mediate rod-core interactions while permitting rod mobility, potentially balancing light capture and photoprotection. Furthermore, we demonstrate that the lateral cylinders of the APC core create alternative energy transfer pathways from the rods to the terminal emitters, ensuring robust excitation delivery under varying light conditions. This redundancy may represent an evolutionary adaptation to *Gloeobacter*'s unique physiology. Our findings provide the high-resolution structural insight into *Gloeobacter* PBS, revealing divergent adaptations in early-branching cyanobacteria. The structure advances understanding of PBS diversity and evolution, offering a framework for bioengineering light-harvesting systems in synthetic biology and biotechnological applications.

Keywords
Phycobiliproteins, light-harvesting complexes, evolution, cryo-EM

Introduction

Light absorption is the first step in photosynthesis, and photosynthetic organisms have evolved various mechanisms to improve its efficiency. Cyanobacteria and red algae use extramembranous antenna megacomplexes called phycobilisomes (PBS), composed of phycobiliproteins (PBPs), to harvest light. PBPs covalently bind bilins, which are open-chain tetrapyrroles (*1*). Bilins absorb light in different spectral regions and transfer energy to the reaction centers of photosystems (*2*–*4*). This compensates for chlorophyll's limited ability to absorb light at short wavelengths (*5*). The efficient funneling of energy absorbed by bilins in PBPs is enabled by structural organization of PBS. PBPs are composed of heterodimeric αβ protomers, often designated "monomers", which oligomerize into ring-like trimers $(\alpha\beta)_3$ and hexamers $(\alpha\beta)_6$. These further assemble with the help of linker proteins that interact with the PBP-rings in their cavities. A typical PBS contains an allophycocyanin (AP) core – two (*6*), three (*7*), or five (*8*) cylinders of $(\alpha\beta)_3$ trimers of AP – and a set of radiating rods – stacks of hexamers formed by αβ protomers of phycocyanin (PC) and, in red algae and some cyanobacteria, phycoerythrin (PE) (*1*, *9*, *10*). AP, PC, and PE differ in the number of bilins bound to their α (one in AP and PC, or two in PE) and β (one in AP, two in PC, and three in PE) subunits. The three main bilin types in cyanobacterial PBS are phycocyanobilin (PCB, found in AP and PC), phycoerythrobilin (PEB), and phycourobilin (PUB) (both found in PE) (*1*), whose collective absorbance covers the visible spectral range that is poorly accessible to chlorophyll.

Known PBS in red algae are hemi-ellipsoidal/block-type (*10*, *11*), whereas in cyanobacteria the prevalent PBS form is hemi-discoidal (*1*). In addition, the paddle-shaped PBS was recently described in the relict cyanobacterium *Anthocerotibacter panamensis* UTEX 3164, which utilizes single PC hexamers and chains of staggered PC hexamers that surround a heptacylindrical AP core instead of rods (*5*). A unique case is the cyanobacterium *Gloeobacter violaceus* PCC 7421 (hereafter, *G. violaceus*), which diverged from the crown cyanobacteria the earliest (~2 billion years ago). *G. violaceus* is characterized by slow metabolism and growth; primitive bacterial-type carotenoid biosynthesis; an inability to stably maintain intracellular pH; and a lack of circadian clock and thylakoids (*12*, *13*). PBS in *G. violaceus* attach to the photosystems localized in the plasma membrane (*12*, *14*). Although bundle-type PBS were first described in 1981 (*12*), they remain the only known PBS morphology without a high-resolution structure.

Results and discussion

1. Cryo-electron microscopy of the bundle-shaped *G. violaceus* PBS

The *G. violaceus* culture exhibited an absorbance spectrum with peaks at 500, 560, and 620 nm, corresponding to PBPs. Small features were also present at 440 and 680 nm, indicating relatively low chlorophyll content (Fig. 1A) (*14*). *G. violaceus* PBS (hereafter, GviPBS), which was isolated via sucrose gradient ultracentrifugation (Fig. 1B), exhibited the same prominent absorbance bands as PBPs, reflecting the presence of PUB (500 nm), PEB (560 nm), and PCB (an excitonic peak at 650 nm and a shoulder at 620 nm) (Fig. 1A). The first two of these are more prevalent in red algae than in cyanobacteria (*1*, *11*). All resolved cyanobacterial PBS structures contain only PCB and no other bilin types (*1*). PBPs could be biochemically separated into three distinct PBP fractions, corresponding to PE, PC, and AP, according to their absorbance and fluorescence spectra (Fig. 1C). The presence of PBPs and all linker

proteins, including *Gloeobacter*-specific multidomain linkers Glr1262 and Glr2806, was confirmed by SDS-PAGE and mass spectrometry (Fig. S1 and Table S1).

To preserve the intact PBS for cryo-EM, we performed ultracentrifugation in parallel gradients of sucrose and glutaraldehyde (GraFix (*15*)) (Fig. S1). To improve the quality of the cryo-EM density map, we focused on a set of substructures for model building and refinement against the masked map. This resulted in six focused maps with resolutions ranging from 2.72 Å to 3.67 Å (Table S2, Fig. S2, S3). The refined substructures were then combined to create a model of the entire PBS and fitted into the composite cryo-EM map (Fig. 1D).

For the GviPBS model, we used the crystallographic structures of $(\alpha\beta)_6$ hexamers of AP (PDB 2vjt), PC (PDB 2vjr) and PE (PDB 2vjh) (Fig. S4) and available electron microscopy and mutagenesis data (*16*). The reconstructed cryo-EM structure revealed three pairs of C2-symmetry-related rods stemming from the top of the pentacylindrical AP core (R1/R1', R2/R2', R3/R3') (Fig. 1D,E). The AP core contains two antiparallel bottom cylinders (A/A'), a top cylinder (B) and two lateral AP $(\alpha\beta)_6$ hexamers (C/C'), which are attached sideways to the top B cylinder. This "3+2" AP core can be superimposed on the pentacylindrical cores in hemi-discoidal PBS from *Anabaena* sp. PCC 7120 (*8*) and *Thermosynechococcus vulcanus* NIES-2134 (*17*) as well as on the lower portion of the heptacylindrical AP core in *A. panamensis* PBS (*5*) (Fig. S5). In GviPBS, the "3+2" core extends to the sides with additional PC $(\alpha\beta)_6$ hexamers (Rx1/Rx1') that lie coaxially with the C/C' $(\alpha\beta)_6$ hexamers of AP and support the rods (Fig. 1D).

The cryo-EM map of the rods dispersed more towards the periphery. Prior work revealed that deleting PE subunits results in a shorter GviPBS morphology, with rods containing three PC $(\alpha\beta)_6$ hexamers in height (*16*). Although two PC $(\alpha\beta)_6$ hexamer stacks could be confidently built in each rod, the density of the third PC hexamer and PE $(\alpha\beta)_6$ hexamers only permitted approximate placement (Fig. 1D). Two $(\alpha\beta)_6$ hexamers of PE (PDB 2vjh) could be docked in the rods R1/R1' and R3/R3', whereas three PE $(\alpha\beta)_6$ hexamers could be docked in the rods R2/R2', for a total of 18 PC and 14 PE $(\alpha\beta)_6$ hexamers in the rods of the composite model. Unexpectedly, the rods formed two bundles (R1-R2-R3 and R1'-R2'-R3') that radiated at ~15° angle so that the first-level PE $(\alpha\beta)_6$ hexamers in rods R2/R2' became ~145 Å apart (Fig. 1D,E). Each rod bundle formed a superhelix with a period of ~1000 Å (Fig. S6). Two maps supporting different bundle conformations allow us to imagine movement of the bundles relative to each other (Fig. S7).

The cryo-EM map permitted confident modeling of seven AP $(\alpha\beta)_6$ hexamers out of eight, while the fourth $(\alpha\beta)_3$ trimer was absent from both bottom AP cylinders (Fig. 1D), similar to *T. vulcanus* PBS (PDB 7vea) (*17*). The absence of the fourth $(\alpha\beta)_3$ trimer may be due to its loss during sample preparation (*1*, *17*, *18*), though light-controlled reconfiguration of the PBS architecture is also possible (*19*). Notably, the missing trimer would contain ApcD, one of the terminal emitters responsible for the energy transfer to the photosystems in other PBS (*6–8*). Mass spectrometry confirmed that both minor PBPs, ApcD (α-chain) and ApcF (β-chain), were present in our GviPBS preparations (Table S1). Since no justification based on density alone was found for an alternative location of ApcD (instead of the common location in the fourth trimer), ApcD was not included in our models. ApcF was modeled in its expected position in the third AP trimer of the basal cylinders (*6–8*), although the corresponding density did not allow us to distinguish it unambiguously from ApcB because the two proteins are highly identical (70.8%). Notably, *A. panamensis* PBS naturally lacks ApcD and ApcF (*5*). Similarly, deleting either gene does not affect the functionality or regulation of PBS in *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis* 6803) (*20*). Thus, we assume that the fourth ApcD-containing AP trimer is dispensable for GviPBS functioning.

The AP core width is narrower than the lower level of the rods (Fig. 1D). However, blurred density proximal to Rx1/Rx1' suggests the tentative placement of Rx2/Rx2' (Fig. S8). Consistent with the mutagenesis data (*16*), we hypothesize that there are up to three Rx hexamers on each side of the core. The third Rx substructure is likely too dynamic to have a distinct density (Fig. S8). The tentative positioning of three Rx hexamers is sterically compatible with the lateral stacking of GviPBS into arrays, similar to those of PBS in crown cyanobacteria (*3*, *7*) (Fig. S8). This scenario is most probable in thylakoid-less *G. violaceus* because it saves space on the plasmalemma, which accommodates all photosynthetic complexes in this organism (*12*, *14*).

2. Linker proteins

The reconstructed GviPBS structure contains 82 $(\alpha\beta)_3$ trimers of PBP and 864 bilins. It is the largest resolved cyanobacterial PBS, with dimensions of 450x550 Å, exceeding those of paddle-like and hemi-discoidal PBS (Fig. 1E). The bundle-shaped morphology is maintained by linker proteins (Fig. 2A), including the multidomain linkers Glr1262 and Glr2806, which are unique to *Gloeobacter* (*16*, *21*, *22*). Each is predicted to have three pfam00427 (REP) domains in its sequence (*16*, *21*). Nevertheless, these proteins are only 58.4% identical to each other and are dissimilar to the functional analogs CpcG and CpcL from *Synechocystis* 6803 (<30% sequence identity). We identified the precise location of Glr1262 and Glr2806 REP domains by cryo-EM density for non-conserved amino acids (Fig. S9, S10). Glr2806 is responsible for attaching Rx to the core (*16*). Its long N-terminal arm interacts with the A/A' and C/C' cylinders, while its REP1 domain binds inside the Rx1/Rx1' cavity (Fig. 2A-C). The REP2-REP3 domains of Glr2806 are expected to tether two additional Rx copies, Rx2/Rx2' and Rx3/Rx3', to the core (Fig. S8) (*16*). Although mass spectrometry could not cover the entire Glr2806 sequence (Table S1), our data unequivocally confirm Glr2806's role in attaching Rx structures to the core (Fig. 2B). Interestingly, Glr2806 of *G. morelensis* MG652769 also contains three REP domains, but the *G. kilaueensis* JS1 homolog lacks one (Fig. S11). This suggests that three Rx hexamers are not necessary for bundle-shaped PBS assembly.

The Glr1262 linker is also predicted to have three pfam00427 domains, which enable the interaction with three PBP hexamers. Each REP domain of Glr1262 is superimposable with Glr2806 REP1 (0.6-1 Å Cα RMSD; Fig. S12). The density helped unequivocally localize Glr1262 on the first level of PC hexamers in the rods, consistent with its predicted position (*16*). The three REP domains (pfam00427) of Glr1262 were identified at vertices of an equilateral triangle, which corresponded to the cavities of three PC hexamers that formed the base of a rod bundle (Fig. 2C). Although incomplete density prevented tracing the entire Glr1262 chain, assigning REP domains to a Glr1262 copy that connects neighboring PC hexamers within the rod bundle appears to be the most reasonable option. Two copies of Glr1262, each with three REP domains, are necessary and sufficient to cluster the six rods symmetrically into two triangular rod bundles. Each triangle is centered at the top side of the C/C' cylinder of AP (Fig. 2C). In the absence of a CpcG homolog in the genome, a rod-core linker found in studied cyanobacteria, Glr1262 apparently plays this role in *Gloeobacter* (*16*, *21*). Indeed, the REP1-REP2 arm (residues 280-350) of Glr1262 interacts with the C/C' cylinder residues and the REP2-REP3 arm interacts with the B cylinder of the AP core (Fig. 2D).

We located the conserved rod linkers, CpcC1 and CpcC2 (*22*), which contain pfam01383 and pfam00427 domains that connect two neighboring PC layers (Fig. 2A, E). According to mass spectrometry, all small capping linkers were present in the sample (Table S1). However, we

could unequivocally localize multiple copies of only the core linker ApcC (Fig. 2A). Their pfam01383 domain is similar to those of other PBS types and to the pfam01383 domain in CpcC1/CpcC2 (Fig. 2E).
The pentacylindrical GviPBS core is assembled by two copies of the core-membrane linker ApcE (or L_{CM}), which has four conserved REP domains (pfam00427) that bind within the cavities of four AP hexamers (Fig. 2F and Fig. S12). Pfam00427 domains of ApcE are conformationally similar to REP domains of Glr2806, Glr1262 and CpcC1/CpcC2 (<1.6 Å Cα RMSD) (Fig. S13). Although ApcE of GviPBS is conformationally similar to other ApcE proteins from PBS with the pentacylindrical cores, it contains unique unstructured protrusions that potentially interfere with the attachment of the fourth ApcD-containing trimer (Fig. 2F). As in other known cyanobacterial PBS structures, ApcE is the only linker in GviPBS with a chromophore. It is canonically located within the N-terminal domain of ApcE, incorporated into the third AP $(\alpha\beta)_3$ trimer of the basal cylinders as the α subunit.

3. Chromophore system

GviPBS is notable for containing not only PCB, but also PEB and PUC chromophores, which give *G. violaceus* its characteristic violet color. PCB is located in AP core and PC hexamers in rods, whereas PEB and PUC are solely located within the upper part of the rods (Fig. 1D). The PEB and PUC chromophores could not be sufficiently resolved in cryo-EM maps; therefore, their conformations and local environments were inferred from the crystal structure of PE (PDB 2vjh). PCB conformations within PC hexamers in rods and AP hexamers in the core were resolved in cryo-EM density map, including the terminal emitting bilin in ApcE (Fig. 3A-E). The three bilin types differ in the number of conjugated double bonds, increasing from PUB to PEB to PCB, which is accompanied by enhanced coplanarity of the pyrroles (Fig. 3F,G). While PCB and PEB are covalently attached to the single conserved cysteines of the α and β subunits, PUB is doubly tethered to two cysteines simultaneously (Fig. 3A-E). The deficit of double bonds between the pyrroles and the two-sided covalent linkage forces PUB to adopt a bent conformation, with A and D rings completely outside the B-C plane (Fig. 3A). This is consistent with the absorbance maximum of ~500 nm for PUB, the shortest among *G. violaceus* PBPs (Fig. 1C) (*9*). PEB has two double bonds in the linkages between the pyrrole rings, with A and D rings displaced on the opposite sides of the B-C plane (Fig. 3B,F), featuring a red-shifted absorbance maximum of 560 nm. PCBs with the three double bonds in the linkages between the four pyrrole rings adopt conformations with different degrees of ring coplanarity and the optional presence of hydrophobic side chains in the vicinity of the D ring. The ApcE bilin conformation has the most coplanar A-B-C-D rings. Its A ring flips from a *ZZZasa* to a *ZZZssa* configuration (*23*), while the D ring remains coplanar due to a π-stacking interaction with the Trp162 side chain (Fig. 3E), missing from PCB-binding sites in other AP or PC hexamers (Fig. 3A-D). Correspondingly, the *ZZZssa* bilin of ApcE is a red-shifted fluorophore with the properties of the terminal emitter (*23*).
The stoichiometry of bilin incorporation (1:1 per αβ protomer of AP, 1:2 per αβ protomer of PC, 2:3 per αβ protomer of PE) establishes a descending bilin concentration gradient from the top to the button of GviPBS (Fig. S14), which may be relevant for intra-PBS excitation energy transfer (EET). EET pathways are typically estimated based on minimal interbilin distances (*5*, *6*, *8*). The bundles of rods position multiple PCB chromophores within 35 Å of each other, which is sufficient for interrod EET. This increases the probability of capturing photons in typical *G. violaceus* habitats (*1*, *14*, *16*). However, the absence of lateral radiating rods in *G.*

violaceus, which enable direct EET to the A/A' core cylinders in crown cyanobacteria (*7*, *8*), suggests peculiarities in rod-core EET in *G. violaceus*.
Remarkably, the minimal distances (~33-35 Å) between the rod and core bilins in GviPBS were found to be between the PCBs of the lowest PC hexamers and the upper PCBs of the AP hexamers of C/C' cylinders (Fig. 3H). The corresponding distances involving PCBs of the top B core cylinder are 37-41 Å. Furthermore, the coaxial position of PC hexamers Rx1/Rx1' with C/C' core cylinders brings several PCBs of those PC hexamers to distances of 27-31 Å from AP bilins (Fig. 3H). This is expected to substantially facilitate EET and indicates the special relevance of C/C' hexamers (a hallmark feature of the pentacylindrical cores) in receiving energy from rods and auxiliary PC hexamers in GviPBS.
In summary, we determined the cryo-EM structure of one of the largest phycobilisomes known, with the unique morphology characterized by the presence of rod bundles. We revealed that the additional C/C' cylinders are a unique structural feature of the pentacylindrical core that add extra EET pathways to A/A' cylinders. In turn, this requires additional solutions to regulate these energy flows by non-photochemical quenching under the conditions of intense insolation, warranting further investigations.

Materials and Methods

Phycobilisome isolation

The strain of *Gloeobacter violaceus* PCC 7421 used in this study was provided by the Collection of microalgae and cyanobacteria IPPAS of the K.A. Timiryazev Institute of Plant Physiology of the Russian Academy of Sciences (collection ID: IPPAS-B470). *G. violaceus* culture was grown without shaking for 2 months in 1 L Erlenmeyer flasks containing 0.5 L of standard BG-11 medium (*24*) buffered with 10 mM HEPES at 22 °C and illuminated with LED (4000 K, 8W) with light intensity 10 μmol photons×s^{-1}×m^{-2}. The cells were harvested by centrifugation at 3300×g for 30 min and washed once in 0.65 M potassium phosphate buffer (pH 7.4) with 0.5 M sucrose, centrifuged and suspended in 6 ml of the same buffer to a concentration of 0.1 g (wet weight) per ml. The cells were disrupted through a french press at 1500 psi (Thermo electron) three times. Triton X-100 was added to the brown lysate to the final concentration of 2% (v/v), and the solution was gently shaken for 20 min under aluminium foil. Unbroken cells and cell debris were removed by centrifugation at 15557×g for 45 min. The blue-violet PBS containing supernatant fraction (830 μl per tube) was loaded onto a step sucrose gradient: 0.6 mL of 2.0 M; 2 mL of 1.0 M; 2 mL of 0.75 M, 2 mL of 0.6 M, and 2.4 ml of 0.5 M sucrose made on 0.75 M potassium phosphate buffer (pH 7.4). The resulting gradients were centrifuged at 288000×g for 6 h using a SW41 Ti rotor on a Beckman L8-80M centrifuge at 25 °C. All buffers contained 3 mM sodium azide and procedures described above were carried out at 25 °C. Intact phycobilisomes (PBS) were recovered from the middle part of the 1 M sucrose layer. Collected samples were stored at 4 °C for immediate use. Long-term preservation (at least 3-4 months) could be achieved by adding ammonium sulfate until the final concentration of 1 M.

Phycobiliprotein separation

Cells were harvested as described above and frozen at –80 °C until needed. Cell pellet was thawed and washed twice with a buffer A (30 mM Tris-HCl (pH 8.1), 5 mM EDTA) and suspended in the same buffer containing 0.1 mM PMSF and 0.26 mg/ml lysozyme, to a concentration of 0.4 g (wet weight) per ml. The cells were disrupted through a french press at 1500 psi (Thermo electron, USA) three times. Brown lysate was centrifuged at 52360×g for

30 min using LG-25M centrifuge (Shuke Instrument, China) at 4 °C to remove cell debris and membrane fragments. The resulting supernatant was fractionated with ammonium sulfate. Solid ammonium sulfate was added to 20% of saturation and the turbid solution was centrifuged at 52360×g for 30 min, which led to partial phycobiliprotein precipitation. Ammonium sulfate was added to 40% saturation and centrifugation was repeated, yielding deep blue supernatant and precipitation of pink protein, which was resuspended in buffer A. Ammonium sulfate was raised to 60% of saturation. Centrifugation was repeated and the blue precipitate was resuspended in buffer A. Pink and blue fractions from 40% and 60% ammonium sulfate saturation, respectively, were dialyzed against 10 mM Tris-HCl pH 7.85 at 4 °C overnight. The dialyzed solutions of pink and blue fractions were clarified by centrifugation at 52360×g for 30 min. Both fractions were further purified in the same way by anion-exchange chromatography on a HiTrapQ HP column (5 ml, Cytiva) as follows. The protein sample was loaded on the column pre-equilibrated with the dialysis buffer, and washed with 20 mM Tris-HCl (pH 7.6) until the flowthrough fraction became completely colorless. Then, bound phycobiliproteins were eluted with a linear gradient of 0-0.5 M NaCl in 20 mM Tris-HCl (pH 7.6). Phycoerythrins from the pink fraction eluted at 140-250 mM NaCl, whereas phycocyanins and allophycocyanins from the blue fraction eluted at 110-260 mM NaCl and 320-500 mM NaCl, respectively. Phycobiliprotein preparations were stored at 4 °C.

GraFix

PBS have a high tendency to dissociate in low-ionic strength buffers and are often chemically stabilized for structural studies (*5*, *15*–*17*). We used the GraFix method (*15*) for mild chemical crosslinking of PBS without undesired particle aggregation. The ultracentrifuged PBS fraction was dialyzed against 0.75 M potassium phosphate buffer (pH 7.4) with 0.4 M sucrose at 4 °C overnight and passed through a filter paper. The resulting solution was concentrated with 30-kDa Vivaspin Turbo 4 centricon (Sartorius, Germany) till the final sample absorbance at 566 nm equal to 4.9. The concentrated sample (150 µl per tube) was loaded onto a step sucrose (M); glutaraldehyde (GA) (v/v %) gradient made with 0.75 M potassium phosphate buffer pH 7.4: 1 mL of 2.0 M; 0.05%; 2 mL of 1.0 M; 0.03%; 2 mL of 0.75 M; 0.01%, 2 mL of 0.6 M; 0.005%, and 2.9 ml of 0.5 M. Control gradient with the PBS sample was made without GA to exclude its adverse effect on the sedimentation pattern. Centrifugation was performed at 288000×g for 6 h using a SW41 Ti rotor on a Beckman L8-80M centrifuge at 25 °C. After centrifugation, crosslinked PBS (GraFix-PBS) fractions were obtained from the 1 M sucrose layer completely separated from the aggregated PBS. Isolated GraFix-PBS fraction was quenched with glycine-KOH pH 7.7 (final concentration 73 mM) and dialyzed against 50 mM Tris-HCl (pH 7) at 4 °C overnight. GraFix-PBS were concentrated with 30-kDa Vivaspin Turbo 4 centricon (Sartorius, Germany) till final sample absorbance at 566 nm equal to 8.074 and used for grid preparation.

Cryo-EM sample preparation

A 3 µl of the sample was applied to the Quantifoil R1.2/1.3 300 mesh Cu grid coated with 2 nm amorphous carbon film, which was not glow-discharged prior to sample application. The grid was plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific, USA) with the following settings: chamber humidity 100 %; chamber temperature 4 °C; blotting time 3 s; blotting force 0. The grid was then stored in liquid nitrogen until use.

Cryo-EM data collection

Cryo-EM data were collected on a Titan Krios 60-300 transmission electron microscope (Thermo Fisher Scientific, USA) equipped with a field emission electron gun X-FEG (Thermo Fisher Scientific, USA), spherical-aberration corrector (CEOS GmbH, Germany), a post-column BioQuantum energy filter (Gatan, USA) and a K3 direct electron detector (Gatan, USA) in counting non-CDS mode using SerialEM 4.055 and 9-hole image-shift data acquisition strategy at the National Research Centre ''Kurchatov Institute''. The microscope was operated at 300 kV with a nominal magnification of 81,000x, corresponding to a pixel size of 0.874 Å at the specimen level, and an electron energy selecting slit of 20 eV. A total dose of 66 $e^-/Å^2$ within a 3.9 s exposure time was fractionated into 60 frames, resulting in an electron dose of 1.1 $e^-/Å^2$ per frame. In total, ~45,000 movies were collected in a nominal defocus range from -0.7 to -1.7 mm with a step of 0.1 mm. Detailed parameters of data acquisition are listed in Table S2.

Cryo-EM data processing

The initial cryo-EM data were pre-processed in Warp ver. 1.0.9 (*25*). Global and local motion estimation, CTF model estimation, particle picking using a specifically re-trained deep CNN BoxNet were performed for all the movies. All images were inspected in a semiautomated manner using Warp's thresholds for defocus (less than 2 mm), estimated resolution (less than 4 Å) and drift (less than 2 Å). A total of 1,364,249 particles were picked from 41,564 selected images and extracted into 960 pixel boxes. The motion-corrected particle stacks were imported into CryoSPARC ver. 4.7 (*26*) for further data processing. Several rounds of 2D classification were performed using downsampled to 480 pixels boxes, 50 classes at each step to maximize the number of true-positive particles in the subset of interest. Reference-free 2D classification revealed an accurate C2 symmetry of the PBS particles. A cleaned subset of 746,972 PBS particles, downsampled to 672 pixels (1.25 Å per pixel), was used for *ab initio* initial map reconstruction and homogeneous refinement. Following rounds of refinement with a C1 and C2 symmetry, per-particle defocus adjustment, and adjustments to the higher-order CTF terms for each exposure group, as well as local refinement of the most stable core part of the PBS, a 2.85 Å resolution cryo-EM density map of the PBS core was obtained (EMD-64815).

To elucidate the structural intricacies of the more dynamic parts of the PBS as well as the structure of the AP core hexamers, five soft masks were created using UCSF Chimera (*26*) based on the locally refined PBS core map. Further local refinement resulted in two maps of the core part: 2.94 Å PBS C (EMD-64816), 2.72 Å PBS Rx (EMD-64812), and three maps of the hexamers of the 1st and the 2nd rods: 3.76 Å PBS R1 (EMD-64817), 2.72 Å PBS R2 (EMD-64813), 3.03 Å PBS R3 (EMD-64814).

To elucidate the structural heterogeneity of the rods, signal subtraction of the core part was performed followed by heterogeneous refinement. Two distinct conformations with diverging and parallel rods were obtained. Subsequent non-uniform refinement yielded a 4.25 Å map for diverging conformation (EMD-64851) and a 4.66 Å map for parallel conformation, with C2 and C1 symmetry imposed, respectively. Schematic representation of data processing is shown in Fig. S2.

Model building and refinement

Crystal structures of AP, PC and PE hexamers from *G. violaceus* (PDB codes: 2vjt, 2vjr and 2vjh, respectively) as well as cryo-EM structures of the ApcE and ApcC proteins from *Anabaena* 7120 (PDB 7eyd) were used for the initial modeling. The atomic models were first docked into the corresponding cryo-EM density maps using UCSF ChimeraX (*27*). The

conformational differences in hexamers were manually adjusted in COOT (*28*). Sequences of the linker proteins were fetched from Uniprot – ApcC (Q7NL78), ApcE (Q7NL81), ApcF (Q7NJA2), Glr1262 (Q7NL64), Glr2806 (Q7NGT2), CpcC1 (Q7NM19), CpcC2 (Q7NGF2). Two ApcF and eight ApcC were built using "mutate residue range" to the corresponding subunits. ApcE was manually adjusted and mutated from the template model. Linker proteins Glr1262, Glr2806, CpcC1 and CpcC2 were built *de novo* by tracing the backbone, assigning the sequence and fitting the side chains. CpcC1 and CpcC2 sequences were extended in N-termini according to the previous work (*22*). All the structures were modelled and fitted using COOT (*28*). The final atomic models were refined in real space using Servalcat Refmac in CCPEM and geometric restraints (*29*). The final atomic models were evaluated using Molprobity (*28*) and the statistics of data collection and model validation were included in Table S2.

Phylogenetic analysis

Construction of the phylogenetic tree was carried out for Glr2806 and Glr1262 homologs using the same protocol. Protein sequences of full-length homologs were retrieved from the NCBI core nucleotide BLAST database (GenBank+EMBL+DDBJ+PDB+RefSeq), update date:2025/05/25. Multiple sequence alignment was performed using Clustal Omega in the EMBL-EBI Job Dispatcher sequence analysis tools framework (*30*). A tree was built using MEGA (*31*) using Maximum Likelihood method and Jones-Taylor-Thornton model (*32*) of amino acid substitutions. Gap positions in the alignment were completely deleted resulting in a final data set. The initial tree for the heuristic search was selected by choosing the tree with the superior log-likelihood between a Neighbor-Joining (NJ) tree (*33*) and a Maximum Parsimony (MP) tree. The NJ tree was generated using a matrix of pairwise distances computed using the Jones-Taylor-Thornton model (*32*). The MP tree had the shortest length among 10 MP tree searches, each performed with a randomly generated starting tree. The number of bootstrap replicates was determined adaptively (*31*).

Acknowledgements

The authors are thankful to Prof. Alexey Amunts for help with preliminary screening of GviPBS samples on grids and to Dr. Rustam H. Ziganshin for help with running LC-MS for identification of GviPBS proteins.

We thank Resource Center of Probe and Electron Microscopy of National Research Center "Kurchatov Institute" (http://rc.nrcki.ru/pages/main/nanozond/facilities/12604/index.shtml). This work has been carried out using computing resources of the federal collective usage center Complex for Simulation and Data Processing for Mega-science Facilities at National Research Center "Kurchatov Institute" (http://ckp.nrcki.ru).

Funding

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (Y.B.S., K.M.B., V.O.P. and N.N.S.), Ministry of Science and Higher Education of the Russian Federation theme no.122042700045-3 (M.A.S.) and the Russian Science Foundation grant 23-74-00062 (A.D.B.).

Author contributions

Conceptualization: NNS
Methodology: ADB, YBS, TNB, EGM, VOP, NNS
Investigation: ADB, YBS, TNB, MAS, EGM, KMB, NNS

Visualization: ADB, YBS, EGM, NNS
Funding acquisition: MAS, VOP, KMB, NNS
Project administration: NNS
Supervision: KMB, NNS
Writing – original draft: NNS
Writing – review & editing: NNS, with input from all authors

Declaration of competing interests
The authors declare no conflicts of interest.

Data availability
The atomic coordinates have been deposited in the Protein Data Bank with the accession codes 9V7J (PBS core), 9V7K (PBS C), 9V7G (PBS Rx), 9V7L (PBS R1), 9V7H (PBS R2) and 9V7I (PBS R3). The cryo-EM maps have been deposited in the Electron Microscopy Data Bank with accession codes EMD-64815 (PBS core), EMD-64816 (PBS C), EMD-64812 (PBS Rx), EMD-64817 (PBS R1), EMD-64813 (PBS R2), EMD-64814 (PBS R3), EMD-64828 (consensus), EMD-64868 (composite), EMD-64851 (diverging rods) and EMD-64852 (parallel rods). Other data are available in the paper or in the Supplementary information.

References

1. D. A. Bryant, C. J. Gisriel, The structural basis for light harvesting in organisms producing phycobiliproteins. *The Plant Cell*, koae126 (2024).

2. L. Chang, X. Liu, Y. Li, C.-C. Liu, F. Yang, J. Zhao, S.-F. Sui, Structural organization of an intact phycobilisome and its association with photosystem II. *Cell Res* **25**, 726–737 (2015).

3. X. Zhang, Y. Xiao, X. You, S. Sun, S.-F. Sui, In situ structural determination of cyanobacterial phycobilisome-PSII supercomplex by STAgSPA strategy. *Nat Commun* **15**, 7201 (2024).

4. X. You, X. Zhang, J. Cheng, Y. Xiao, J. Ma, S. Sun, X. Zhang, H.-W. Wang, S.-F. Sui, In situ structure of the red algal phycobilisome–PSII–PSI–LHC megacomplex. *Nature* **616**, 199–206 (2023).

5. H.-W. Jiang, H.-Y. Wu, C.-H. Wang, C.-H. Yang, J.-T. Ko, H.-C. Ho, M.-D. Tsai, D. A. Bryant, F.-W. Li, M.-C. Ho, M.-Y. Ho, A structure of the relict phycobilisome from a thylakoid-free cyanobacterium. *Nat Commun* **14**, 8009 (2023).

6. Z. Zheng, C. Ma, H. Wang, G. Wang, C. Dong, N. Gao, J. Zhao, The structure of phycobilisome with a bicylindrical core from the cyanobacterium Synechococcus elongatus PCC 7942. bioRxiv [Preprint] (2025). https://doi.org/10.1101/2025.04.28.650843.

7. M. A. Domínguez-Martín, P. V. Sauer, H. Kirst, M. Sutter, D. Bína, B. J. Greber, E. Nogales, T. Polívka, C. A. Kerfeld, Structures of a phycobilisome in light-harvesting and photoprotected states. *Nature* **609**, 835–845 (2022).

8. L. Zheng, Z. Zheng, X. Li, G. Wang, K. Zhang, P. Wei, J. Zhao, N. Gao, Structural insight into the mechanism of energy transfer in cyanobacterial phycobilisomes. *Nat Commun* **12**, 5497 (2021).

9. D. A. Bryant, G. Cohen-Bazire, A. N. Glazer, Characterization of the biliproteins of Gloeobacter violaceus chromophore content of a cyanobacterial phycoerythrin carrying phycourobilin chromophore. *Arch. Microbiol.* **129**, 190–198 (1981).

10. J. Ma, X. You, S. Sun, X. Wang, S. Qin, S.-F. Sui, Structural basis of energy transfer in Porphyridium purpureum phycobilisome. *Nature* **579**, 146–151 (2020).

11. J. Zhang, J. Ma, D. Liu, S. Qin, S. Sun, J. Zhao, S. F. Sui, Structure of phycobilisome from the red alga Griffithsia pacifica. *Nature* **551**, 57–63 (2017).

12. G. Guglielmi, G. Cohen-Bazire, D. A. Bryant, The structure of Gloeobacter violaceus and its phycobilisomes. *Arch. Microbiol.* **129**, 181–189 (1981).

13. T. Tsuchiya, S. Takaichi, N. Misawa, T. Maoka, H. Miyashita, M. Mimuro, The cyanobacterium Gloeobacter violaceus PCC 7421 uses bacterial-type phytoene desaturase in carotenoid biosynthesis. *FEBS Letters* **579**, 2125–2129 (2005).

14. R. Rippka, J. Waterbury, G. Cohen-Bazire, A cyanobacterium which lacks thylakoids. *Arch. Microbiol.* **100**, 419–436 (1974).

15. B. Kastner, N. Fischer, M. M. Golas, B. Sander, P. Dube, D. Boehringer, K. Hartmuth, J. Deckert, F. Hauer, E. Wolf, H. Uchtenhagen, H. Urlaub, F. Herzog, J. M. Peters, D. Poerschke, R. Lührmann, H. Stark, GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat Methods* **5**, 53–55 (2008).

16. H. Wang, Z. Zheng, L. Zheng, Z. Zhang, C. Dong, J. Zhao, Mutagenic analysis of the bundle-shaped phycobilisome from Gloeobacter violaceus. *Photosynth Res* **158**, 81–90 (2023).

17. K. Kawakami, T. Hamaguchi, Y. Hirose, D. Kosumi, M. Miyata, N. Kamiya, K. Yonekura, Core and rod structures of a thermophilic cyanobacterial light-harvesting phycobilisome. *Nat Commun* **13**, 3389 (2022).

18. C. J. Gisriel, G. Shen, G. W. Brudvig, D. A. Bryant, Structure of the antenna complex expressed during far-red light photoacclimation in *Synechococcus* sp. PCC 7335. *Journal of Biological Chemistry* **300**, 105590 (2024).

19. E. J. Dodson, J. Ma, M. Suissa Szlejf, N. Maroudas-Sklare, Y. Paltiel, N. Adir, S. Sun, S.-F. Sui, N. Keren, The structural basis for light acclimation in phycobilisome light harvesting systems systems in Porphyridium purpureum. *Commun Biol* **6**, 1210 (2023).

20. D. Jallet, M. Gwizdala, D. Kirilovsky, ApcD, ApcF and ApcE are not required for the Orange Carotenoid Protein related phycobilisome fluorescence quenching in the cyanobacterium Synechocystis PCC 6803. *Biochimica et biophysica acta* **1817**, 1418–27 (2012).

21. D. W. Krogmann, B. Pérez-Gómez, E. B. Gutiérrez-Cirlos, A. Chagolla-López, L. González de la Vara, C. Gómez-Lojero, The presence of multidomain linkers determines the bundle-shape structure of the phycobilisome of the cyanobacterium Gloeobacter violaceus PCC 7421. *Photosynth Res* **93**, 27–43 (2007).

22. E. Berta Gutiérrez-Cirlos, B. Pérez-Gómez, D. W. Krogmann, C. Gómez-Lojero, The phycocyanin-associated rod linker proteins of the phycobilisome of *Gloeobacter violaceus* PCC 7421 contain unusually located rod-capping domains. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1757**, 130–134 (2006).

23. K. Tang, W.-L. Ding, A. Höppner, C. Zhao, L. Zhang, Y. Hontani, J. T. M. Kennis, W. Gärtner, H. Scheer, M. Zhou, K.-H. Zhao, The terminal phycobilisome emitter, LCM: A light-harvesting pigment with a phytochrome chromophore. *Proceedings of the National Academy of Sciences* **112**, 15880–15885 (2015).

24. R. Rippka, J. Deruelles, J. B. Waterbury, M. Herdman, R. Y. Stanier, Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Microbiology* **111**, 1–61 (1979).

25. D. Tegunov, P. Cramer, Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* **16**, 1146–1152 (2019).

26. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296 (2017).

27. E. C. Meng, T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **32**, e4792 (2023).

28. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Cryst D* **66**, 486–501 (2010).

29. K. Yamashita, C. M. Palmer, T. Burnley, G. N. Murshudov, Cryo-EM single-particle structure refinement and map calculation using Servalcat. *Acta Cryst D* **77**, 1282–1291 (2021).

30. F. Madeira, N. Madhusoodanan, J. Lee, A. Eusebi, A. Niewielska, A. R. N. Tivey, R. Lopez, S. Butcher, The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Research* **52**, W521–W525 (2024).

31. K. Tamura, G. Stecher, S. Kumar, MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* **38**, 3022–3027 (2021).

32. D. T. Jones, W. R. Taylor, J. M. Thornton, The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275–282 (1992).

33. N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425 (1987).

34. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).

Figures and legends

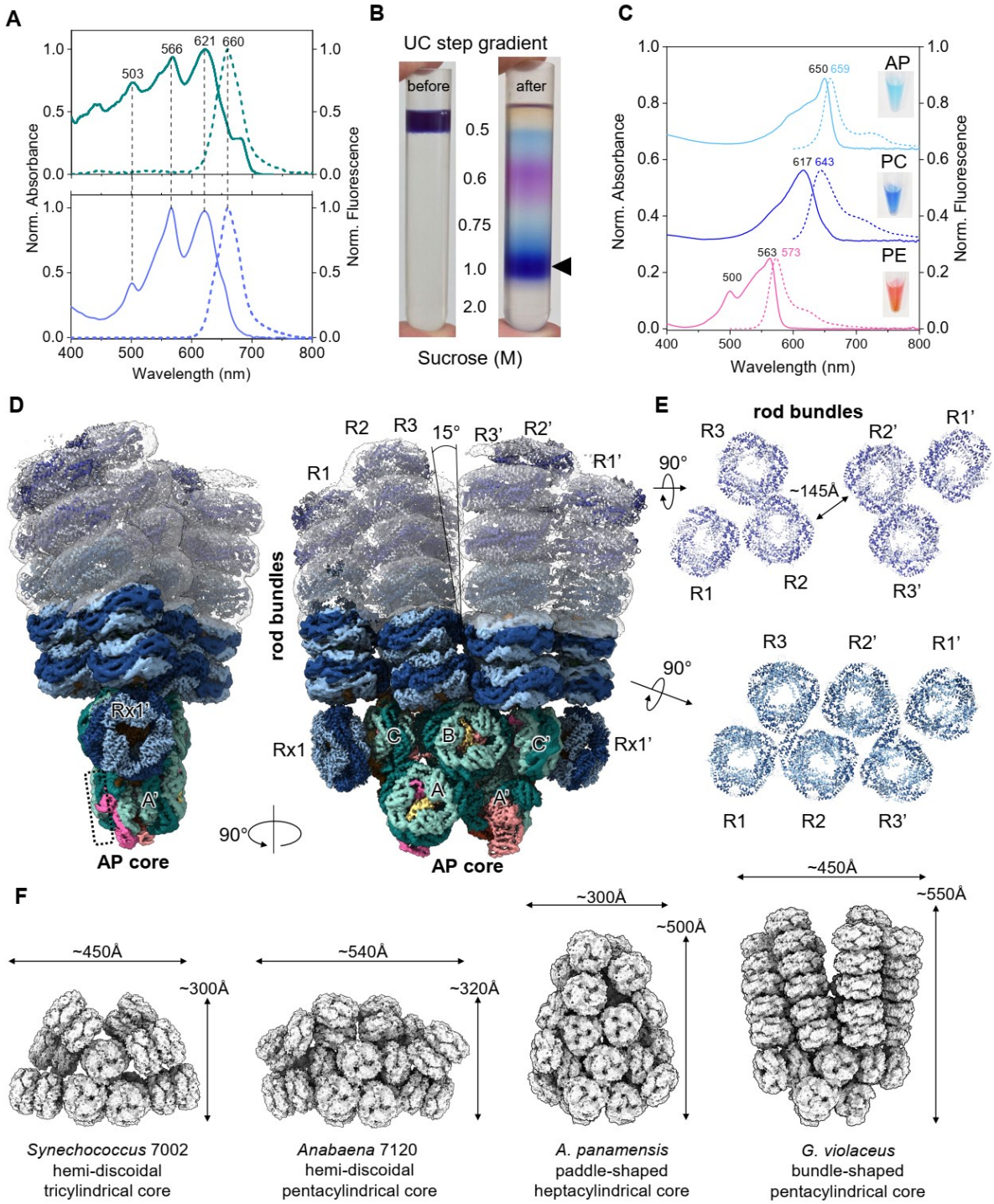


Fig. 1. The bundle-shaped phycobilisome from *G. violaceus*. **A**. Absorption (solid line) and fluorescence (dashed line) spectra of *Gloeobacter* cells (top) and isolated PBS (bottom). An integrating sphere was used to measure the absorption spectrum of the cells. In both cases, fluorescence was excited at 630 nm. **B**. PBS isolation by ultracentrifugation in sucrose gradient as shown. Arrow indicates the PBS zone. **C**. Absorption (solid line) and fluorescence (dashed line) spectra of the individual allophycocyanin (AP), phycocyanin (PC) and phycoerythrin (PE). Excitation was at 475 (PE) or 575 nm (PC, AP). The samples are shown in insert. **D**. Composite cryo-EM map of the bundle-shaped PBS showing the location of the

AP core (A/A', B and C/C' cylinders) and rod bundles. **E**. Sections across the rods on the two levels indicated by arrows, showing diverging three-rod bundles R1-R2-R3 and R1'-R2'-R3'. **F**. Cryo-EM structures representing the relative sizes of different cyanobacterial PBS – hemi-discoidal PBS from *Synechococcus* and *Anabaena* (*8*), paddle-shaped PBS from *A. panamensis* (*5*) and bundle-shaped PBS from *G. violaceus* PBS (this work).

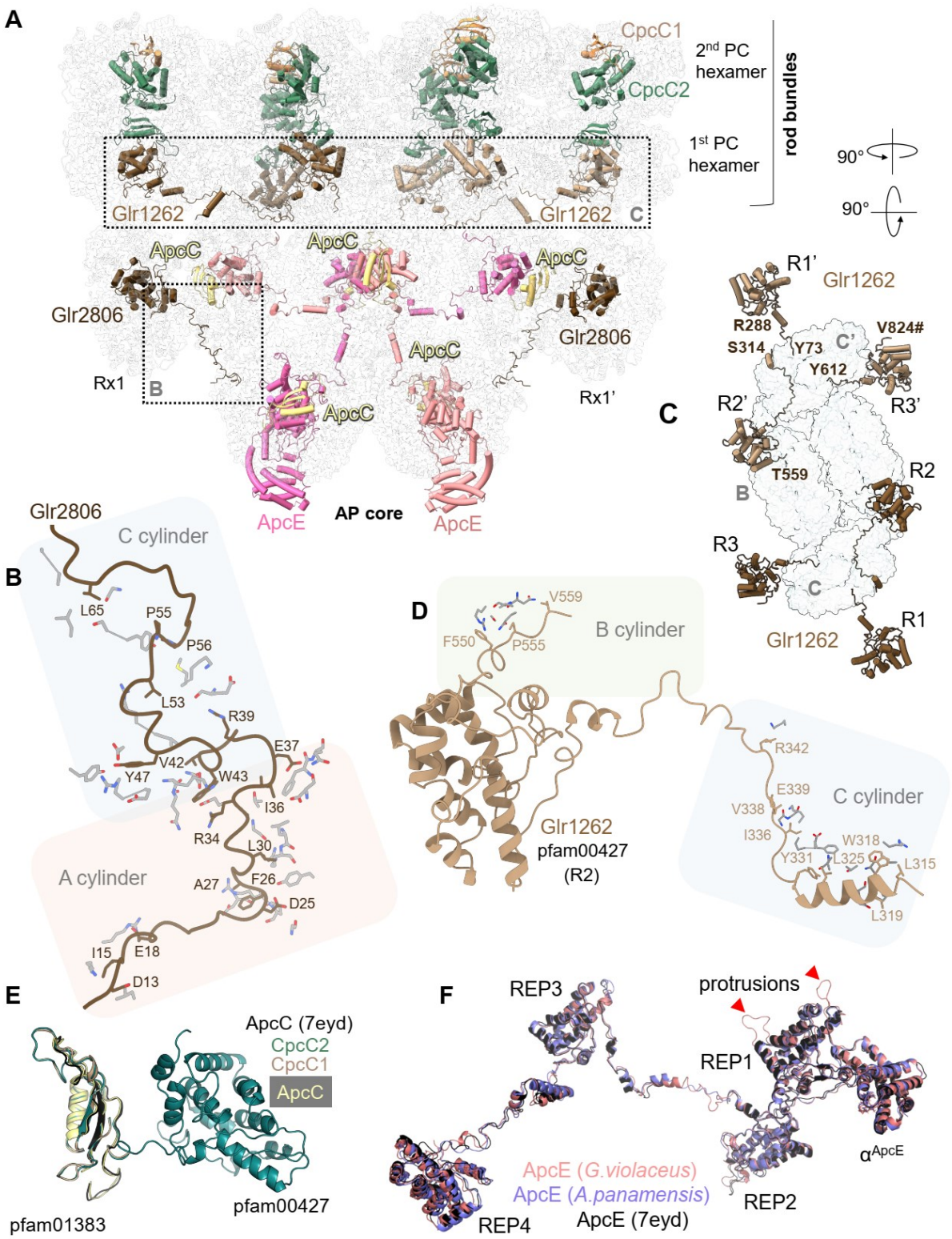


Fig. 2. Linker proteins of GviPBS. **A**. Front view of the AP core and two PC levels in the rods. PBP hexamers are shown as semitransparent ribbons, linker proteins are in cartoon representation. All linkers are labeled except for CpcC1/CpcC2 for which only one copy is labeled. Parts of the model for focused analysis in separate panels are framed. **B**. Contacts formed by the N-terminal arm of Glr2806 with the core cylinders A/A' and C/C'. **C**. Topology of the two copies of Glr1262 on top of the core showing relative positions of the repeat domains

REP1, REP2, REP3 that attach bases of the rods R1/R1', R2/R2', R3/R3', respectively, by clustering them as two rod bundles. # indicates the C terminal of the Glr1262 chain. **D**. Contacts made by the arms of Glr1262 with the core cylinders B and C/C'. **E**. Superposition of ApcC capping linker of G. violaceus with ApcC from Anabaena PBS (PDB 7eyd) and with the equivalent pfam001383 domains of CpcC1 and CpcC2 rod linkers. **F**. Superposition of ApcE from *G. violaceus*, *A. panamensis* and *Anabaena* 7120 showing the location of the REP1-4 domains and the bilin-binding α^{ApcE} domain. The profound difference is the presence of large unstructured protrusions exclusively in REP1 of *G. violaceus* ApcE (indicated by red arrows).

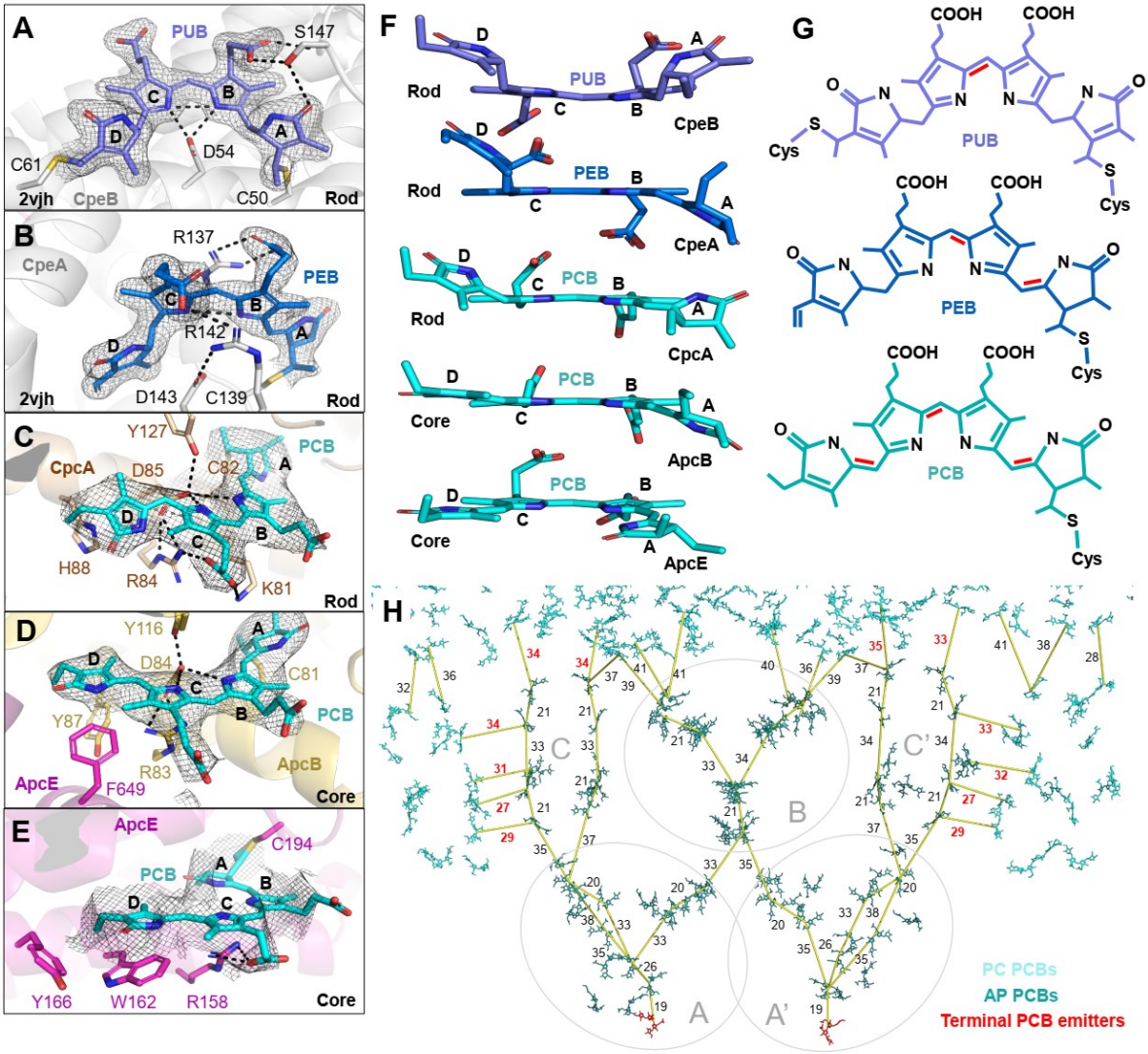


Fig. 3. Bilins of GviPBS. **A-E.** Representative PUB, PEB and PCB bilin conformations, shown with the polar contacts to the surrounding protein residues and the respective electron density (1σ contour). PUB and PEB binding sites are imaged using the crystal structure of PE (PDB 2vjh), PCB sites are from cryo-EM data obtained in this work. **F**. Coplanarity of the bilins shown in panels **A-E**. Pyrrole rings A-B-C-D are labeled. **G**. Structures of the three bilin types found in GviPBS. Red lines mark the distinctive double bonds connecting the pyrroles. **H**. Estimation of the EET pathways from PC to AP and within AP in GviPBS. Distances connecting mass centers of the selected PCBs are indicated in Å by yellow lines. Remarkable shortest distances (≤35 Å) between PC and AP bilins are indicated by red font. Terminal emitting bilins of ApcE are shown by red sticks.

Supplementary information

Supplementary Tables

Table S1. Mass spectrometry-based identification of PBS proteins from *G. violaceus*.

Protein	Category	NCBI Reference	Uniprot ID	Mw, Da	Sequence coverage (MALDI)	Expect value (MALDI)	Sequence coverage (LC-MS)
ApcA	PBPs	WP_011141246.1	Q7NL80	17537.1	59%	2.6e-004	86.3%
ApcB	PBPs	WP_011141247.1	Q7NL79	17211.7	-	-	91.3%
ApcC	linkers	WP_011141248.1	Q7NL78	7754.0	52%	1.1e-008	52.2%
ApcD	PBPs	WP_011141181.1	Q7NLE3	18032.8	-	-	87.6%
ApcE	PBPs/linkers	WP_011141245.1	Q7NL81	129836.6	57%	2.8e-055	72.1%
ApcF	PBPs	WP_011141928.1	Q7NJA2	17433.9	-	-	80.1%
CpcA	PBPs	WP_011141185.1	Q7M7F7	17662.9	82%	1.4e-005	94.4%
CpcB	PBPs	WP_011141184.1	Q7M7C7	18459.9	79%	2.8e-012	97.7%
CpcC1	linkers	WP_164928640.1	Q7NM19	31043.1	77%	3.5e-014	77.2%
CpcC2	linkers	WP_231848246.1	Q7NGF2	30877.9	65%	6.0e-006	66.2%
CpcD-1	linkers	WP_011141267.1	Q7NL59	7769.9	24%	1.4e-011	90.0%
CpcD-2	linkers	WP_011141266.1	Q7NL60	8169.4	48%	4.5e-007	72.4%
CpeA	PBPs	WP_011141189.1	Q7NLD7	17658.0	59%	3.1e-004	94.5%
CpeB	PBPs	WP_011141190.1	Q7NLD6	18427.2	33%	2.0e-002	83.6%
CpeC	linkers	WP_011141263.1	Q7NL63	31825.2	75%	2.2e-013	87.4%
CpeD	linkers	WP_011141264.1	Q7NL62	28400.2	81%	8.9e-014	67.8%
CpeE	linkers	WP_011141265.1	Q7NL61	28354.9	53%	1.4e-010	66.9%
Glr1262	linkers	WP_011141262.1	Q7NL64	92019.8	53%	5.6e-032	74.9%
Glr2806	linkers	WP_011142800.1	Q7NGT2	81422.5	44%	1.1e-007	43.2%

Red font indicates polypeptides not covered by MALDI MS, but reliably confirmed by LC-MS.

Table S2. Cryo-EM data collection, refinement and validation statistics.

	PBS core	PBS C	PBS Rx	PBS R1	PBS R2	PBS R3
PDB	**9V7J**	**9V7K**	**9V7G**	**9V7L**	**9V7H**	**9V7I**
EMDB	EMD-64815	EMD-64816	EMD-64812	EMD-64817	EMD-64813	EMD-64814
Data collection						
Grid	Quantifoil R1.2/1.3 300 mesh Cu					
Microscope	Titan Krios					
Energy filter, Camera	BioQuantum, K3					
Energy selecting slit width, eV	20					
Imaging mode	Counting, non-CDS					
Accelerating voltage (kV)	300					
Condenser apertures C1, C2, um	2000, 100					
Spot size	5					
Illumination area, µm	1.1					
Nominal magnification	85000					
Calibrated pixel size (Å)	0.874					
Total exposure time (sec)	3.9					
Total electron dose (e^-/Å^2)	66					
Nominal defocus range (µm)	-0.7 to -1.7					
Number of frames per stack (movie)	60					
Number of stacks (movies) collected	~45000					
Data processing						
Number of stacks (movies) used	41 564					
Original box size (px)	960					
Initial number of particles	1 364 249					
Final number of particles	746972	746972	746972	746972	207644	746972
Pixel size for final map (Å)	1.25	1.25	1.25	1.25	1.25	1.25
Symmetry imposed	C2	C1	C1	C1	C1	C1
Map resolution (Å)	2.85	2.94	2.72	3.76	2.72	3.03
B-factor for sharpening	77.9	70.8	73.8	96.0	60.2	81.4
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143
Refinement						
Initial models (PDB codes)	2VJT, 7EYD	2VJT	2VJR	2VJR	2VJR	2VJR
Model resolution (Å)	2.85	2.94	2.72	3.76	2.72	3.03
Model composition						
Non-hydrogen atoms	123484	17887	17599	36440	36476	36438
Protein residues	15766	2275	2207	4570	4575	4567
Ligands	84	12	12	24	24	24
B factors (Å^2)						
Protein	56.7	49.5	41.7	122.9	46.5	56.4
Ligands	51.9	35.3	39.4	112.1	42.5	51.5
R.m.s. deviations						
Bond lengths (Å)	0.01	0.01	0.01	0.01	0.01	0.01
Bond angles (Å)	1.89	1.64	1.58	1.45	1.74	1.69
Validation						
MolProbity score	2.47	2.68	1.67	2.76	2.80	2.82

Clashscore	15.59	7.43	7.43	23.53	21.91	20.20
Poor rotamers (%)	7.27	19.60	2.23	1.75	2.50	3.15
Ramachandran plot						
Favored (%)	83.2	96.9	99.1	81.8	84.9	85.8
Allowed (%)	16.4	2.6	0.8	15.6	12.9	11.6
Disallowed (%)	0.4	0.5	0.1	2.6	2.2	2.6

Supplementary figures

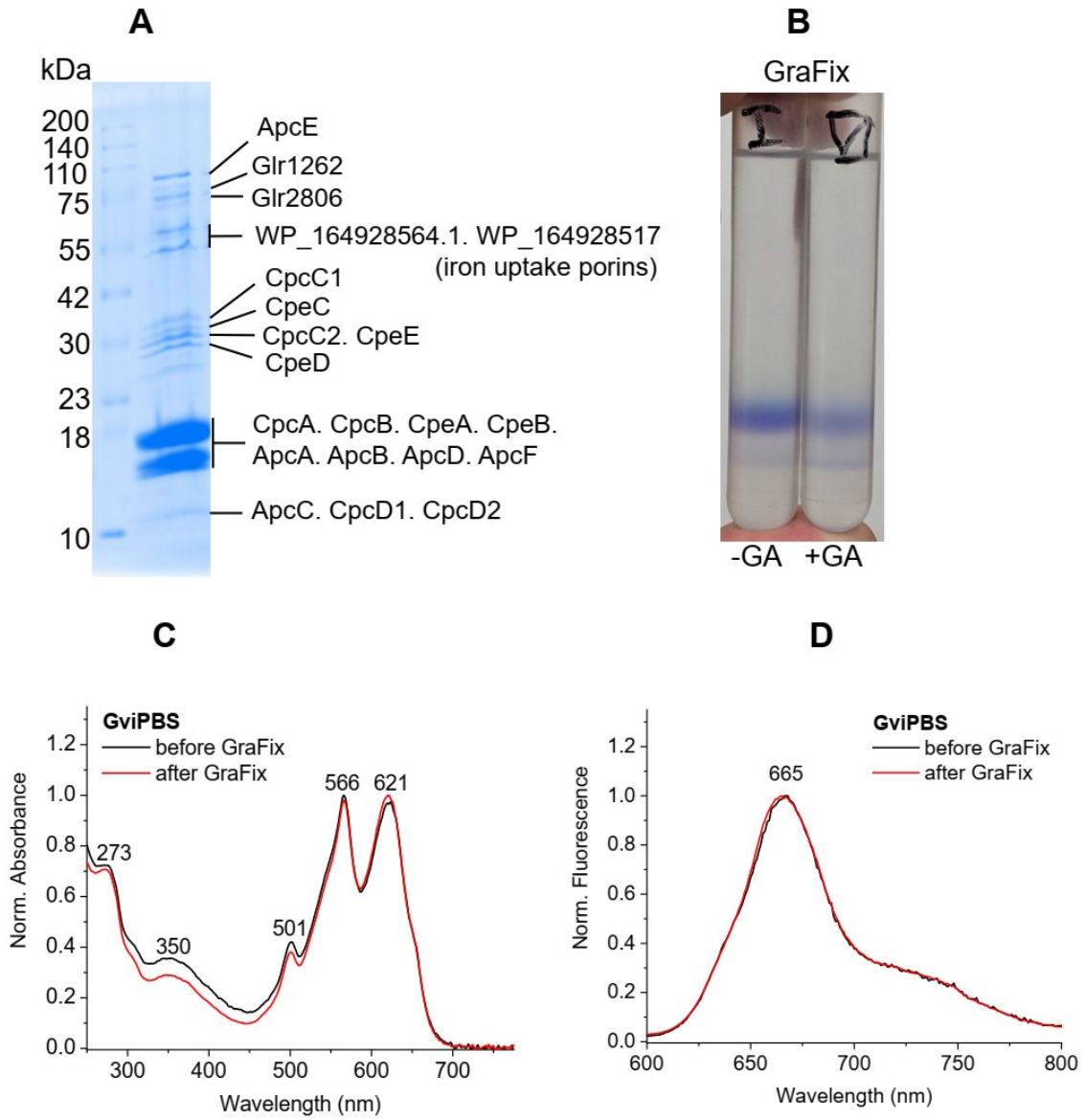


Fig. S1. *G. violaceus* PBS purification and characterization. **A**. SDS-PAGE analysis of the protein constituents of GviPBS. **B**. Ultracentrifugation tubes showing that the sedimentation pattern of GviPBS did not change in the presence of glutaraldehyde (GA) during GraFix (*15*). Absorbance (**C**) and fluorescence (**D**) spectra of the GviPBS sample obtained by ultracentrifugation before and after GraFix showing the preservation of the GviPBS assembly.

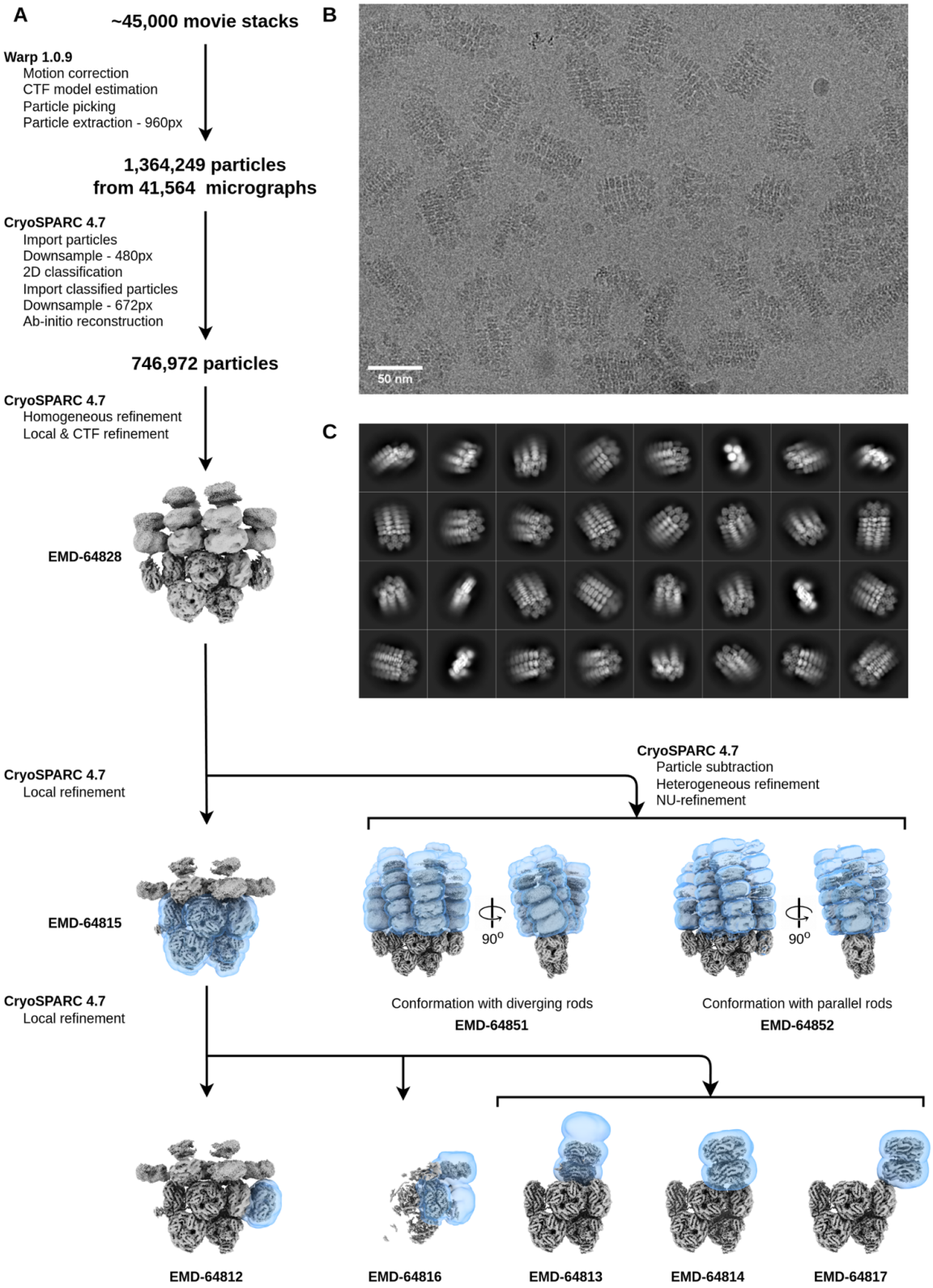


Fig. S2. Cryo-EM data processing workflow and 3D reconstruction. **A**. The workflow for the cryo-EM data processing. Cryo-EM density maps of GviPBS are shown in grey, masks used for the local refinement are shown in transparent blue. Each map is labeled with the corresponding EMDB code. **B**. A representative motion-corrected electron micrograph of GviPBS. The scale bar corresponds to 50 nm. **C**. 2D classification of the cryo-EM data demonstrating different views of the GviPBS.

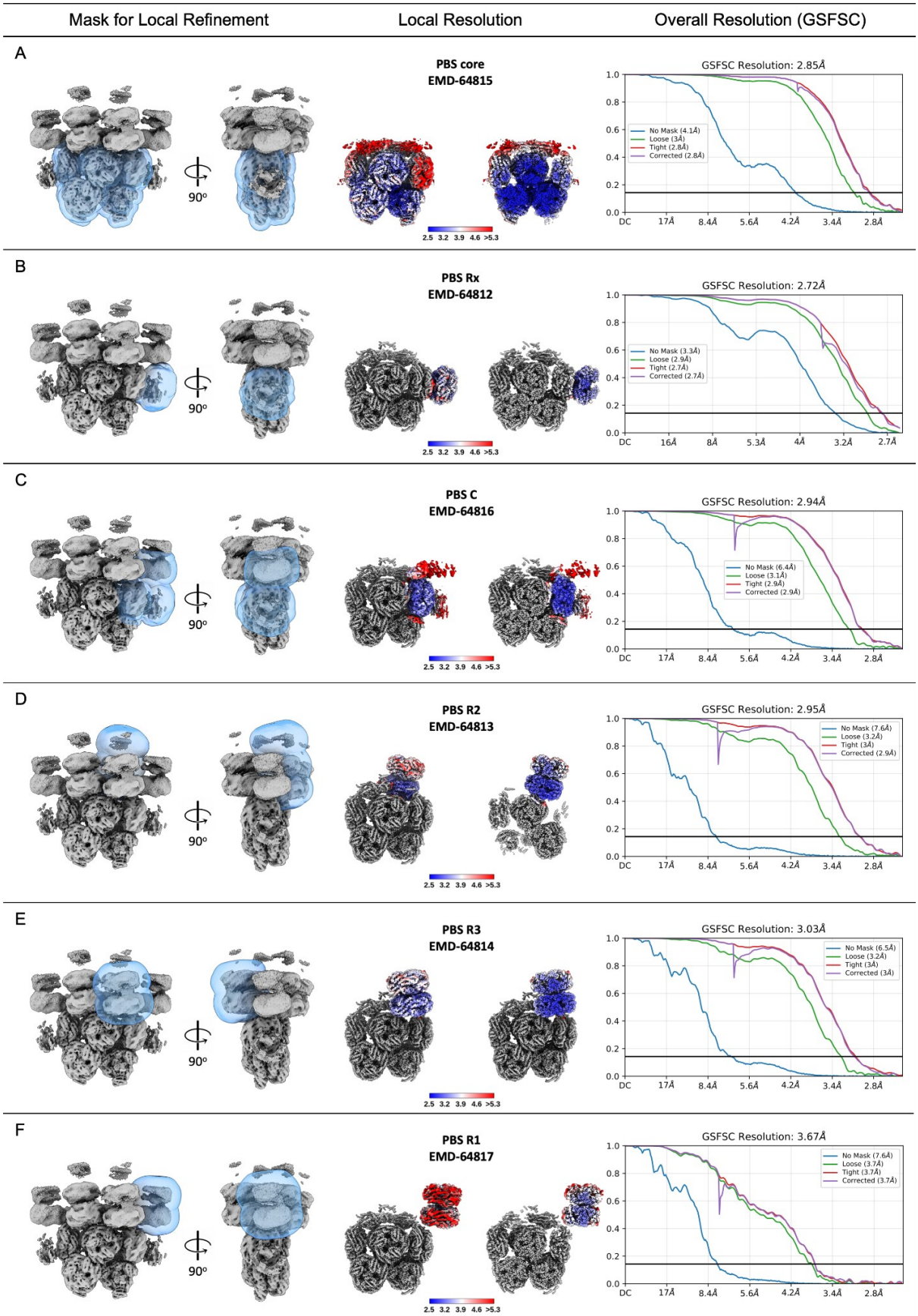


Fig. S3. Overall and local resolution estimation for each local refinement. Left panel: The masks used for the local refinement. The consensus cryo-EM density map is shown as gray, and the mask used for local refinement is shown as transparent blue. Middle panel: Local

resolution estimation for corresponded parts of PBS. The consensus cryo-EM density map is shown as gray, the locally refined cryo-EM density maps are colored according to the local resolution calculated by cryoSPARC, the color key is in Angstroms. Right panel: Gold-standard FSC plots for the locally refined reconstructions. The global resolutions were estimated by gold-standard FSC (threshold at 0.143).

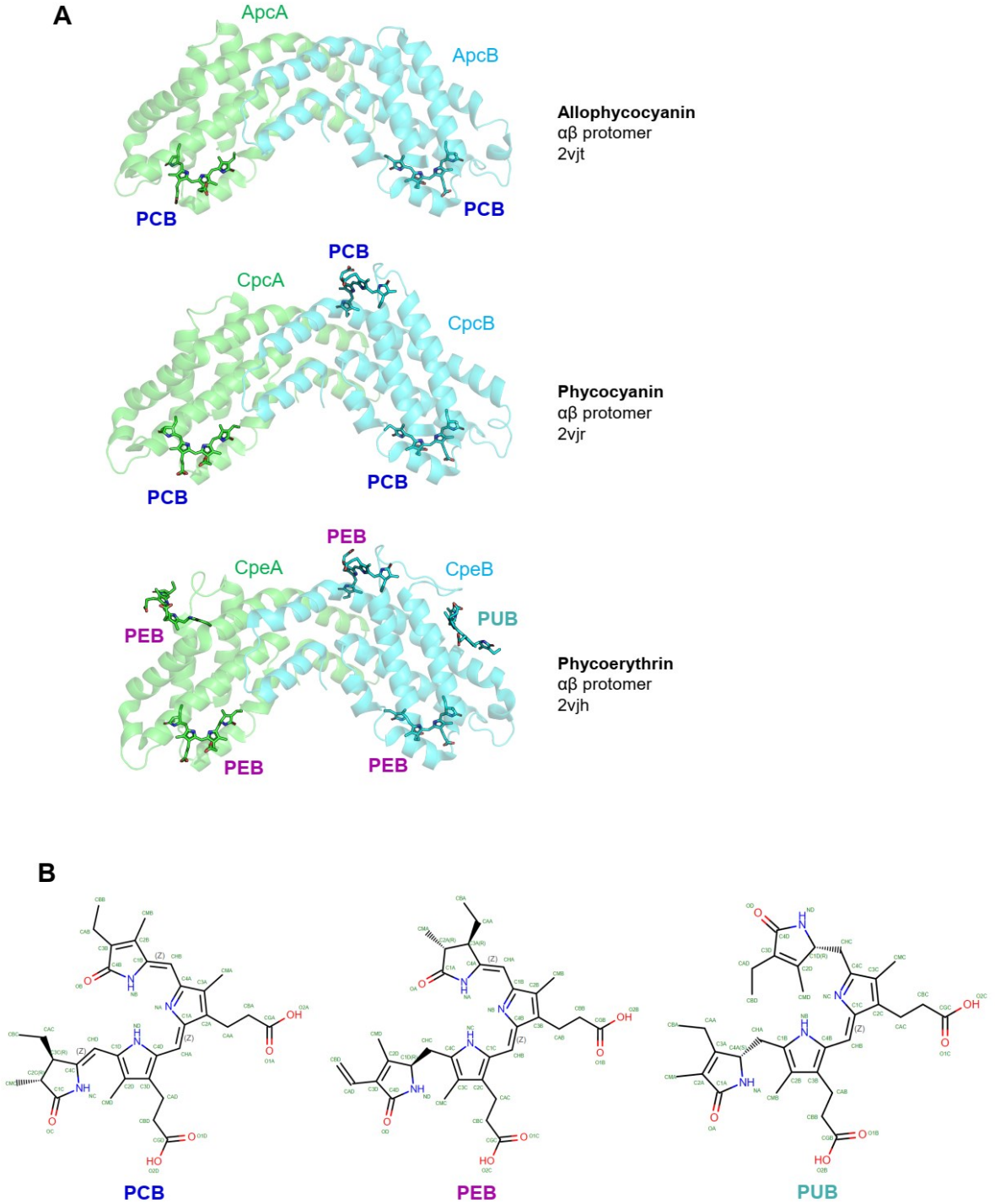


Fig. S4. Crystallographic structures of αβ protomers of allophycocyanin (AP), phycocyanin (PC) and phycoerythrin (PE) from *G. violaceus* with three main chromophores covalently attached to these PBPs. PCB – phycocyanobilin, PEB – phycoerythrobilin, PUB – phycourobilin. PDB codes are indicated. **A**. Cartoon representation of the PBP protomers with bilins shown as sticks. **B**. Chemical structures of the bilins from *G. violaceus* PBS.

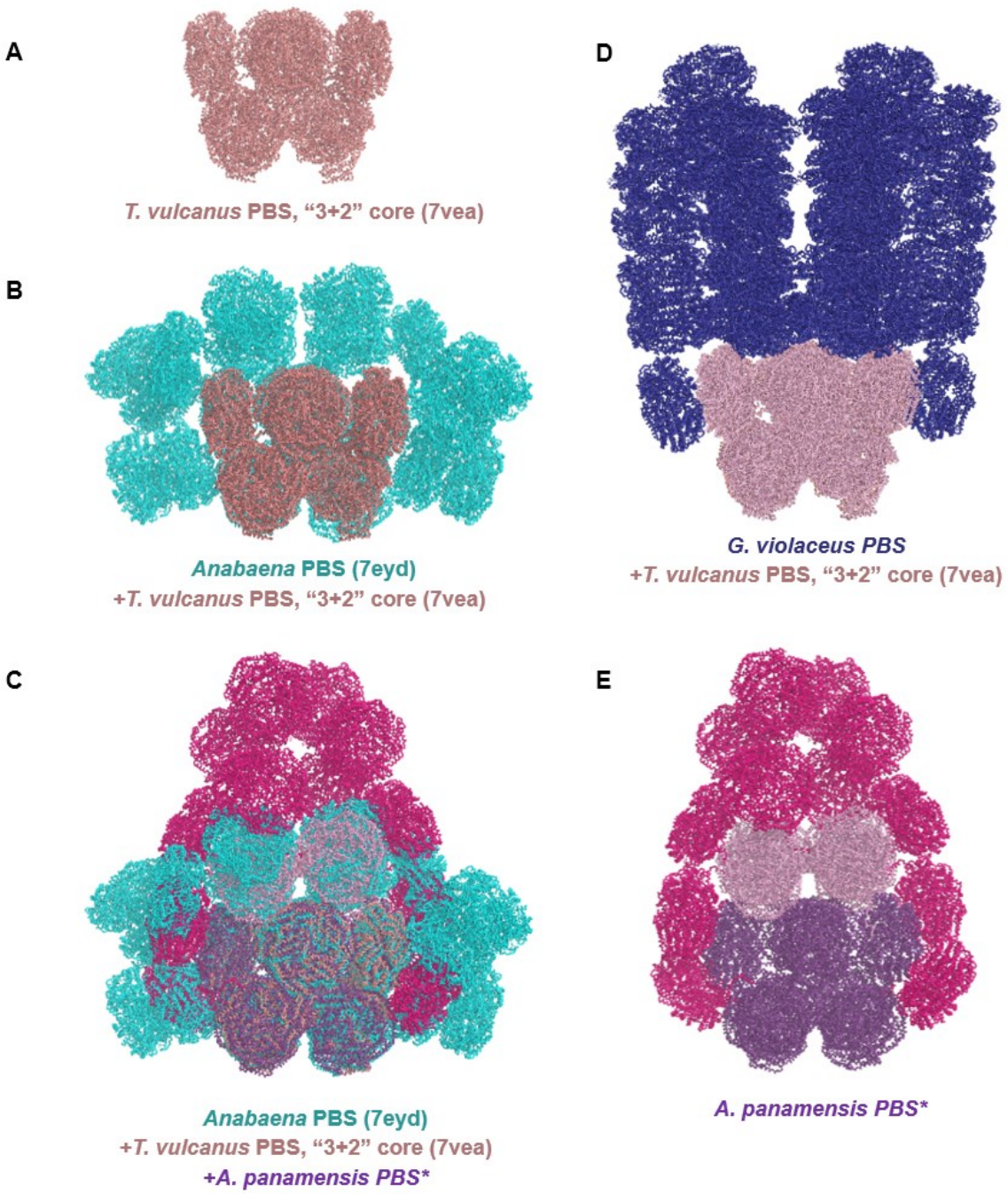


Fig. S5. The pentacylindrical core as the common substructure of PBS with different morphologies. **A**. The pentacylindrical core of *T. vulcanus* PBS. **B**. The pentacylindrical core of *Anabaena* PBS (*8*) shown overlaid with that of *T. vulcanus* (*17*). **C**. Overlay of the structures from panel B with the structure of the paddle-shaped PBS from *A. panamensis* (*5*). **D**. GviPBS with the pentacylindrical core (this work). **E**. The paddle-shaped PBS from *A. panamensis* (*5*).

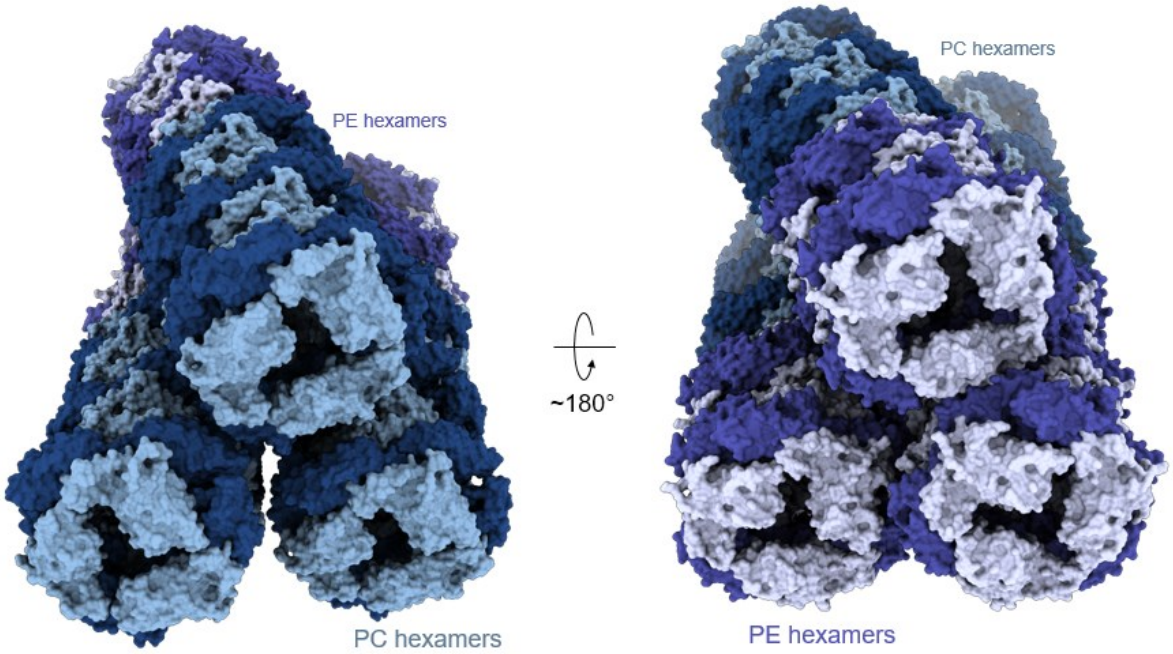


Fig. S6. Helicity of rod bundles in GviPBS. Alpha and beta PBP subunits are shown by dark and light tints of color, respectively. Two views of a rod bundle showing non-parallel and helical orientation of the rods. The period of a superhelix is estimated as ~1000 Å.

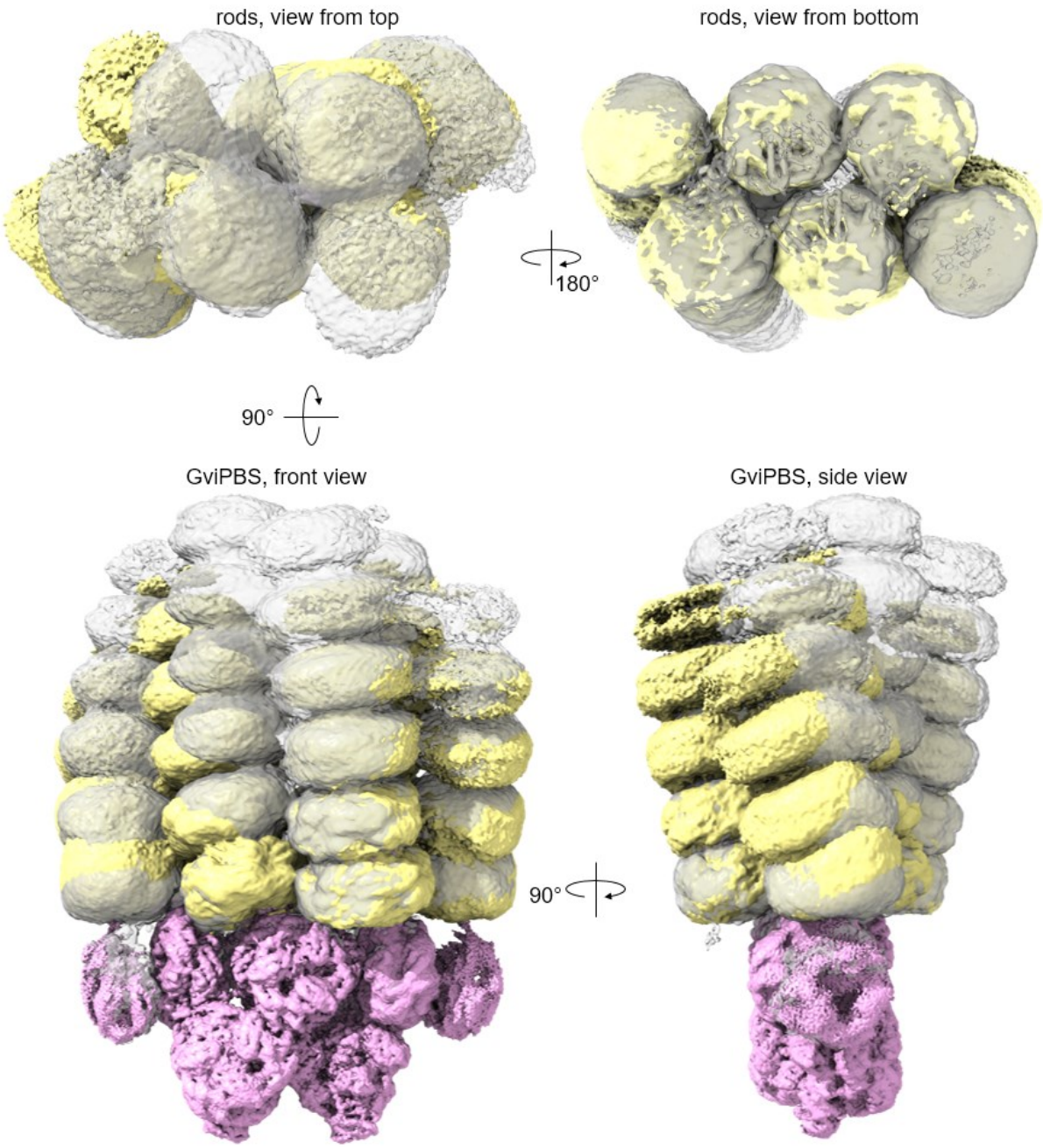


Fig. S7. Conformational heterogeneity in rod bundles in GviPBS. Shown are different views of the superposition of two cryo-EM maps (light yellow and grey) corresponding to the two different conformational states of the rod bundles. Pink map corresponds to the AP core with Rx1/Rx1' hexamers of PC; this map is omitted on the top views for clarity.

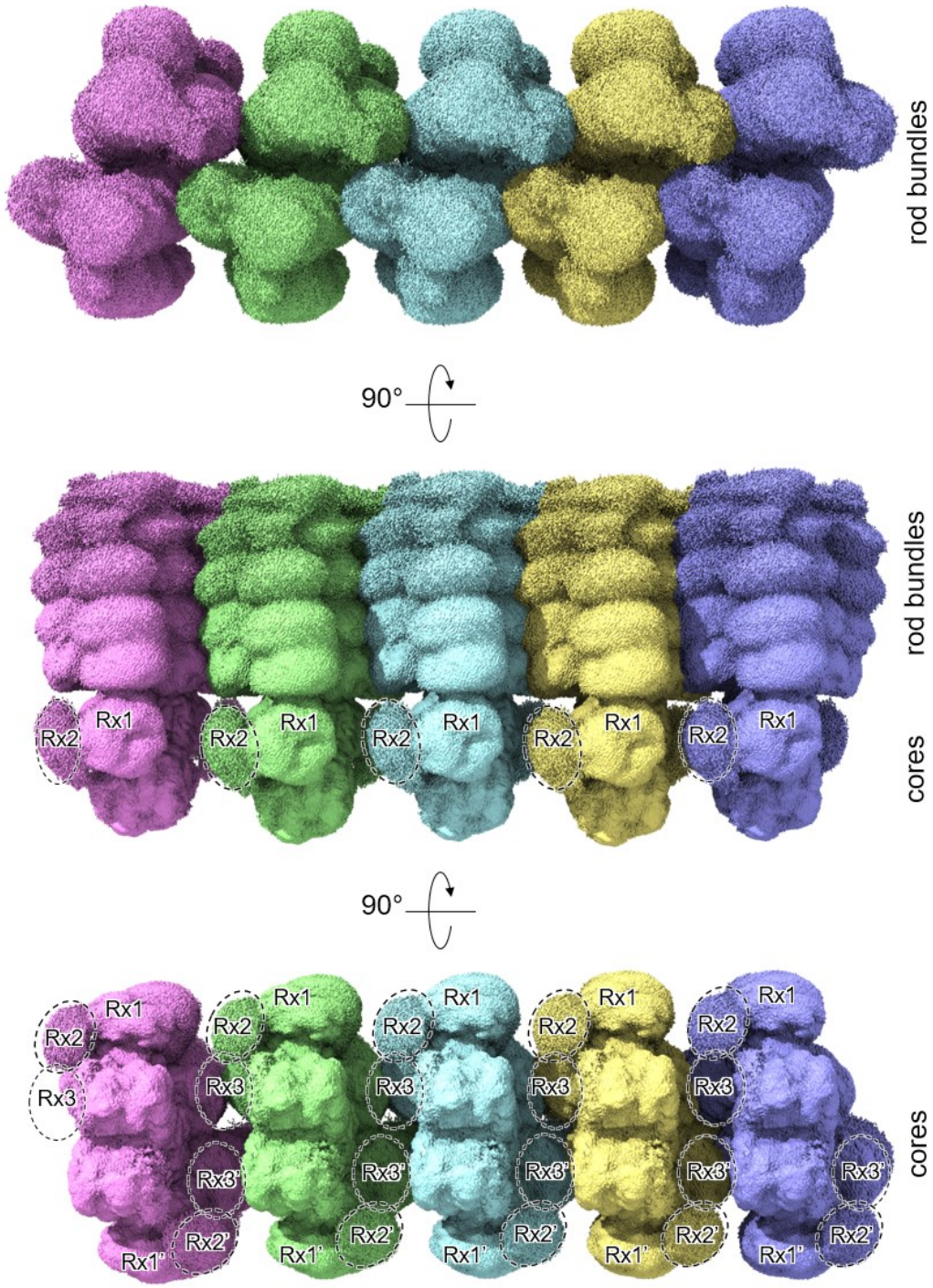


Fig. S8. Tentative model showing the principle of GviPBS assembly into arrays. Five stacked copies of the GviPBS cryo-EM map are shown from different views. Note that the poorly resolved density in the vicinity of Rx1/Rx1' suggests the position of the Rx2/Rx2' hexamers. The third hexamers Rx3/Rx3' are too dynamic and have no density, although their approximate location can be predicted from the presence of the third REP domain in the Glr2806 rod-core linkers (*16*). The complete GviPBS assemblies with six Rx hexamers (Rx1/Rx1', Rx2/Rx2', Rx3/Rx3') belting the AP core can attach to each other in arrays, which would effectively use limited space in the plasma membrane in thylakoid-less *G. violaceus* cells, consistent with earlier considerations and observations (*1*, *12*).

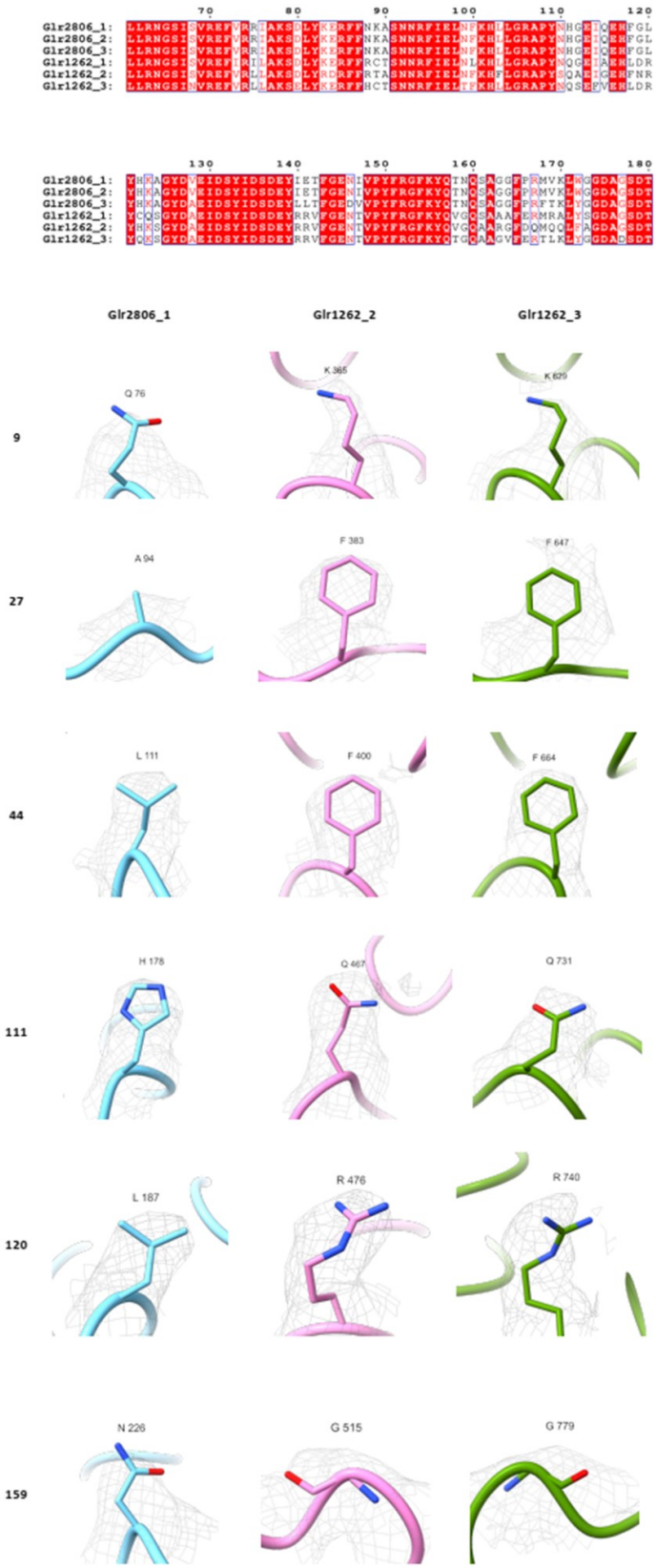


Fig. S9. Identification of Glr1262 and Glr2806 linkers by cryo-EM density for the positions specifically conserved only within REP domains of either protein. Top, multiple sequence alignment of REP domains of the two proteins. Bottom, representative demarcating positions supported by cryo-EM density.

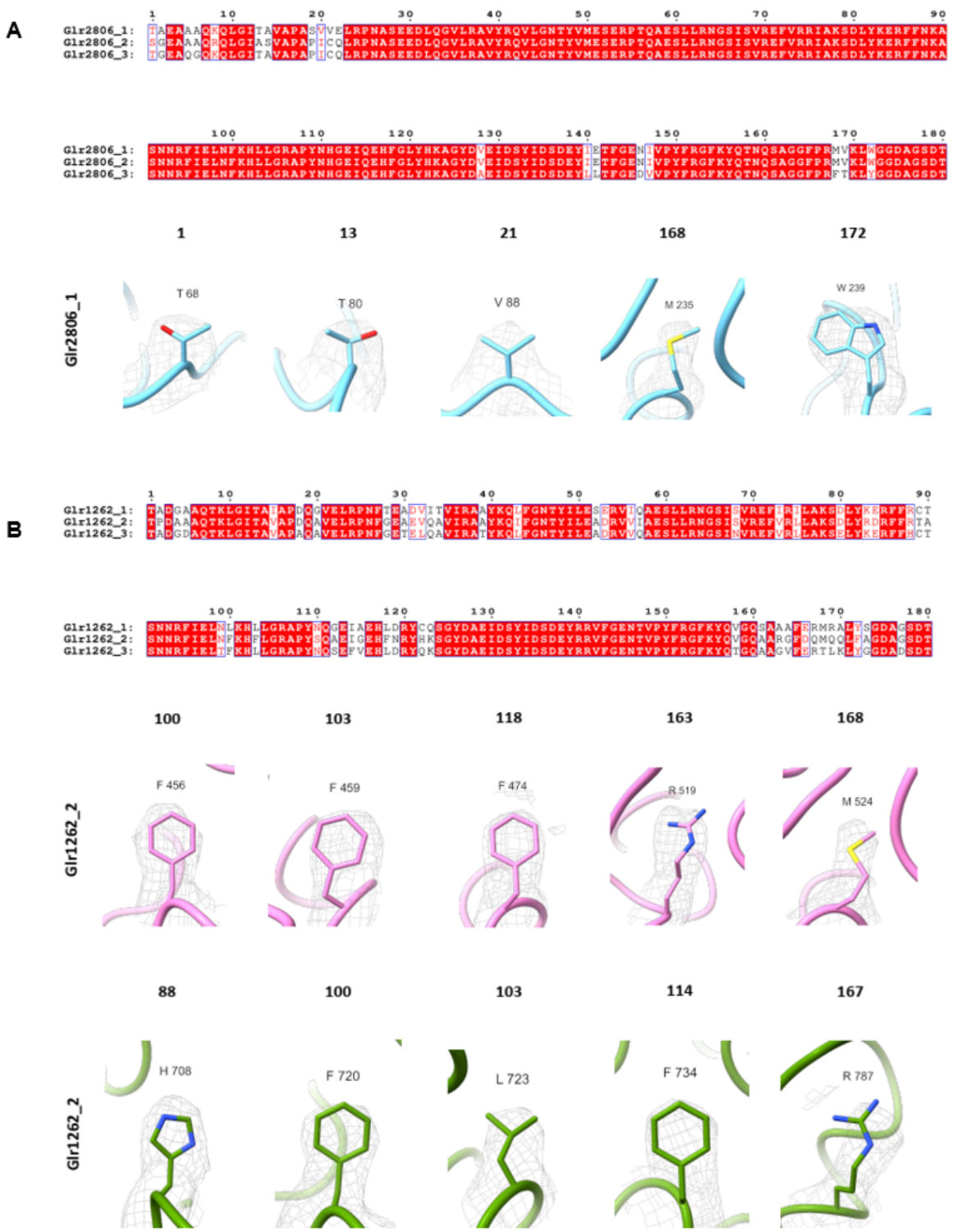


Fig. S10. Identification of REP domains within either Glr1262 or Glr2806 linkers by cryo-EM density for the nonconserved positions. **A**. Top, multiple sequence alignment of REP domains of Glr2806. Bottom, representative demarcating positions supported by cryo-EM density. **B**. Top, multiple sequence alignment of REP domains of Glr1262. Bottom, representative demarcating positions supported by cryo-EM density.

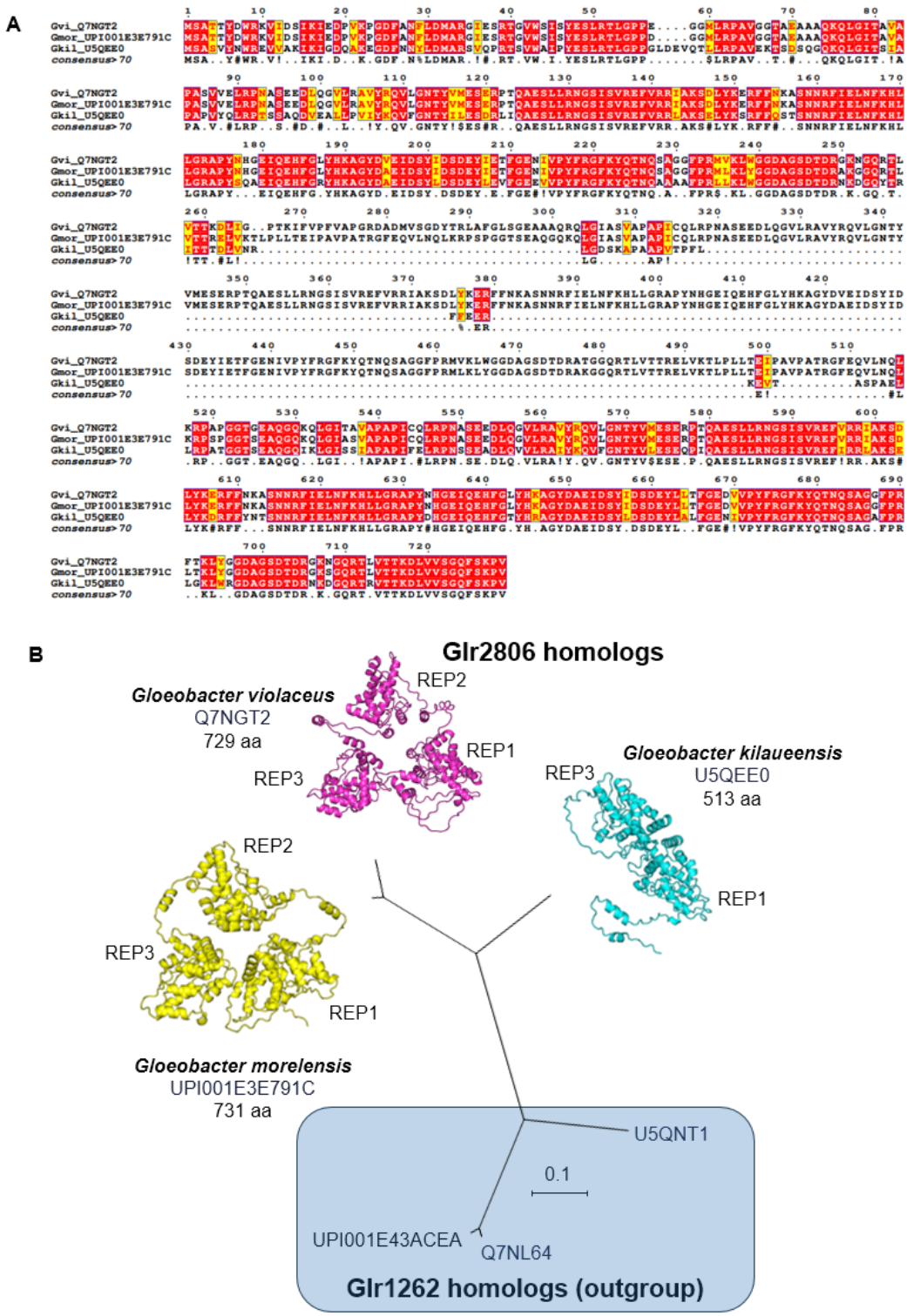


Fig. S11. Analysis of the number of REP domains in Glr2806 linkers of *Gloeobacter* species. **A**. Multiple sequence alignment of the Glr2806 homologs from *G. violaceus*, *G. morelensis* and *G. kilaueensis* showing that the latter has only two REP domains out of three found in other two homologs. **B**. Phylogenetic tree of Glr2806 homologs built using Glr1262 homologs from the same organisms as an outgroup. Scale bar represents the number of per residue substitutions. Alphafold3 (*34*) models of the three Glr2806 homologs analyzed are shown along with the corresponding REP domains numbered.

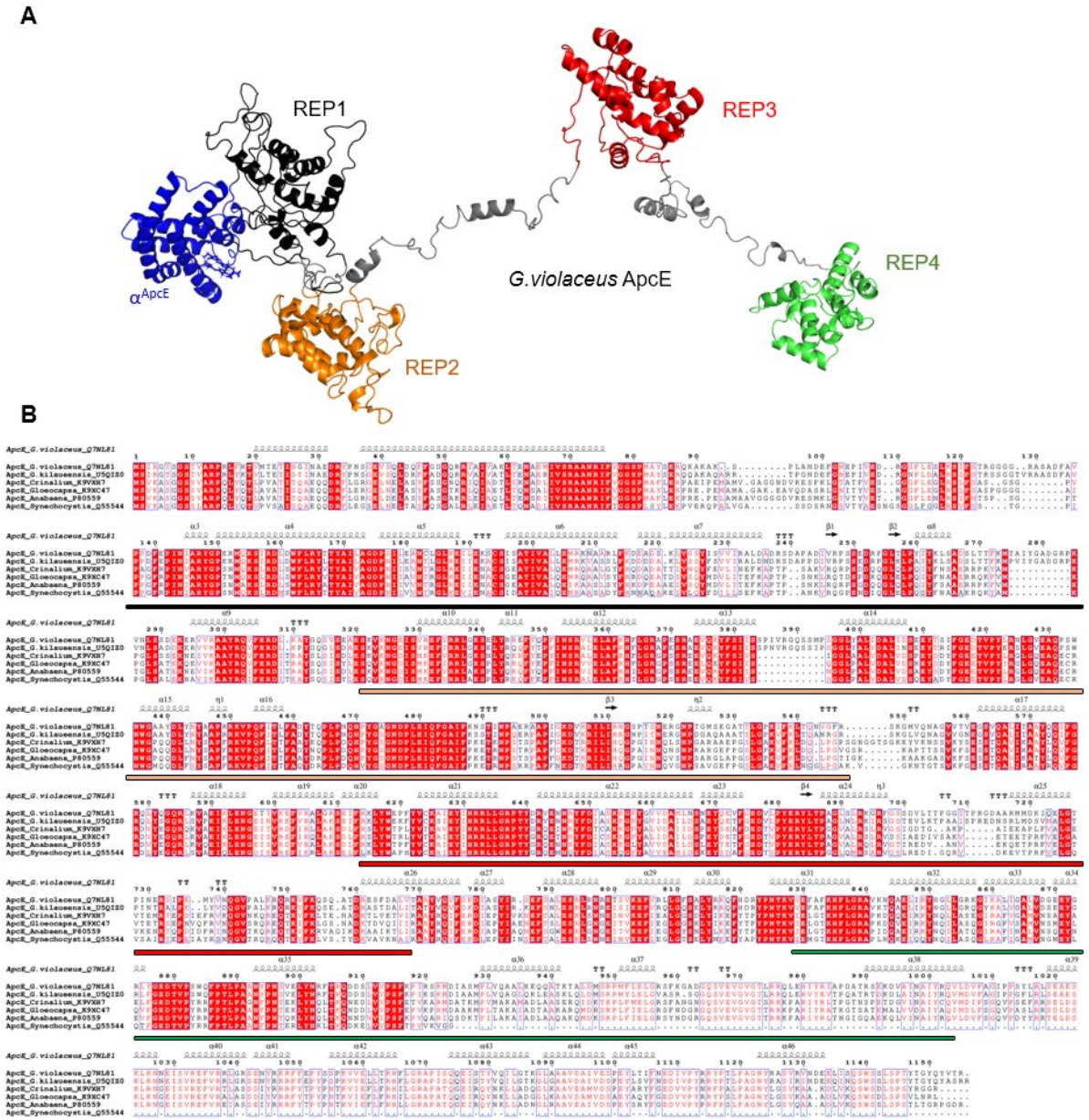
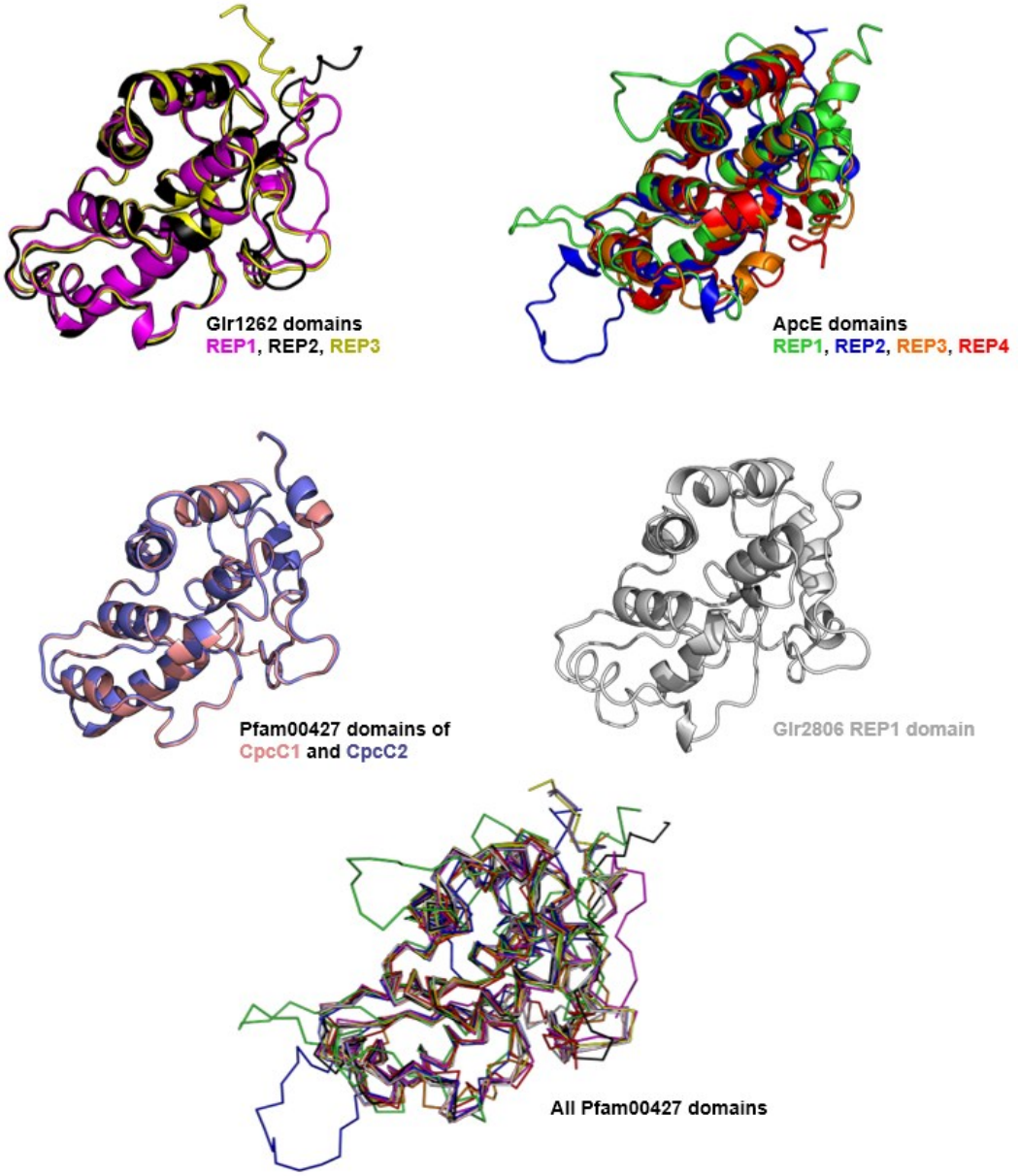


Fig. S12. Domain structure of ApcE proteins. **A**. One copy of ApcE from *G. violaceus* showing the four REP domains following the N-terminal bilin-binding domain α^{ApcE}. **B**. Multiple sequence alignment of ApcE linker proteins from *G. violaceus*, *G. kilaueensis*, *Crinalium epipsammum* PCC 9333, *Gloeocapsa* sp. PCC 7428, *Anabaena* 29413 and *Synechocystis* 6803. Given that the number of REP domains in ApcE proteins indicate the number of cylinders in the respective PBS core, all taken species possess the pentacylindrical PBS cores (four REP domains), except for *Synechocystis* 6803, which possesses the tricylindrical PBS core (the last REP domain is missing). The location of REP domains are indicated above the alignment with the color coding corresponding to panel A.

A



B

	Glr2806 REP1	Glr1262 REP1	Glr1262 REP2	Glr1262 REP3	ApcE REP1	ApcE REP2	ApcE REP3	ApcE REP4	CpcC1	CpcC2
Glr2806 REP1	x	1.034 (178)	0.624 (166)	0.698 (179)	1.34 (114)	1.124 (126)	0.936 (115)	1.169 (130)	0.794 (163)	0.827 (165)
Glr1262 REP1		x	0.813 (173)	0.657 (181)	1.749 (126)	1.045 (97)	0.803 (95)	0.928 (114)	1.002 (159)	0.934 (155)
Glr1262 REP2			x	0.498 (165)	1.552 (116)	1.009 (121)	0.928 (98)	1.006 (110)	0.801 (155)	0.755 (149)
Glr1262 REP3				x	1.117 (104)	0.987 (107)	0.846 (106)	1.094 (126)	0.832 (172)	0.756 (163)
ApcE REP1					x	1.149 (107)	0.743 (101)	0.963 (103)	1.636 (116)	1.680 (117)
ApcE REP2						x	0.764 (118)	0.833 (104)	1.136 (102)	1.168 (104)
ApcE REP3							x	0.741 (126)	1.002 (108)	0.999 (108)
ApcE REP4								x	0.934 (104)	0.888 (102)
CpcC1									x	0.196 (179)
CpcC2										x

Fig. S13. Pfam00427 domains of GviPBS linker proteins. **A**. Structural alignment of three REP domains of Glr1262, ApcE, CpcC1/CpcC2 and Glr2806 proteins. **B**. Cα RMSD values characterizing the similarity of the Pfam00427 domains found in GviPBS proteins. The values in Å were calculated upon structural alignment of a given domain pair. Numbers in parentheses indicate the number of the aligned atoms.

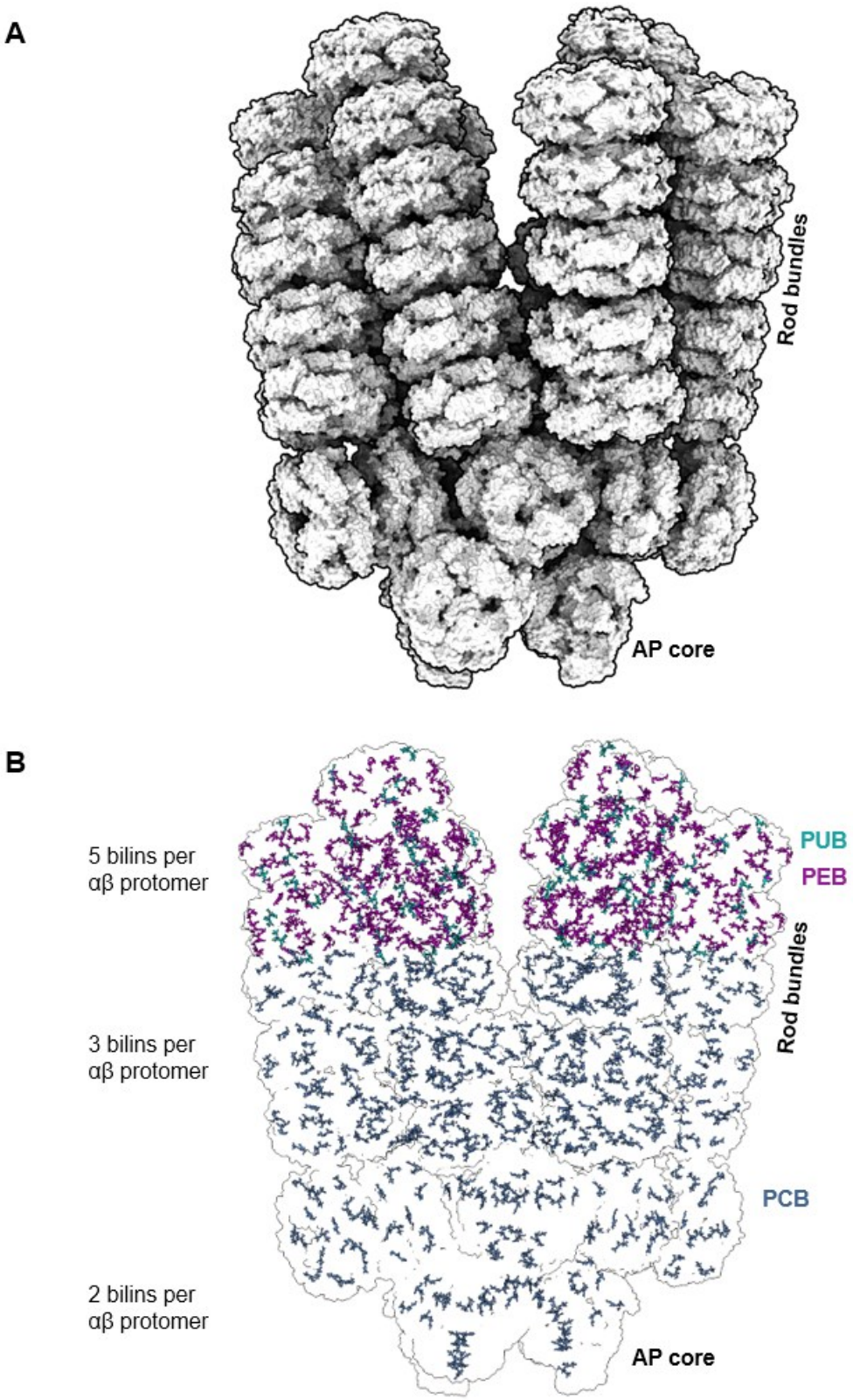


Fig. S14. GviPBS bilins exhibit a descending concentration gradient from the rods to the core. **A**. GviPBS model in surface representation. **B**. Silhouette of the GviPBS model showing the location of PCB, PEB and PUB. Given the different numbers of bilins bound in PE, PC and AP, the gradient is formed.