Non-conventional octameric structure of C-phycocyanin

Takuo Minato1,2,3,9, Takamasa Teramoto4,9, Naruhiko Adachi5, Nguyen Khac Hung1,2, Kaho Yamada1,2, Masato Kawasaki5,6, Masato Akutsu5, Toshio Moriya5, Toshiya Senda5,6, Seiji Ogo1,2,7, Yoshimitsu Kakuta4,8✉ & Ki-Seok Yoon1,2,7✉

C-phycocyanin (CPC), a blue pigment protein, is an indispensable component of giant phycobilisomes, which are light-harvesting antenna complexes in cyanobacteria that transfer energy efficiently to photosystems I and II. X-ray crystallographic and electron microscopy (EM) analyses have revealed the structure of CPC to be a closed toroidal hexamer by assembling two trimers. In this study, the structural characterization of non-conventional octameric CPC is reported for the first time. Analyses of the crystal and cryogenic EM structures of the native CPC from filamentous thermophilic cyanobacterium Thermoelleptolyngbya sp. O–77 unexpectedly illustrated the coexistence of conventional hexamer and novel octamer. In addition, an unusual dimeric state, observed via analytical ultracentrifugation, was postulated to be a key intermediate structure in the assemble of the previously unobserved octamer. These observations provide new insights into the assembly processes of CPCs and the mechanism of energy transfer in the light-harvesting complexes.
ight-harvesting in photosynthesis is a crucial initial step in the conversion of light energy into chemical energy; therefore, structural investigation of light-harvesting complexes is important for gaining an understanding of energy transfer mechanisms and developing effective systems of artificial photosynthesis. Phycobilisomes (PBSs) are light-harvesting, antenna complexes that are located on the thylakoid membranes of cyanobacteria, red algae, and glaucophytes; they play a key role in the transfer of energy to photosystems I and II. The most common hemisidoidal-type PBSs consist of core and peripheral rod moieties; the core consists of cylindrical substructures of allophycocyanins (APCs), whereas the rod consists of stacked phycocyanins (PCs) and phycoerythrins (PEs). These phycobiliproteins (PBPs) possess different types of open-chain tetapyrrolochromophores covalently linked by thioether bonds, resulting in the different absorption properties; the maximum absorptions in APC, PC, and PE are observed at approximately 650–655, 615–620, and 540–570 nm, respectively, which enables PBSs to increase the range of photon absorption from sunlight.

Recently, cryogenic electron microscopy (cryo-EM) analyses have unexpectedly revealed the coexistence of conventional hexameric (αβ)6 states at low phosphate buffer or low CPC concentrations, while the previously reported crystal models of TlCPC showed absorption bands at 277, 353, and 615 nm, while the fluorescence spectrum showed an emission band at 644 nm; these are characteristic bands of CPC (Supplementary Fig. 2). The circular dichroism spectrum of TlCPC showed negative Cotton effects with maxima at 220 and 209 nm and a positive Cotton effect with a maximum at 193 nm, which are characteristic bands of α-helix structures (Supplementary Fig. 3). These results were compatible with those previously reported for PCs and suggested that TlCPC had been successfully isolated.

**Results and discussion**

**Isolation and characterization of TlCPC.** Based on genome sequence analysis, O-77 possesses CPC-, APC-, and linker protein-encoding genes but entirely lacks PE-encoding genes, which is similar to several cyanidophyta and cyanobacteria. The genome encodes single genes for CpcA (BAU24407) and CpcB (BAU24408), encoding the α and β subunits for CPC. Since the Apce gene in O-77 (BAU24084), which encodes the α subunit of APC with a linker in the core of PBS, contains four repeat domains (pfam00427 domains in the Pfam database), the PBS complex in O-77 is expected to possess pentacylindrical APC core. It should be noted that a whole structure of this PBS type has, to date, rarely been investigated at high resolution.

Initially, TlCPC was isolated using a conventional column chromatography technique (see Methods section for details). In SDS-PAGE of TlCPC, approximately 17.5-kDa and 22.5-kDa bands, assignable to the α and β subunits of TlCPC, respectively, were observed (Supplementary Fig. 1). The observed 17.5-kDa band agreed well with the calculated mass of the α subunit (17.4 kDa), whereas the 22.5 kDa band was larger than expected (β subunit: 17.9 kDa). The UV–vis spectrum of TlCPC showed absorption bands at 277, 353, and 615 nm, while the fluorescence spectrum showed an emission band at 644 nm; these are characteristic bands of CPC (Supplementary Fig. 2). The circular dichroism spectrum of TlCPC showed negative Cotton effects with maxima at 220 and 209 nm and a positive Cotton effect with a maximum at 193 nm, which are characteristic bands of α-helix structures (Supplementary Fig. 3). These results were compatible with those previously reported for PCs and suggested that TlCPC had been successfully isolated.

**X-ray crystallographic analysis of TlCPC-6.** To elucidate the oligomeric state of TlCPC, the crystallization of TlCPC was examined under various crystallization conditions, which resulted in the formation of many blue single crystals (Supplementary Table 1), where crystal systems were determined in the early data processing stages. The initial screening of X-ray crystallographic analyses revealed that single crystals belonged to the space groups of P2 (No. 2), P21 (No. 4), and P21212 (No. 18), which represented the conventional hexameric (αβ)6 states. TlCPC crystallized in the primitive orthorhombic P21212 was further refined and then the hexameric structure (TlCPC-6) with D3 symmetry was determined at a resolution of 1.65 Å (Fig. 1a, c; Table 1), where the asymmetric unit contained a hexamer. In a comparison with the crystal model of Spirulina platensis CPC (SpCPC; PDB ID: 1GHO), the crystals of which contained two hexamers in the asymmetric unit, the root mean square deviation (rmsd) value was 0.68 Å over 1,993 Ca atoms, supporting the formation of a well-established hexameric (αβ)6 state. The α and β subunits of TlCPC-6 shared 68–86% and 65–87% sequence identities, respectively, with the previously reported crystal models of PCs. The X-ray crystallographic analysis of TlCPC-6 showed that phycocyanobilins (PCBs), i.e., the chromophores in
CPCs, were covalently bonded at the conserved Cys residues of \( \alpha \text{Cys84} \) (a84PCB), \( \beta \text{Cys82} \) (b82PCB), and \( \beta \text{Cys153} \) (b153PCB) (Supplementary Fig. 4). The conformations of PCBs were essentially identical to those in CPCs from *Synechococcus elongatus* (ScCPC; PDB ID: 1JBO)\(^{19} \), which were previously determined at high resolution (1.45 Å) with a low \( R_{\text{free}} \) value (Supplementary Fig. 5). The identical conformations of aromatic rings in PCBs also supported the typical absorption and emission spectra observed in TlCPC (Supplementary Fig. 2). Methylation of \( \beta \text{Asn72} \) is a highly conserved post-translational modification in all reported CPCs with few exceptions\(^{20,26} \) and is thought to play a crucial role in highly efficient energy transfer\(^{44} \). In TlCPC-6, the methyl group at the \( \beta \text{Asn72} \) residue could be clearly modeled into an electron density map (Supplementary Fig. 6).

Determining the crystal packing of CPCs is important for estimating the intra- and inter-CPC energy transfer pathways because orientationally aligned flat-shaped CPCs that are assembled face-to-face in crystals can be regarded as a motif of layered CPCs in the rod moieties of PBSs, even in the absence of linker proteins\(^{15,21,22} \). Approximately 90% of the previously reported single crystals from CPCs belong to the space groups of \( P_2_1 \) (No. 4), \( R_3_2 \) (No. 155), and \( P_6_3 \) (No. 173), and most showed orientationally aligned hexameric CPCs in crystals. Only two previous reports indicated that CPCs were crystallized in the orthorhombic space group, with these crystals containing two types of orientation\(^{27,30} \). Although TlCPC-6 was also crystallized in the primitive orthorhombic space group (\( P_2_1 2_1 2_1 \)), hexameric CPCs were orientationally aligned in the crystal. In TlCPC-6, the buried surface area between hexamers along the \( C \) axis, as calculated by the PISA program\(^{13} \), was 609 Å\(^2 \) per hexamer, which indicated weak interactions similar to layered CPCs in the rods of PBS (Fig. 2a). Rod-like layered structures of orientationally aligned CPCs, of which \( C_3 \) axes are shared, have also been observed in CPCs from *Microchaete diplosiphon* (MdCPC; PDB ID: 1CPC)\(^{14} \), *Thermosynechococcus vulcanus* (TvCPC; PDB ID: 1KTP)\(^{18} \), and *Thermosynechococcus elongatus* BP-1 (TeCPC; PDB ID: 3LOF). In TlCPC-6, the distances of intraprotein PCB pairs were as follows: a84PCB–b82PCB (20.0 Å), b153PCB–b153PCB (25.8 Å), and a84PCB–a84PCB (26.9 Å) (Supplementary Fig. 7). Additionally, the distances of interprotein PCB pairs were as follows: b82PCB–b82PCB (26.2 Å) and a84PCB–a84PCB (34.4 Å). These values were close to those of the reported CPCs, indicating that the estimated main energy pathways between hexamers were via a84PCB and b82PCB. It is noteworthy that the structure of a double-layered \( \left[ \left( \alpha \beta \right)_6 \right]_2 \) unit in TlCPC-6 crystal closely resembles that of a double-layered PE–PC unit in PBS from *P. purpureum* (PDB ID: 6KGX) (Supplementary Fig. 8).

### Table 1 Crystallographic data collection and refinement statistics.

| (PDB ID) | TlCPC-6 (7EFW) | TlCPC-8 (7EFV) |
|----------|----------------|----------------|
| **Data collection** | | |
| Space group | \( P_2_1 2_1 2_1 \) (No. 18) | \( 4_3 2 \) (No. 211) |
| Cell dimensions | | |
| \( a, b, c \) (Å) | 60.1, 187.4, 210.1 | 230.0, 230.0, 230.0 |
| \( \alpha, \beta, \gamma \) (°) | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 |
| Resolution (Å) | 50.0–1.65 (1.76–1.65) | 50.0–2.77 (2.93–2.77) |
| \( R_{\text{free}} \) | 21.2 (106.2) | 37.8 (420.0) |
| \( C_{\text{c1/2}} \) | 99.3 (49.5) | 99.8 (55.7) |
| I/σ | 9.5 (1.53) | 11.39 (0.91) |
| Completeness (%) | 92.6 (65.1) | 100.0 (100.0) |
| Redundancy | 5.8 (2.9) | 61.6 (61.7) |
| **Refinement** | | |
| Resolution (Å) | 49.18–1.65 | 49.03–2.77 |
| No. reflections | 264075 | 26620 |
| \( R_{\text{work}}/R_{\text{free}} \) (%) | 17.0/19.0 | 19.9/23.8 |
| **No. atoms** | | |
| Protein | 15354 | 5021 |
| Ligand | 774 | 258 |
| Water | 2676 | — |
| **B-factors** | | |
| Protein | 16.4 | 68.8 |
| Ligand | 16.8 | 68.9 |
| Water | 29.3 | — |
| **R.m.s. deviations** | | |
| Bond lengths (Å) | 0.005 | 0.003 |
| Bond angles (°) | 0.985 | 0.829 |
| **Ramachandran plot** | | |
| Favored (%) | 98.3 | 98.0 |
| Allowed (%) | 1.7 | 2.0 |
| Disallowed (%) | 0 | 0 |

*Values in parentheses are for a highest-resolution shell.
supporting that layered CPCs in crystals are structurally correlated with rod moieties of PBSs in vivo.

Taken together, the aforementioned results show that TiCPC crystallized in the space group $P_{2_1}2_12_1$ was a conventional hexameric ($\alpha\beta$)$_6$ structure with a layered crystal packing; thus, the distinctive features of TiCPC-6 were hardly observed in comparison with previously reported CPCs.

**X-ray crystallographic analysis of TiCPC-8.** During X-ray crystallographic analyses, we unexpectedly found crystals that belonged to the extraordinary space group of $I$ centered cubic $I432$ (No. 211) (Supplementary Table. 1), in which CPCs crystallized in a cubic space group have not previously been reported. X-ray crystallographic analysis revealed that the structure of TiCPC in the space group of $I432$ at a resolution of 2.77 Å was not a conventional hexamer but rather a novel octamer (TiCPC-8) (Fig. 1b, d; Table 1). This crystal model contained a dimer ($\alpha\beta$)$_2$ in the asymmetric unit and consisted of eight ($\alpha\beta$)$_8$ monomers to give an extended closed toroidal structure with a $D_4$ symmetry. To date, all previously reported PBPs, including APCs, PCs, and PEs, have consisted of basic trimeric ($\alpha\beta$)$_3$ units with $C_3$ symmetry; in contrast, TiCPC-8 possessed a 4-fold rotation axis that could not be directly assembled from trimers or hexamers with 3-fold rotation axes. Notably, TiCPC-8 is the first example of a PBP with a 4-fold rotation axis as well as an octameric CPC.

The inner diameter of ring-shaped TiCPC-8 (45 Å) increased in comparison with that of TiCPC-6 (25 Å) due to the extended toroidal structure of TiCPC-8 (Fig. 1a, b). The structural differences in the ($\alpha\beta$) monomer unit between TiCPC-6 and TiCPC-8 (rmsd value of 2.06 Å over 321 Ca atoms) were larger than those between TiCPC-6 and Ti$\nu$CPC (rmsd value of 0.33 Å over 329 Ca atoms) due to the slight structural conformation changes. In the ($\alpha\beta$) monomer, the helices A, B, E, and F in the $\alpha$ subunit ($A_\alpha$, $B_\alpha$, $E_\alpha$, and $F_\alpha$), the helices A, B, C, and E in the $\beta$ subunit ($A_\beta$, $B_\beta$, $C_\beta$, and $E_\beta$), and the loops between these helices ($A/B_\alpha$, $E/F_\alpha$, $A/B_\beta$, and $B/C_\beta$) interacted with each other to form a bent structure (Supplementary Fig. 9). The dihedral angle of the ($\alpha\beta$) monomer (aGln70–aArg30–βGly70) in TiCPC-6 was approximately 110°, whereas the equivalent angle in TiCPC-8 was approximately 123° (Fig. 1e, f). In addition, the angle between two adjacent monomers (aGln70–βPro69–βPro69) in TiCPC-6 was approximately 52°, whereas the equivalent angle in TiCPC-8 was approximately 90° (Fig. 1g, h). These results indicate that the structural flexibilities of the subunits and their interfaces presumably affect the assembly process of hexameric ($\alpha\beta$)$_6$ and octameric ($\alpha\beta$)$_8$ states even though both crystal models possess the same amino acid sequence.

In the crystal of TiCPC-8, two octameric ($\alpha\beta$)$_8$ units were stacked along the $C_4$ axis to form a double-layered orientationally aligned structure of [(($\alpha\beta$)$_8$)$_2$], which were further assembled at right angles to each other. Consequently, the overall crystal model of TiCPC-8 possessed a unique framework with exceedingly large void spaces of approximately $100 \times 100 \times 100$ Å$^3$ (Fig. 2b; Supplementary Fig. 10). From an engineering perspective, the topology of this structure resembles zeolite A as well as several types of metal-organic frameworks and polyoxometalates$^{44,45}$; thus, the single crystals of TiCPC-8 could potentially be utilized as catalysts and adsorbents to take advantage of the expected large specific surface area.

Considering [(($\alpha\beta$)$_8$)$_2$] units in the crystal, the layered structures could be regarded as octameric versions of rods (Fig. 2b). The overall conformations of PCBs in TiCPC-8 were essentially identical to those in TiCPC-6 (Supplementary Fig. 5). In TiCPC-8, the distances of intraprotein PCB pairs were as follows: a84PCB–β82PCB (20.7 Å), β153PCB–β153PCB (21.7 Å), and a84PCB–α84PCB (31.6 Å) (Supplementary Fig. 11). These values were similar to those in TiCPC-6, whereas the distance of interprotein PCB pairs of β82PCB–β82PCB (31.7 Å) was increased relative to the equivalent distance in TiCPC-6 (26.2 Å) because a contact area was not observed between octamers by the PISA program, presumably due to the demand for octamers to fit the unique crystal packing of TiCPC-8. Nevertheless, the distances of the interprotein PCB pairs were similar to the shortest distances between chromophores in PE-PE (24.4 Å) and PE-PC (23.8 Å) of PpPBS, indicating the potential intra- and inter-protein energy transfer in TiCPC-8.

As the crystal shapes of TiCPC-6 and TiCPC-8 were different from each other (Supplementary Fig. 12), CPC solutions derived from hexameric ($\alpha\beta$)$_6$ and octameric ($\alpha\beta$)$_8$ states could be prepared by picking up crystals and dissolving them separately. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometric analysis of these two solutions showed essentially the same two signals assignable to the $\alpha$ and $\beta$ subunits of TiCPC (Supplementary Fig. 13). In addition, CD spectra of these solutions also showed almost the same ellipticity in a range of 190–250 nm (Supplementary Fig. 14a), and the thermal denaturation midpoints were 71 °C for both solutions (Supplementary Fig. 14b). It should be noted that oligomeric structures were disassembled in these solutions because of the low concentration of CPC$^{28–32}$. These results strongly supported that the subunits composition and physical properties of TiCPC-6 and TiCPC-8 were the same and post-translational modifications did not occur except for the methylation and PCB chromophorylation (Supplementary Figs. 5 and 6). Crystals seemed to be grown one type of morphology in the same drop, and the residual crystallization solution were much clearer than the initial states one type of morphology in the same drop, and the residual crystallization solution were much clearer than the initial states.

**Cryo-EM structure of TiCPC.** To avoid the effect of crystallization, the presence of hexameric ($\alpha\beta$)$_6$ and octameric ($\alpha\beta$)$_8$ states was further investigated via cryo-EM analysis. The...
potassium phosphate buffer solution of TlCPC was applied to a holey carbon grid and then flash-frozen to prepare a cryo-grid for structural determination (Supplementary Note 1). Cryo-EM micrographs and selected reference-free 2D class averages from single particle analysis (Fig. 3b, c; Supplementary Figs. 15–18), which resulted in reconstructions of hexameric (αβ)₆ and octameric (αβ)₈ states in the frozen solution. These results strongly indicated that the concentrated solution of the native TlCPC contained hexamers and octamers. The cryo-EM model of TlCPC-6 was essentially isostructural to the cryo-EM models of PCs in PBSs from G. pacifica (GpPC; PDB ID: 5Y6P) and P. purpureum (PpPC; PDB ID: 6KGX) (GpPC: rmsd value of 1.06Å over 1,887 Ca atoms; PpPC: rmsd value of 0.96Å over 1,932 Ca atoms) (Supplementary Fig. 20). Although low-resolution EM structures from CPC and high-resolution cryo-EM structures from PCs in PBSs have been reported, the cryo-EM models of TlCPC-6 and TlCPC-8 are the first examples of high-resolution cryo-EM structures from CPC.

The local resolution maps of TlCPC-6 and TlCPC-8 showed relatively low resolutions at helices Cα, Dα, Aβ, Eβ, Cβ, Dβ, Fβ, and Gβ and in the loops D/E, D/E, F/Gα, and H/Iβ (Supplementary Fig. 21). These helices and loops also possessed relatively high B-factors in the crystal models of TlCPC-6 and TlCPC-8 (Supplementary Fig. 22). Therefore, the structural fluctuations of these regions are relatively large regardless of the effects of

---

**Fig. 3 Cryo-EM analyses of TlCPC-6 and TlCPC-8.** Representative motion-corrected electron micrograph of TlCPCs (a). Typical reference-free 2D class averages from single-particle images of TlCPC-6 (b) and TlCPC-8 (c). Top and side views of cryo-EM density maps of TlCPC-6 at 3.06Å resolution (d and f) and TlCPC-8 at 3.71Å resolution (e and g). Contour levels of TlCPC-6 and TlCPC-8 are shown at 0.06 and 0.04, respectively.

---

**Table 2 Cryo-EM data collection, refinement, and validation statistics.**

| (EMDB/PDB ID) | TlCPC-6 (EMD-31090/7EH8) | TlCPC-8 (EMD-31089/7EH7) |
|---------------|---------------------------|--------------------------|
| **Data collection and processing** | | |
| Microscope    | Talos Arctica             | Talos Arctica            |
| Voltage (kV)  | 200                       | 200                      |
| Detector      | Falcon 3EC                | Falcon 3EC               |
| Magnification | 120,000                   | 120,000                  |
| Pixel size (Å) | 0.88                      | 0.88                     |
| Automation software | EPU                      | EPU                      |
| Total exposure (e/Å²) | 50                        | 50                       |
| Exposure rate (e/Å²) | 1.02                     | 1.02                     |
| Number of frames | 49                       | 49                       |
| Defocus range (µm) | -1, -15, -2, -2.5         | -1, -15, -2, -2.5        |
| Number of collected micrograph | 2,036                    | 2,036                    |
| Number of particles for Class2D | 283,980                  | 129,653                  |
| Number of particles for Class3D | 72,655                   | 23,643                   |
| Number of particles for Refine3D | 28,120                   | 10,402                   |
| Symmetry imposed | D₁                        | D₄                       |
| Map resolution (Å) | 3.06                      | 3.71                     |
| FSC threshold | 0.143                     | 0.143                    |
| Map resolution range (Å) | 2.91-3.38               | 3.47-4.49                |
| **Refinement** | | |
| Initial model used | 7EFW                      | 7EFV                     |
| (PDB code) | | |
| Map-to-model resolution (Å) | 3.05                      | 3.75                     |
| FSC threshold | 0.5                       | 0.5                      |
| Model resolution range (Å) | 2.91-3.38 | 3.47-4.49 |
| **Model composition** | | |
| Non-hydrogen atoms | 15,773                    | 20,904                   |
| Protein residues | 2004                      | 2656                     |
| Ligands | 18                        | 24                       |
| **B factors (Å²)** | | |
| Protein | 38.8                      | 51.9                     |
| Ligand | 39.4                      | 51.2                     |
| **Map-model CC** | | |
| CC (mask) | 0.90                      | 0.86                     |
| CC (box) | 0.77                      | 0.80                     |
| CC (peaks) | 0.75                      | 0.75                     |
| CC (volume) | 0.85                      | 0.83                     |
| **R.m.s. deviations** | | |
| Bond lengths (Å) | 0.003                    | 0.004                    |
| Bond angles (°) | 1.051                    | 1.298                    |
| **Validation** | | |
| MolProbity score | 1.3                       | 1.7                      |
| Clash score | 5.5                       | 8.7                      |
| Poor rotamer (%) | 0.79                     | 1.69                     |
| **Ramachandran plot** | | |
| Favored (%) | 98.2                     | 97.8                     |
| Allowed (%) | 1.8                      | 2.2                      |
| Disallowed (%) | 0                       | 0                        |
crystal packings (X-ray) and damages induced by freezing processes (cryo-EM).

**Geometries of TICPC-6 and TICPC-8.** Interestingly, the X-ray crystallographic, cryo-EM, and MALDI-TOF mass analyses of TICPC clearly indicate that different oligomeric states of TICPC-6 and TICPC-8 were assembled from the same monomeric (αβ) units as described above. Investigating the structures of homologous proteins with different oligomeric states is crucial to elucidating their functions in metabolism as well as the mechanisms of protein assembly and evolution. In general, quaternary structures are conserved among proteins with the same or very high sequence identities because the misassembly of oligomers in vivo is implicated in physiological disorders. However, especially in closed circular oligomers with cyclic (Cn) or dihedral (Dn) symmetry (n > 2), it has been reported that different oligomeric states from identical subunits, including protective antigens of toxins (n = 7, 8, 38, and 44), portal proteins (n = 12 and 13), flagellar motors (n = 32–36), and trp RNA-binding attenuation protein (n = 11 and 12), play important roles in various biological regulations to control diameter and curvature of ring-shaped proteins. These studies have shown that slight conformational changes in secondary and tertiary structures enable the control of rotation angles between adjacent subunits (Δα, ideally equal to (360°/n–360°/m)°) for oligomers with n- and m-fold rotation axes (n < m) in contrast, only theoretical studies have been conducted in oligomers with a small number of n-fold rotation axes (n = 3 and 4) because of the requirement of large structural changes (Δα = 30°) (Supplementary Fig. 23). In this context, TICPC-6 and TICPC-8 with their respective D3 and D4 symmetries are the first examples of identical subunits being assembled into different cyclic oligomers with 3- and 4-fold rotation axes.

**Analytical ultracentrifugation of TICPC.** To date, size-exclusion chromatography and analytical ultracentrifugation have mainly been utilized to analyze the oligomeric states of CPC in solutions. Thoren et al. reported that a monomeric (αβ) state in solution could be obtained by size-exclusion chromatography, whereas a homogeneous elution of trimeric (αβ)3 or hexameric (αβ)6 states could not be obtained partly due to the aforementioned trimmer–hexamer equilibrium of diluted CPC samples in the column. On the other hand, analytical ultracentrifugation has long been a powerful tool by which to estimate the oligomeric states of CPCs in solutions. Berns and MacColl previously determined the sedimentation coefficients of CPCs under various conditions and established a widely accepted interpretation as follows: monomeric (αβ), trimeric (αβ)3, and hexameric (αβ)6 species were observed at sedimentation coefficient distributions of 3.7, 7, and 11 S, respectively.

Despite this established consensus, there has been an exceptional observation of dimeric (αβ)2 states: in a report by Neufeld and Riggs, analytical ultracentrifugation of the CPC from *S. elongatus* PCC 7942 (ScCPC, previously known as *Anacystis nidulans*) showed monomer–dimer–hexamer equilibrium in solutions. Although X-ray crystallographic analysis later revealed that the ScCPC was a hexameric state in crystallographic structures, dimeric (αβ)2 states in solutions could be obtained by extended cyclic structures were possible because the flexible interfaces acted as “hinges”, which enabled control of the curvatures of protein rings although TICPC-6 and TICPC-8 possessed the same amino acid sequences. In addition, Adir et al. reported that the hydrogen bond network between αAsp28, βAsn35, and βPCB153 was critical to form hexameric structure in *T. vulcanus* CPC, while TICPC possesses βSer35 that is rarely observed in the previously reported CPC crystals: (1) the direct hydrogen bond between βAsn35–β153PCB was observed in TICPC, whereas the indirect hydrogen bond between βSer35–β153PCB via a water molecule was observed in *T. vulcanus*, and (2) one of the two indirect hydrogen bonds between βAsn35–αAsp28 in *T. vulcanus* was not observed in TICPC-6, suggesting the destabilization of hexameric structures in TICPC (Supplementary Fig. 26a). Moreover, the indirect hydrogen bond between αLys32–β153PCB in TICPC is important to stabilize (αβ) monomer, whereas αGlul32 in TICPC rather stabilizes a dimeric structure (αβ)2 by interacting with the neighboring β subunit in the dimeric interface (Supplementary Figs. 26b and 27a).

To investigate the solution states of diluted TICPC, analytical ultracentrifugation was performed in the present study. Analytical ultracentrifugation of TICPC (0.92 mg/mL) in potassium phosphate buffer (pH 7.0, 10 mM) gave a major sedimentation coefficient distribution at 5.24 S (s20,w = 5.56 S) and minor sedimentation coefficient distributions at 3.15 S (s20,w = 3.33 S), 7.97 S (s20,w = 8.45 S), and 11.32 S (s20,w = 12.00 S) (Supplementary Fig. 24). According to the reported s values, observed species distributed at 3.15, 5.24, and 11.32 S were assignable to monomeric (αβ), dimeric (αβ)2, and hexameric (αβ)6 states, respectively, whereas the species distributed at 7.97 S is observed for the first time here. The molecular weights of these species were calculated using the Svedberg equation as follows: 37.1 kDa (for 3.15 S), 79.8 kDa (for 5.24 S), 149.5 kDa (for 7.97 S), and 253.1 kDa (for 11.32 S). The obtained values of 37.1 and 79.8 kDa were in agreement with the calculated masses of the monomer (37.3 kDa) and dimer (74.7 kDa), respectively, whereas the value of 253.1 kDa was slightly larger than expected (hexamer: 224.1 kDa). Interestingly, the obtained weight of the previously unobserved distribution at 7.97 S (149.5 kDa) was similar to the calculated mass of the tetramer (149.4 kDa). It should be noted that Iso et al. previously reported the presence of tetramers in a solution. Overall, these results indicate that the unusual dimeric (αβ)2 state in the solution was a key intermediate structure in the assembly of the newly observed octameric (αβ)8 state.

**Assembly mechanism of TICPC.** Based on the results reported thus far, an assembly mechanism for TICPC is proposed. The crystal and cryo-EM models of TICPC-6 and TICPC-8 clearly showed slight structural differences of monomer–monomer interfaces (helices Gα, Gβ, and Eβ and loop D/Eβ) and α–β subunit interfaces (helices Aα, Bβ, Eα, Fα, Aβ, Bγ, Cγ, and Eγ) between TICPC-6 and TICPC-8. The superimposition of TICPC-6 and TICPC-8 monomers showed that slight structural differences existed at helices Aβ, Bβ, Aγ, Bγ, Cγ, and Dγ and loops A/Bα, D/Eα, A/Bβ, and D/Eβ (Fig. 1e–h; Supplementary Fig. 25). A possible basis for these structural changes were high B-factors and low local resolutions observed by the X-ray crystallographic and cryo-EM analyses, respectively (Supplementary Figs. 21 and 22), suggesting the relatively flexible protein backbones at these interfaces. Therefore, extended cyclic structures were possible because the flexible interfaces acted as “hinges”, which enabled control of the curvatures of protein rings although TICPC-6 and TICPC-8 possessed the same amino acid sequences. In addition, Adir et al. reported that the hydrogen bond network between αAsp28, βAsn35, and βPCB153 was critical to form hexameric structure in *T. vulcanus* CPC, while TICPC possesses βSer35 that is rarely observed in the previously reported CPC crystals: (1) the direct hydrogen bond between βAsn35–β153PCB was observed in TICPC, whereas the indirect hydrogen bond between βSer35–β153PCB via a water molecule was observed in *T. vulcanus*, and (2) one of the two indirect hydrogen bonds between βAsn35–αAsp28 in *T. vulcanus* was not observed in TICPC-6, suggesting the destabilization of hexameric structures in TICPC (Supplementary Fig. 26a). Moreover, the indirect hydrogen bond between αLys32–β153PCB in TICPC is important to stabilize (αβ) monomer, whereas αGlul32 in TICPC rather stabilizes a dimeric structure (αβ)2 by interacting with the neighboring β subunit in the dimeric interface (Supplementary Figs. 26b and 27a). These unique residues might be responsible for unusual octameric state.

The PISA program calculated the solvation free energies of oligomeric (αβ)2, (αβ)3, (αβ)4, (αβ)6, and (αβ)8 states (ΔGsol) as −87.4, −162.3, −204.1, −352.3, and −420.3 kcal/mol, respectively. These values increased as the accessible surface area of each oligomer increased, where the structure of the dimer was assumed to be one-third of the hexamer [contact% = buried surface area / (accessible surface area + buried surface area) = 26%] because a dimer in two-thirds of a trimer is expected to be unstable in solution (contact% = 4%) (Supplementary Fig. 27). The dissociation free energies of oligomeric (αβ)2, (αβ)3, (αβ)4, (αβ)6, and (αβ)8 states into monomeric (αβ) states (ΔGdiss) were also calculated as 4.8, 20.9, 160, 54.6, and 20.9 kcal/mol, respectively, clearly indicating that hexamers were thermodynamically more stable than octamers. These results also revealed that trimers were
thermodynamically more stable than tetramers, which is consistent with findings that all previously observed CPCs in crystals have been trimers or hexamers. Therefore, the formation of the hexamer from monomers is thought to be the major assembly route for CPCs.

By utilizing the $\Delta G^0_{\text{diss}}$ values, the Gibbs free energies of formation ($\Delta G^0_f$) between oligomers were calculated to further investigate the assembly process (Fig. 4). Interestingly, the $\Delta G^0_f$ value from tetramers into an octamer was positive (11.1 kcal/mol), which indicated that octamers are difficult to assemble from tetramers, whereas the $\Delta G^0_f$ value from dimers into an octamer was negative. These results supported the proposed monomer–dimer–octamer assembly process as a plausible process by which TlCPC-8 is formed. Therefore, the formation of unusual dimeric ($\alpha\beta$)$_2$ states may be kinetically more favorable than the formation of trimers or tetramers in TlCPC, which supports the conclusion that dimers are key intermediate structures in the assembly of the newly observed octameric ($\alpha\beta$)$_8$ state in TlCPC.

In summary, an unusual octameric ($\alpha\beta$)$_8$ state in TlCPC was presumably formed by (1) adjusting the dihedral angles of flexible monomeric ($\alpha\beta$) units, (2) assembling monomers into kinetically favored dimeric ($\alpha\beta$)$_2$ states, and (3) assembling four dimers into a cyclic structure. Although the key parameter for switching the hexameric and octameric states of TlCPC remains under investigation, one possible factor is that crystallization of the solution, freezing of the solution, and analysis of the solution by ultracentrifugation were performed soon after isolating native CPCs from O-77 at 4–20 °C and kinetically controlled. These in vitro results expand the current perspective of the PBP assembly process in vivo.

**Conclusion**

In this study, a previously unobserved octameric state of CPC that differed from the conventional hexameric state was discovered for the first time. The existence of non-conventional octameric CPC was confirmed by X-ray crystallographic analysis and the first high-resolution cryo-EM analysis of the native CPC from O-77. Although the monomers in both hexameric and octameric states were essentially identical, the structural analysis indicated that slight conformation changes in potentially flexible interfaces and unique amino acid residues enabled the formation of different oligomeric states from the same primary structures. The unusual dimeric states observed via analytical ultracentrifugation were presumably key intermediate structures in the assembly of the octamer in TlCPC; however, the proposed monomer–dimer–octamer assembly process might be observed in other CPCs because TlCPC shares relatively high sequence identities ($\alpha$: 68–86%; $\beta$: 65–87%) with the previously reported hexameric CPCs in crystals. Since the assembly of identical subunits into symmetrically different oligomeric states is rare and interesting in biology, the experimental results reported in this study are important for understanding not only the assembly process of oligomers but also the evolution process of proteins. Although the biological significance of the octameric structure remains to be clarified, we will focus on it in future studies. We believe that this study provides new insights into the assembly processes of CPCs and PBSs both in vivo and in vitro and sheds new light on the mechanism of energy transfer in the light-harvesting complexes of cyanobacteria.
Methods

Chemicals. Potassium hydroxide, potassium dihydrogen phosphate, sodium chloride, 2-amino-2-hydroxymethyl-1,3-propanediol (Tris), and hydrochloric acid were purchased from FUJIFILM Wako Pure Chemical Corporation and used as received.

Purification of CPC. Cells of O-77 were grown according to a previously reported procedure. The cells were harvested and homogenized with 10-mM Tris/HisCl buffer at pH 8.0 (buffer A) using NZ-1000 (EYELA, Japan) and then disrupted in an ice bath by sonication three times (2-min sonication at 30 W with a 2-min break) using an Ultrasonic Disruptor UD-200 (TOMY SEIKO, Japan). Cell debris and unbroken cells were removed by centrifugation (5000 g, 20 min, 277 K) using a Himac CR20Gi (HITACHI, Japan), and the resulting supernatant was centrifuged at 100,000 g for 1 h using an Optima L-90K (Beckman Coulter, USA). The supernatant of soluble cell extracts was then loaded onto a DEAE Sepharose fast flow column (SK 26/20; GE Healthcare Life Sciences) preequilibrated with buffer A by washing with 500 mL of the same buffer at a flow rate of 10 mL min⁻¹. The CPC-containing solution was eluted at 0.15–0.24 M NaCl with 10–19 mM ionic conductivity using buffer A alone and buffer A containing 1.0 M NaCl (buffer B) as eluents. Fractions containing CPC were combined and diluted threefold with buffer A and the resulting solution was loaded onto a Q Sepharose high-performance column (HR 16/10; GE Healthcare Life Sciences) preequilibrated with buffer A by washing with 150 mL of the same buffer at a flow rate of 4 mL min⁻¹. Blue CPC-containing solution was eluted at 0.15–0.20 M NaCl with 13–17 mM ionic conductivity using buffers A and B. The fractions containing CPC were combined and concentrated using Amicon Ultra-15 50 kDa (Merck, Germany). The resulting concentrated CPC was purified using a Superdex 200 prep grade column (HR 16/50; GE Healthcare Life Sciences) and a potassium phosphate buffer solution (pH 7.0, 10 mM) as an eluent. The concentration of the purified CPC solution was determined by the following equation:

\[
\text{Concentration of CPC (mg/mL)} = \frac{[A_{415} - 0.474[A_{528}]]}{5.34},
\]

where A_415 and A_528 are the absorbance intensities at 615 and 652 nm, respectively, which were obtained by measuring the UV–vis spectrum.

Spectral analysis. Circular dichroism was measured using Chirascan (Applied Photophysics, UK) with a 0.1-cm quartz cell at 20 °C. The UV–vis spectrum was measured on JASCO V-670 (JASCO, Japan) with a 1-cm quartz cell. The fluorescence spectrum with an excitation wavelength at 436 nm was measured using a Fluoromax-4 Spectrofluorometer (Horiba, Japan) with a 0.1-cm quartz cell at 20 °C. The UV vis spectrum.

Crystallization, data collection, and structural determination and refinement. CPC was crystallized using the sitting-drop vapor diffusion method at 20 °C.itting drops containing 200 mL of protein solution mixed with 200 mL of reservoir solution. Two types of crystals (TICPC-6 and TICPC-8) were obtained (Supplementary Table 1 and Supplementary Fig. 12). For TICPC-6 crystals (rhombus or rod), the reservoir condition contained 8% Tacsimate (pH 5.0) and 20% polyethylene glycol 3,350. Prior to data collection, crystals were transferred to a cryoprotectant solution containing polyethylene glycol 3,350 and then flash-cooled to −180 °C. For TICPC-8 crystals (cube), the reservoir condition contained 1.6-M sodium chloride, 8% polyethylene glycol 6,000, and 20% glycerol, i.e., a cryoprotectant solution. The crystals were then flash-cooled to −180 °C. The cubic crystals of O-77 were determined by the X-ray crystallographic analysis were obtained at 42 °C which is the growth temperature of O-77 in this study.

X-ray diffraction data were collected at beamline BL45XU at Spring-8 (Hyogo, Japan) and 100 K with a wavelength of 1.00 Å. These data were processed using the ZOOS system. Phases were determined by molecular replacement using the program Phaser and search models of monomers of TcCP (PDB ID: 1T7V). These Models were built using the program COOT, and the program Phenix.refine was used for refinement. The crystal models displayed good geometry when analyzed by MolProbity. In the Ramachandran plots, 98.3% and 1.7% of the residues constituting TICPC-6 were in the most favored and allowed regions, whereas 98.0% and 2.0% of the residues constituting TICPC-8 were in the most favored and allowed regions, respectively.

Cryo-EM sample preparation, data collection, and data processing. For cryo-grid preparation, 3 μL of TICPC samples (6.7 and 10 mg/mL) in potassium phosphate buffer solution (pH 7.0, 10 mM) were applied onto a holey carbon grid (Quantifoil, Cu, R1.2/1.3, 300 mesh). The grid was rendered hydrophilic by a 30-s glow-discharge in the air (11 mA current) with PIB-10 (Vacuum Device Inc., Ibaraki, Japan). The grid was then blotted for 5 sec (blot force 15) at 18 °C and 100% humidity before being flash-frozen in liquid ethane using Vitrobot Mark IV (Thermo Fisher Scientific, Waltham, MA, USA). For automated data collection, 2,036 micrographs were acquired on a Talos Arctica (Thermo Fisher Scientific) microscope operating at 200 kV in nanoprobe mode and using EPU software. The movie micrographs were collected on a 4k × 4k using a Falcon 3EC direct electron detector (electron counting mode) at a nominal magnification of 120,000 (0.88 Å/pixel). Forty-nine movie fractions were recorded at an exposure of 1.02 electrons per Å² per fraction, which corresponded to a total exposure of 50 e⁻/Å². The defocus steps used were −1.0, −1.5, −2.0, and −2.5 μm. The movies were processed by MotionCor2, Gctf, and RELION3. The cryo-EM models were visualized by UCSF chimera. See Supplementary Note 1 and Supplementary Figs. 15–18 for the details of the cryo-EM data processing.

Summary reporting. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Crystal structures are deposited at the Protein Data Bank with accession codes 7EFW (TICPC-6) and 7EFU (TICPC-8). Cryo-EM maps are deposited in the Electron Microscopy Data Bank under accession codes EMD-31900 (TICPC-6) and EMD-31089 (TICPC-8). Structure coordinates related to the cryo-EM maps are deposited at the Protein Data Bank with accession codes 7EFH (TICPC-6) and 7EF7 (TICPC-8). Any remaining information can be obtained from the corresponding authors upon reasonable request.

Received: 29 April 2021; Accepted: 5 October 2021; Published online: 29 October 2021

References

1. Green, B. R., Anderson, J. M. & Parson, W. W. Photosynthetic membranes and their light-harvesting antennas. In Light-Harvesting Antennas in Photosynthesis (eds. Green, B. R. & Parson, W. W.) 1–28 (Springer, 2003).
2. Croce, R. & van Amerongen, H. Natural strategies for photosynthetic light harvesting. Nat. Chem. Biol. 10, 492–501 (2014).
3. Cheng, Y. -C. & Fleming, G. B. Dynamics of light harvesting in photosynthesis. Annu. Rev. Phys. Chem. 60, 241–262 (2009).
4. Tachibana, Y., Vayssieres, L. & Durant, J. R. Artificial photosynthesis for solar water-splitting. Nat. Photon. 6, 511–518 (2012).
5. Blankenship, R. E. Bioenergy applications and artificial photosynthesis. In Molecular Mechanisms of Photosynthesis. 237–248 (Wiley, 2014).
6. Grossman, A. R., Schierer, M. R., Chiang, G. G. & Collier, J. L. The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiol. Rev. 57, 725–749 (1993).
7. MacColl, R. Cyanobacterial phycobilisomes. J. Struct. Biol. 124, 311–334 (1998).
13. Glazer, A. N. Light harvesting by phycobilisomes. Biochim. Biophys. Acta Bioenerg. 1861, 148047 (2020).

28. Eisenberg, I. et al. Concentration-based self-assembly of phycocyanin. Protein Expr. Purif. 42, 248–255 (2005).

34. Nakamori, H., Yatabe, T., Yoon, K.-S. & Ogo, S. Purification and characterization of an oxygen-evolving photosystem II from Leptolyngbya sp. strain O-77. J. Biosci. Bioeng. 118, 119–124 (2014).

35. Yoon, K.-S., Nguyen, N. T., Tran, K. T., Tsuji, K. & Ogo, S. Nitrogen fixation genes and nitrogenase activity of the non-heterocystous cyanobacterium Thermosynechococcus vulcanus PCC 7924. Biotechnol. Appl. Biochem. 55, 234–339 (2017).

36. Minato, T., Teramoto, T., Kakuta, Y., Ogo, S. & Yoon, K.-S. Biochemical and structural characterization of a thermostable Dps protein with His-type ferroxidase centers and outer metal-binding sites. FEBS Open Bio 10, 1219–1229 (2020).

41. Chang, L. et al. Structural organization of an intact phycobilisome and its association with photosystem II. Cell Res 25, 726–737 (2015).

42. Swanson, R. V. & Glazer, A. N. Phycobiliprotein methylation. effect of the γ-N-methylasparagine residue on energy transfer in phycocyanin and the phycobilisomes. J. Mol. Biol. 214, 787–796 (1990).

44. Hayashi, H., Çôte, A. P., Furukawa, H., O’Keefe, M. & Yagi, O. M. Zeolite A imidazolate frameworks. Nat. Mater. 6, 501–506 (2007).

50. Chen, C.-S. et al. How to change the oligomeric state of a circular protein assembly: switch from 11-subunit to 12-subunit TRAP suggests a general mechanism. Proc. Natl Acad. Sci. U.S.A. 109, 2348–2353 (2012).

53. Cornish-Bowden, A. J. & Koshland, D. E. The quaternary structure of proteins and its consequences. Adv. Enzymol. 44, 247–275 (1977).

56. Young, H. S., Dang, H., Lai, Y., DeRosier, D. J. & Khan, S. Variable symmetry in Salmomonella typhimurium flagellar motors. Biophys. J. 84, 571–577 (2003).

57. Tian, H., Wang, J., Chang, L., Jonas, D. & Fang, Y. Structure and function of the antenna protein from the thermophilic cyanobacterium Synechococcus elongatus. EMBO J. 15, 1205–1209 (1996).

58. Arndt, N. M. & Marler, H. J. 105–160 (Novo Science Publishers, Inc., 2009).

59. Boezi, E., Bolognesi, D. & Maraldi, T. The proposal of identical subunits in the core of the TRAP complex. Acta Crystallogr. D. Struct. Biol. 62, 324–329 (2016).
Acknowledgements

Thermoleptolyngbya sp. O-77 was named after Saori Ogo, whose assistance was invaluable to the hunt for a new thermophilic cyanobacterium. This paper is dedicated to her memory. This work was supported in part by JST CREST Grant Number JPMJCR18R2 (S.O.), JSPS KAKENHI Grant Numbers JP18H02091 (K.-S.Y) and JP18H00191 (T. Minato), MEXT Leading Initiative for Excellent Young Researchers Grant Number JPAXS030200400 (T. Minato), JSPS Core-to-Core Program, the World Premier International Research Center Initiative (WPI), Japan, and the Platform Project for Supporting Drug Discovery and Life Science Research (BINDER) from the Japan Agency for Medical Research and Development (AMED) under Grant Number JP21am0101071 (supporting no. 2493). T. Minato was supported by the JSPS through a Research Fellowship for Young Scientists. We thank the staff members of the beamline BL45XU facilities at SPring-8 for their help with X-ray data collections. The cryo-EM data were collected at the Cryo-EM facility in KEK (Ibaraki, Japan).

Author contributions

T. Minato, T.T., Y.K., S.O. and K.-S.Y. devised the study. T. Minato, N.K.H., K.Y. and K.-S.Y. isolated and characterized the protein. T.T. and Y.K. performed X-ray crystallographic and cryo-EM structural analyses. N.A., M.K., M.A., T. Moriya and T.S. performed cryo-EM data collection, processing, and cryo-EM structural analysis. T. Minato and T.T. wrote the paper.

Competing interests

The authors declare no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-021-02767-x.

Correspondence and requests for materials should be addressed to Yoshimitsu Kakuta or Ki-Seok Yoon.

Peer review information Communications Biology thanks Kai-Hong Zhao and the other, anonymous, reviewers for their contribution to the peer review of this work. Primary Handling Editors: Joanna Timmins and Anam Akhtar. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021