STRUCTURAL BIOLOGY COMMUNICATIONS

Volume 76 (2020)

Supporting information for article:

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

Sven Falke, Christian Feiler, Henry Chapman and Iosifina Sarrou

S1. Materials and methods for protein identification via LC-MS/MS

Coomassie-stained SDS-PAGE protein bands were excised and the molecules were reduced by the addition of 200 µl of a 10 mM DTT solution in 100 mM ammonium bicarbonate and incubation at 56°C for 30 min. Proteins were alkylated by the addition of 200 µl of a 55 mM chloroacetamide (CAA) in 100 mM AmBiC and incubation for 20 min. Trypsin dissolved in trypsin resuspension buffer (0.1 µg/µl; Promega, USA) was diluted with ice-cold 50 mM ammonium bicarbonate buffer to achieve a final concentration of 1 ng/µl and then incubated with the gel pieces over night at 37°C. Gel pieces were sonicated, centrifuged and washed with 50% ACN and 1% formic acid. The supernatant was dried in a speedvac and reconstituted in 0.1% (v/v) formic acid.

Peptides were analyzed by LC-MS/MS on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scentific, USA) as previously described (Sridharan *et al.*, 2019). Precursors were isolated using the quadrupole with a window of 1.2 m/z and fragmentation was triggered by HCD in fixed collision energy mode with fixed collision energy of 30%. MS2 spectra were acquired in ion trap normal mode. Acquired data were analyzed using IsobarQuant (Franken *et al.*, 2015) and Mascot V2.4 (Matrix Science) using a reverse UniProt FASTA *T. elongates* database (UP000000440) including common contaminants. A minimum of two unique peptides with a peptide length of at least seven amino acids and a false discovery rate below 0.01 were required on the peptide and protein level.

Table S1 Identification of the TeCyt c_6 (gene name: petJ1) via LC-MS and related values.

protein ID	P0934_IS2|P0A3X9 - isoform: P0934_IS2	P0934_IS2|P0A3X9 - isoform: P0A3X9
Description	P0A3X9petJ	CYC6_THEEB Cytochrome c6
MW	11761.67	11761.67
Top3	6.368601	6.368601
Ssm	56	56
Usm	9	9
Upm	9	9
Max_score	120	120
Total score	839	839
Sequence coverage [%]	51.3	51.3

MKKRFISVCA IAIALLVSLT PAALAADLAN GAKVFSGNCA ACHMGGGNVV MANKTLKKEA

LEQFGMYSED AIIYQVQHGK NAMPAFAGRL TDEQIQDVAA YVLDQAAKGW AG

Figure S1 Identification of the purified target protein as TeCyt c_6 via LC-MS. Residues indicated in green are covered and were identified (Uniprot code: P0A3X9).

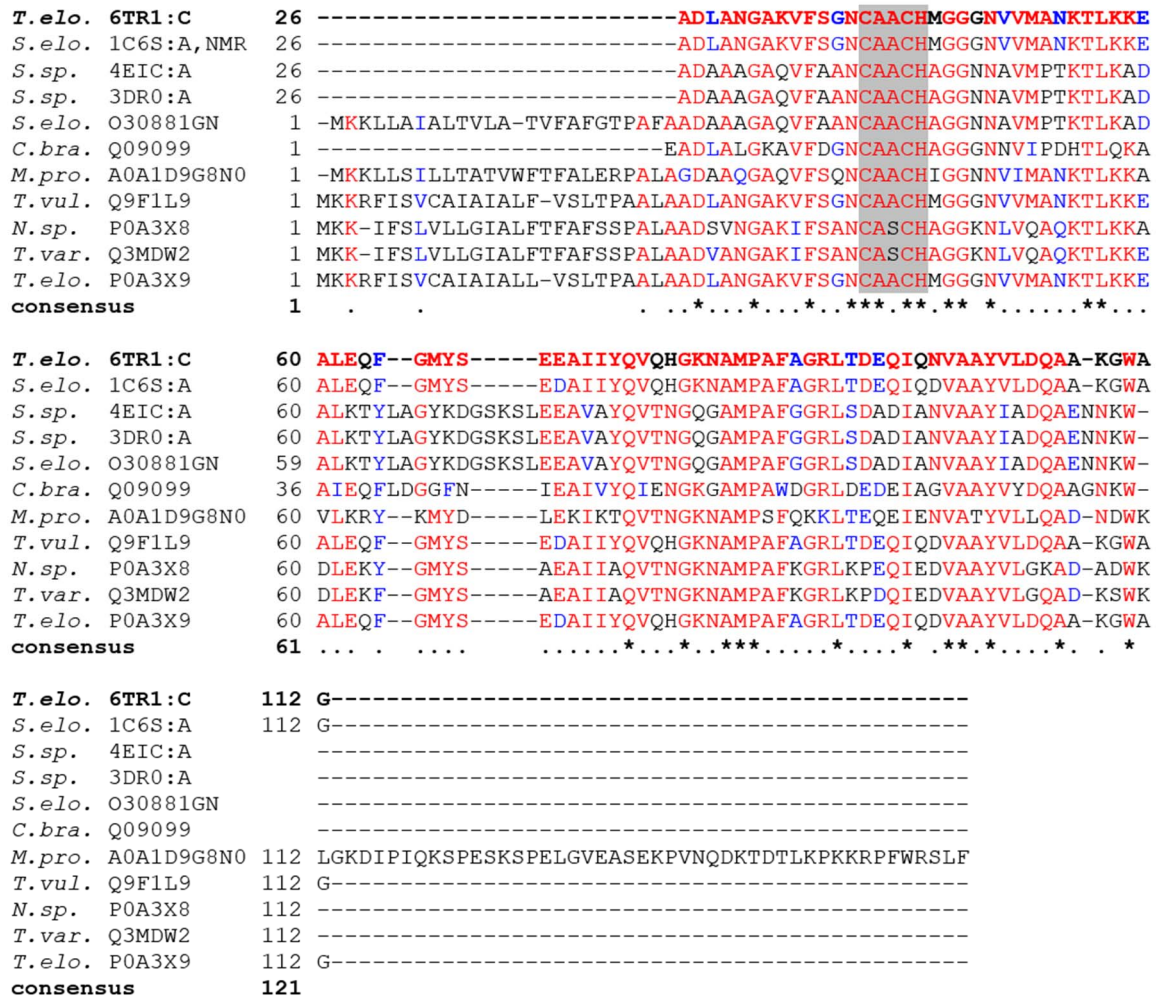

Figure S2 Sequence alignment (Clustal Omega; Sievers *et al.*, 2011) of TeCyt c_6 with related cytochromes from cyanobacteria and eukaryotic algae. The N-terminal signal peptide is considered for the numbering of the amino acid positions, i.e. 25 amino acids for TeCyt c_6, which are missing in the crystal structures (Uniprot code: P0A3X9). The respective UniProt or if available PDB accession code is indicated; the conserved heme binding motif CXXCH is highlighted in grey.

Table S2 List of selected closely related c-type cytochromes in the protein data bank, their crystal geometry and sequence identities.

Protein	**PDB code**	Space group	Unit cell: a,b,c [Å]; α,β,γ, [°]	Cell volume [10^6 Å^3]	Molecules per ASU	Seq. identity with PDB 6TR1 [%]
TeCyt c_6	**6TR1**	H3	94.8, 94.8, 160.22; 90, 90, 120	1.24699	3	100
TeCyt c_6	**6TSY**	C2	106.02, 109.91, 55.4; 90, 100.89, 90	0.63393	6	98
Synechococcus sp. (BP-1) Cyt c_6	**1C6S** (NMR)	-	-	-	-	98
Synechococcus sp. (PCC 7002) Cyt c_6	**4EIC**	$P2_1$	31.86, 27.69, 44.07; 90, 101.1, 90	0.038151	1	58
Synechococcus sp. (PCC 7002) Cyt c_6	**3DR0**	$P3_2$	82.88, 82.88, 28.28; 90, 90, 120	0.16823	3	58
Synechococcus sp. (WH8102) Cyt c_6B	**4KMG**	C2	106.11, 28.98, 24.68; 90, 92.3, 90	0.07583	1	35
Nostoc sp. Cyt c_6	**4GYD**	$P2_12_12$	77.72, 79.8, 80.15; 90, 90, 90	0.49709	6	66
Phormidium laminosum Cyt c_6	**2V08**	$P6_3$	57.36, 57.36, 89.55; 90,90,120	0.25515	2	75
Arthrospira maxima Cyt c_6	**1KIB**	$I4_132$	237, 237, 237; 90, 90, 90	13.31205	8	47

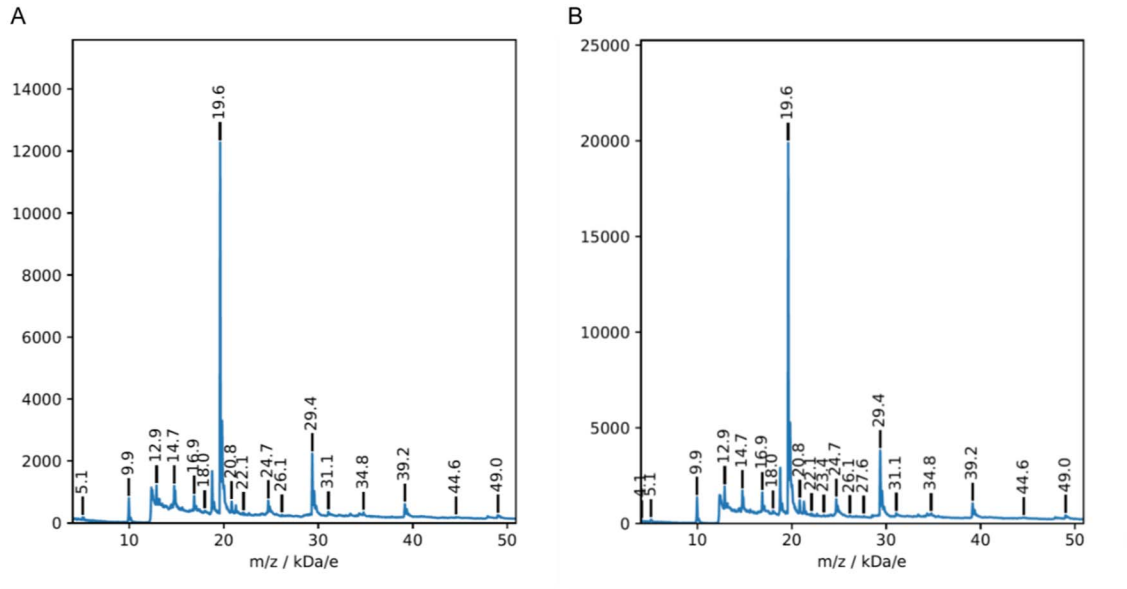

Figure S1 MALDI-TOF spectra obtained for 4 pmol (**A**) and 6 pmol (**B**) of TeCyt c_6.

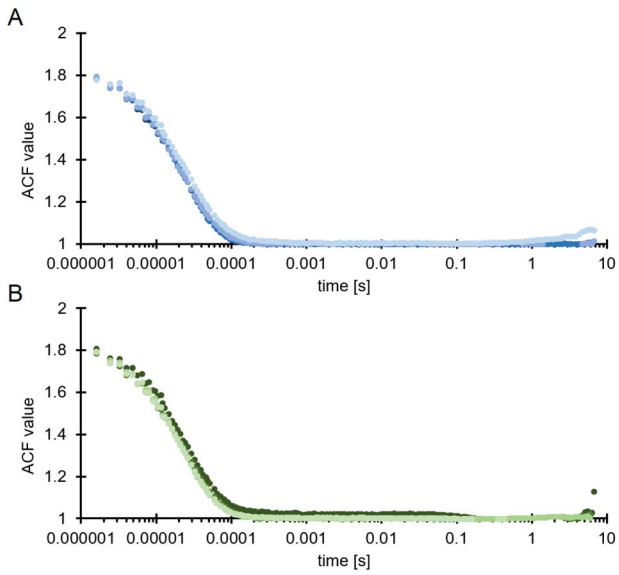

Figure S2 Individual auto correlation functions (ACFs) of the performed DLS experiments using protein concentrations of 4 mg ml^{-1} (**A**) 10 mg ml^{-1} (**B**) in purification buffer. Four ACFs were recorded over 10 s each for both sample solutions, averaged and fitted to determine an averaged hydrodynamic radius of 3.1 ± 0.1 nm (**A**) and 3.2 ± 0.2 nm (**B**) respectively.

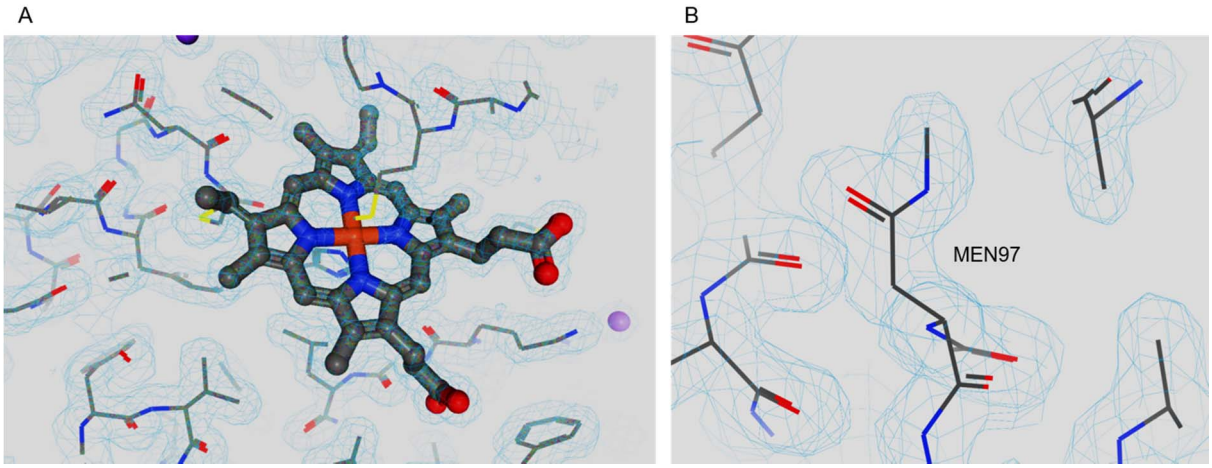

Figure S3 2Fo-Fc electron density map of the heme co-factor (**A**) and *N*-methyl asparagine (MEN; **B**) contoured at 1σ level as incorporated in the structure of TeCyt c_6 (chain C, PDB code 6TR1).

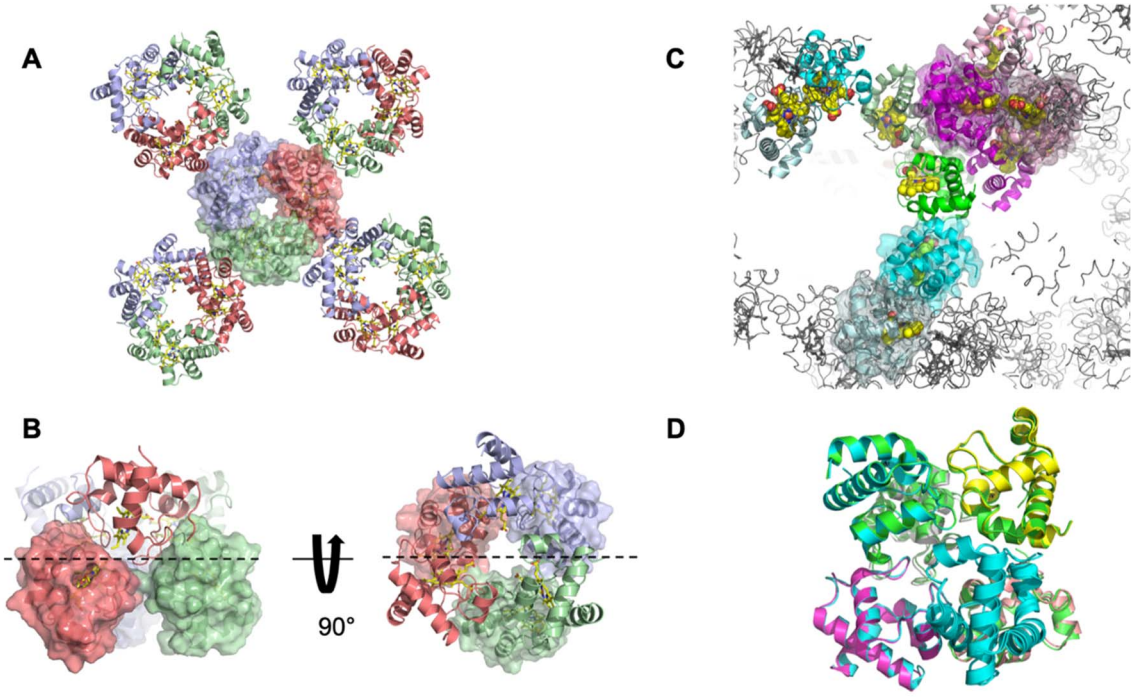

Figure S4 Crystal packing analysis. (**A**) Crystal packing of the C2 crystal form. The content of the ASU is shown in surface representation, each protein chain is coloured individually. (**B**) Content of the ASU. The six protein chains form a barrel-shaped hexameric structure with an open central pore. The individual homodimeric building blocks are color-coded with one monomer shown in cartoon- and one in surface representation. The individual co-factors are coloured in yellow. (**C**) Crystal packing of the H3 space group. The three chains represented within the ASU are stained in brilliant shade and the respective symmetry related dimer partner in the same pale colour. The surface representation indicates the chains involved in the assembly of the core structure. The green chain mediates crystal contacts between the individual barrel structures. The respective symmetry neighbouring molecules, spanning the trimeric ring structure are stained in pale. (**D**) Both barrel-shaped structures are superimposed. The structure 6TSY is stained in green and blue while the chains of the structure 6TR1 are individually coloured.