STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

# Crystal structures of native cytochrome $c_6$ from *Thermosynechococcus elongatus* in two different space groups and implications for its oligomerization

Sven Falke,[a] Christian Feiler,[b] Henry Chapman[c] and Iosifina Sarrou[c]*

[a]Institute for Biochemistry and Molecular Biology, University of Hamburg, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany, [b]Helmholtz-Zentrum Berlin für Materialien und Energie, Albert-Einstein-Strasse 15, 12489 Berlin, Germany, and [c]Center for Free-Electron Laser Science, c/o DESY, Notkestrasse 85, 22607 Hamburg, Germany. *Correspondence e-mail: iosifina.sarrou@cfel.de

Received 10 March 2020
Accepted 23 July 2020

Edited by N. Sträter, University of Leipzig, Germany

**Keywords:** cyanobacteria; photosynthesis; oligomerization; mass spectrometry; redox model protein.

**PDB references:** cytochrome $c_6$ from *Thermosynechococcus elongatus*, 6tr1; 6tsy

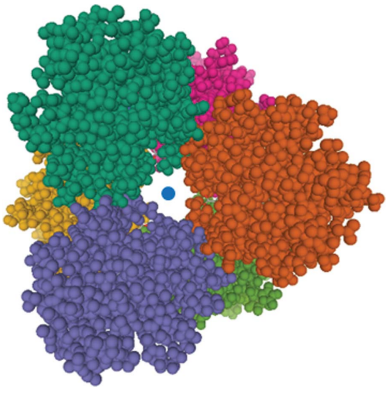
**Supporting information**: this article has supporting information at journals.iucr.org/f

Native cytochrome $c_6$ was purified from an extract of strain BP-1 of the thermophilic cyanobacterium *Thermosynechococcus elongatus*. The protein was crystallized, and with only slight modifications of the buffer and vapour-diffusion conditions two different space groups were observed, namely *H*3 and *C*2. Both crystal structures were solved; they contained three and six molecules per asymmetric unit and were refined to 1.7 and 2.25 Å resolution, respectively. To date, the structure of native cytochrome $c_6$ from *T. elongatus* has only been reported as a monomer using NMR spectroscopy, *i.e.* without addressing putative oligomerization, and related structures have only previously been solved using X-ray crystallography after recombinant gene overexpression in *Escherichia coli*. The reported space groups of related cyanobacterial cytochrome $c_6$ structures differ from those reported here. Interestingly, the protein–protein interfaces that were observed utilizing X-ray crystallography could also explain homo-oligomerization in solution; specifically, trimerization is indicated by infra-red dynamic light scattering and blue native gel electrophoresis in solution. Trimers were also detected by mass spectrometry. Furthermore, there is an indication of post-translational methylation in the crystal structure. Additionally, the possibility of modifying the crystal size and the redox activity in the context of photosynthesis is shaping the investigated cytochrome as a highly suitable model protein for advanced serial crystallography at highly brilliant X-ray free-electron laser sources.

## 1. Introduction

Cytochrome $c_6$ (Cyt $c_6$) is a key protein in the light reaction of photosynthesis. It is a small water-soluble protein that is located on the luminal side of the thylakoid membrane. Cyt $c_6$ dynamically interacts within the membrane, forming a protein complex with cytochrome $b_6f$, which has to be sufficiently specific to allow rapid electron transfer towards photosystem I (PSI; Díaz-Moreno *et al.*, 2014). At the same time, the complex needs a high dissociation rate to enable rapid turnover in order not to limit the flow of electrons through the redox chain (Moser *et al.*, 1992). The electron-transfer pathway of Cyt $c_6$ and its interaction sites with the large membrane complexes are still not well understood (Kölsch *et al.*, 2018). Therefore, new structural information is of great interest.

Cyt $c_6$ is a typical 10 kDa single-heme *c*-type cytochrome with the cofactor covalently bound to the cysteine residues in a conserved C*xx*CH motif (Supplementary Figs. S1 and S2). The central iron ion is hexacoordinated in a protoporphyrin derivative with histidine and methionine residues acting as



axial ligands, as also revealed by the structures of cyanobacterial and green algal Cyt $c_6$ (Beissinger *et al.*, 1998; Banci *et al.*, 1998; Sawaya *et al.*, 2001; Worrall *et al.*, 2007; Bialek *et al.*, 2009). Further, there is an indication that a functional class I *c*-type cytochrome of a chloroplast forms functionally relevant small oligomers (Kerfeld *et al.*, 1995). Selected closely related Cyt $c_6$ sequences are aligned for comparison and presented in Supplementary Fig. S2.

Focusing on the cyanobacteria, a high-resolution X-ray structure of recombinant cyanobacterial Cyt $c_6$ from *Synechococcus* sp. PCC 7002 (Bialek *et al.*, 2009; PDB entry 3dr0) and the NMR structure of the native protein from *Thermosynechococcus elongatus* (Beissinger *et al.*, 1998; PDB entry 1c6s) have been reported, but so far no three-dimensional crystal structure of a native cyanobacterial Cyt $c_6$ has been described and potential oligomerization has not been addressed. Consequently, a detailed structural investigation of Cyt $c_6$ from its native source will help to understand the dynamic regulatory protein–protein interactions in PSI in more detail and will add another piece to a more detailed

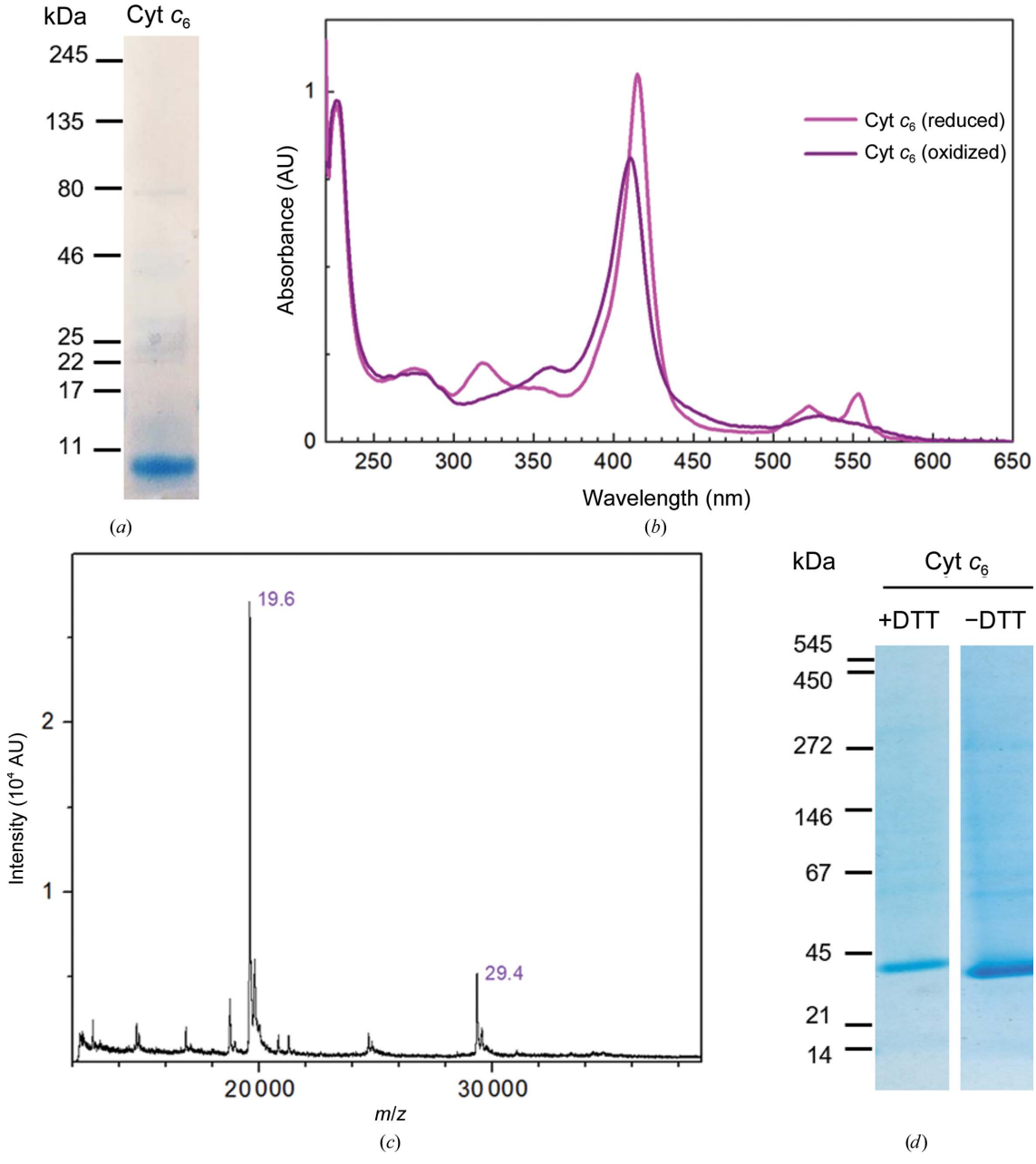



**Figure 1**
Purification and characterization of TeCyt $c_6$. (*a*) The purity of the protein was verified by SDS–PAGE. (*b*) Absorbance properties were characterized by UV–Vis absorbance spectroscopy; the cytochrome was oxidized by the addition of 1 m*M* potassium ferricyanide. (*c*) Additionally, MALDI-TOF MS indicates a dimeric state of the protein (19.6 kDa) as well as a trimeric state (29.4 kDa) (see also Supplementary Fig. S3). (*d*) In solution, homogenous oligomerization is indicated by bnPAGE in the presence and absence of 10 m*M* DTT at a protein concentration of 0.5 mg ml$^{-1}$. Considering a nonglobular oligomerization state of the protein, the molecular-weight standard provides a rough mass estimation. Trimerization is further supported by IR-DLS (Supplementary Fig. S4). The molecular masses of the standard proteins are indicated in kDa.

**Table 1**
Crystallization conditions.

| | TeCyt $c_6$ (data set 1; PDB entry 6tsy) | TeCyt $c_6$ (data set 2; PDB entry 6tr1) |
|---|---|---|
| Method | Vapour diffusion, sitting drop | Vapour diffusion, hanging drop |
| Plate type | 96-well plate | 24-well plate |
| Temperature (K) | 293 | 277 |
| Protein concentration (mg ml$^{-1}$) | 18 | 40 |
| Buffer composition of protein solution | 20 m*M* Tris pH 8.0 | 20 m*M* MES pH 6.5 |
| Composition of reservoir solution | 2 *M* $(NH_4)_2SO_4$, 40 m*M* $KNO_3$ | 2 *M* $(NH_4)_2SO_4$, 40 m*M* $KNO_3$ |
| Drop volume and mixing ratio | 500 nl protein solution, 500 nl reservoir solution | 500 nl protein solution, 500 nl reservoir solution |
| Volume of reservoir (µl) | 50 | 500 |

structural and functional understanding of the photosynthetic machinery of cyanobacteria (Kölsch *et al.*, 2018; Kłodawska *et al.*, 2020). Cyanobacteria are of exceptionally high economic and ecological value owing to their high abundancy, their photosynthetic activity in the oceans to consume and reduce $CO_2$ as well as their applications in biotechnology and agriculture (Kumar *et al.*, 2019).

Here, we report the crystallization and high-resolution crystal structures of native *T. elongatus* cytochrome $c_6$ (TeCyt $c_6$) in two different space groups, which contain three and six molecules per asymmetric unit, respectively, potentially indicating the presence of trimers of cytochromes, which were also observed by mass spectrometry, blue native gel electrophoresis (bnPAGE) and infra-red dynamic light scattering (IR-DLS). Structural differences between the solved and related structures and interfaces are discussed below. Additionally, slight modifications of the protocols described here produce crystals with dimensions of less than 10 µm and with a high solvent content, which provide a valuable starting point in the preparation for upcoming soaking and time-resolved serial crystallography investigations of Cyt $c_6$ and its interaction with PSI and the Cyt $b_6f$ complex at currently available X-ray free-electron laser sources.

## 2. Materials and methods

### 2.1. Macromolecule production and sample-quality verification

TeCyt $c_6$ was purified from an extract of the cyanobacterium *T. elongatus* strain BP-1 grown in a photobioreactor (PSI, Czech Republic) essentially as described by Shin *et al.* (1984). However, in contrast to this protocol, a DEAE chromatography column equilibrated with 20 m*M* MES pH 6.5 was used connected to an ÄKTA purifier (GE Healthcare, USA). TeCyt $c_6$ does not bind to the column material under these conditions and is collected pure for further use. The integrity and purity of the target protein obtained by chromatography were verified via SDS–PAGE analysis using 4–15% Mini-PROTEAN TGX precast protein gels (Bio-Rad, Germany). In addition to mass-spectrometry experiments (Sections 2.2 and S1, Fig. 1*c* and Supplementary Fig. S3), UV–Vis absorption measurements were performed using a Varian 50 Bio single-beam spectrophotometer (Agilent Technologies, USA; Fig. 1*c*). The oxidation of TeCyt $c_6$ was achieved by the addition of 1 m*M* potassium ferricyanide (Koike & Katoh, 1979).

### 2.2. MALDI-TOF MS

A Bruker autoflex speed MALDI-TOF/TOF instrument was utilized to determine the mass of TeCyt $c_6$. The protein dissolved in 20 m*M* MES pH 6.5 was mixed with sinapinic acid (SA) matrix and spotted onto the MALDI target plate for ionization. The data were processed using the *flexAnalysis* 3.4.76.0 and *flexControl* 3.4.135 software.

### 2.3. IR-DLS and bnPAGE

A multi-well plate-reading DLS instrument (SpectroLight 600, XtalConcepts, Germany) equipped with an infra-red laser (wavelength 785 nm) and the dedicated software were used to investigate the hydrodynamic radius and dispersity of the protein without being influenced by its fluorescence. Droplets with 4 µl volume each in a Terazaki plate (NUNC, Denmark), covered with silicone oil to prevent evaporation, were prepared for the measurements (Supplementary Fig. S4). For bnPAGE, sample solutions of 0.5 mg ml$^{-1}$ Cyt $c_6$ containing 50 m*M* NaCl were applied to 3–12% acrylamide bis-Tris-buffered gradient gels. Materials and bnPAGE buffers were obtained from Serva (Germany) and used according to the instructions from the manufacturer. Gels were subsequently stained using Coomassie Brilliant Blue for evaluation.

### 2.4. Crystallization

Pure protein was subjected to initial robot-assisted vapour-diffusion crystallization trials using 400 distinct crystallization solutions. Conditions were further optimized by modifying the incubation temperature, buffer pH, protein concentration, precipitant concentration and additive salts. Crystals obtained under two different conditions as specified in Table 1 and using the same reservoir composition were exposed to X-rays. Micrographs of the mature crystals are shown in Fig. 2.

### 2.5. Data collection and processing

Diffraction data were collected using synchrotron radiation (Table 2). Prior to data collection and flash-cooling, crystals were dipped into reservoir solution containing 20%(*v*/*v*) glycerol for cryoprotection. For the crystal utilized to collect data set 2 twinning by merohedry was indicated and considered (twinning fractions: 0.522 for $h, k, l$ and 0.478 for $k, h, -l$).

**Table 2**
Data collection and processing.

Values in parentheses are for the outer shell.

| | TeCyt $c_6$ (data set 1; PDB entry 6tsy) | TeCyt $c_6$ (data set 2; PDB entry 6tr1) |
|---|---|---|
| Diffraction source | Beamline P11, PETRA III, DESY, Hamburg, Germany | Beamline P13, PETRA III, DESY, Hamburg, Germany |
| Wavelength (Å) | 1.0332 | 0.9762 |
| Temperature (K) | 100 | 100 |
| Detector | PILATUS 6M, Dectris | PILATUS 6M, Dectris |
| Crystal-to-detector distance (mm) | 200 | 189 |
| Rotation range per image (°) | 0.1 | 0.1 |
| Total rotation range (°) | 200 | 120 |
| Exposure time per image (s) | 0.070 | 0.060 |
| Space group | *C*2 | *H*3 |
| *a*, *b*, *c* (Å) | 106.0, 109.9, 55.4 | 94.8, 94.8, 160.22 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90, 100.9, 90 | 90, 90, 120 |
| Mosaicity (°) | 0.148 | 0.573 |
| Resolution range (Å) | 41.76–2.25 (2.33–2.25) | 73.07–1.70 (1.80–1.70) |
| Total No. of reflections | 87025 (9118) | 170023 (32575) |
| No. of unique reflections | 28365 (2409) | 55146 (4873) |
| Completeness (%) | 95.5 (80.7) | 99.9 (99.7) |
| Multiplicity | 3.2 (3.4) | 3.1 (3.1) |
| $\langle I/\sigma(I)\rangle$ | 9.9 (2.0) | 8.5 (2.7) |
| $CC_{1/2}$ (%) | 99.9 (99.9) | 98.8 (85.5) |
| $R_{r.i.m.}$† | 0.18 (1.02) | 0.19 (0.55) |
| Overall *B* factor from Wilson plot (Å$^2$) | 30.8 | 31.7 |

† Estimated as $R_{r.i.m.} \simeq R_{merge}[N/(N-1)]^{1/2}$, where $N$ is the data multiplicity.

### 2.6. Structure solution and refinement

As specified in Table 3, diffraction data were processed and the protein model quality was verified using the *PDB-REDO* server (Joosten *et al.*, 2014) or *MolProbity* (Williams *et al.*, 2018). *XDS* (Kabsch, 2010) or *XDSAPP* (Sparta *et al.*, 2016) was used for data processing and indexing, *CCP*4 software (*SCALA*, *MOLREP* and *Phaser*; Vagin & Teplyakov, 2010; McCoy *et al.*, 2007; Winn *et al.*, 2011) was applied to reduce and scale the diffraction data as well as to perform molecular replacement, and the *Phenix* application *phenix.refine* (data set 1; Afonine *et al.*, 2012; Liebschner *et al.*, 2019) and *REFMAC* (data set 2; Murshudov *et al.*, 2011) were used to refine the coordinates. The obtained coordinates of the TeCyt $c_6$ structures were deposited in the Protein Data Bank (PDB) as entries 6tr1 and 6tsy.

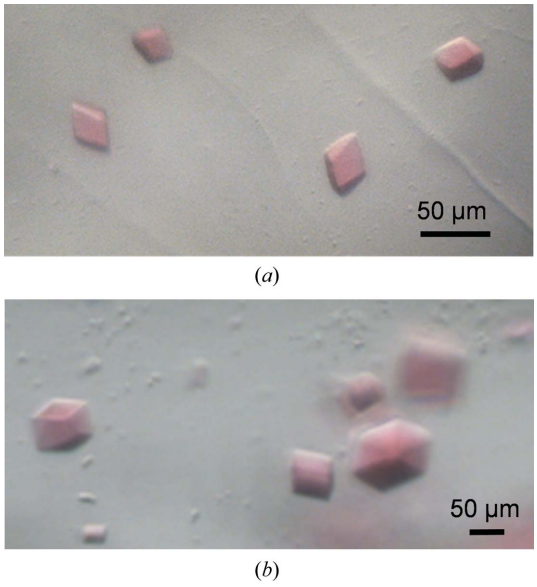



**Figure 2**
'Brick-shaped' crystals of TeCyt $c_6$ obtained by vapour-diffusion crystallization with a characteristic pink colour and sizes between 30 and 100 µm in all three dimensions. (*a*) The crystals utilized for the collection of data set 1 (Table 1; space group *C*2) grew to full size within one day. (*b*) The crystals utilized to obtain data set 2 (space group *H*3) grew to full size within 21 days at 4°C.

## 3. Results and discussion

UV–Vis spectroscopy and LC-MS/MS confirmed the identity of the protein (Fig. 1 and Supplementary Fig. S1). The obtained crystals that were exposed to X-rays are displayed in Fig. 2. The resulting structure of TeCyt $c_6$ isolated from the BP-1 strain possesses the typical conserved pure $\alpha$-helical cytochrome $c_6$ fold and heme cofactor (Supplementary Fig. S5*a*) as clearly observed in both crystallographic data sets. Loops and short turns interconnect the six $\alpha$-helices. As expected, the main chains of both of the TeCyt $c_6$ structures reported here and the previously determined NMR structure (Beissinger *et al.*, 1998) are almost identical, with r.m.s.d. values of around 1 Å (Fig. 3*a*). The respective heme cofactors of all TeCyt $c_6$ structures reported are very similar and superimpose well (Fig. 3*b*), with rotated carboxypropyl side chains of the heme compared with the NMR data set (PDB entry 1c6s). Other deviating loops and the C-terminus position of the TeCyt $c_6$ X-ray structures in comparison to the NMR structure could be related to local flexibility of the NMR model and observed higher *B* factors for the crystal structure. Compared with the structures of TeCyt $c_6$, a strongly deviating loop of two related cytochrome X-ray structures is highlighted in Fig. 3(*a*) and consists of amino acids 65–75.

**Table 3**
Structure solution and refinement.

Values in parentheses are for the highest shell.

| | TeCyt $c_6$ (data set 1; PDB entry 6tsy) | TeCyt $c_6$ (data set 2; PDB entry 6tr1) |
|---|---|---|
| Resolution range (Å) | 41.76–2.25 (2.33–2.25) | 36.53–1.70 (1.80–1.70) |
| Completeness (%) | 95.5 (80.7) | 99.9 (99.7) |
| No. of reflections, working set | 28267 (2386) | 55146 (2864) |
| No. of reflections, test set | 1413 (118) | 2864 (207) |
| Final $R_{cryst}$ (%) | 18.7 (40.6) | 18.6 (20.1) |
| Final $R_{free}$ (%) | 25.0 (40.2) | 20.8 (27.0) |
| Cruickshank DPI | 0.263 | 0.067 |
| No. of non-H atoms | | |
| Protein | 3829 | 1922 |
| Ion | 5 | 5 |
| Ligand (heme) | 263 | 134 |
| Water | 177 | 55 |
| Total | 4269 | 2116 |
| R.m.s. deviations | | |
| Bonds (Å) | 0.013 | 0.032 |
| Angles (°) | 1.28 | 2.804 |
| Average *B* factors (Å$^2$) | | |
| Protein | 38.8 | 21.4 |
| Ion | 65.0 | 36.7 |
| Ligand, heme | 28.9 | 14.8 |
| Water | 36.2 | 25.3 |
| Ramachandran plot | | |
| Most favoured (%) | 95 | 95 |
| Allowed (%) | 5 | 5 |

Also, the presence of specific post-translational methylation may depend on particular culture conditions and light stress, as indicated for related cytochromes and other algal proteins (Kerfeld *et al.*, 1995). It typically involves methyltransferases (Clarke & Tamanoi, 2006). In particular, one diffraction data set indicated an *N*-4-methyl-L-asparagine at amino acid position 97 of chain *C* (PDB entry 6tr1; Supplementary Fig. S5*b*), which is reported as a glutamine residue elsewhere but can also be occupied by asparagine in related cytochrome sequences from cyanobacteria. Additionally, glutamate was identified in position 70 (PDB entry 6tr1; Supplementary Fig. S2) and is conserved among different related cyanobacteria, but was identified as an aspartate in the NMR structure (PDB entry 1c6s; Beissinger *et al.*, 1998), which has the highest sequence identity. Consequently, a minor sequence inhomogeneity between different preparations of TeCyt $c_6$ from the native source is indicated here after careful refinement of the structures. Post-translational modifications or sequence polymorphisms might be underestimated in sequence analyses of proteins obtained from their natural sources (Whitelegge *et al.*, 2007). They can, in some cases, be responsible for an altered protein crystal structure (Xin & Radivojac, 2012) or potentially even crystal geometry, if the polymorphisms are surface-exposed, including different oxidation states of a cytochrome (Kerfeld *et al.*, 1995).

The geometry and the molecular packing of the TeCyt $c_6$ structures (PDB entries 6tr1 and 6ts1) are shown in Fig. 4 and Supplementary Fig. S6. The two crystal space groups are *H*3 and *C*2, and the asymmetric unit contains three molecules in the former case and six molecules in the latter (Fig. 4). This phenomenon is not rare. During the crystallization process, parameters such as the temperature, protein flexibility, pH-dependent surface charge, oxidation state and solution

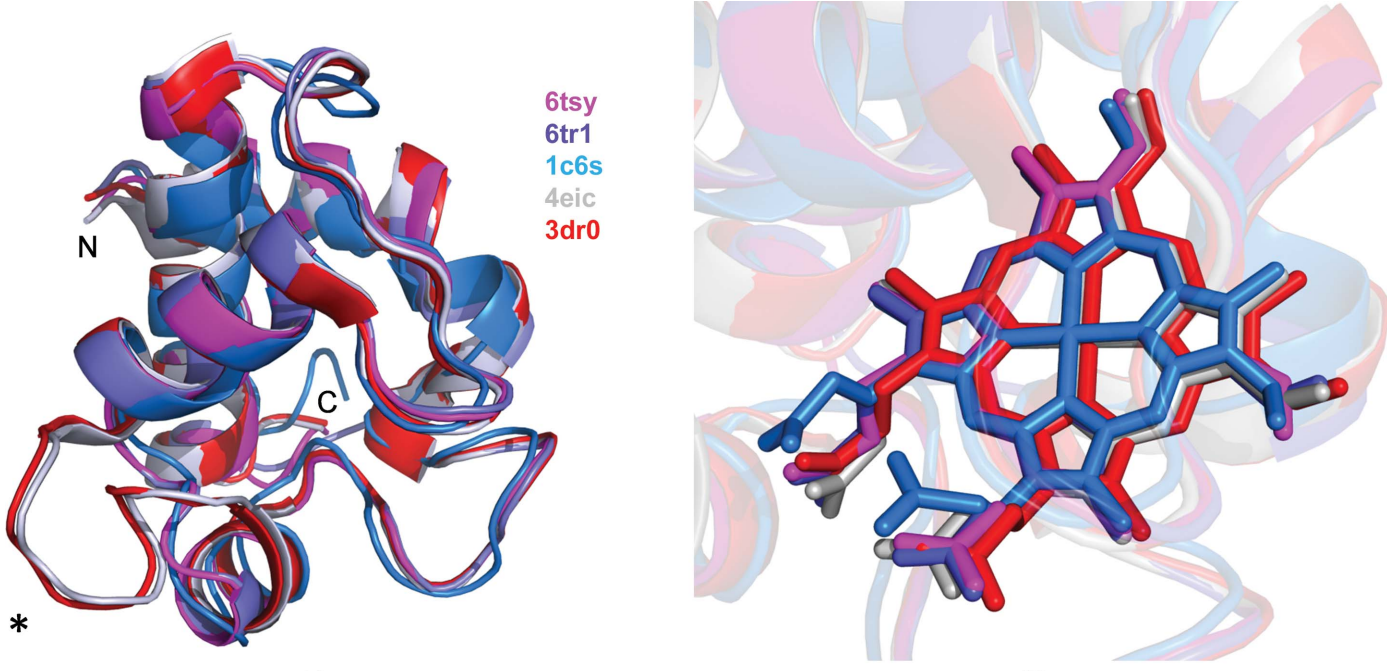



**Figure 3**
Superimposition of the main chain (*a*) and heme group (*b*) of different related cytochrome $c_6$ structures. The different structures are colour-coded by their respective PDB code. Along with the model obtained by NMR spectroscopy (PDB entry 1c6s), the two crystal structures presented here do not contain the insertion ($^{65}$LAGYKDGSKSL$^{75}$) labelled by an asterisk, but instead contain a much shorter almost straight connecting loop (Supplementary Fig. S2). The structure and position of the heme group are widely conserved, with a minor difference in the position of the attached carboxy groups in the model derived by NMR coloured light blue.

additives can alter molecular interactions and could lead to different packing symmetries (Kondrashov *et al.*, 2008; Rashin *et al.*, 2014) despite the use of an identical precipitant solution. A change in the space group was likewise observed for the structure of a recombinant Cyt $c_6$ in the PDB depositions 3dr0 (Bialek *et al.*, 2009) and 4eic (S. Krzywda, W. Bialek, M. Jaskolski & A. Szczepaniak, unpublished work) (Supplementary Table S2). As there is no associated publication for PDB deposition 4eic, a discussion of the very small unit cell with greater than 50% sequence identity of the protein compared with PDB entry 6tr1 (Supplementary Table S2) remains lacking. The three molecules per asymmetric unit in PDB entry 3dr0 form an elongated trimer. One of the two protein–protein interfaces of the central protein chain in the asymmetric unit (interface areas of 345 and 352 Å$^2$) is much larger compared with the other one (interface areas of 139 and 129 Å$^2$) according to *PDBSUM* (Laskowski *et al.*, 2018). This has a certain similarity to PDB entry 6tr1 (interface areas of 413 and 454 Å$^2$ and of 152 and 157 Å$^2$), including a similar number of interface hydrogen bonds. In contrast, the residues involved in the interfaces are partly significantly different for PDB entries 3dr0, 4eic, 6tr1 and 6tsy.

Nonetheless, the elongated assembly of three monomers in the asymmetric unit of PDB entry 6tr1, with space group *H*3, is only an incomplete snapshot of the molecular assembly in the crystal. An expanded view indicates a distinct higher-ordered oligomeric state that relates PDB entries 6tr1 and 6tsy (Fig. 4 and Supplementary Fig. S6).

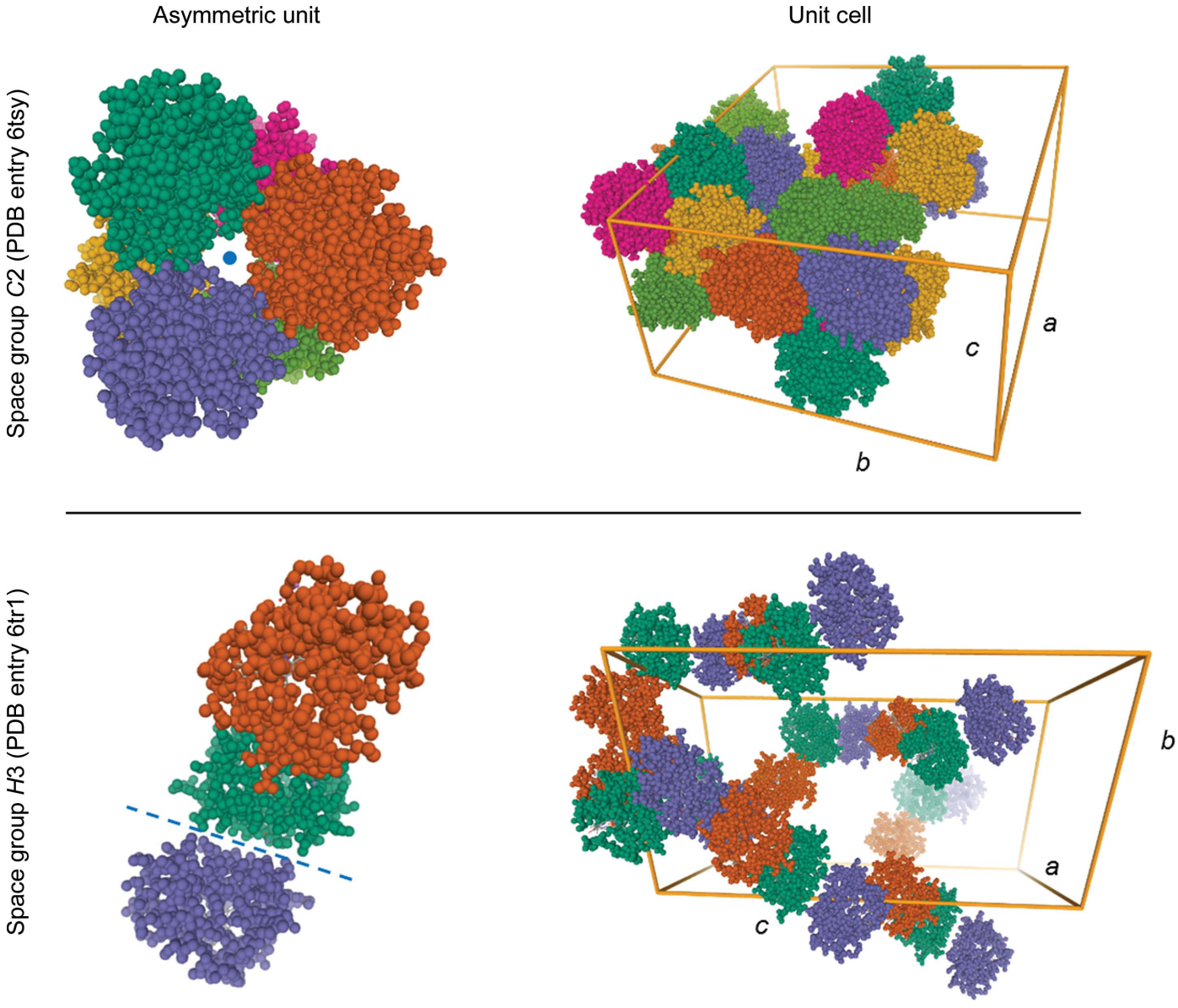



**Figure 4**
Geometry and comparison of the molecular packing of the TeCyt $c_6$ structures with PDB codes 6tsy and 6tr1. Symmetry axes are indicated in blue. In space group *C*2 six molecules per asymmetric unit were identified in a 'ring-like' assembly with a central threefold symmetry axis (upper panel). In space group *H*3 the asymmetric unit consists of three molecules, which are arranged like 'twisted chains' in the crystal (lower panel).

The three monomers within the asymmetric unit in space group *H*3 play individual roles in the crystal packing. The 'outer' chains shown in orange and blue (Fig. 4, lower panel; PDB entry 6tr1) homodimerize with a symmetry-related molecule and describe a presumably physiologically relevant interface. Within this assembly, the monomers exhibit their cofactor in a face-to-face manner. As validated by *PISA* (*Proteins, Interfaces, Structures and Assemblies*; Krissinel & Henrick, 2005) analysis, this assembly is physiologically pertinent, and stable protein homodimers assemble into a higher ordered oligomeric state, as also observed by MALDI-TOF mass spectrometry. These elongated building blocks trimerize in a slightly tilted fashion and constitute a barrel-shaped structure. One monomer in each staggered dimer is post-translationally modified and carries an asparagine methylation. This observed alteration is exclusive to the functionally relevant dimers. It cannot be observed either as a double-methylated homodimer or in the central molecule in the asymmetric unit (Fig. 4, lower panel), which is not involved in forming these dimers. The latter unit forms a solid homotrimer and tops the dimer barrel in a mushroom-head style, and exclusively mediates crystal contacts between the stacked trimeric rings (Supplementary Fig. S6). Specifically, the calculated stability for all observed protein interfaces is comparable and indicates a trimeric assembly, as supported by interface analysis using *PISA* according to Krissinel & Henrick (2005). Therefore, as concluded by *PISA* analysis, the physiologically relevant interface is most likely not present within one asymmetric unit. The asymmetric unit needs to be expanded to envision the biologically relevant interfaces. Crystallographic symmetry axes relate the respective monomers in the oligomers and constitute the appropriate interfaces.

The trimeric arrangement described for the higher symmetry space group *H*3 can similarly be recognized in the crystal packing in the monoclinic space group (*C*2) with the same building blocks (Fig. 4, upper panel; PDB entry 6tsy). Although the overall architectures of both structures appear to be similar, the most obvious difference is the packing density. Compared with the high solvent content of >70% for the higher symmetry space group, the packing in space group *C*2 is much more compact. Each monomer contributes to an active dimer and composes the monohexameric structure; they present their heme cofactors towards each other and expose them to the central part of the barrel. Unlike in space group *H*3, crystal contacts are mediated by back-to-back protein interactions involving each monomer of the assembly simultaneously. The individual residues involved in putative relevant oligomerization interfaces would ideally be mutated to probe and further verify their contribution to the interface in solution.

No indication of a post-translational modification could be identified within the model electron density of PDB entry 6tsy. *PISA* analysis revealed a stable interaction between the homodimeric structures. The monomers do foster a second relevant protein interaction in both reported structures. Individual dimers also assemble an offset trimeric ring structure, which includes several hydrogen bonds and salt bridges, rendering the assembly stable. Despite the slight differences, both barrel core structures are similar in overall shape and superpose with an r.m.s.d. value of 0.5 Å.

Another known case in which the same protein was crystallized in two different space groups using very similar but not identical crystallization conditions is the structure of reduced Cyt $c_6$ from *Synechococcus* sp. PCC 7002, deposited with PDB codes 4eic and 3dr0. The former was crystallized in the monoclinic space group $P2_1$ and the latter in the hexagonal space group $P3_2$. The monomers in the monoclinic unit cell do not indicate a biologically relevant interface, and an analysis of specific interactions did not reveal a presumably stable quaternary structure. The protein chains within the high-symmetry space group are constructed along the crystallographic rotation axis and do not resemble any arrangement found in our structures.

Three molecules per asymmetric unit were also observed for recombinant Cyt $c_6$ from *Synechococcus* sp. PCC 7002 in space group *P*3 (Bialek *et al.*, 2009); however, it did not form the same homotrimeric assembly as indicated in Fig. 4. MALDI-TOF mass spectrometry further indicated small oligomers of up to trimers independent of the crystallization solution and crystal-packing forces (Fig. 1*b*), which was also verified in solution using IR-DLS (Supplementary Fig. S4) and bnPAGE (Fig. 1*d*). The hydrodynamic particle radii obtained are 3.1 $\pm$ 0.1 and 3.2 $\pm$ 0.2 nm, respectively. For an ideal globular compact protein particle with a hydrodynamic radius of 3.1 nm, the molecular weight would be estimated as 40 kDa according to Cantor & Schimmel (1980), which is similar to the molecular weight estimated via the standard proteins in bnPAGE. However, the TeCyt $c_6$ trimers observed in the crystal structures significantly deviate from a globular shape and might result in an overestimation of the trimer mass. The maximum diameter of the trimeric assembly in the asymmetric unit in space group *H*3 is approximately 7.5 nm, which may fit the determined hydrodynamic radii, taking the elongated shape and hydrate shell into account. DLS data and bnPAGE, under the conditions tested, also indicate a specific protein–protein affinity and noncovalent oligomerization in solution. These data do not show an equilibrium between different oligomeric states, which might be based on rather weak concentration-dependent interactions of the monomeric protein. Further, the relative amount of trimeric TeCyt $c_6$ detected by mass spectrometry is approximately constant when using two different amounts of protein (Supplementary Fig. S3). Despite the consistent observations of a trimer, a dependence of the oligomerization on ionic strength is possible. This could be analyzed in more detail based on the distribution of electrostatic surface potential and identity of the ions, for example using light-scattering techniques. A Cyt $c_6$ homolog from the eukaryotic alga *Chlamydomonas reinhardtii* specifically dimerized at a strongly increased ionic strength in the crystallization solution (Kerfeld *et al.*, 1995). For horse heart cytochrome *c*, oligomerization via domain swapping of the C-terminal helix has been reported, even with the formation of longer polymers (Hirota *et al.*, 2010). Both of

these examples contribute to the diversity of different *c*-type cytochrome oligomers and their surface properties in solution.

Specific protein oligomerization in nature can be beneficial by improving the protein stability, by facilitating catalytic activity or for efficient storage, amongst other reasons (Ali & Imperiali, 2005). In this context, the minimum separation distance between the heme Fe atoms in the TeCyt $c_6$ crystals is approximately 16 Å in space group *H*3 and approximately 15 Å in space group *C*2, which agree with the reported values of 10–25 Å for typical distances in electron-transfer reactions (Kadenbach, 1993; Bertini *et al.*, 2006; Breuer *et al.*, 2014). Therefore, the oligomerization, as seen in solution, might resemble previous studies on the activity of algal Cyt $c_6$ (Kerfeld *et al.*, 1995; Schnackenberg *et al.*, 1999). The oligomerization might affect the midpoint potential of the heme and could cover the respective heme site after reduction to prevent the backward reaction. There are numerous structural and biophysical studies of Cyt $c_6$ from chloroplasts isolated from different species of green algae (Kerfeld *et al.*, 1995; Schnackenberg *et al.*, 1999; Dikiy *et al.*, 2002; Howe *et al.*, 2006), but systematic comparative studies of native cytochromes isolated from cyanobacteria are somewhat lacking.

Further trimeric assemblies with potential biological relevance and hydrophobic interfaces have previously been identified, for instance for Cyt $c_6$ from *C. reinhardtii*, *Monoraphidium braunii* and the cyanobacterium *Arthrospira maxima* (Kerfeld *et al.*, 1995, 2002; Frazão *et al.*, 1995); interestingly, however, these are not superimposable with each other. It was hypothesized when comparing this obvious variety of reported nonsuperimposable crystallographic assemblies and oligomers of Cyt $c_6$ homologues in eukaryotic and prokaryotic species that the diversity in assembly may indicate the involvement of different oligomers in storage with specificity for individual species (Kerfeld *et al.*, 2002), rather than in facilitating the electron transfer directly as a small oligomer. Selected cyanobacterial Cyt $c_6$ structures are listed in Supplementary Table S2.

In recent years, a large amount of work has been performed to investigate the mechanism of PSI reduction by Cyt $c_6$ and the copper-binding plastocyanin. However, essential aspects of the respiratory and photosynthetic processes in cyanobacteria remain unknown (Pessarakli, 2016; Kłodawska *et al.*, 2020). With this report, we contribute two X-ray structures that will allow further discussion of the dimerization and trimerization of native Cyt $c_6$. Possible trimerization will be dissected further in future investigations of electron transfer and the interaction interfaces with PSI and Cyt $b_6f$. Such experiments should include time-resolved structural aspects of the redox reaction in parallel to electrochemical approaches.

## 4. Related literature

The following references are cited in the supporting information for this article: Franken *et al.* (2015), Sievers *et al.* (2011) and Sridharan *et al.* (2019).

## Acknowledgements

We thank the Sample Preparation and Characterization (SPC) Facility of EMBL Hamburg for access to their mass spectrometer. The authors further thank Stephan Niebling, EMBL Hamburg for assistance with the mass-spectrometry measurements. We acknowledge the EMBL Proteomics Core Facility for mass spectrometry. We acknowledge DESY and EMBL (Hamburg, Germany) and the facilities of PETRA III, beamlines P11 and P13. Open access funding enabled and organized by Projekt DEAL.

## References

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Ali, M. H. & Imperiali, B. (2005). *Bioorg. Med. Chem.* **13**, 5013–5020.
Banci, L., Bertini, I., De la Rosa, M. A., Koulougliotis, D., Navarro, J. A. & Walter, O. (1998). *Biochemistry*, **37**, 4831–4843.
Beissinger, M., Sticht, H., Sutter, M., Ejchart, A., Haehnel, W. & Rösch, P. (1998). *EMBO J.* **17**, 27–36.
Bertini, I., Cavallaro, G. & Rosato, A. (2006). *Chem. Rev.* **106**, 90–115.
Bialek, W., Krzywda, S., Jaskolski, M. & Szczepaniak, A. (2009). *FEBS J.* **276**, 4426–4436.
Breuer, M., Rosso, K. M. & Blumberger, J. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 611–616.
Cantor, C. R. & Schimmel, P. R. (1980). *Biophysical Chemistry*. San Francisco: W. H. Freeman.
Clarke, S. G. & Tamanoi, F. (2006). Editors. *The Enyzmes*, Vol. 24. Burlington: Academic Press.
Díaz-Moreno, I., Hulsker, R., Skubak, P., Foerster, J. M., Cavazzini, D., Finiguerra, M. G., Díaz-Quintana, A., Moreno-Beltrán, B., Rossi, G.-L., Ullmann, G. M., Pannu, N. S., De la Rosa, M. A. & Ubbink, M. (2014). *Biochim. Biophys. Acta*, **1837**, 1305–1315.
Dikiy, A., Carpentier, W., Vandenberghe, I., Borsari, M., Safarov, N., Dikaya, E., Van Beeumen, J. & Ciurli, S. (2002). *Biochemistry*, **41**, 14689–14699.
Franken, H., Mathieson, T., Childs, D., Sweetman, G. M. A., Werner, T., Tögel, I., Doce, C., Gade, S., Bantscheff, M., Drewes, G., Reinhard, F. B. M., Huber, W. & Savitski, M. M. (2015). *Nat. Protoc.* **10**, 1567–1593.
Frazão, C., Soares, C. M., Carrondo, M. A., Pohl, E., Dauter, Z., Wilson, K. S., Hervás, M., Navarro, J. A., De la Rosa, M. A. & Sheldrick, G. M. (1995). *Structure*, **3**, 1159–1169.
Hirota, S., Hattori, Y., Nagao, S., Taketa, M., Komori, H., Kamikubo, H., Wang, Z., Takahashi, I., Negi, S., Sugiura, Y., Kataoka, M. & Higuchi, Y. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 12854–12859.
Howe, C. J., Schlarb-Ridley, B. G., Wastl, J., Purton, S. & Bendall, D. S. (2006). *J. Exp. Bot.* **57**, 13–22.
Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. (2014). *IUCrJ*, **1**, 213–220.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kadenbach, B. (1993). *Angew. Chem.* **105**, 800–801.
Kerfeld, C. A., Anwar, H. P., Interrante, R., Merchant, S. & Yeates, T. O. (1995). *J. Mol. Biol.* **250**, 627–647.
Kerfeld, C. A., Sawaya, M. R., Krogmann, D. W. & Yeates, T. O. (2002). *Acta Cryst.* D**58**, 1104–1110.
Kłodawska, K., Kovács, L., Vladkova, R., Rzaska, A., Gombos, Z., Laczkó-Dobos, H. & Malec, P. (2020). *Photosynth. Res.* **143**, 251–262.
Koike, H. & Katoh, S. (1979). *Plant Cell Physiol.* **20**, 1157–1161.
Kölsch, A., Hejazi, M., Stieger, K. R., Feifel, S. C., Kern, J. F., Müh, F., Lisdat, F., Lokstein, H. & Zouni, A. (2018). *J. Biol. Chem.* **293**, 9090–9100.
Kondrashov, D. A., Zhang, W., Aranda, R. IV, Stec, B. & Phillips, G. N. Jr (2008). *Proteins*, **70**, 353–362.

Krissinel, E. & Henrick, K. (2005). *Computational Life Sciences*, edited by M. R. Berthold, R. C. Glen, K. Diederichs, O. Kohlbacher & I. Fischer, pp. 163–174. Berlin, Heidelberg: Springer.
Kumar, J., Singh, D., Tyagi, M. B. & Kumar, A. (2019). *Cyanobacteria*, pp. 327–346. London: Academic Press.
Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S. & Thornton, J. M. (2018). *Protein Sci.* **27**, 129–134.
Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S. & Dutton, P. L. (1992). *Nature*, **355**, 796–802.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Pessarakli, M. (2016). *Handbook of Photosynthesis*, 3rd ed. Boca Raton: CRC Press.
Rashin, A. A., Domagalski, M. J., Zimmermann, M. T., Minor, W., Chruszcz, M. & Jernigan, R. L. (2014). *Acta Cryst.* D**70**, 481–491.
Sawaya, M. R., Krogmann, D. W., Serag, A., Ho, K. K., Yeates, T. O. & Kerfeld, C. A. (2001). *Biochemistry*, **40**, 9215–9225.
Schnackenberg, J., Than, M. E., Mann, K., Wiegand, G., Huber, R. & Reuter, W. (1999). *J. Mol. Biol.* **290**, 1019–1030.
Shin, M., Sakihama, N., Koike, H. & Inoue, Y. (1984). *Plant Cell Physiol.* **25**, 1575–1578.
Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
Sparta, K. M., Krug, M., Heinemann, U., Mueller, U. & Weiss, M. S. (2016). *J. Appl. Cryst.* **49**, 1085–1092.
Sridharan, S., Kurzawa, N., Werner, T., Günthner, I., Helm, D., Huber, W., Bantscheff, M. & Savitski, M. M. (2019). *Nat. Commun.* **10**, 1155.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Whitelegge, J. P., Zabrouskov, V., Halgand, F., Souda, P., Bassilian, S., Yan, W., Wolinsky, L., Loo, J. A., Wong, D. T. W. & Faull, K. F. (2007). *Int. J. Mass Spectrom.* **268**, 190–197.
Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (2018). *Protein Sci.* **27**, 293–315.
Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 235–242.
Worrall, J. A. R., Schlarb-Ridley, B. G., Reda, T., Marcaida, M. J., Moorlen, R. J., Wastl, J., Hirst, J., Bendall, D. S., Luisi, B. F. & Howe, C. J. (2007). *J. Am. Chem. Soc.* **129**, 9468–9475.
Xin, F. & Radivojac, P. (2012). *Bioinformatics*, **28**, 2905–2913.