Current limits of structural biology: the transient interaction between cytochrome c_6 and photosystem I

Authors:

A. Kölsch[1], C. Radon[2], M. Golub[3], A. Baumert[2], J. Bürger[4,5], T. Mielke[4], F. Lisdat[6], A. Feoktystov[7], J. Pieper[3], A. Zouni[1], P. Wendler[2]

[1] Department of Biology, Humboldt–Universität zu Berlin, Philippstrasse 13, 10115 Berlin, Germany.

[2] Institute of Biochemistry and Biology, Department of Biochemistry, University of Potsdam, Karl-Liebknecht Strasse 24-25, 14476 Potsdam-Golm, Germany.

[3] Institute of Physics, University of Tartu, Wilhelm Ostwaldi 1, 50411 Tartu, Estonia

[4] Max-Planck Institute of Molecular Genetics, Ihnestrasse 63-73, 14195 Berlin, Germany.

[5] Charité, Institut für Medizinische Physik und Biophysik, Charitéplatz 1, 10117 Berlin

[6] Institute of Applied Life Sciences, Technical University of Applied Sciences Wildau, Hochschulring 1, 15745 Wildau, Germany.

[7] Forschungszentrum Jülich GmbH, Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Lichtenbergstr. 1, 85748 Garching, Germany

Corresponding authors: petra.wendler@uni-potsdam.de, athina.zouni@hu-berlin.de

Abstract

Trimeric photosystem I from the cyanobacterium *Thermosynechococcus elongatus* (*Te*PSI) is an intrinsic membrane protein, which converts solar energy into electrical energy by oxidizing the soluble redox mediator cytochrome c_6 (Cyt c_6) and reducing ferredoxin. Here, we use cryo-electron microscopy and small angle neutron scattering (SANS) to characterize the transient binding of Cyt c_6 to *Te*PSI. The structure of *Te*PSI cross-linked to Cyt c_6 was solved at a resolution of 2.9 Å and shows additional cofactors as well as side chain density for 84% of the peptide chain of subunit PsaK, revealing a hydrophobic, membrane intrinsic loop that enables binding of associated proteins. Due to the poor binding specificity, Cyt c_6 could not be localized with certainty in our cryo-EM analysis. SANS measurements confirm that Cyt c_6 does not bind to *Te*PSI at protein concentrations comparable to those for cross-linking. However, SANS data indicate a complex formation between *Te*PSI and the non-native mitochondrial cytochrome from horse heart (Cyt c_{HH}). Our study pinpoints the difficulty of identifying very small binding partners (less than 5% of the overall size) in EM structures when binding affinities are poor. We relate our results to well resolved co-structures with known binding affinities and recommend confirmatory methods for complexes with K_M values higher than 20 µM.

Keywords

Photosystem I, cryo EM, *Thermosynechococcus elongatus,* small angle neutron scattering, cytochrome c_6, electron transfer, photo-biotechnology

Introduction

Thylakoid membranes of the thermophilic cyanobacterium *Thermosynechococcus elongatus* harbor four major protein complexes: photosystem II (PSII), the cytochrome b_6f complex (Cyt b_6f), photosystem I (PSI) and the ATP synthase. PSII performs the light driven oxidation of water to oxygen, and the reduction of plastoquinone. The Cyt b_6f-complex transfers one electron from plastoquinol to cytochrome c_6 (Cyt c_6), which delivers it to the primary electron donor of PSI, P_{700}. P_{700} performs its charge separation after light induction and transfers the electron via an electron transport chain to the terminal iron sulfur cluster (F_B) from where it is taken to Ferredoxin (Fd). Fd then transfers the electron to Ferredoxin-NADP$^+$-Reductase (FNR) to generate NADPH. Unlike in higher plants, the electron transfer between Cyt b_6f and PSI in *T. elongatus* is only mediated by Cyt c_6, which needs to exhibit sufficient promiscuity to have functional interactions with PSI as well as Cyt b_6f. On the other hand, promiscuity can lead to low binding affinities. Indeed, the binding affinity between PSI and Cyt c_6 is characterized by K_M values of 30 µM to 300 µM [1], most likely translating into a very transient interaction when the electron is offloaded to PSI.

PSI is of great biotechnological interest, as it can be incorporated into an electrochemical system to yield a stand-alone biohybrid photoelectrochemical cell that converts light energy into electrical energy [2–5]. Besides its high quantum yield PSI is also remarkably stable and thus, can show long term operation in such systems. However, the performance of such devices is limited by several factors – an important one is the electron transport between the electrode and PSI by mediator molecules. Here even small redox proteins can be used. In order to design electrode surfaces that efficiently interact with PSI, a structural characterization of the protein-protein binding interface between Cyt c_6 and PSI is needed.

The high-resolution structure of *Te*PSI was solved to 2.5 Å in 2001 by cryo-X-ray diffraction (XRD) [6]. PSI forms a trimer with a total mass of more than 1 MDa. The crystal structure reveals 12 subunits per monomer; three subunits are located on the cytoplasmic side and nine subunits are membrane intrinsic. PsaA and PsaB are the two largest subunits, which contain

most of the cofactors. In total, TePSI contains 96 chlorophylls, two phylloquinones, three 4Fe4S clusters, 22 carotenoids, four lipids and one Ca^{2+} ion per monomer. The electron transport chain, consisting of P_{700} and the acceptors A_0, A_1, as well as the three 4Fe4S clusters F_X, F_A, and F_B, is located in PsaA, PsaB and PsaC. The membrane-extrinsic subunits PsaC, PsaD and PsaE form the Fd binding site [7]. Recently, the location of the cytoplasmic binding site of Fd was solved at 4.2 Å resolution by co-crystallization of Fd with TePSI [8]. The affinity of Fd to PSI ranges from 50 to 500 nM at physiological pH values [9]. Multiple attempts have been undertaken to characterize the transient interaction between PSI and Cyt c_6 on the luminal side of the complex. Co-crystallization of PSI with Cyt c_6 has remained unsuccessful [1]. The binding affinity between Cyt c_6 and PSI has been determined to 30 µM in Tricine buffer and 200 mM $MgSO_4$, pH 8. Rigid body docking combined with electrostatic calculations resulted in a model of the likely active binding site [1]. Yet, rigid body docking also suggests that there is more than one Cyt c_6 binding site located close to P_{700}. Intriguingly, a cryo-EM analysis of PSI from *Chlamydomonas reinhardtii* unintendingly resolved residual bound Pc as a weak, low-resolution density connected to a patch of lysines near PsaF [10]. This binding site does not fully agree with the model obtained from electrostatic calculations and rigid body docking of Cyt c_6 to TePSI.

In this work, we set out to structurally characterize the interaction between PSI and Cyt c_6 using single particle cryo-EM. A strength of this method is that a dataset can be sorted into dominant and auxiliary populations, thereby enabling low affinity binding partners or less populated conformations of a protein complex to be visualized. This approach is routinely applied in 2- dimensional and 3-dimentionsal image processing procedures [11–14] using focused refinements or classifications. However, large size differences between the binding partners and very small ligand sizes can render sorting procedures impossible, because the signal strength is too low to be distinguished from background signal. Partial signal subtraction from experimental images is designed to overcome this problem and enables signal separation at secondary structure level [15]. So far, distinct orientations of a 30 kDa domain within a complex of up to 1 MDa could successfully be differentiated [15,16]. Undeniably,

structural and conformational heterogeneity of a protein complex depends on the binding affinities between transient binding partners. Several studies and databases aim at correlating structural features with affinity data [17] (https://www.bindingdb.org/, http://bindingmoad.org/, http://www.pdbbind.org.cn/). Despite this comprehensive data pool predictions on whether a binding event can structurally be trapped remain difficult. Often, a high resolution cryo-EM map has to be obtained in order to determine whether a small (less than 5% of the overall weight) binding partner or ligand is bound productively to the complex.

Here, we present the cryo-EM analysis of 1MDa *Te*PSI crosslinked with 10 kDa Cyt c_6, its only native electron donor. Despite 3D reconstructions being solved to 3.2 Å for the trimeric complex without applied symmetry, 3.0 Å for the complex with applied trimeric symmetry and 2.9 Å for the symmetry expanded and masked out monomer, Cyt c_6 could not be localized unequivocally in our cryo-EM structures. We compare the obtained structures with the *Te*PSI crystal structure and describe the density distribution in 3D classifications. In addition, we use SANS measurements to characterize the complex formation of PSI with Cyt c_6 and c_{HH} under similar condition as the cross-linking experiments. Finally, we relate these results to binding affinities of protein-protein complexes with known structures and deduce a general guidance for structural analysis of transiently formed protein complexes.

Results

Cryo-EM structure of the PSI Cyt c_6 protein complex

Cyt c_6 has a molecular weight of 10 kDa and binds to PSI with low affinity (K_M=30 µM), as described in [1]. High concentrations of free Cyt c_6, which are necessary for high saturation of the PSI binding site, lead to a crowding of the small proteins on the EM grid. Consequently, for structural analysis, Cyt c_6 was covalently crosslinked to PSI and only crosslinked PSI complexes were included to the dataset. Crosslinking of the charged side chains was done using the carboxylic acid activator EDC (a carbodiimide). Subsequently, proteins were purified via the 6x His-tag on Cyt c_6 and Ni-NTA chromatography. PSI, which was not crosslinked to

Cyt c_6, eluted in the flow through, while crosslinked PSI eluted 0.5 column volumes prior to Cyt c_6 (Figure S1a). Free Cyt c_6 remaining in the solution was separated from crosslinked particles in a centrifugal concentrator (100000 MWCO). When the protein complex was subjected to cryo-EM, the crosslinked particles appear as trimeric PSI complexes in the 2D class averages (Figure S1 b, c). The trimeric PSI was solved to 3.2 Å without applied symmetry (in the following, referred to as C1 map) and 3.0 Å with applied C3 symmetry (C3 map, Figure 1 and Figure S1).

All 12 subunits are resolved to a local resolution of at least 3.6 Å in both maps (Figure S1 f), and in the C1 map the three monomers are highly symmetric. When comparing the three monomers in the C1 map, only loops at the surface show differences in the backbone with more than 1 Å Angstrom displacement. The root mean square deviation between the three monomers is 0.26 Å (Figure S2). Although the local resolution at the luminal side of PSI is 2.9-3.2 Å, no density corresponding to the crosslinked Cyt c_6 could be identified in either of the two maps.

Since the stoichiometry of the crosslinked cytochrome to the trimeric TePSI is unknown, symmetry expansion, partial signal subtraction, and 3D classification were used to identify monomer particle subsets with extra density that could correspond to Cyt c_6 [11,18] (Figure S1d). The symmetry expanded data set, which consists of images of each monomer of every PSI trimer, yielded a map of 2.9 Å resolution. This map was subjected to 3D classifications with various masks encompassing potential Cyt c_6 binding sites on TePSI (Figure 2). In focused 3D classifications with a mask encompassing either the entire stromal or luminal side, no classes with extra density were observed (data not shown). 3D classifications using differently positioned masks revealed numerous classes with extra density on the luminal as well as cytoplasmic side. The by far most populated classes (>88% of the data) in all 3D classifications show only TePSI density. Thus, despite crosslinking no discernable Cyt c_6 density was found in the reconstructions.

However, many low occupancy classes (3-6% of the dataset) display extra density at various contact points with the photosystem. Some of the classes appear on the stromal side of PSI, which would not get in contact with Cyt c_6 in the cell. The most populated classes appear above the interface between PsaA and PsaB luminal to P700, near a patch of charged residues in PsaF (Figure 2A and B). 3D refinement of these classes did not yield reconstructions resolved to a higher resolution than seen in the focussed 3D classifications. Due to the low resolution, no structural details on the binding partners and interface can be deduced. In a recent structure of PSI and light harvestings complex I from the green algae, *Chlamydomonas reinhardtii*, Pc density was found at the corresponding location on PsaF, despite no effort to crosslink or even purify Pc with the complex [10]. A contact point above PsaA and PsaB was predicted by docking [1].

Model of the PsaK subunit

In the *Te*PSI crystal structure, all subunits except PsaK were well resolved. Also, in the recently published Cryo-EM PSI structure from *Fischerella thermalis* PCC 7521 resolved at 3.19 Å resolution, all 11 subunits could be identified except for the PsaK subunit [19]. PsaK is a small, transmembrane subunit, which is located radially outwards of the PSI trimer, in the interface between two monomers. It participates in state transition and binding of additional chlorophyll containing proteins, such as the light harvesting complex II (LHCII) in eukarya and the iron stress-induced protein A (IsiA) in cyanobacteria [20,21]. Aside from the first four N-terminal residues, PsaK was resolved at 3.2 Å resolution in our monomeric EM structure. In contrast to the *Te*PSI crystal structure [6], in which two truncated transmembrane helices could be modeled as a poly-alanine backbone, the EM structure shows side chain density for 84 % of the residues of PsaK. Loop 32-55 of PsaK is highly flexible and no side chains are visible for residues 44 to 50. The latter residues are highly hydrophobic (sequence: PIALPAL, further referred to as PIALPAL-patch) and extend into the membrane volume (Figure 3A). The first residue after the PIALPAL-patch that is outside the membrane plane is Glu52. Glu52 stabilizes the patch by forming a salt bridge with Arg31.

Differences in Cofactor occupancy of PSI

Comparing the cofactors of the electron transfer chain of the monomeric crystal structure and the EM derived models, no difference in localization or spacing is present (Figure S3A). In the symmetry expanded monomer, 97 chlorophylls, two phylloquinones, three 4Fe4S clusters, 24 carotenoids, six lipids and one Ca^{2+} ion were assigned to the EM map. Five of these cofactors are not present in the crystal structure model of *T. elongatus* (Figure S3B). Out of the five cofactors, two lipids (Dipalmitoyl-Phosphatidyl-Glycerol, LHG) and two β-carotenes are also present in the crystal structure of the cyanobacterium *Synechocystis* sp. (pdb code 5oy0, [22]). The last cofactor is a chlorophyll, which is bound to PsaF by π-stacking with the conserved Tyr 72 at 4 Å distance (Figure S3C, D).

Small-angle neutron scattering measurements with *Te*PSI

In order to investigate the binding of Cyt c_6 to *Te*PSI under similar conditions as the cross-linking experiments, we recorded SANS data of the complex in a water-based solution at physiological temperature (Figure 4A,B). We started the analysis with a check of aggregation in the PSI sample solution. All measured curves have a classical Guinier behavior at small q values and are thus free of aggregation (not shown). So far, the findings resemble our earlier results concerning the solution structure of *Te*PSI [23]. However, the distance distribution function (Figure 4C) calculated from the SANS data reveals that *Te*PSI is slightly expanded in solution at physiological temperature. The radius of gyration (Rg) is 75.2 Å, which is larger than the Rg of 68.2 Å and 68.3 Å calculated from the crystal structure [6] and the cryo-EM structure, respectively.

In the next step we added Cyt c_6 to the sample. Previous experiments have shown that Cyt c from horse heart (Cyt c_{HH}) can also deliver an electron to *Te*PSI. Therefore, we probed Cyt c_{HH} as well as endogenous Cyt c_6 in our SANS experiments. We performed the measurements under conditions at which maximum affinity could be achieved, namely a K_M of 10 μM for the binding between *Te*PSI and Cyt c_{HH} and a K_M of 30 μM for *Te*PSI and Cyt c_6 [1,6,23]. The shift

of the SANS curve towards lower scattering vectors q upon addition of Cyt c_{HH} as well as the concomitant change in the P(r) function (Figure 4B, C) suggest binding under the conditions tested. This finding is similar to the results for another photosynthetic complex by Tiede and colleagues [24], who reported the binding of Cyt c_2 to the bacterial reaction center (bRC) (Table S1). A structure reconstitution of the TePSI-Cyt c_{HH} complex based on the SANS data and a simulated structure of the TePSI-Cyt c_{HH} complex as modelled in [1] is shown in Figure 4E. The fit function corresponding to the reconstituted structure of the TePSI-Cyt c_{HH} complex is shown in Figure 4B (lower full line) and compared to a fit of the TePSI data only (upper full line). Both fits reproduce the experimental data reasonably well. In contrast, there is no shift of the SANS curve towards lower scattering vectors q in the case of endogenous Cyt c_6 (Figure 4D). Therefore, we conclude, that the transient interaction between PSI and Cyt c_6 was not retained during the SANS experiment.

Discussion

Differences between EM and crystal structure

Five cofactors were identified in the EM map, which are not present in the TePSI crystal structures. These cofactors are located at the outer edges of the monomers, which are more flexible than the rigid core regions. The lipids and carotenoids are also present in the crystal structure of Synechocystis sp. with only slight re-orientations. One planar density was assigned as chlorophyll CLA1304 and is not present in any known crystallographic or EM structure of PSI. Due to the low local resolution in the corresponding volume, the identity of this cofactor cannot be solved with certainty. The putative cofactor is parallel to Tyr72 in 4 Å distance indicating a π-stacking. The tyrosine arene is located central to the porphyrin ring, comparable to the interaction of Phe70 with CLA1229 from PsaB. Tyr72 is highly conserved with the sidechain orientation being identical in all known PSI models (from cyanobacteria to plant). As this is the only binding surface for the porphyrin ring, the chlorophyll can only be weakly bound, which is reflected in its weak density. Functionally, CLA1304 may be advantageous for excitation transfer from neighboring proteins to PSI. In Synechocystis sp. it

was found, that the excitation energy transfer from the antenna protein IsiA to PSI occurs via one of three chlorophylls (CLA 8,14 and 17) [21]. CLA1304 from *T. elongatus* is in a hypothetical Mg-Mg distance of 12.8 Å from one of these three chlorophylls (CLA 8), enabling an additional excitation energy transfer path from IsiA to PSI (Fig S4). This additional path from IsiA to PsaF may be necessary as *T. elongatus* lacks the chlorophyll CLA1302 in PsaF, which participates in the excitation energy transfer in *Synechocystis* sp.

Several PsaK structures from eukaryotic algae, cyanobacteria and from plant are already known [22,25,26]. We solved the structure of PsaK from *T. elongatus* using cryo-EM. The cytoplasmic loop, which connects the two transmembrane helices of PsaK, strongly diverges between different species. The loop structures of two different PsaK subunits from cyanobacterium *Synechocystis* sp. have been solved by crystallography and cryo-EM [21,22]. Although they vary in sequence, they adopt similar conformations. In contrast to *Synechocystis* sp., *T. elongatus* contains only one *psaK* gene and the PsaK PIALPAL-patch contains only hydrophobic residues. Interestingly, it adopts a different conformation to the corresponding sections in *Synechocystis* sp., which also contain hydrophilic residues. The conformation of *T. elongatus* PsaK is stabilized by a salt bridge between Glu52 and Arg31. The same salt bridge is formed in PsaK and PsaG of eukaryotic PSI [20,26], while the corresponding charged residues are not present in any of the *psaK* genes of *Synechocystis* (*psaK1*, see [26] or *psaK2*, see [21]). In *Synechocystis* sp., this patch binds to the poorly conserved C-terminus of IsiA [21]. IsiA forms a ring around PSI, consisting of 18 subunits and 306 chlorophylls. The protein-backbone structure of *isiA* is highly conserved between cyanobacterial species [21,27], and given an identical orientation of the IsiA-protein in *T. elongatus*, the PIALPAL-patch of *T. elongatus* could bind to IsiA as well.

More likely, the PIALPAL-patch can stabilize PSI-PSI interactions. Atomic force microscopy (AFM) analysis of thylakoid membranes from *T. elongatus* revealed that PSI-trimers are in hexagonal packing, while the membrane from *Synechocystis* is less ordered [28]. One of the interaction sites in native thylakoid membranes is found between PsaK of one PSI trimer and PsaJ and PsaF of the neighboring PSI trimer (Figure 4B). Here, the PIALPAL-patch can bind

beneath the cytoplasmic helices of PsaF and PsaJ. Lys108 from PsaF does not form a salt bridge with neighboring residues and could interact with Glu52 from PsaK to further strengthen the binding. Therefore, PsaK from *T. elongatus* might have adapted to promote the binding of both, IsiA as well as neighboring PSI proteins.

Cyt c_6 binding to *Te*PSI

Electron transfer from Cyt c_6 to PSI presumably occurs at a hydrophobic site close to P_{700}. The localization and orientation of Cyt c_6 in the PSI binding site was postulated based on rigid body docking analysis combined with electrostatic calculations [1]. However, the computational analysis showed additional Cyt c_6-docking sites with similar binding energy at the center of the luminal surface of PSI. In most cyanobacteria, Cyt c_6 can be interchanged against Pc. The reaction with multiple partners often prevents the optimization of the binding site [29], and indeed, the interaction between cyanobacterial PSI and its electron donor does not show a high specificity [30]. It rather follows a collisional mechanism [31], while a transient complex is formed by eukaryotic PSI. In eukarya, Pc is first bound to PsaF as determined by flash absorption spectroscopy [32], then relocates to interact with the hydrophobic surface close to P_{700} after illumination [33]. Recently, density originating from Pc could be identified at PsaF in the EM map of PSI from *Chlamydomonas reinhardtii* [10]. Although Pc is located at its rather high affinity PsaF binding site, the additional density does not show secondary structure elements and its size is merely ~1/10th of the Pc volume. Focused 3D classification of Pc did not improve the local resolution, likely due to its flexibility.

Despite crosslinking Cyt c_6 to *Te*PSI and purifying the complex via the 6x-his tag on Cyt c_6, putative Cyt c_6 density was only found in a small portion of the dataset. Ultimately, putative Cyt c_6 density was only uncovered through masked classification of a focused refinement of symmetry expanded particles. The resolution and definition of the putative Cyt c_6 density is poor and does not even allow for rigid body docking of the Cyt c_6 NMR structure. We conclude that the effect of Cyt c_6 being crosslinked to *Te*PSI in a rather unspecific manner outweighs a potential structural flexibility during binding. Obviously, electron transfer from Cyt c_6 to *Te*PSI occurs with low specificity and low binding affinity. We already determined that Cyt c_6 binds

*Te*PSI with a Km of 30 µM to 300 µM depending on ionic strength and pH, but that it does not co-crystallize with PSI [1]. The SANS measurements in this study do not detect complex formation either. Apparently, the interaction is too transient to capture it with cross-linking and cryo-EM, SANS or X-ray crystallography under all conditions tested.

PSI and Cyt c_{HH} on the other hand bind with a K_M of 5 µM to 45 µM. The higher binding affinity between the proteins is reflected by complex formation during our SANS measurements and their ability to co-crystallize [1], albeit Cyt c_{HH} is not visible in the electron density of the crystal. Despite the slightly better binding parameters, a complex between *Te*PSI and its artificial binding partner Cyt c_{HH} has not been resolved to high resolution yet. A detailed description of the interaction is still missing. It has to be mentioned that Cyt c_{HH} enables electron transfer to *Te*PSI, but much less efficient as the native Cyt c_6, which has a more than 5 times higher k_{cat} for this reaction [1].

Structural data on transient interactions

Protein structure analysis of transient protein-protein interactions adds important information about the character of interacting residues and the shape of the interaction surface. However, it is not trivial to trap these interactions structurally. We have analyzed the transient interaction between different cytochromes and PSI from cyanobacterium *T. elongatus* with SANS, cryo-X-ray crystallography [1] as well as cryo-EM. Weak densities at different positions of the PSI EM structure indicate some binding of Cyt c_6, the only electron donor of *Te*PSI *in vitro*. The SANS experiments only show binding of an artificial electron donor, Cyt c_{HH}, to PSI at low resolution. Albeit EM image processing procedures allow for sophisticated sorting of different populations in the dataset, they reach a limit where binding affinities are too low to support stable complex formation. Given that we crosslinked Cyt c_6 to PSI before investigating the cryo-EM structure, it is unlikely that crosslinking coupled to mass spectrometry (XLMS) would lead to a more precise description of the binding interface.

A database search of 2594 successfully solved structures of protein-protein complexes deposited to pdbbind showed that only 8% of the data display K_D values greater than 20 µM. Only 6 structures in the dataset derive from EM experiments with K_D ranging from 2.9 nM to

13.4 µM. Despite a wealth of binding databases for small molecules, it is challenging to match EM structures of protein-protein complexes with the appropriate affinity values. Table 1 shows a manual compilation of affinity data for some protein-protein structures. However, it remains difficult to predict whether a small, low affinity binding partner can be visualized in EM reconstructions. While crystallographic approaches can use high ligand concentrations in soaking experiments, small proteins at high concentrations add to background noise in EM experiments. Both, EM and crystallization conditions can thus differ significantly from the conditions used for affinity measurements. Further problems when correlating binding data with the chance of success of structure determination are the relatively small sample size of structures with recorded binding data, a lack of standardization of affinity measurements, and the error prone manual annotation of the data. The three protein-protein complexes deposited with the worst K_D of 48 mM to 212 mM in the dataset were reported with very different K_D values of 48 nM to 1.6 µM in the publications cited. Lastly, depending on the method used for assessing the binding affinity drastically diverging K_D values for the same interaction can be obtained. Deriving the kinetic data from reduction of P_{700}^+ *in vitro* leads to higher K_D values than measuring the interaction directly, for instance by using isothermal titration calorimetry [32](Table S1). Overall, none of the databases at our disposal allow to extract information that is reliable enough to predict the likelihood of trapping a low affinity interaction structurally by just knowing the kinetic details of the complex. We recommend to approach the structural elucidation of a small (<5% of the molecular weight) binding partner of a protein complex by implementing the following check list:

1. Optimisation of binding parameters: find optimal pH, temperature, buffer. Chances of success increase with K_M values of 20 µM or lower.

2. Survey binding specificity using cross-linking and mass spectrometry approaches.

3. Structure determination using cryo-EM, X-ray crystallography or NMR.

Conclusions

Our results show that Cyt c_6 bound PSI has been resistant to structural elucidation with the methods available. Conclusions on the cyanobacterial Cyt c_6 binding site on PSI might be obtained by comparison with binding of Cyt c_{HH}, by investigating mutant complexes, or by switching to a system from another organism that shows better K_M values. All of these approaches require a thorough optimization of the binding parameters, including solution conditions of the particular structural experiment. A broader systematic study that correlates binding affinities of small proteins to protein complexes of interest would be highly desirable in order to gauge chances of success to structurally resolve the interaction. Unfortunately, only very few successfully solved structures report reliable binding affinities, so that gathering these data is cumbersome and error prone.

Materials and Methods

Isolation of proteins

*Te*PSI and Cyt c_6 were isolated as described previously [1]. PSI originated from *T. elongatus* BP-1 and was stored in form of crystals at 4 °C until use [1]. Cyt c_6 was expressed heterologously in *Escherichia coli*, containing a C-terminal 6xHis-tag. The concentration of Cyt c_6 was measured in the presence of 5 mM ascorbic acid at 553 nm with ε_{553}=25 mM^{-1}*cm^{-1} [33] and the concentration of *Te*PSI reaction centers was measured with ε_{680}=5.5 µM^{-1}cm^{-1} [34].

Crosslinking

10 µM P$_{700}$ was crosslinked with 100 µM Cyt c_6 at 20 °C for two hours in the dark. The buffer contained 5 mM sodium ascorbate, 5 mM 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC), 10 mM N-hydroxysuccinimide (NHS), 100 mM NaCl, 0.02% β-dodecyl-maltoside (DDM), 25 mM Tricine pH 8.0. The reaction was quenched by addition of 100 mM ammonium acetate. The sample was diluted tenfold with binding buffer (500 mM NaCl, 0.02% DDM, 20 mM imidazole, 20 mM KH$_2$PO$_4$, adjusted to pH 7.5 with NaOH) and applied to a 1 ml Ni-NTA

column (HisTrap™ HP, GE healthcare). *Te*PSI and Cyt c_6 were eluted by a gradient of imidazole to 250 mM. An exemplary run is shown in Figure S1A. No significant loss of binding capacity due to the ascorbic acid was observed. The green fractions were pooled and the buffer was exchanged against 25 mM Tricine, pH 8.0, 25 mM NaCl, 0.02% DDM in a centrifugal concentrator (Vivaspin 2, Sartorius, 100000 MWCO). The large pore size was selected so that free Cyt c_6 (11 kDa) was washed out together with the buffer.

EM data collection

A 3.5 µL sample of 150 µg/ml *Te*PSI crosslinked with Cyt c_6 was applied to a glow discharged Quantifoil 3/3 300 mesh grid and incubated for 45 seconds before being blotted and plunged into liquid ethane using a Vitrobot plunge-freezing device. Cryo-EM images were acquired at 300 kV on a Tecnai G2 Polara equipped with a Gatan K2 summit direct electron detector. A total of 5176 micrographs were collected at a nominal magnification of 31000 with a defocus ranging from -0.69 to -2.6 µm and sampled at 0.628 Å per pixel. 25 frames were recorded over 5 seconds, giving an accumulated dose of 32 e⁻/Å^2.

EM image processing

All image processing was performed in RELION-3.0 [35], Contrast transfer functions were determined by CTFFIND4 [36] and movie frame alignment and dose-weighting of 25 frames was performed using MotionCor2 [37]. Micrographs with strong astigmatism, at overfocus or very low defocus, or with strong ice contamination were discarded, yielding 3929 images for further processing. 514 manually selected particles were subjected to 2D classification and one class average was used as a template for automated particle selection, resulting in an initial dataset of 511,337 particles. The dataset was cleaned using two successive rounds of 2D classification within RELION, resulting in a dataset of 196,882 particles (Figure S1D). An initial model was generated in RELION before conducting a 3D refinement using the cleaned dataset. The resulting 3D model was used as a reference for 3D classification into five classes. The only reasonable class of 175999 particles was subjected to 3D refinement

yielding reconstructions at 3.26 Å and 3.06 Å without symmetry and with C3 symmetry applied, respectively. The resolution of the maps was improved to 3.17 Å (without symmetry) and 2.98 Å (with symmetry) by both Bayesian polishing or per-particle CTF refinement, however no further improvement occurred when both were performed subsequently.

Since Cyt c_6 could not be identified in either the reconstruction with or without applied symmetry, symmetry expansion and partial signal subtraction were combined, so that Cyt c_6 could be searched for within each individual protomer [11,35,38–40]. The particles were first expanded according to C3 symmetry. Using Chimera [41], a map of two protomers was generated and used to perform signal subtraction on the symmetry expanded particle dataset using the symmetrized map as a reference. Subsequently, the stack of subtracted images was used for 3D structure refinement to 2.88 Å (map3D). Bayesian polishing did not improve the resolution of map3D.

Map3D was then subjected to numerous masked 3D classifications into four classes without alignment, with T values ranging from 4-40. Parameter T is related to the relative weight of the experimental data; it is documented that masked 3D classifications can benefit from higher T-values, such as 10-40 [11]. Soft masks of different shapes and sizes were focused on various areas on both the cytoplasmic and luminal sides of the PS core. All 3D classifications were run until less than 0.41% of the images in the dataset changed assignment of optimal classes between iterations (rlnChangesOptimalClasses ≤ 0.41).

EM model refinement

The 2.5 Å resolution crystal structure of *T. elongatus* (pdb code 1jb0) was aligned with the monomeric EM reconstruction in UCSF chimera [41]. Manual adjustment was performed with COOT [42] followed by real_space_refine from the phenix program suite version 1.15 [43]. COOT and real_space_refine were repeated three times. The refined monomer model was aligned with the trimeric C3- and C1-maps and adjusted with three rounds of COOT followed by phenix.real_space_refine. The data collection and model statistics are summarized in Table S2. Images were generated in UCSF chimera or PyMOL [44].

Sample preparation for small angle scattering

SANS measurements of *Te*PSI with Cyt c_6 were performed in 25 mM Tricine pH 8, 200 mM MgSO$_4$ and 0.02 % DDM. Interaction of *Te*PSI with Cyt c_{HH} was measured in 25 mM Tricin pH 8, 25 mM NaCl and 0.02 % DDM. Buffers were prepared in both, 99.9 % D$_2$O (Sigma Aldrich) and in H$_2$O and were mixed to achieve 5 % D$_2$O, the match point of the hydrophobic chains of the detergent molecules, so that the SANS data reflect the protein contribution only. For 99.9 % D$_2$O, the pD was adjusted from a measured, apparent pH with the correction factor of 0.4. Buffer exchange for *Te*PSI was achieved by washing crystals 3 times with 0.02 % DDM and 5 % D$_2$O and dissolving the crystals in the final buffer solution. For Cyt c_6, buffer was exchanged by washing the protein 3 times in a centrifugal concentrator (Vivaspin 6, 5000 MWCO, Sartorius). Cyt c_{HH}-powder (Sigma Aldrich, 99 % purity) was dissolved in the final buffer solution. The proteins were adjusted to final concentrations of 3 mM Chlorophyll (*Te*PSI), 130 µM (Cyt c_{HH}) and 300 µM (Cyt c_6).

SANS data collection

The SANS experiments were carried out at the KWS-1 small-angle diffractometer (JCNS at MLZ Garching, Germany) [45]. The neutron wavelength used in the present experiment was 5 Å. The Q-range probed was 0.006 to 0.45 Å^{-1} by using two sample-to-detector distances of 8 and 20 m, respectively. The samples were dissolved in buffer solutions containing 5% D$_2$O, i.e. the match point of the hydrophobic chains of the detergent molecules, so that the SANS data reflect the protein contribution only. All samples were kept in standard 1 mm Hellma cells at a constant temperature of 15 °C. The data reduction procedure includes the correction according to detector sensitivity mask. The data treatment was carried out using the QtiKWS program (Pipich, V. Qtikws. www.qtikws.de; accessed Oct 3, 2017).

SANS data treatment

The scattering intensity of an ideally diluted solution of monodisperse particles is given by the master equation [46]:

$$I(q) = n\Delta\rho^2 V^2 P(q)S(q) ,\qquad\qquad (1)$$

where q is the scattering vector, n is the number particles, $\Delta\rho$ is the difference in scattering length density (SLD) between the particles and the solvent, and V is the volume of the particles. P(q) is the form factor, which is a function of the averaged shape and the averaged size of the scattering particles. The effective structure factor is given in the formula as S(q).

The SANS data of a solution of monodisperse particles are directly related to the distance distribution function P(r) by Inverse Fourier transform (IFT) [47]. Using the software package GNOM [47], the P(r) function and the particle maximum dimension Dmax can be determined from the SANS data based on the IFT method. For the IFT analysis, we used the limited q-range up to 0.1 Å 1.

The software routine CRYSON [48] are used to calculate theoretical SANS curves based on pdb structures of proteins and compare them with the experimental data. In turn, the overall shape of a protein complex is obtained from the SANS data using the ATSAS reconstitution tool based on a reverse Monte Carlo minimization approach developed by the group of D. Svergun [49]. All reconstituted structures were averaged over 20 iterations. For each iteration it was taken into account that trimeric *Te*PSI possess an oblate form and a P3 symmetry. We also used the Pymol program [44] to compare the result of the ATSAS analysis with the known pdb structure of the PSII core complex.

Accession numbers

The EM maps of the *Te*PSI Monomer, *Te*PSI C3 symmetry trimer and *Te*PSI C1 symmetry trimer are deposited under accession codes EMD-10557, EMD-10558 and EMD-10559, respectively. Atomic coordinates of *Te*PSI Monomer, *Te*PSI C3 symmetry trimer and *Te*PSI

C1 symmetry trimer have been deposited in the Protein Data Bank under accession codes 6TRA, 6TRC and 6TRD, respectively. Other data are available from the corresponding author upon reasonable request.

Acknowledgments

Author Contributions

A.K. established and performed the enzyme purification. A.K., F.L. and A.Z. designed the cross-linking study. A.K. performed the cross linking. EM specimen preparation and data-collection were carried out by C.R., P.W., J.B., and T.M.. EM data analysis, interpretation, and presentation were carried out by A.B., C.R., A.K., and P.W. SANS data were recorded and analyzed by M.G., A.F. and J.P.. Project conception and overall supervision was provided by A.Z. and P.W.. P.W. and A.K. wrote the manuscript, with contributions from all other authors.

References

[1]   A. Kölsch, M. Hejazi, K.R. Stieger, S.C. Feifel, J.F. Kern, F. Müh, F. Lisdat, H. Lokstein, A. Zouni, Insights into the binding behavior of native and non-native cytochromes to

Photosystem I from Thermosynechococcus elongatus, J. Biol. Chem. (2018).

https://doi.org/10.1074/jbc.RA117.000953.

[2]   P.N. Ciesielski, F.M. Hijazi, A.M. Scott, C.J. Faulkner, L. Beard, K. Emmett, S.J.

Rosenthal, D. Cliffel, G.K. Jennings, Photosystem I–Based biohybrid

photoelectrochemical cells, Bioresour. Technol. 101 (2010) 3047–3053.

[3]   F. Zhao, S. Hardt, V. Hartmann, H. Zhang, M.M. Nowaczyk, M. Rögner, N. Plumeré, W.

Schuhmann, F. Conzuelo, Light-induced formation of partially reduced oxygen species

limits the lifetime of photosystem 1-based biocathodes, Nat. Commun. 9 (2018) 1–9.

[4]   K.R. Stieger, S.C. Feifel, H. Lokstein, M. Hejazi, A. Zouni, F. Lisdat, Biohybrid

architectures for efficient light-to-current conversion based on photosystem I within

scalable 3D mesoporous electrodes, J. Mater. Chem. A. 4 (2016) 17009–17017.

[5]   V.M. Friebe, R.N. Frese, Photosynthetic reaction center-based biophotovoltaics, Curr.

Opin. Electrochem. 5 (2017) 126–134.

[6]   P. Jordan, P. Fromme, O. Klukas, H.T. Witt, W. Saenger, N. Krauss, X-Ray

Crystallographic Structure Analysis of Cyanobacterial Photosystem I at 2.5 A

Resolution, Sci. Access. 3 (2001).

[7]   P. Sétif, N. Fischer, B. Lagoutte, H. Bottin, J.-D. Rochaix, The ferredoxin docking site of

photosystem I, Biochim. Biophys. Acta (BBA)-Bioenergetics. 1555 (2002) 204–209.

[8]   H. Kubota-Kawai, R. Mutoh, K. Shinmura, P. Sétif, M.M. Nowaczyk, M. Rögner, T.

Ikegami, H. Tanaka, G. Kurisu, X-ray structure of an asymmetrical trimeric ferredoxin–

photosystem I complex, Nat. Plants. 4 (2018) 218–224. https://doi.org/10.1038/s41477-

018-0130-0.

[9]   P. Barth, B. Lagoutte, P. Sétif, Ferredoxin reduction by photosystem I from

Synechocystis sp. PCC 6803: toward an understanding of the respective roles of

subunits PsaD and PsaE in ferredoxin binding, Biochemistry. 37 (1998) 16233–16241.

[10]   M. Suga, S.-I. Ozawa, K. Yoshida-Motomura, F. Akita, N. Miyazaki, Y. Takahashi,

Structure of the green algal photosystem I supercomplex with a decameric light-

harvesting complex I, Nat. Plants. 5 (2019) 626–636. https://doi.org/10.1038/s41477-

019-0438-4.

[11]   S.H.W. Scheres, Processing of Structurally Heterogeneous Cryo-EM Data in RELION, 1st ed., Elsevier Inc., 2016. https://doi.org/10.1016/bs.mie.2016.04.012.

[12]   S.H.W.W. Scheres, Classification of structural heterogeneity by maximum-likelihood methods, in: Methods Enzymol., Elsevier, 2010: pp. 295–320. https://doi.org/10.1016/S0076-6879(10)82012-9.

[13]   D. Lyumkis, A.F. Brilot, D.L. Theobald, N. Grigorieff, Likelihood-based classification of cryo-EM images using FREALIGN, J. Struct. Biol. 183 (2013) 377–388.

[14]   P.A. Penczek, J. Frank, C.M.T. Spahn, A method of focused classification, based on the bootstrap 3D variance analysis, and its application to EF-G-dependent translocation, J. Struct. Biol. 154 (2006) 184–194.

[15]   X.C. Bai, E. Rajendra, G. Yang, Y. Shi, S.H.W. Scheres, Sampling the conformational space of the catalytic subunit of human g-secretase, Elife. 4 (2015) 1–19. https://doi.org/10.7554/eLife.11182.

[16]   T.H.D. Nguyen, W.P. Galej, X. Bai, C.G. Savva, A.J. Newman, S.H.W. Scheres, K. Nagai, The architecture of the spliceosomal U4/U6.U5 tri-snRNP, Nature. 523 (2015) 47–52. https://doi.org/10.1038/nature14548.

[17]   P.L. Kastritis, I.H. Moal, H. Hwang, Z. Weng, P.A. Bates, A.M.J.J. Bonvin, J. Janin, A structure-based benchmark for protein–protein binding affinity, Protein Sci. 20 (2011) 482–491. https://doi.org/10.1002/pro.580.

[18]   J.T. Huiskonen, Image processing for cryogenic transmission electron microscopy of symmetry-mismatched complexes, Biosci. Rep. 38 (2018). https://doi.org/10.1042/BSR20170203.

[19]   C. Gisriel, G. Shen, V. Kurashov, M.-Y. Ho, S. Zhang, D. Williams, J.H. Golbeck, P. Fromme, D.A. Bryant, The structure of Photosystem I acclimated to far-red light illuminates an ecologically important acclimation process in photosynthesis, Sci. Adv. 6 (2020) eaay6415. https://doi.org/10.1126/sciadv.aay6415.

[20]   X. Pan, J. Ma, X. Su, P. Cao, W. Chang, Z. Liu, X. Zhang, M. Li, Structure of the maize

photosystem I supercomplex with light-harvesting complexes I and II, Science (80-.). 360 (2018) 1109 LP – 1113. https://doi.org/10.1126/science.aat1156.

[21] H. Toporik, J. Li, D. Williams, P.-L. Chiu, Y. Mazor, The structure of the stress-induced photosystem I–IsiA antenna supercomplex, Nat. Struct. Mol. Biol. 26 (2019) 443–449. https://doi.org/10.1038/s41594-019-0228-8.

[22] T. Malavath, I. Caspy, S.Y. Netzer-El, D. Klaiman, N. Nelson, Structure and function of wild-type and subunit-depleted photosystem I in Synechocystis, Biochim. Biophys. Acta - Bioenerg. (2018). https://doi.org/10.1016/j.bbabio.2018.02.002.

[23] M. Golub, M. Hejazi, A. Kölsch, H. Lokstein, D.C.F. Wieland, A. Zouni, J. Pieper, Solution structure of monomeric and trimeric photosystem I of Thermosynechococcus elongatus investigated by small-angle X-ray scattering, Photosynth. Res. 133 (2017) 163–173.

[24] D.M. Tiede, K. Littrell, P.A. Marone, R. Zhang, P. Thiyagarajan, Solution structure of a biological bimolecular electron transfer complex: characterization of the photosynthetic reaction center-cytochrome c2 protein complex by small angle neutron scattering, J. Appl. Crystallogr. 33 (2000) 560–564. https://doi.org/10.1107/S0021889800099891.

[25] X. Su, J. Ma, X. Pan, X. Zhao, W. Chang, Z. Liu, X. Zhang, M. Li, Antenna arrangement and energy transfer pathways of a green algal photosystem-I–LHCI supercomplex, Nat. Plants. 5 (2019) 273–281. https://doi.org/10.1038/s41477-019-0380-5.

[26] Y. Mazor, A. Borovikova, I. Caspy, N. Nelson, Structure of the plant photosystem I supercomplex at 2.6 Å resolution, Nat. Plants. 3 (2017) 17014.

[27] P. Cao, D. Cao, L. Si, X. Su, L. Tian, W. Chang, Z. Liu, X. Zhang, M. Li, Structural basis for energy and electron transfer of the photosystem I–IsiA–flavodoxin supercomplex, Nat. Plants. 6 (2020) 167–176. https://doi.org/10.1038/s41477-020-0593-7.

[28] C. MacGregor-Chatwin, M. Sener, S.F.H. Barnett, A. Hitchcock, M.C. Barnhart-Dailey, K. Maghlaoui, J. Barber, J.A. Timlin, K. Schulten, C.N. Hunter, Lateral segregation of photosystem I in cyanobacterial thylakoids, Plant Cell. (2017). https://doi.org/10.1105/tpc.17.00071.

[29] P.B. Crowley, M. Ubbink, Close encounters of the transient kind: protein interactions in the photosynthetic redox chain investigated by NMR spectroscopy, Acc. Chem. Res. 36 (2003) 723–730.

[30] M. Hervás, A. Díaz-Quintana, C.A. Kerfeld, D.W. Krogmann, A. Miguel, J.A. Navarro, Cyanobacterial photosystem I lacks specificity in its interaction with cytochrome c 6 electron donors, Photosynth. Res. 83 (2005) 329–333.

[31] H. Hatanaka, K. Sonoike, M. Hirano, S. Katoh, Small subunits of Photosystem I reaction center complexes from Synechococcus elongatus. I. Is the psaF gene product required for oxidation of cytochrome c-553?, BBA - Bioenerg. 1141 (1993) 45–51. https://doi.org/10.1016/0005-2728(93)90187-K.

[32] M. Hippler, F. Drepper, W. Haehnel, J.-D.D. Rochaix, The N-terminal domain of PsaF: precise recognition site for binding and fast electron transfer from cytochrome c6 and plastocyanin to photosystem I of Chlamydomonas reinhardtii, Proc. Natl. Acad. Sci. 95 (1998) 7339–7344. https://doi.org/10.1073/pnas.95.13.7339.

[33] M. Hervás, M. De La Rosa, G. Tollin, A comparative laser-flash absorption spectroscopy study of algal plastocyanin and cytochrome c552 photooxidation by photosystem I particles from spinach, Eur. J. Biochem. 203 (1992) 115–120. https://doi.org/10.1111/j.1432-1033.1992.tb19835.x.

[34] F. Müh, A. Zouni, Extinction coefficients and critical solubilisation concentrations of photosystems I and II from Thermosynechococcus elongatus, Biochim. Biophys. Acta (BBA)-Bioenergetics. 1708 (2005) 219–228.

[35] J. Zivanov, T. Nakane, B.O. Forsberg, D. Kimanius, W.J.H. Hagen, E. Lindahl, S.H.W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3, Elife. 7 (2018) e42166. https://doi.org/10.7554/eLife.42166.

[36] A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs, J. Struct. Biol. 192 (2015) 216–221. https://doi.org/10.1016/J.JSB.2015.08.008.

[37] S.Q. Zheng, E. Palovcak, J.-P. Armache, K.A. Verba, Y. Cheng, D.A. Agard,

MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy, Nat. Methods. 14 (2017) 331–332. https://doi.org/10.1038/nmeth.4193.

[38] S.L. Ilca, A. Kotecha, X. Sun, M.M. Poranen, D.I. Stuart, J.T. Huiskonen, Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes, Nat. Commun. (2015). https://doi.org/10.1038/ncomms9843.

[39] D.O. Passos, M. Li, R. Yang, S. V Rebensburg, R. Ghirlando, Y. Jeon, N. Shkriabai, M. Kvaratskhelia, R. Craigie, D. Lyumkis, Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome, Science. 355 (2017) 89–92. https://doi.org/10.1126/science.aah5163.

[40] M. Zhou, Y. Li, Q. Hu, X.-C.C. Bai, W. Huang, C. Yan, S.H.W.W. Scheres, Y. Shi, Atomic structure of the apoptosome: mechanism of cytochrome c- and dATP-mediated activation of Apaf-1, Genes Dev. 29 (2015) 2349–2361. https://doi.org/10.1101/gad.272278.115.

[41] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612. https://doi.org/10.1002/jcc.20084.

[42] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2010) 486–501. https://doi.org/10.1107/S0907444910007493.

[43] P.D. Adams, P. V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2010) 213–221. https://doi.org/10.1107/S0907444909052925.

[44] Schrodinger LLC, The PyMOL Molecular Graphics System, Version 1.8, (2015).

[45] A. V Feoktystov, H. Frielinghaus, Z. Di, S. Jaksch, V. Pipich, M.-S. Appavou, E. Babcock, R. Hanslik, R. Engels, G. Kemmerling, KWS-1 high-resolution small-angle

neutron scattering instrument at JCNS: current state, J. Appl. Crystallogr. 48 (2015) 61–70.

[46]   D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503. https://doi.org/10.1107/S0021889892001663.

[47]   P. V Konarev, V. V Volkov, A. V Sokolova, M.H.J. Koch, D.I. Svergun, PRIMUS: a Windows PC-based system for small-angle scattering data analysis, J. Appl. Crystallogr. 36 (2003) 1277–1282. https://doi.org/10.1107/S0021889803012779.

[48]   D. Svergun, C. Barberato, M.H.J. Koch, CRYSOL– a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates, J. Appl. Crystallogr. 28 (1995) 768–773. https://doi.org/10.1107/S0021889895007047.

[49]   M. V Petoukhov, D. Franke, A. V Shkumatov, G. Tria, A.G. Kikhney, M. Gajda, C. Gorba, H.D.T. Mertens, P. V Konarev, D.I. Svergun, New developments in the ATSAS program package for small-angle scattering data analysis, J. Appl. Crystallogr. 45 (2012) 342–350.

[50]   F. Sommer, F. Drepper, M. Hippler, The luminal helix l of PsaB is essential for recognition of plastocyanin or cytochrome c6 and fast electron transfer to photosystem I in Chlamydomonas reinhardtii, J. Biol. Chem. 277 (2002) 6573–6581. https://doi.org/10.1074/jbc.M110633200.

[51]   C.C. Moser, P.L. Dutton, Cytochrome c and c2 binding dynamics and electron transfer with photosynthetic reaction center protein and other integral membrane redox proteins, Biochemistry. 27 (1988) 2450–2461. https://doi.org/10.1021/bi00407a031.

[52]   H.L. Axelrod, E.C. Abresch, M.Y. Okamura, A.P. Yeh, D.C. Rees, G. Feher, X-ray Structure Determination of the Cytochrome c2: Reaction Center Electron Transfer Complex from Rhodobacter sphaeroides, J. Mol. Biol. 319 (2002) 501–515. https://doi.org/https://doi.org/10.1016/S0022-2836(02)00168-7.

[53]   J.E. Erman, L.B. Vitello, Yeast cytochrome c peroxidase: mechanistic studies via protein engineering, Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol. 1597 (2002)

193–220. https://doi.org/https://doi.org/10.1016/S0167-4838(02)00317-5.

[54] H. Pelletier, J. Kraut, Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c, Science (80-.). 258 (1992) 1748 LP – 1755. https://doi.org/10.1126/science.1334573.

[55] X. Pan, D. Cao, F. Xie, F. Xu, X. Su, H. Mi, X. Zhang, M. Li, Structural basis for electron transport mechanism of complex I-like photosynthetic NAD(P)H dehydrogenase, Nat. Commun. 11 (2020) 610. https://doi.org/10.1038/s41467-020-14456-0.

[56] X. Qiu, Y. Lei, P. Yang, Q. Gao, N. Wang, L. Cao, S. Yuan, X. Huang, Y. Deng, W. Ma, T. Ding, F. Zhang, X. Wu, J. Hu, S.-L. Liu, C. Qin, X. Wang, Z. Xu, Z. Rao, Structural basis for neutralization of Japanese encephalitis virus by two potent therapeutic antibodies, Nat. Microbiol. 3 (2018) 287–294. https://doi.org/10.1038/s41564-017-0099-x.

[57] A. Kedrov, S. Wickles, A.H. Crevenna, E.O. van der Sluis, R. Buschauer, O. Berninghausen, D.C. Lamb, R. Beckmann, Structural Dynamics of the YidC:Ribosome Complex during Membrane Protein Biogenesis, Cell Rep. 17 (2016) 2943–2954. https://doi.org/10.1016/j.celrep.2016.11.059.

[58] J.Y. Kang, P.D.B. Olinares, J. Chen, E.A. Campbell, A. Mustaev, B.T. Chait, M.E. Gottesman, S.A. Darst, Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex, Elife. 6 (2017) e25478. https://doi.org/10.7554/eLife.25478.

[59] H. Wang, X. Yan, H. Aigner, A. Bracher, N.D. Nguyen, W.Y. Hee, B.M. Long, G.D. Price, F.U. Hartl, M. Hayer-Hartl, Rubisco condensate formation by CcmM in β-carboxysome biogenesis, Nature. 566 (2019) 131–135. https://doi.org/10.1038/s41586-019-0880-5.

[60] O. Mueller-Cajar, M. Stotz, P. Wendler, F.U. Hartl, A. Bracher, M. Hayer-Hartl, Structure and function of the AAA+ protein CbbX, a red-type Rubisco activase, Nature. 479 (2011) 194–199.

[61] J.Y. Bhat, G. Miličić, G. Thieulin-Pardo, A. Bracher, A. Maxwell, S. Ciniawsky, O.

Mueller-Cajar, J.R. Engen, F.U. Hartl, P. Wendler, M. Hayer-Hartl, Mechanism of Enzyme Repair by the AAA+ Chaperone Rubisco Activase, Mol. Cell. 67 (2017) 744-756.e6. https://doi.org/10.1016/j.molcel.2017.07.004.

[62]   S. Jo, T. Kim, V.G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM, J. Comput. Chem. 29 (2008) 1859–1865.

Tables

**Table 1** | Binding affinities of protein-protein complexes with associated structural data.

large Protein	small protein	organism	affinity (method)	structure
PSI	Flavodoxin	*Synechococcus elongatus* PCC 7942	0.02 µM (SPR) [27]	EM, 6KIF 3.3 Å [27]*
PSI	Ferredoxin	*Thermosynechococcus elongatus* BP-1	0.8 µM (FAS) [8]**	XRD, 5ZF0 4.2 Å [8]**
PSI	Plastocyanin	*Chlamydomonas reinhardtii*	83 µM (enzymatic, FAS) [50]	EM, EMDB-9853, not resolved [10]
PSI	Cyt c_6	*Chlamydomonas reinhardtii*	116 µM (enzymatic, FAS) [50]	no
bacterial reaction center	Cyt c_2	*Rhodobacter sphaeroides*	1 µM (FAS) [51]	XRD, 1L9B 2.4 Å, 1L9J 3.2 Å [52]
bacterial reaction center	Cyt c_2	*Rhodobacter sphaeroides*	1 µM (FAS) [51]	SANS [24]
Cytochrome *c* Peroxidase	Cyt *c*	*Yeast*	0.01 µM – 1 µM (ITC) [53]	XRD, 2PCC 2.3 Å [54]
Cytochrome *c* Peroxidase	Cyt *c*	*Yeast/Horse*	0.1 µM (absorption change) [54]	XRD, 2PCB 2.8 Å [54]
NDH-1	Ferredoxin	*Thermosynechococcus elongatus*	1 µM (SPR) [55]	EM, 6KHI 3.2 Å [55], 6L7O 3.2 Å [not yet published]
Japanese encephalitis virus	2F2 Fab fragment	*Mus musculus*	6.5 nM (SPR) [56]	EM, 5YWO 4.7 Å [56]
Japanese encephalitis virus	2H4 Fab fragment	*Mus musculus*	2.9 nM (SPR) [56]	EM, 5YWP 4.6 Å [56]
70S Ribosome	YidC insertase	*Escherichia coli*	0.2 µM (FCS) [57]	EM, 5M5H 4.5 Å [57]
RNA polymerase	Transcription termination factor	*Escherichia coli*	0.31 µM (enzymatic) [58]	EM, 6ALG* 3.7 Å [58]
Rubisco	M35 reduced	*Synechococcus elongatus* PCC 7942	0.2 µM (turbidity assay) [59]	EM, 6HBC 2.78 Å [59]
Rubisco	Rca (Rubisco activase)	*R. sphaeroides*	3.1 µM (enzymatic) [60]	EM, EMD-3701* not resolved 7.56 Å [61]

* Cross-linked complex; ** affinity measurement with wt Ferredoxin, structure with Gallium Ferredoxin; SPR: surface plasmon resonance; FAS: flash absorption spectroscopy; ITC: isothermal titration calorimetry; FCS: fluorescence correlation spectroscopy
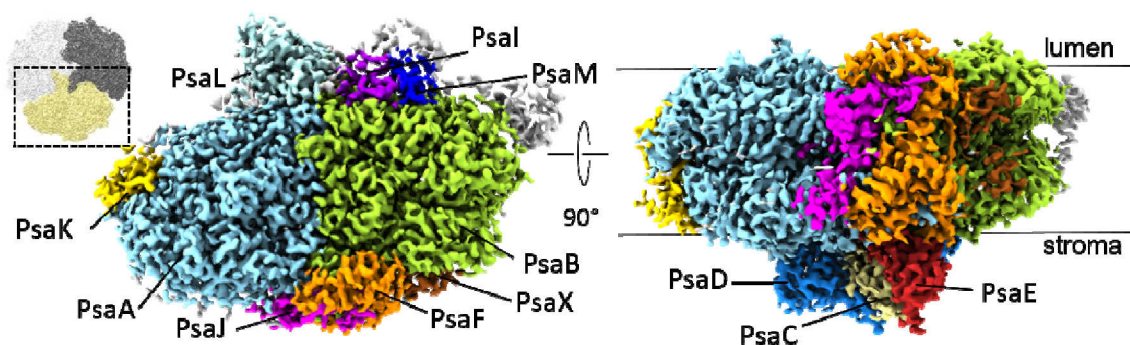
Figure 1 | Cryo-EM structure of the monomeric *Te*PSI crosslinked to Cyt c_6 derived from the symmetry expanded dataset shown from the luminal side (left) and from along the membrane plane (right). The inset on the far left shows the position of the monomer in the PSI trimer as seen from the luminal side.
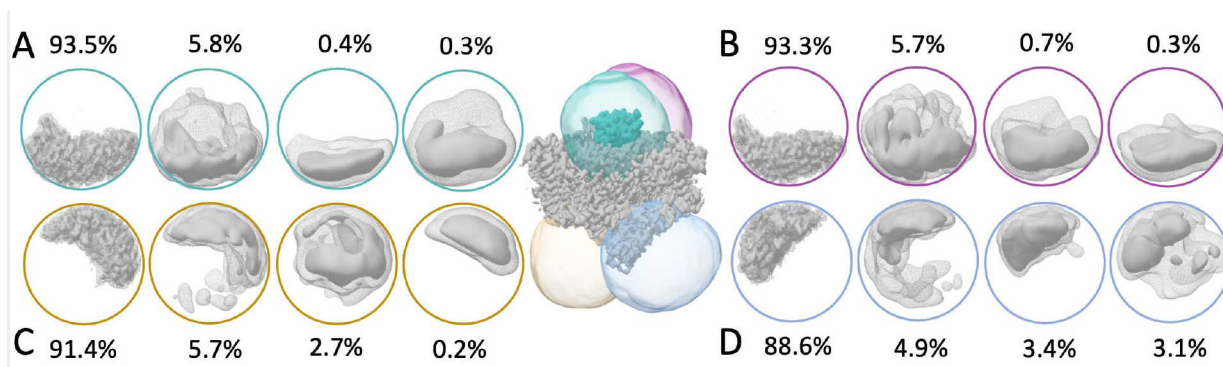

Figure 2 | Focused 3D classifications on the PSI monomer map using differently placed and sized soft masks. The model in the centre depicts monomeric PSI bound to Cyt c_6 with mask locations on luminal (**A**, teal mask and **B**, magenta mask) and stroma sides (**C**, gold mask and **D**, light blue mask). The densities derived from 3D classifications of the symmetry expanded monomer dataset are shown colour coded by mask used and contoured at 5 sigma (mesh) and 20 sigma (surface). The class occupancy for each class of each focussed classification is given.
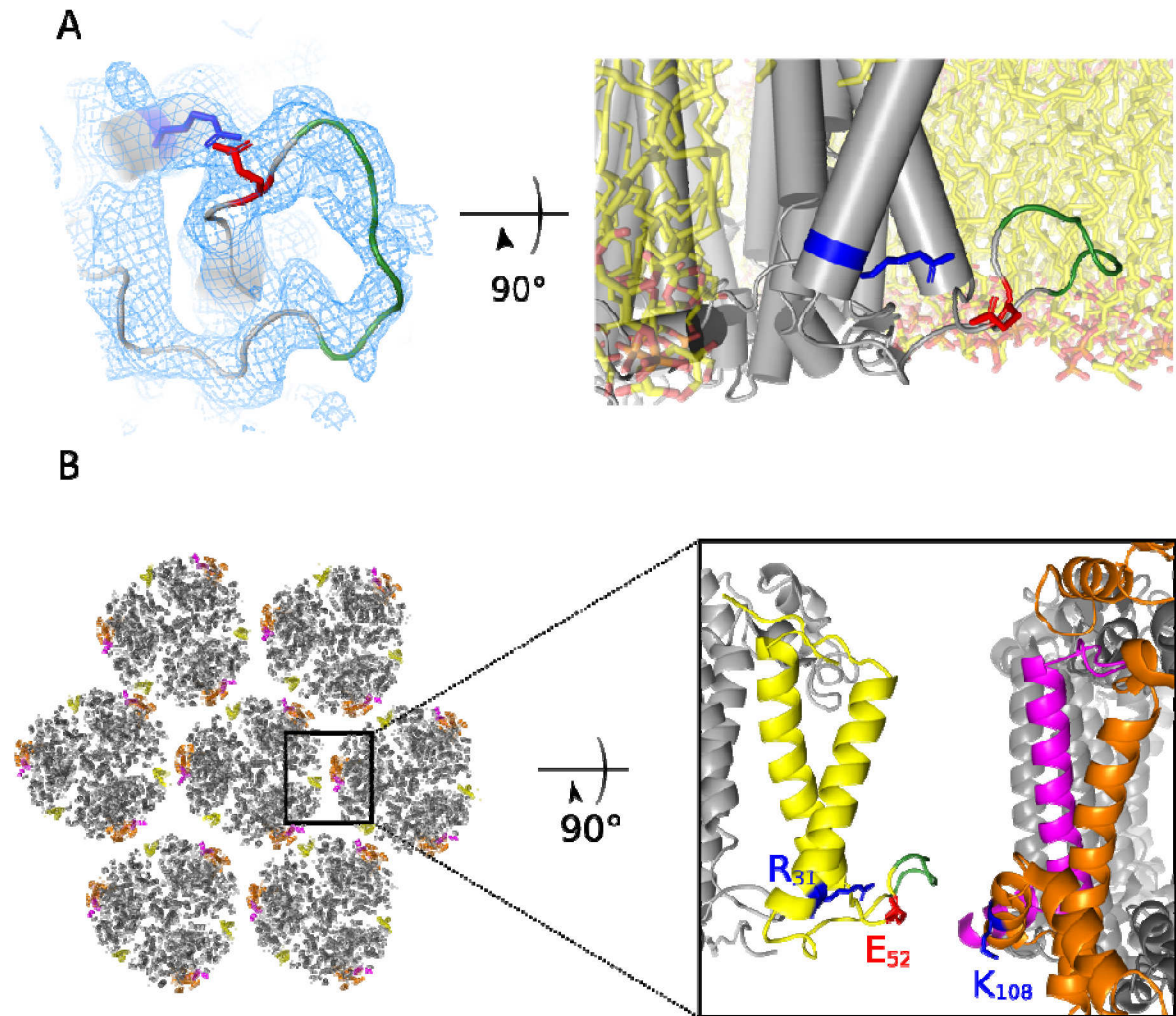
Figure 3 | The structure of PsaK loop 32-55. **A,** View onto cytoplasmic side of the model with symmetry expanded monomeric EM map displayed as blue mesh (left) and view from membrane onto the same section of PsaK with modelled palmitoyloleoyl-phosphatidylglycerid- (POPG)-membrane (right). The PIALPAL-patch 44-50 is highlighted in green, and the salt bridge forming residues Arg31 and Glu52 are coloured blue and red, respectively. POPG membrane was simulated using CHARMM membrane builder [62]. The map is plotted at a B-factor of -173 Å ^{2}. **B,** Model of trimer-trimer interface of *T. elongatus* thylakoid membrane. Hexagonal packing of PSI trimers (left) is based on the AFM images as shown in [28]. Close up of the interface between two neighbouring trimers (right). The subunits at the interface are PsaF (orange), PsaJ (magenta) and PsaK (yellow). The PIALPAL-patch 44-50 of PsaK is highlighted in green. PsaK Glu52 [2], Arg31 and PsaF Lys108 (blue) are shown in stick representation.
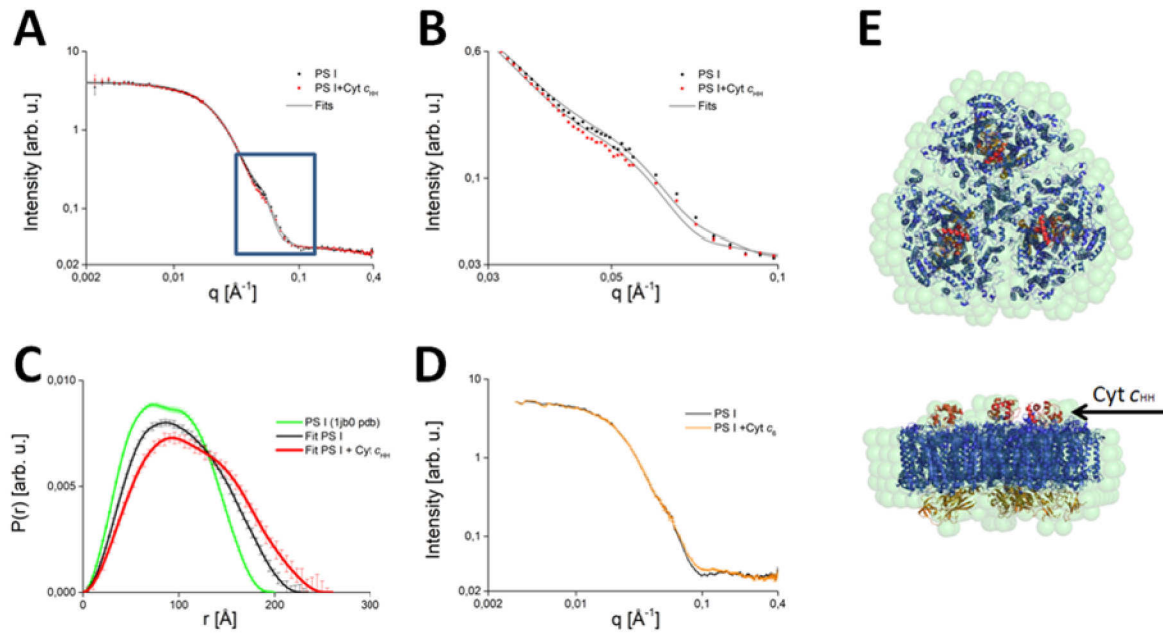
Figure 4 | SANS measurements of *Te*PSI in the presence of Cyt c$_{HH}$ or Cyt c$_6$. **A**, SANS data of trimeric *Te*PSI (black line) and of *Te*PSI-Cyt c$_{HH}$ complexes (red line) obtained at a contrast of 5% D2O. **B**, Magnification of the SANS data shown in Panel A for the q-region, where a different signal is observed for trimeric *Te*PSI (black line) and of *Te*PSI-Cyt c$_{HH}$ complexes (red line), respectively. **C**, Pair distance distribution function P(r) obtained from the SANS data of trimeric *Te*PSI (red line) and of *Te*PSI-Cyt c$_{HH}$ complexes (green line), respectively. For the comparison, we also present the P(r) function calculated from the crystal structure of *Te*PSI (pdb code 1jb0). **D**, SANS data of trimeric *Te*PSI (black line) and of a mixture of *Te*PSI-Cyt c$_6$ complexes (red line) obtained at a contrast of 5% D$_2$O. **E**, Comparison of the structure of the *Te*PSI-Cyt c$_{HH}$ complex reconstructed from the SANS data using the ATSAS routine (light blue spheres) with a crystal structure of the PSI-cytochrome complex taken from [1].