Structural and functional analysis of photosystem complexes using cryo-electron microscopy

「クライオ電子顕微鏡を用いた光化学系複合体の構造と機能解析」

March 2022

Huaxin Yu

于 華 鑫

Leading Program in Doctoral Education

Graduate School of Life Science

University of Hyogo

兵庫県立大学大学院生命理学研究科

博士課程教育リーディング大学院

論 文 内 容 の 要 旨

論文題目　Structural and functional analysis of photosystem complexes using cryo-electron microscopy

「クライオ電子顕微鏡を用いた光化学系複合体の構造と機能解析」

論文提出者　Huaxin Yu (于 華鑫)

Photosynthesis is a light energy storage process by which photosynthetic organisms capture light energy and use them to drive a series of cellular processes to provide foods and fuels on earth. This process is mainly achieved by two large membrane-embedded complexes, photosystem II (PSII) and photosystem I (PSI), together with other related proteins and their complexes in cyanobacteria, algae, and plants.

PSII is a multi-subunit pigment-protein enzyme which catalyzes the water oxidation and oxygen evolution reaction. There are a variety of technologies to analyze its structure. The crystal structure of PSII dimer has been resolved at a resolution of 1.9 Å using SPring-8 synchrotron radiation X-rays in 2011, in which all the metal atoms of the Mn_4CaO_5 cluster, together with the coordinated ligands were identified and located. The damage-free structure was further determined at 1.95 Å resolution using an X-ray free electron laser (XFEL) in 2015. However, the native structure in solution may be slightly different from that in the crystal, which is still not solved. Fortunately, modern cryo-electron microscopy (cryo-EM) has progressed tremendously owing to the advances such as direct electron detectors, automated data collection setup, and phase-contrast plates, which make it possible to preserve biological specimens near their physiological state within a thin vitreous ice film and allow direct observations of their structures without crystallization. It has become one of the most powerful tools to elucidate complicated conformations of proteins and their complexes, and enables analysis of biologically relevant functions across the photosynthetic complexes.

PSII generally functions as a dimer to drive its biological functions, but a part of PSII exists as a monomer in vivo. PSII dimer assembly and repairment following irreversible damage by visible light is a stepwise process through several intermediate states, which are achieved by the collaborations of protein subunits, pigments, and lipids. PSII monomer is formed prior to the dimer formation, and its structure is helpful to investigate the important factors required for the dimer

formation. The structure of PSII monomer from a thermophilic cyanobacterium *Thermosynechococcus elongatus* (*T. elongatus*) has been solved by X-ray crystallography at 3.6 Å resolution, and apo-PSII monomer that lacks the three extrinsic proteins and the Mn_4CaO_5 cluster from a mesophilic cyanobacterium *Synechocystis* sp. PCC 6803 has also been solved by cryo-EM. However, the factors that are important for the dimer formation are not clear at present, due to either the lower resolution or different species of the cyanobacteria used. In this study, we solved the structure of *T. vulcanus* PSII monomer at 2.78 Å resolution using cryo-EM and compared it with the crystal structure of a single monomer of the *T. vulcanus* PSII dimer, crystal structure of *T. elongatus* PSII monomer, and cryo-EM structure of *Synechocystis* sp. PCC 6803 apo-PSII monomer lacking the water-oxidation complex and some peripheral subunits. From these comparisons, we found that a *β*-carotene and two SQDGs located in the monomer-monomer interface of PSII dimer were absent in our monomer structure, suggesting their important roles for the dimer formation. In addition, the extrinsic subunit PsbO, which was reported to interact with CP47 from the adjacent monomer, exhibits disordered structure in a wide region, which may also indicate its importance for the stabilization of the dimer.

The water oxidation of PSII proceeds through the Kok cycle consisting of five sequential oxidation state transitions (S_0 to S_4) by which the oxygen-evolving complex (OEC) of PSII oxidizes water molecules and generates O_2. The catalytic center of water oxidation, the Mn_4CaO_5 cluster, has a low-spin (S_2^{LS}, open cubane structure) and a high-spin (S_2^{HS}, closed cubane structure) conformation in the S_2 state. The D1-V185 residue of PSII has been shown to be an important residue participating in the PSII water oxidation reaction, and the directed mutation of this residue to threonine (V185T) has been shown to slow down the proton release from the Mn_4CaO_5 cluster and to have a higher proportion of the S_2^{HS} configuration. In this work, we purified and characterized the D1-V185T mutant PSII dimer from the thermophilic cyanobacterium *T. elongatus* and detected a significantly lower oxygen-evolving activity than that of the wild strain. We have collected the negative staining EM image of the purified PSII dimer particles, and are planning to analyze its structure by cryo-EM.

The cryo-EM structural analysis uses a single particle analysis (SPA) technique to reconstruct biological macromolecular 3D structure from a huge amount of 2D snapshots captured by microscope. Several software packages are developed for SPA, including SPIDER, EMAN2, FREALIGN, RELION, cryoSPARC, and THUNDER, etc. Currently, RELION and cryoSPARC

stand out from them as the mainly used tools in the cryo-EM structural analysis owing to their excellent performances, but there are still some differences between them, and their performances need to be compared and evaluated. In this study, these two software packages are compared based on a PSII high-resolution dataset in many aspects, such as resolution, B-factor, quality of map, and calculation speed, etc. The results showed that, while cryoSPARC give rise to a nominally higher resolution than that of RELION with the same dataset, the electron potential map generated by RELION showed more features that may represent the structure. One advantage of the cryoSPARC software is its higher speed and less computer resources; therefore, one way to process the cryo-EM data may be to use a mixed procedure in which, cryoSPARC is first used to process the data, followed by 3D refinement with RELION, which may give a reasonably good electron potential map at a higher speed.

CONTENTS

ABBREVIATIONS

ATP: Adenosine triphosphate

BCR: β-carotene

β-DDM: n-dodecyl-β-D-maltoside

Chl: Chlorophyll

Cryo-EM: Cryo-electron microscopy

Cyt: Cytochrome

DGDG: Digalactosyldiacyl glycerol

E_m: Redox potential

Fd: Ferredoxin

FNR: Fd:$NADP^+$ reductase

FSC: Fourier shell correlation

LHC: Light-harvesting complex

LHG: Dipalmitoylphosphatidyl glycerol

LMG: Distearoylmonogalactosyl diglyceride

MGDG: Distearoylmonogalactosyl diglyceride

NADP: Nicotine amide dinucleotide phosphate

OEC: Oxygen-evolving complex

PBS: Phycobilisome

Pc: Plastocyanin

PG: Dipalmitoylphosphatidyl glycerol

Pheo: Pheophytin

PQ: Plastoquinone

PS: Photosystem

RC: Reaction center

RRX: Cryptoxanthin

SPA: Single particle analysis

SQDG: Sulfoquinovosyl diacylglycerol

Q_A/Q_B: Primary quinone acceptor/Secondary quinone acceptor

XRD: X-ray diffraction

Y_Z: Tyrosine 161 in D1 subunit of PSII

GENERAL INTRODUCTION

1. Photosynthesis

Photosynthesis is a biological process which captures light energy and uses them to convert carbon dioxide and water into organic compounds for supplying food and energy to almost all living organisms on the earth, with molecular oxygen released into the atmosphere as a by-product[1,2]. There are two types of photosynthesis classified based on the oxygen-releasing or not, namely, anoxygenic and oxygenic photosynthesis. Some groups of bacteria undergo anoxygenic photosynthesis, whereas oxygenic photosynthesis takes place in plants, algae, and cyanobacteria in which the harvested light energy is converted into chemical energy which is coupled with the water oxidation, oxygen evolution, and electron transfer reactions. The whole photosynthesis is abbreviated in the following equations.

$$2\ H_2O + 4\ photons \longrightarrow O_2 + 4\ e^- + 4\ H^+$$

$$4\ e^- + 4\ H^+ + CO_2 \longrightarrow [CH_2O]_n + H_2O$$

These reactions are catalyzed by two photosystems, photosystem I (PSI) and photosystem II (PSII), both of which are multi-subunit, pigment-protein complexes embedded in the thylakoid membranes. The absorptions of light energy are carried out by pigments and antenna complexes including chlorophylls (Chls), carotenoids, phycobilisomes (PBSs), and various light-harvesting complexes (LHC). PBSs are arrays of pigments bound to water-soluble proteins and function in cyanobacteria, whereas LHCs are trans-membrane proteins and the energy trapped by LHCII and LHCI is transferred to PSII and PSI of eukaryotic algae and plants, respectively.

The absorbed light energy induces photosynthetic electron transfer, which starts from the excitation of primary electron donor Chls P680 in PSII and P700 in PSI. In PSII, the reaction center P680 including four Chls, P_{D1}, P_{D2}, Chl_{D1} and Chl_{D2} are excited and become $P680^{\bullet+}$ after charge separation to form $Chl_{D1}^{+}Phe_{D1}^{-}$, and then forming the $P_{D1}^{+}Phe_{D1}^{-}$ radical pair. P_{D1}^{+} then oxidizes Y_Z (D1-Tyr161), which sequentially oxidizes the Mn_4CaO_5 cluster, resulting in water oxidation and oxygen evolution. On the other hand, the electron released from Phe_{D1}^{-} is then transferred to the primary quinone electron acceptor, Q_A and then the second quinone electron acceptor, Q_B. Q_B accepts electrons twice from Q_A to become Q_B^{2-}, which is protonated twice and detached from its binding site (D1 protein) at the stromal side of the thylakoid membrane. The reduced PQH_2 enters the PQ pool, which is oxidized by the cytochrome *b*$_6$*f* complex (Cyt *b*$_6$*f*). One of the electrons

from PQH_2 participates in the Q-cycle, and another electron is transferred to the soluble electron transfer carrier plastocyanin (Pc) or cytochrome c_6 (Cyt c_6) depending on species. Meanwhile, the protons released into the lumenal side contribute to the formation of a proton gradient across the thylakoid membrane, and this proton uptake-release cycle is called Q-cycle. The proton gradient across the thylakoid membrane drives ATP synthesis by the ATP synthase (ATPase) (Fig. 0-1).

The reduced Pc and Cyt c_6 then donates electrons to PsaF, a PSI subunit located in the lumenal side of the membrane. The positively charged region of PsaF rich in lysine residues can interact with the negatively charged region of Pc and Cyt c_6 to facilitate the electron transfer. The electrons received are transferred to the PSI reaction center P700, which is excited to P700* upon the light energy capture from PSI antennas. P700* then transfers electrons to the PSI electron acceptor side, and sequentially to another soluble protein, ferredoxin (Fd). Fd donates electrons to Fd:NADP reductase (FNR), which in turn catalyzes the reduction of oxidized nicotine amide dinucleotide phosphate ($NADP^+$) to NADPH (reduced NADP). The NADPH and ATP thus synthesized are utilized to fix carbon dioxide into carbohydrates in the Calvin-Benson cycle. An overall view of electron and proton transfer pathways in photosynthesis is depicted in Fig. 0-1.

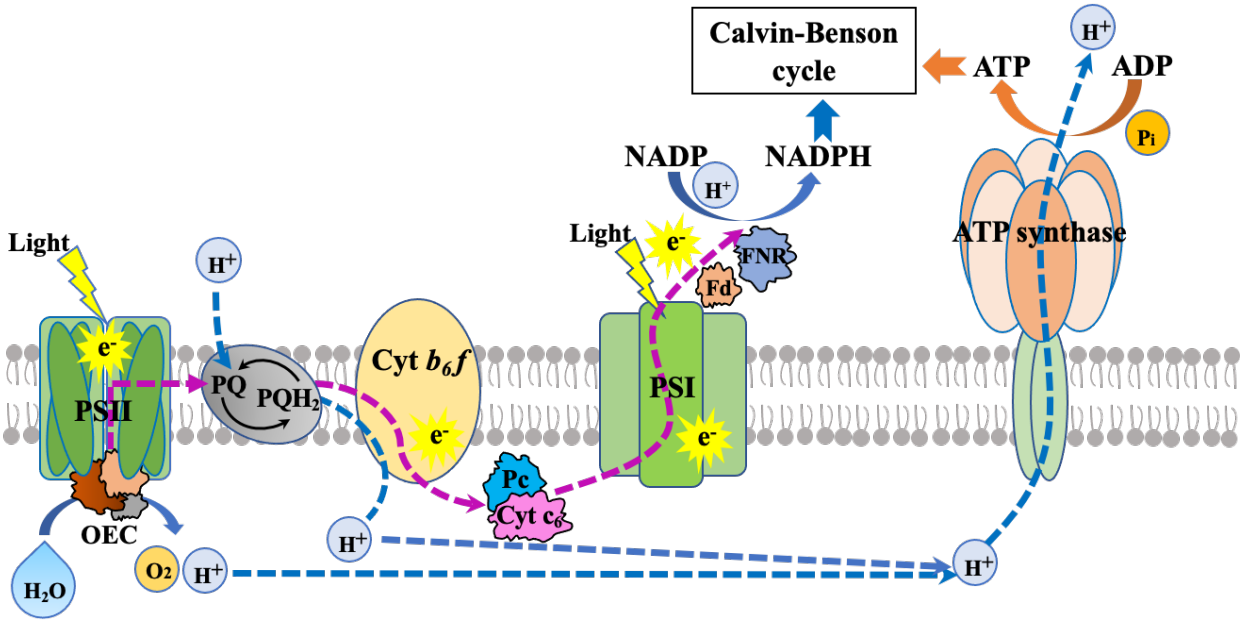


Fig. 0-1. Overview of the electron and proton transfer pathways in photosynthesis. Pathways of electron and proton flow directions are represented in dashed magenta and blue lines, respectively.

Fig. 0-2 shows an energetic picture of the electron transfer pathway, called Z-scheme, incorporating the major reactions in oxygenic photosynthesis. PSII and PSI play important roles in the series of reactions because they can use light energy to deliver the electrons. In particular, PSII, which is the main target of the present study, is the only known biological machine utilizing the light energy to split water into molecular oxygen, protons, and electrons. Thus, studies on PSII may help us to develop clean energy and solve the greenhouse effect in the future.

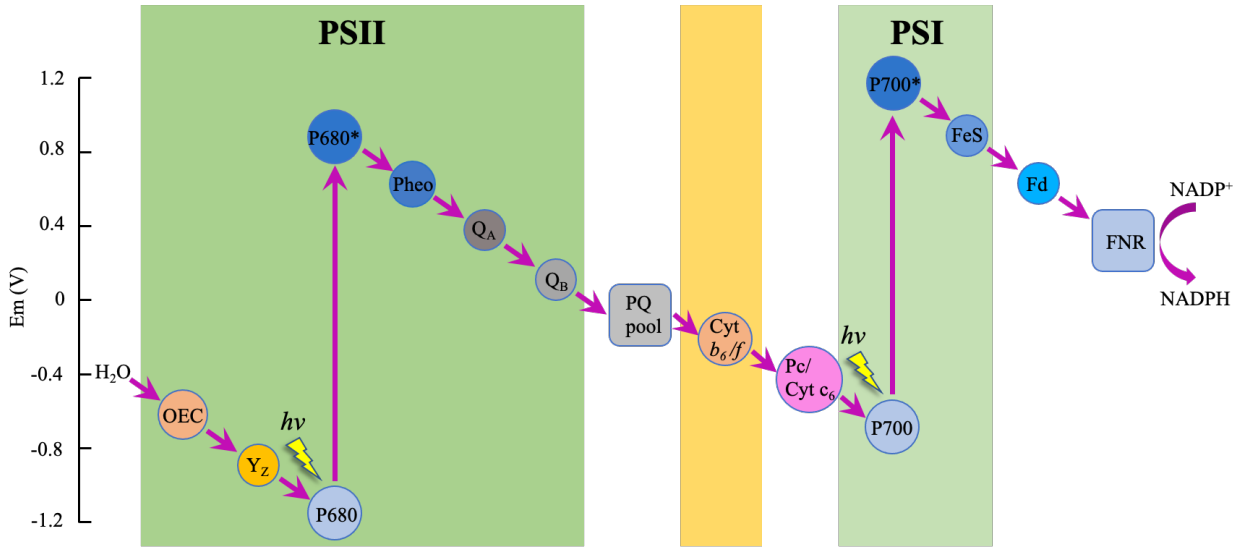


Fig. 0-2. The Z-scheme for electron transfer in oxygenic photosynthesis.

2. Photosystem II

Owing to the critical function of PSII in water splitting and oxygen generation which are indispensable for aerobic lives on earth, extensive studies have been devoted to investigating the structure and function of PSII. The mechanism of water oxidation has been a long and hot topic for many years. The crystal structure of PSII from thermophilic cyanobacterium *Thermosynechococcus elongatus* (*T. elongatus*) was first reported at 3.8 Å resolution in 2001[3]. Another PSII crystal structure from a similar species, thermophilic cyanobacterium *Thermosynechococcus vulcanus* (*T. vulcanus*) was solved at 3.7 Å in 2003[4]. Along with the optimization of the crystallization method, the resolution gradually improved to 3.0 Å in 2005[5] and 2.9 Å resolution in 2009[6], which enabled the positioning of all subunits as well as most cofactors. However, the accurate positions of metal atoms in the oxygen-evolving complex (OEC), a Mn_4CaO_5 cluster was not well defined owing to the ambiguous electron density map obtained at the resolution achieved. A breakthrough was made in 2011, which solved the crystal structure of

PSII dimer from *T. vulcanus* in the dark-stable state at 1.9-Å resolution using X-ray diffraction (XRD) with synchrotron radiation[7] (Fig. 0-3). All atoms in the Mn_4CaO_5 cluster, together with the coordinated ligands, are identified for the first time, which allows elucidation of the mechanism of the water-splitting reaction.

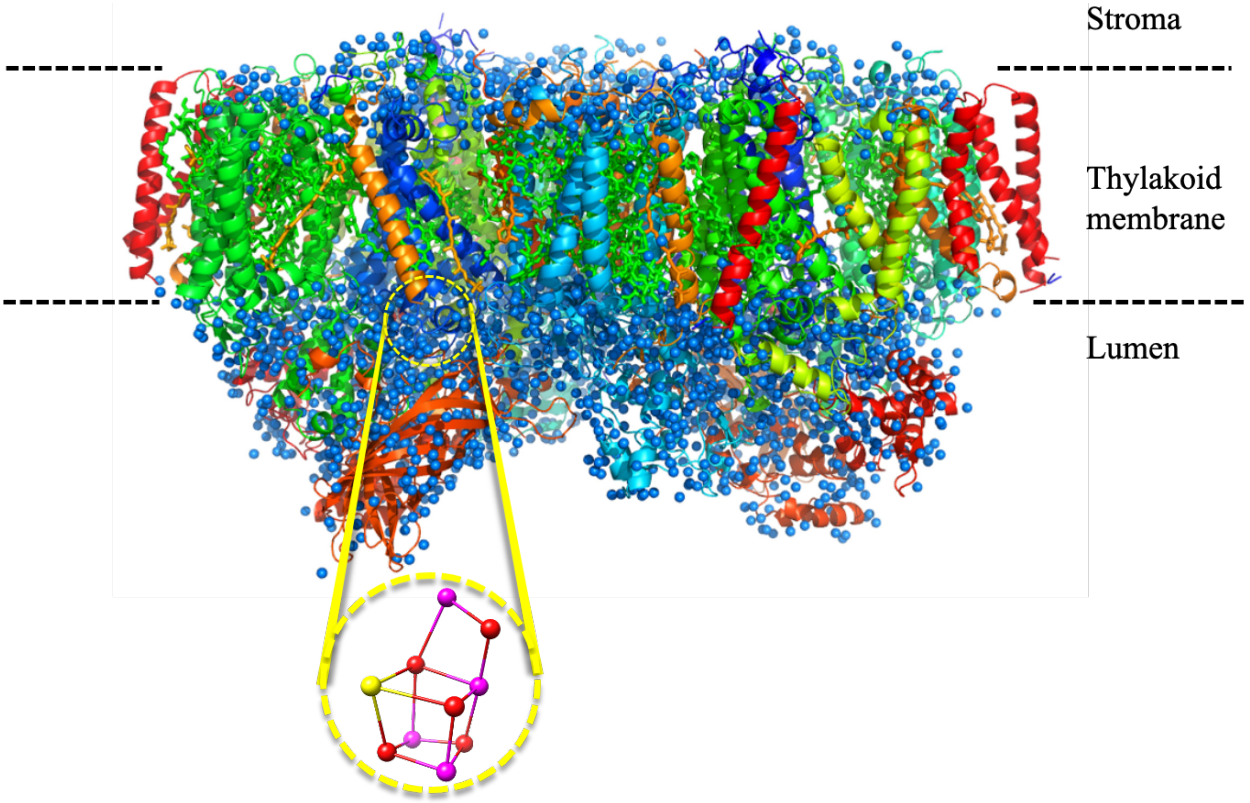


Fig. 0-3. Overall structure of PSII dimer from *T. vulcanus* and OEC at 1.9 Å resolution (PDB code: 3WU2). Side view of the PSII dimer with all subunits colored differently. The location of OEC in a monomer is circled with a yellow dashed line, and an enlarged view of OEC is shown.

PSII core complex has a molecular weight of 350 kDa for a monomer, and its dimer has a total molecular mass of 700 kDa. Cyanobacteria PSII consists of 20 protein subunits per monomer (PsbY was not observed in the 1.9-Å resolution structure), including 17 transmembrane subunits such as the heterodimeric photosynthetic reaction center D1 and D2 subunits, internal antenna components CP47 and CP43, etc., and three extrinsic subunits, PsbO, PsbU, and PsbV protruded into the thylakoid lumen (Table 0-1). In addition, each PSII monomer contains 35 chlorophylls (Chls), 11 β-carotenes, more than 20 lipids, two plastoquinones, two haems, one non-haem iron, a Mn_4CaO_5 cluster, etc. More than 2,800 water molecules are found in a PSII dimer, many of which

participate in forming well defined hydrogen-bond networks that are necessary for the inlet of substrate water and egress of product protons during the water-splitting reaction (Table 0-2). The missing PsbY subunit is later found in one of the two monomers in the crystal structures of Sr-substituted PSII[8] and native PSII solved by a damage-free X-ray free electron laser[9]. This suggests that the PsbY subunit is associated with PSII loosely, and is lost from one of the two monomers due to crystallographic restraints within the PSII crystal. This is one of the disadvantages of crystallization and suggests that the physiological state of molecules may not be represented. Resolution is another bottleneck in X-ray structural analysis. In addition, crystallization requires a highly purified sample in a large quantity, and membrane proteins are usually difficult to be purified in a large quantity. Therefore, it is necessary to investigate the PSII structure in physiological state using more advanced techniques.

Table 0-1. PSII subunits contained in the cyanobacterial dimer complex.

PDB chain	Gene	Protein	Mw (Da)	Amino acids	Function
A, a	*psb* A	D1	39,766	360	core RC
B, b	*psb* B	CP47	56,017	505	Inner antenna
C, c	*psb* C	CP43	50,245	461	Inner antenna; Ligand for Mn_4CaO_5 cluster
D, D	*psb* D	D2	39,378	352	core RC
E, e	*psb* E	Cytochrome b559 α	9,573	84	PSII assembly at early steps; Photoprotection
F, f	*psb* F	Cytochrome b559 β	5,065	45	PSII assembly at early steps; Photoprotection
H, h	*psb* H	Psb H	7,223	65	PSII stabilization and assembly; Photoprotection
I, i	*psb* I	Psb I	4,405	38	PSII dimerization; Maintain PSII structure and function under high light
J, j	*psb* J	Psb J	4,105	40	Assembly of water splitting complex
K, k	*psb* K	Psb K	4,047	37	Plastoquinone binding; Maintain PSII dimeric form
L, l	*psb* L	Psb L	4,297	37	Assembly of PSII; Maintain PSII dimeric form
M, m	*psb* M	Psb M	4,015	36	PSII dimer stabilization; Electron transfer at Q_B site
T, t	*psb* Tc	Psb T	3,875	32	Recovery of photo damaged PSII; PSII dimerization
X, x	*psb* X	Psb X	5,280	50	Binding or turnover of quinone molecules at Q_B site
R	*psb* Y	Psb Y	4,600	41	pH-dependent stabilizing protein for PSII; Docking protein for PSII extrinsic proteins
Z, z	*psb* Z	Psb Z	6,778	62	Linker between LHCII and PSII core
Y, y	*psb* 30	Ycf12	5,037	46	Stabilize PSII dimer; Prevent Cytb559 from converting to low potential form under high light
O, o	*psb* O	33 kDa	29,703	272	Mn_4CaO_5 cluster stabilization; Environment optimization for water oxidation; D1 turnover
U, u	*psb* U	12 kDa	14,945	134	Protecting shield for Mn_4CaO_5 cluster; Stabilize the interaction between PSII and phycobilisome
V, v	*psb* V	Cytochrome b559	18,027	163	Providing Ca^{2+}/Cl^- to OEC

Table 0-2. Numbers of cofactors in a PSII dimer (PDB code: 3WU2).

	Monomer A	Monomer B
OEC (Mn_4CaO_5)	1	1
Ca (Ca of OEC)	4 (5)	3 (4)
Cl	3	3
Heme	2	2
Non-heme iron	1	1
Chlorophyll *a*	35	35
Pheophytin	2	2
BCR (β-carotene)	11	11
BCT (bicarbonate)	1	1
MGDG	6	6
DGDG	5	5
PG	5	5
SQDG	4	4
Plastoquinone	2	2
H_2O	1,438	1,357
Detergents	15	14
Unknown diglyceride	3	3
Unknown alkyl chains	16	16
Glycerol	19	18

4. Structural analysis of the PSII monomer

How PSII is assembled into a multi-subunit, multi-cofactor complex and how the photoinhibition mechanism works (repair following irreversible damage by visible light) are still open questions. Up to now, this process is reported to be assisted by the collaboration of over 20 auxiliary factors which guide the several intermediate states to produce a mature complex[10].

In cyanobacteria, PSII assembly starts from the formation of the D1/D2 heterodimer reaction centre (RC) complex from the precursor form of D1 (pD1) and D2. This process is assisted by Ycf48 after catalyzed by the D1 C-terminal protease A (CtpA)[11,12]. Next, Psb28 helps RC complex to recruit CP47 to form the RC47 complex, in which the intermediate D1 is further processed into its mature state by CtpA[13,14]. CP43 binds to the complex with the assistance of Psb27 and several small subunits[15] to form the Psb27-PSII monomer, which is then attached by mature extrinsic subunits PsbO, PsbU and PsbV and activated. After completion of the assembly, Psb28 and Psb27 are released from the complex, and two monomers are assembled to form an active dimer[16-19].

The assembly of OEC is also a multistep process that requires the light-dependent oxidation of Mn^{2+} to build up the μ-oxo bridges between the Mn atoms (photoactivation)[20,21]; however, its

structural details are not well understood. A kinetic model of photoactivation is established, referred to as "two quantum model"[22], in which a single high-affinity Mn^{2+}-binding site[23] exists. After binding of Mn^{2+} to this site, the bound Mn^{2+} is oxidized to Mn^{3+} followed by a slow light-dependent phases, by which the remaining Mn^{2+} are oxidized and incorporated.

PSII monomer is formed prior to the dimer, and its structure is helpful to analyse the important factors required for the dimer formation, especially the residues located in the monomer-monomer interface may provide us inspirations on how two monomers are linked to form a dimer. However, the structure of the PSII monomer from the thermophilic cyanobacterium *T. elongatus* is analyzed at 3.6 Å resolution[24], which is not enough to allow identification of the molecules that are present in the monomer-monomer interface of the dimer but absent in the monomer. The apo-PSII monomer structure has also been analyzed from *Synechocystis* sp. PCC 6803 by cryo-EM[25]; however, it is in the apo-form lacking the extrinsic proteins and OEC. Thus, we analyzed the PSII monomer structure from *T. vulcanus* at a higher resolution and compared it with the structure of the dimer from the same organism as well as the monomer structure solved by X-ray crystallography and the apo-PSII monomer structure from *Synechocystis* sp. PCC 6803 solved by cryo-EM.

4. PSII D1-V185T

The 1.9-Å crystal structure reveals that Mn_4CaO_5 cluster is surrounded by a number of water molecules, some of them are coordinated with the Mn_4CaO_5 cluster directly and others are located in the water channels from the Mn_4CaO_5 cluster to the PSII luminal side[26,27]. These channels may provide access for substrate water to the Mn_4CaO_5 cluster[28], and also seems to function in proton transfer[29]. To analyse the roles of the water channels, a set of point mutations was generated at the water-containing cavities at the interface between D1 and the Mn_4CaO_5 cluster. Three mutations of Val185 in the D1 subunit, D1-Val185Asn, D1-Val185Thr, and D1-Val185Phe, are detected to have a strong effect on the water oxidation mechanism[30-32]. The D1-V185T mutant from *T. elongatus* was demonstrated to impact the state transition of OEC and decrease oxygen evolution dramatically[32]. The structure of PSII mutant V185T is not solved, but it would help us to explain functional changes occurred in the mutant.

5. Cryo-electron microscopy

With the existence of some limitations of crystal structure, a modern technique, cryo-electron microscopy (cryo-EM) has become one of the most powerful tools to elucidate complicated conformations of proteins and their complexes, and biologically relevant functions across the photosystem complexes. It has developed rapidly in recent 20 years owing to the improvements of hardware and computing technology including direct electron detectors, automated data collection system, and phase-contrast plates, etc[33-35]. These advances enable us to observe specimens in near their native state in a small volume and without crystallization.

Cryo-EM structural analysis involve single particle analysis (SPA) to reconstruct biological macromolecular 3D structure from huge amount of 2D snapshots captured by microscope. A brief process of data acquisition using SPA is illustrated in Fig. 0-4. Several μL of purified sample is dropped onto the cryo-EM grid, after the excessive solution is blotted off, the sample are embedded in a thin layer of vitreous ice film through rapid plunging into liquid ethane to freeze the grid. The prepared grid is then loaded into the microscope, and the target holes are positioned after screening. An incident electron beam is applied onto the target hole, and the image projections are captured by the detector. Generally, thousands of 2D snapshots of the sample need to be collected for subsequent 3D reconstructions. As the molecules are oriented randomly within the ice, these images can represent the sample at various angles, and can be recombined into a high-resolution electron potential map using a series of image processing software packages (Fig. 0-5A). Finally, 3D atomic model can be constructed based on the obtained map by structural analysis. Fig. 0-5B shows a picture of CRYO ARM^{TM} 300 JEOL microscope in RIKEN Spring-8 Yonekura lab, which is equipped with a cold field emission gun, an in-column Omega energy filter, a side-entry liquid nitrogen cooling stage and an automated specimen exchange system. With the use of this microscope, a 1.95-Å PSII dimer structure was solved in 2021[36], which is comparable to that of the crystal structure.

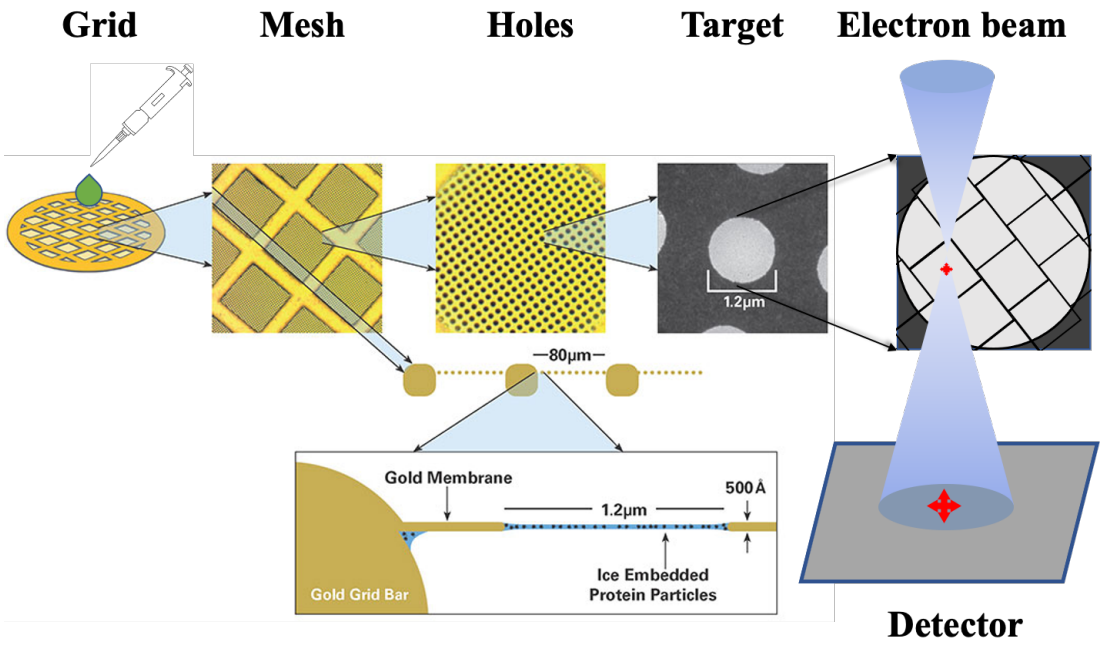


Fig. 0-4. An example of stepwise process of cryo-EM data acquisition using SPA.
An enlarged view of particles embedded in the vitreous ice supported by grid mesh is shown.

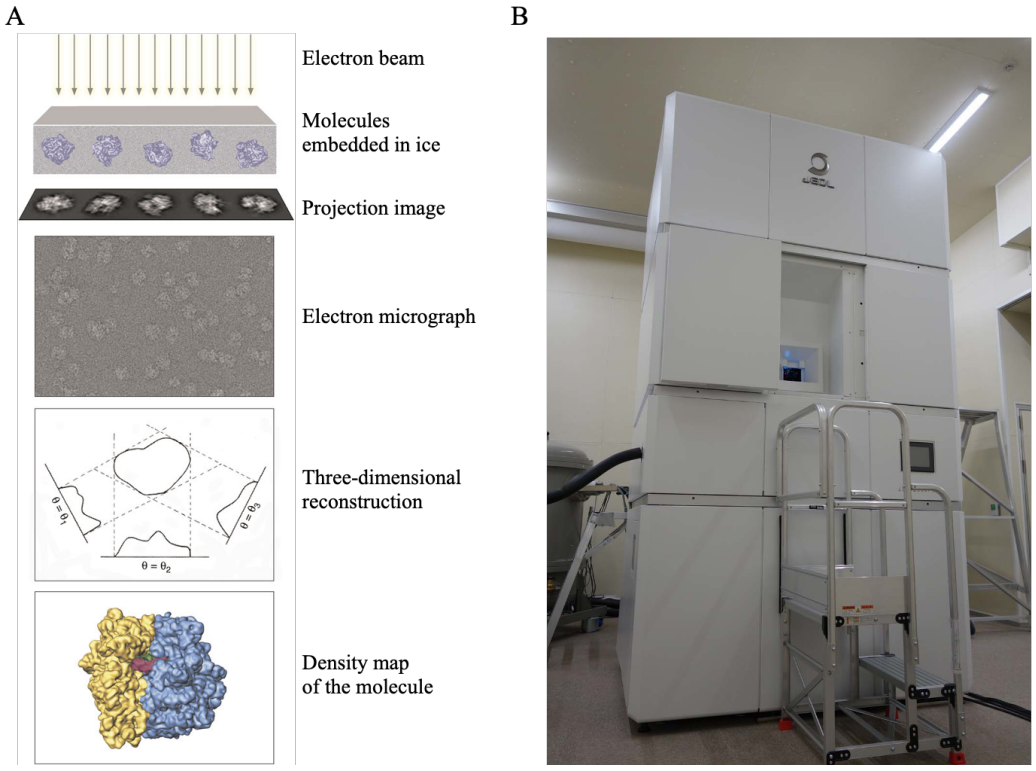


Fig. 0-5. A. Principle of particle imaging and data processing in SPA. B. CRYO ARM^{TM} 300 JEOL microscope in RIKEN Spring-8 Yonekura lab.

6. RELION and cryoSPARC

The acquired raw micrographs from cryo-EM need to be further processed to reconstruct a 3D map. A general processing workflow is depicted in Fig. 0-6. In brief, the motion correction and contrast transfer function (CTF) estimation primarily calibrate the particle movements and the height within ice, and numerous particles are picked from the corrected images for the reconstruction. After 2D and 3D classifications, the particle sets in consistent with our wish are selected for 3D refinement and local refinement. Based on the refined map, several rounds of classifications will be performed until a reasonably high-quality map is obtained. This calculation work is attributed to the efforts of several image-processing software packages developed for SPA, such as SPIDER[37], EMAN2[38], FREALIGN[39], RELION[40], cryoSPARC[41], and THUNDER[42], etc. Among them, RELION and cryoSPARC stand out as the mainly used tools owing to their excellent performances, but there are still some differences between them, and their performances need to be evaluated.

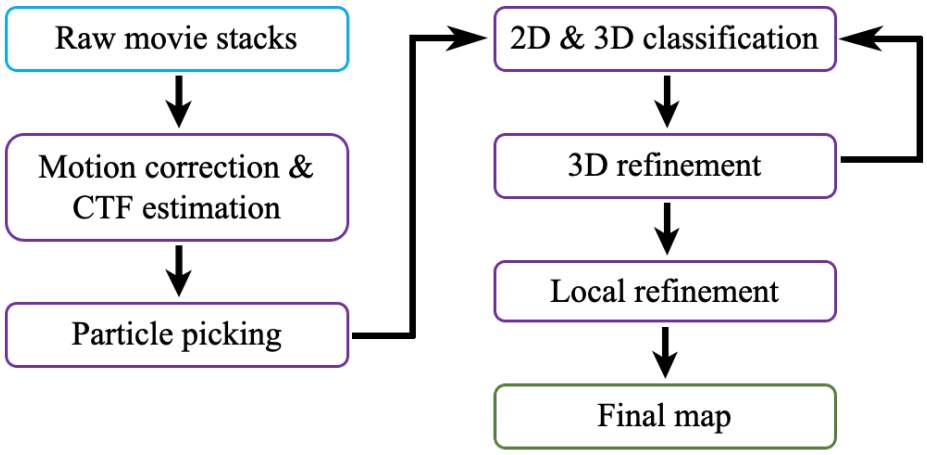


Fig. 0-6. A general image-processing workflow for SPA.

7. Research contents in this study

In chapter 1, PSII monomer from *T. vulcanus* was purified and its structure was solved at 2.78 Å using cryo-EM[43]. The solved structure was compared with other related PSII monomer structures to inspect the key factors for the PSII dimerization. From these comparisons, a *β*-carotene, two SQDGs and the extrinsic subunit, PsbO were found to be important factors in maintaining the PSII dimer conformation.

In chapter 2, we cultured bulk volumes of the V185T mutant strain and successfully purified PSII dimer from the cells. An extremely low oxygen-evolving activity was found in the purified sample, which may indicate the important role of V185 in the oxygen evolution process. For further examinations, a large quantity of particle images of the purified sample was recorded using negative staining EM. We are going to analyze its structure in S_1 state using cryo-EM and following states, to investigate the implications of this mutation on the state transitions.

In chapter 3, RELION and cryoSPARC are compared based on a PSII high-resolution data set in different views, such as resolution and B-factor, quality of map, calculation speed, etc. The results showed that, while cryoSPARC gives nominally higher resolutions than those of RELION with the same dataset, the electron potential map generated by cryoSPARC is poorer than RELION. One advantage of the cryoSPARC software is its faster computing speed and less resource costs, therefore, one strategy to process the cryo-EM data may be to use a mixed procedure, in which the data is first processed with cryoSPARC followed by further processing with RELION, which may give rise to a reasonably good electron potential map in a short time.

References

1. Dismukes, G. C. *et al.* The origin of atmospheric oxygen on Earth: The innovation of oxygenic photosynthesis. *Proc Natl Acad Sci USA* **98**, 2170-2175 (2001).
2. Blankenship, R. E. Molecular Mechanisms of Photosynthesis. 2nd ed. (Wiley/Blackwell, Chichester, West Sussex, UK, 2014).
3. Zouni, A. *et al.* Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 Å resolution. *Nature* **409**, 739-743 (2001).
4. Kamiya, N. & Shen, J.-R. Crystal structure of oxygen-evolving photosystem II from *Thermosynechococcus vulcanus* at 3.7-Å resolution. *Proc Natl Acad Sci USA* **100**, 98-103 (2003).
5. Loll, B., Kern, J., Saenger, W., Zouni, A. & Biesiadka, J. Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II. *Nature* **438**, 1040-1044 (2005).
6. Guskov, A. *et al.* Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. *Nat Struct Mol Biol* **16**, 334-342 (2009).

7. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55-60 (2011).
8. Koua, F. H. M., Umena, Y., Kawakami, K. & Shen, J.-R. Structure of Sr-substituted photosystem II at 2.1 Å resolution and its implications in the mechanism of water oxidation. *Proc Natl Acad Sci USA* **110**, 3889-3894 (2013).
9. Nadtochenko, V. A., Semenov, A. Y. & Shuvalov, V. A. Formation and decay of P680 (P_{D1}-P_{D2})$^+$ $Pheo_{D1}^-$ radical ion pair in photosystem II core complexes. *Biochim Biophys Acta Bioenerg* **1837**, 1384-1388 (2014).
10. Zabret, J. *et al.* Structural insights into photosystem II assembly. *Nat Plants* **7**, 524-538 (2021).
11. Komenda, J. *et al.* Cleavage after residue Ala352 in the C-terminal extension is an early step in the maturation of the D1 subunit of photosystem II in *Synechocystis* PCC 6803. *Biochim Biophys Acta Bioenerg* **1767**, 829-837 (2007).
12. Komenda, J. *et al.* The cyanobacterial homologue of HCF136/YCF48 is a component of an early photosystem II assembly complex and is important for both the efficient assembly and repair of photosystem II in *Synechocystis* sp. PCC 6803. *J Biol Chem* **283**, 22390-22399 (2008).
13. Dobakova, M., Sobotka, R., Tichy, M. & Komenda, J. Psb28 protein is involved in the biogenesis of the photosystem II inner antenna CP47 (PsbB) in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol* **149**, 1076-1086 (2009).
14. Boehm, M. *et al.* Subunit composition of CP43-less photosystem II complexes of *Synechocystis* sp. PCC 6803: implications for the assembly and repair of photosystem II. *Philos Trans R Soc B Biol Sci* **367**, 3444-3454 (2012).
15. Komenda, J. *et al.* The Psb27 assembly factor binds to the CP43 complex of photosystem II in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol* **158**, 476-486 (2012).
16. Mamedov, F., Nowaczyk, M. M., Thapper, A., Rogner, M. & Styring, S. Functional characterization of monomeric photosystem II core preparations from *Thermosynechococcus elongatus* with or without the Psb27 protein. *Biochemistry* **46**, 5542-5551 (2007).
17. Nowaczyk, M. M. *et al.* Psb27, a cyanobacterial lipoprotein, is involved in the repair cycle of photosystem II. *Plant Cell* **18**, 3121-3131 (2006).
18. Roose, J. L. & Pakrasi, H. B. The Psb27 protein facilitates manganese cluster assembly in photosystem II. *J Biol Chem* **283**, 4044-4050 (2008).

19. Huang, G. *et al.* Structural insights into a dimeric Psb27-photosystem II complex from a cyanobacterium *Thermosynechococcus vulcanus*. *Proc Natl Acad Sci USA* **118** (2021).
20. Bao, H. & Burnap, R. L. Photoactivation: The light-driven assembly of the water oxidation complex of photosystem II. *Front Plant Sci* **7**, 578 (2016).
21. Becker, K., Cormann, K. U. & Nowaczyk, M. M. Assembly of the water-oxidizing complex in photosystem II. *J Photochem Photobiol B* **104**, 204-211 (2011).
22. Radmer, R. & Cheniae, G. M. Photoactivation of the manganese catalyst of O_2 evolution. II. A two-quantum mechanism. *Biochim Biophys Acta Bioenerg* **253**, 182-186 (1971).
23. Nixon, P. J. & Diner, B. A. Aspartate-170 of the photosystem-II reaction center polypeptide-D1 is involved in the assembly of the oxygen-evolving manganese cluster. *Biochemistry* **31**, 942-948 (1992).
24. Broser, M. *et al.* Crystal structure of monomeric photosystem II from *Thermosynechococcus elongatus* at 3.6-Å resolution. *J Biol Chem* **285**, 26255-26262 (2010).
25. Gisriel, C. J. *et al.* Cryo-EM structure of monomeric photosystem II from *Synechocystis* sp. PCC 6803 lacking the water-oxidation complex. *Joule* (2020).
26. Murray, J. W. & Barber, J. Structural characteristics of channels and pathways in photosystem II including the identification of an oxygen channel. *J Struct Biol* **159**, 228-237 (2007).
27. Ho, F. M. & Styring, S. Access channels and methanol binding site to the CaMn4 cluster in Photosystem II based on solvent accessibility simulations, with implications for substrate water access. *Biochim Biophys Acta Bioenerg* **1777**, 140-153 (2008).
28. Vassiliev, S., Comte, P., Mahboob, A. & Bruce, D. Tracking the flow of water through photosystem II using molecular dynamics and streamline tracing. *Biochemistry* **49**, 1873-1881 (2010).
29. Ishikita, H., Saenger, W., Loll, B., Biesiadka, J. & Knapp, E. W. Energetics of a possible proton exit pathway for water oxidation in photosystem II. *Biochemistry* **45**, 2063-2071 (2006).
30. de Lichtenberg, C. *et al.* The D1-V185N mutation alters substrate water exchange by stabilizing alternative structures of the Mn_4Ca-cluster in photosystem II. *Biochim Biophys Acta Bioenerg*, 148319 (2020).

31. Dilbeck, P. L., Bao, H., Neveu, C. L. & Burnap, R. L. Perturbing the water cavity surrounding the manganese cluster by mutating the residue D1-Valine 185 has a strong effect on the water oxidation mechanism of photosystem II. *Biochemistry* **52**, 6824-6833 (2013).
32. Sugiura, M. *et al.* Probing the role of Valine 185 of the D1 protein in the Photosystem II oxygen evolution. *Biochim Biophys Acta Bioenerg* **1859**, 1259-1273 (2018).
33. Kuhlbrandt, W. The resolution revolution. *Science* **343**, 1443-1444 (2014).
34. McMullan, G., Faruqi, A. R., Clare, D. & Henderson, R. Comparison of optimal performance at 300 keV of three direct electron detectors for use in low dose electron microscopy. *Ultramicro* **147**, 156-163 (2014).
35. Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152-156 (2020).
36. Kato, K. *et al.* High-resolution cryo-EM structure of photosystem II reveals damage from high-dose electron beams. *Commun Biol* **4**, 382 (2021).
37. Shaikh, T. R. *et al.* SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat Protoc* **3**, 1941-1974 (2008).
38. Tang, G. *et al.* EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol* **157**, 38-46 (2007).
39. Grigorieff, N. Frealign: An exploratory tool for single-particle cryo-EM. *Methods Enzymol* **579**, 191-226 (2016).
40. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7** (2018).
41. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).
42. Hu, M. *et al.* A particle-filter framework for robust cryo-EM 3D reconstruction. *Nat Methods* **15**, 1083-1089 (2018).
43. Yu, H. X. *et al.* Cryo-EM structure of monomeric photosystem II at 2.78 Å resolution reveals factors important for the formation of dimer. *Biochim Biophys Acta Bioenerg* **1862** (2021).

CHAPTER 1

Structural analysis of photosystem II monomer at 2.78 Å resolution using cryo-electron microscopy

Abstract

Photosystem II (PSII) functions mainly as a dimer to catalyze the light energy conversion and water oxidation reactions. However, monomeric PSII also exists and functions in vivo in some cases. The crystal structure of monomeric PSII has been solved at 3.6 Å resolution, but it is still not clear which factors contribute to the formation of the dimer. Here, we solved the structure of PSII monomer at a resolution of 2.78 Å using cryo-electron microscopy (cryo-EM). From our cryo-EM density map, we observed apparent differences in pigments and lipids in the monomer-monomer interface between the PSII monomer and dimer. One *β*-carotene and two sulfoquinovosyl diacylglycerol (SQDG) molecules are found in the monomer-monomer interface of the dimer structure but not in the present monomer structure, although some SQDG and other lipid molecules are found in the analogous region of the low-resolution crystal structure of the monomer, or cryo-EM structure of an apo-PSII monomer lacking the extrinsic proteins from *Synechocystis* sp. PCC 6803. In the current monomer structure, a large part of the PsbO subunit was also found to be disordered. These results indicate the importance of the *β*-carotene, SQDG and PsbO in formation of the PSII dimer.

Introduction

PSII monomer is formed prior to the dimer and own an equivalent oxygen-evolving activity to half of the activity of dimer. Its structure is heuristic to understand PSII assembly process, but has not been solved at a high resolution. A crystal structure of monomeric PSII from thermophilic cyanobacterium *Thermosynechococcus elongatus* (*T. elongatus*) was reported at 3.6 Å resolution in 2010[1]. This resolution is good for locating protein subunits and most of the cofactors, but the electron densities for some cofactors is not continuous, and their localizations are partly based on the incomplete groups and their analogous positions in the PSII dimer. In particular, in the interface between the two monomers, it is expected to have different groups and structures between the dimer and monomer, but it is not clear if these groups are present, and if yes, where they are located. The structure of the PSII monomer was also solved from a mesophilic cyanobacterium *Synechocystis* sp. PCC 6803 at 2.58 Å resolution by cryo-electron microscopy (cryo-EM), but it lacks the three extrinsic proteins and the Mn_4CaO_5 cluster, and thus is an inactive one (named apo-PSII monomer here). Therefore, how the two monomers associate to form a dimer and what factors contribute to this formation are still not clear.

Recently, a dimeric Psb27-PSII complex from a PsbV deletion mutant of thermophilic cyanobacterium *Thermosynechococcus vulcanus* (*T. vulcanus*) was solved by cryo-EM[2], in which a possible process for PSII repair/assembly was proposed. The highly active dimeric PSII complex is transformed into a less active dimeric PSII and then to an inactive one upon light damage, with the three extrinsic proteins and the Mn_4CaO_5 cluster released. Subsequently, Psb27 is attached, and the dimer is split into two monomers, followed by the replacement of the damaged D1 by a newly synthesized premature D1. After the processing of the C-terminal region of D1, Psb27 is detached, followed by the assembly of the Mn_4CaO_5 and re-attachment of the three extrinsic proteins. After completion of the assembly, the PSII monomer is linked to form an active dimer. This mechanism provides us with a rather clear view on the sequences occurred during photoactivation, assembly of protein subunits, and dimerization of PSII in vivo. However, the factors that contribute to the dimerization remain elusive.

In this study, we solved the structure of monomeric, native PSII from *T. vulcanus* by cryo-EM single particle analysis at a resolution of 2.78 Å. The overall structures of the core subunits D1 and D2 and major Chl-binding subunits CP47 and CP43 are nearly identical to those of the PSII dimer, whereas some extrinsic subunits showed a higher flexibility. Recently, several reports

mentioned that carotenoids, sulfoquinovosyl diacylglycerols (SQDGs) and some other ligands on the monomer-monomer interface may have some correlations with the process of PSII dimerization[3-6]. Carotenoids are essential mediators of interactions between PSII monomers, and lipids may regulate the equilibrium between dimerization and dissociation of the dimer. Mutants completely deficient in carotenoid, ΔcrtH/B or ΔcrtB, of *Synechocystis* sp. PCC 6803 are devoid of active PSII reaction centers[5-7], and these strains do not show oxygen-evolving activity. Inactivation of an SQDG synthase in *T. elongatus* resulted in a complete loss of SQDG but some increases of other lipids, notably dipalmitoylphosphatidyl glycerol (PG) molecules, and the two SQDGs located at the monomer–monomer interface were found to be replaced by other lipids (most possibly PGs). This replacement decreased the stability of the PSII dimer, resulting in an increase in the amount of PSII monomer in the mutant[3]. Herein, we mainly focus on the differences of the cofactors, in particular BCRs and SQDGs, between the structures of monomer and dimer, and their possible roles in mediating the dimer formation. Our results indicate that the monomer lacks one BCR and at least one of the two SQDGs in the monomer-monomer interface, which suggest their importance in the formation of the PSII dimer.

Materials and Methods

Purification of PSII monomer from *T. vulcanus*

Liquid culture of *T. vulcanus* was grown at 50 ˚C with bubbling of air containing 3 % CO_2, under constant LED light. Thylakoid membranes were prepared by freeze-thaw after treatment with lysozyme and DNase[8]. Crude PSII was prepared by solubilizing the thylakoid membrane with lauryldimethylamine oxide[9]. Monomeric PSII complexes were further purified with Q-Sepharose High Performance Column (Cytiva) according to[8,10,11]. The purified PSII monomer sample was suspended in a storage buffer containing 5 % glycerol, 20 mM MES-NaOH (pH 6.0), 20 mM NaCl, and 3 mM $CaCl_2$, and stored in liquid nitrogen until data collection. All purification procedures were performed under dim green light.

Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was performed according to[12], with a 16 % acrylamide gel containing 7.5 M urea. The samples were solubilized with sample buffer containing 2 % lithium lauryl sulfate, 60 mM dithiothreitol, and 60 mM Tris-HCl (pH 8.5) at 60 ˚C for 40 min, before being loaded onto the gel.

Blue native-PAGE (BN-PAGE) was performed according to[13] with a 7.5 % acrylamide gel. The samples were first treated with 1 % n-dodecyl-*β*-D-maltoside (*β*-DDM) and centrifuged at 22,000 g for 10 min to remove insolubilized materials. The supernatant was further treated with 2 % *β*-DDM at 25 ˚C for 10 min, and mixed with Coomassie Brilliant Blue G-250 before being applied onto the gel.

Cryo-EM data collection

In order to obtain a better image quality with cryo-EM, the sample prepared was centrifuged at 14,000 g, 4 °C twice with a dilution buffer containing 20 mM MES (pH 6.0), 5 mM $CaCl_2$, and 0.03 % *β*-DDM using Amicon Ultra-100, to remove glycerol in the solution. For cryo-EM observations, a 3-μl aliquot of the monomeric PSII (2.15 mg Chl mL^{-1}) was applied to a holey carbon grid (Quantifoil R1.2/1.3 Cu 200 mesh grids, Microtools GmbH, Berlin, Germany) that had been treated by gold-sputtering[14,15] and glow-discharged. The grid was incubated for 30 s in the chamber of an FEI Vitrobot Mark IV and then blotted for 4.5 s once. Subsequently, 4 more sets of the 3-μl aliquots of the sample were applied, with each incubated for 10 s, and blotted for 5 s. The grid was immediately plunged into liquid ethane, cooled by liquid nitrogen and then transferred into a cryo-electron microscope (CRYO ARM 300, JEOL, Tokyo, Japan) equipped with a cold-field emission gun, an in-column type energy filter, and a direct electron detection camera (K3, Gatan). The microscope was operated at 300 keV and a nominal magnification of 60,000. Inelastic scattering electrons were removed using an energy slit of 20 eV. In total, 7,515 electron counting movies were recorded in a pixel size of 0.8219 Å, under a total electron dose of 70 electrons $Å^{-2}$ for each movie. Each exposure of 4.0 s was dose-fractionated into 50 frames. The nominal defocus range was -0.9 to -1.8 μm. All movies were collected with a JEOL Automatic Data Acquisition System (JADAS)[15].

Cryo-EM image processing

The movie frames were aligned and summed using the motion correction function implemented in RELION3.1[16] to obtain a final dose weighted image. Estimation of the contrast transfer function (CTF) was performed using the CTFFIND-4.1 program[17]. All of the following processes were performed using RELION3.1. For structural analyses of PSII monomer, 1,496,989 particles were automatically picked from 7,515 micrographs, and then subjected to two rounds of

particle screening by reference-free 2D classification and 3D classification without imposing any symmetry. Accordingly, 242,522 particles were selected from good classes of 2D and 3D classifications. The structure of the PSII dimer (PDB: 3WU2)[18] was employed for the initial model for the first 3D classification with a 60-Å low-pass filter. A total of 173,875 particles were selected after the subtraction of detergent micelle and the following 3D classification. Finally, estimation of per-particle defocus and beam tilt parameters were performed using the CTF refinement and then further refined by Bayesian polishing. The map was refined by a post-refinement procedure at an overall resolution of 2.78 Å. The resolution was estimated by the gold standard Fourier shell correlation (FSC) curve with a cut-off value of 0.143[19]. Local resolution was estimated using RELION3.1[16].

Model building and refinement

The 2.78-Å cryo-EM map was used for model building of the PSII monomer. First, one monomer of the PSII dimer structure was manually fitted into the 2.78-Å cryo-EM map using UCSF Chimera[20], and then the peripheral regions and individual chains were inspected and adjusted manually with Coot[21]. The complete PSII monomer structure was then refined with phenix.real_space_refine[22] with geometric restraints for the protein-cofactor coordinations. The final model was further validated with MolProbity[23]. The statistics for all data collection and structure refinement are summarized in Table 1-1. Figures were prepared by Pymol[24], UCSF Chimera[20] and ChimeraX[25]. The root-mean square deviation (RMSD) values and distances were calculated with lsqkab and contact of the CCP4 package[21], respectively.

Table 1-1. Cryo-EM data collection, refinement parameters, and validation statistics for PSII monomer.

Data collection and processing	
Magnification	60,000
Voltage (kV)	300
Electron exposure (e^{-} Å^{-2})	70.22
Defocus range (μm)	-0.9 to -1.8
Pixel size (Å)	0.8219
Symmetry imposed	C1
Initial particle images (no.)	1,496,989
Final particle images (no.)	173,875
Map resolution (Å)	2.78
FSC threshold	0.143
Map resolution range (Å)	2.5 to 25
Refinement	
Initial model used (PDB code)	3WU2
Model resolution (unmasked, Å)	2.79
Model resolution (masked, Å)	2.73
FSC threshold	0.5
Model-sharpening B factor (Å^2)	-74.75
Model composition	
Non-hydrogen atoms	22,845
Protein residues	2,567
Ligands	67
B factors (Å^2)	
Protein	68.77
Ligands	51.63
R.m.s. deviations	
Bond lengths (Å)	0.019
Bond lengths (Å)	1.955
Validation	
Molproblity	1.57
Clashscore	9.41
Rotamer outliers (%)	0.47
Ramachandran plot	
Favored (%)	96.62
Allowed (%)	3.38
Disallowed (%)	0

Results

Overall structure of the PSII monomer

Biochemical analysis of the purified PSII monomer showed that it contains all known cyanobacterial PSII subunits, and a small amount of PSII dimer is present as contamination (Fig. 1-1), but this contamination does not affect the result of our cryo-EM analysis. With this purified sample, images were recorded by the CRYO ARM 300 microscope, and the cryo-EM map was finally reconstituted from 173,875 particles at a resolution of 2.78 Å (Fig. 1-2, 1-3, 1-4A, Table 1-1). The local resolution map is shown in Fig. 1-3B, which shows that most of the regions of the map has a high quality that allows us to assign the side chains of most amino acid residues and other prosthetic groups such as Chls and haems.

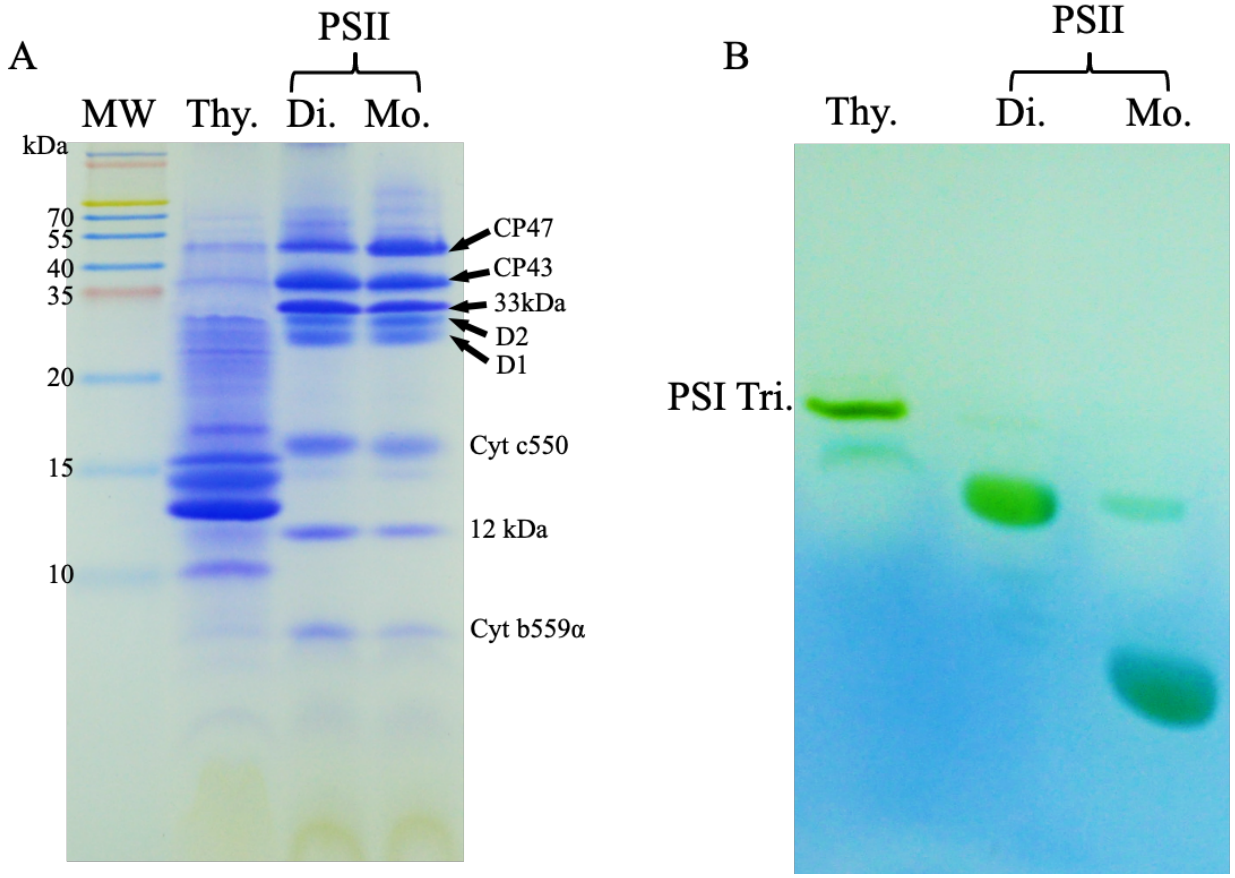


Fig. 1-1. Analyses of the PSII monomer from *T. vulcanus*.

A. SDS-PAGE of *T. vulcanus* thylakoid membrane (Thy.), PSII dimer (Di.), and PSII monomer (Mo.). PSII subunits are indicated by arrows. B. BN-PAGE of *T. vulcanus* thylakoid membrane (Thy.), PSII dimer (Di.), and PSII monomer (Mo.).

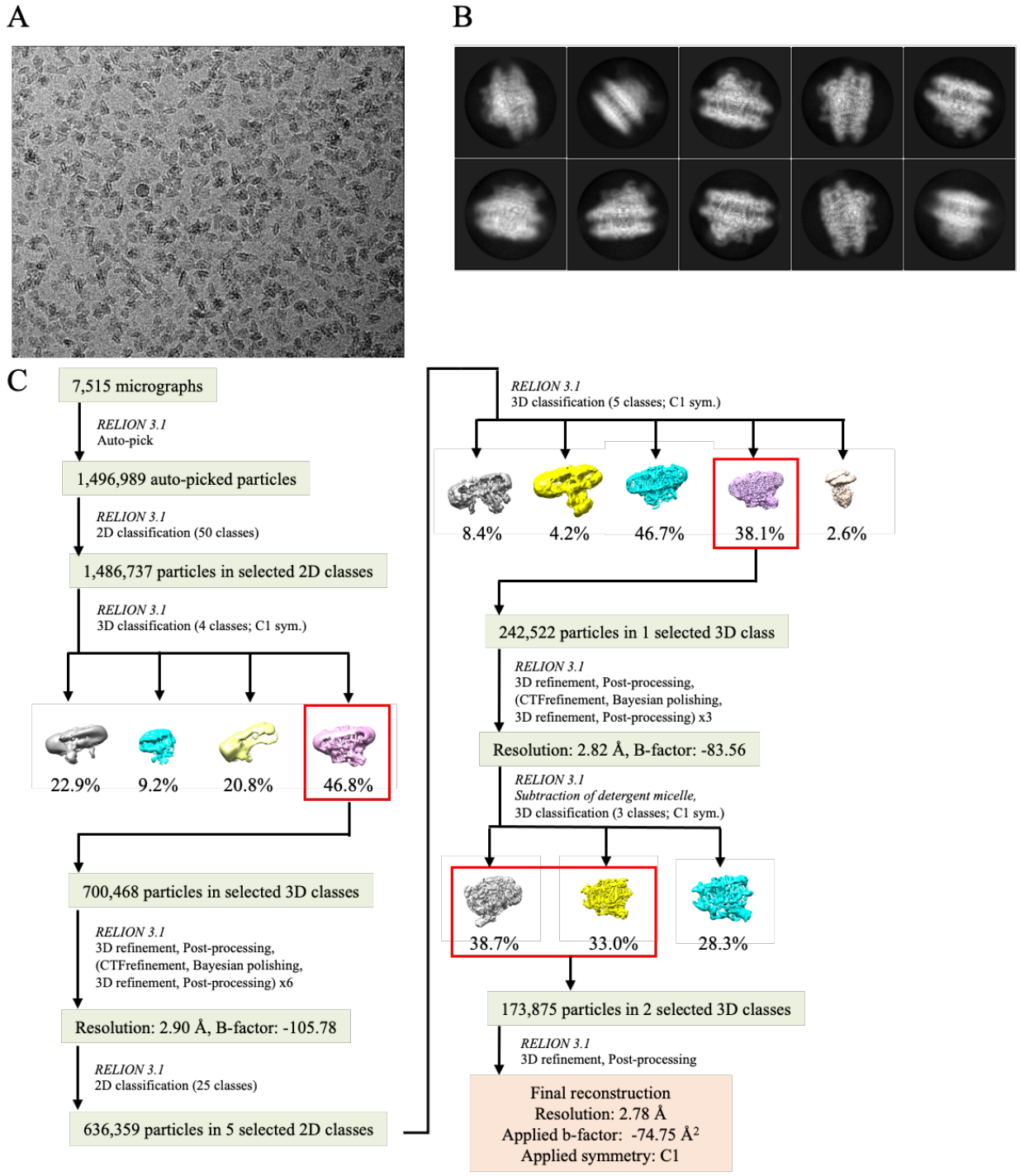


Fig. 1-2. Single particle cryo-EM analysis of PSII monomer.

A. A representative cryo-EM micrograph at a pixel size of 0.82 Å. B. Representative 2D classes for the PSII monomer particles. C. Flowchart for the classification scheme of the PSII monomer.

The PSII monomer structure was reconstructed at 2.78 Å resolution from 173,875 particles. See the Materials and Methods section for further details.

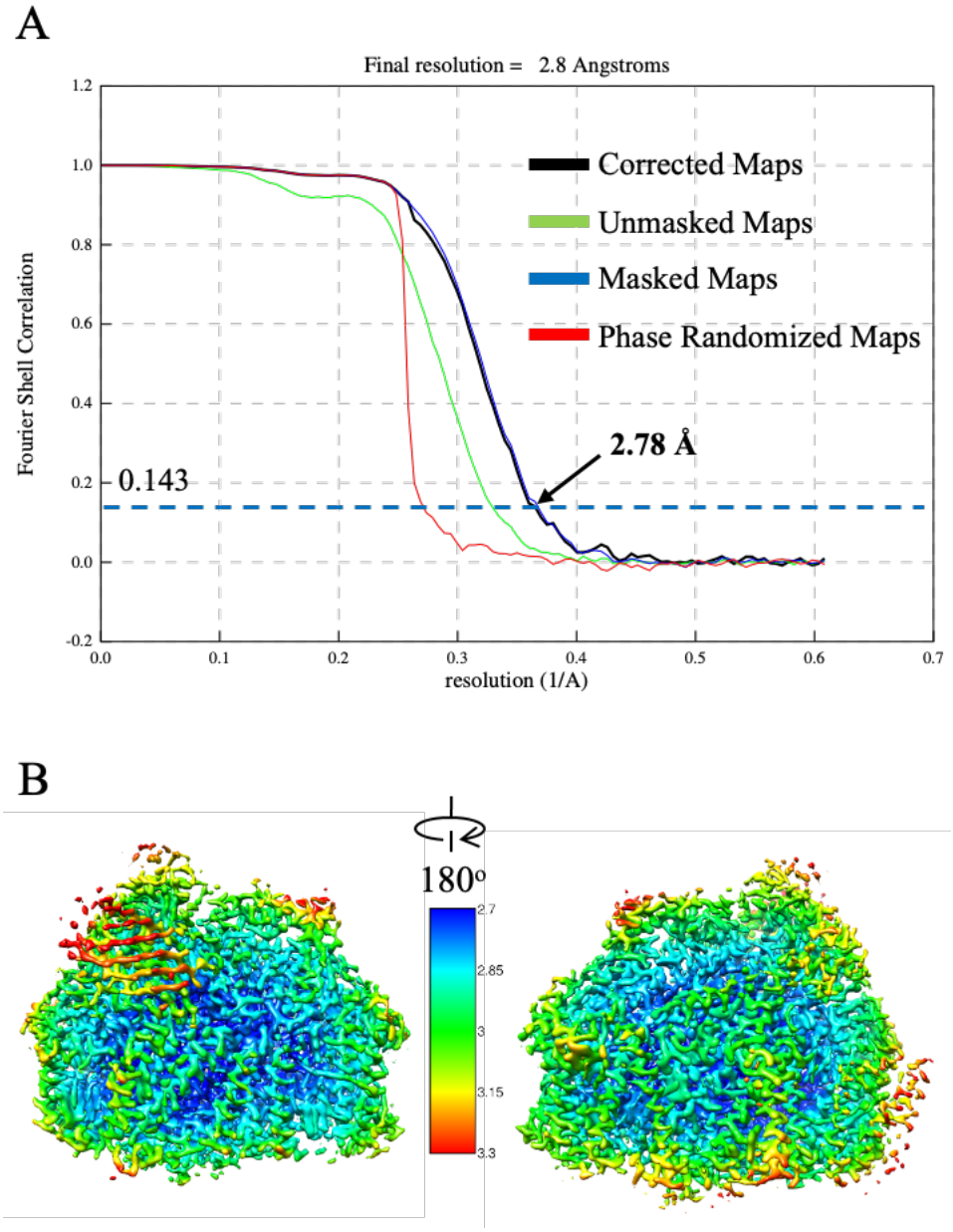


Fig. 1-3. Evaluation of the resolution of the cryo-EM map of PSII monomer.

A. The FSC curve calculated between independently refined half maps of the reconstruction. B. Local resolution maps of the PSII monomer.

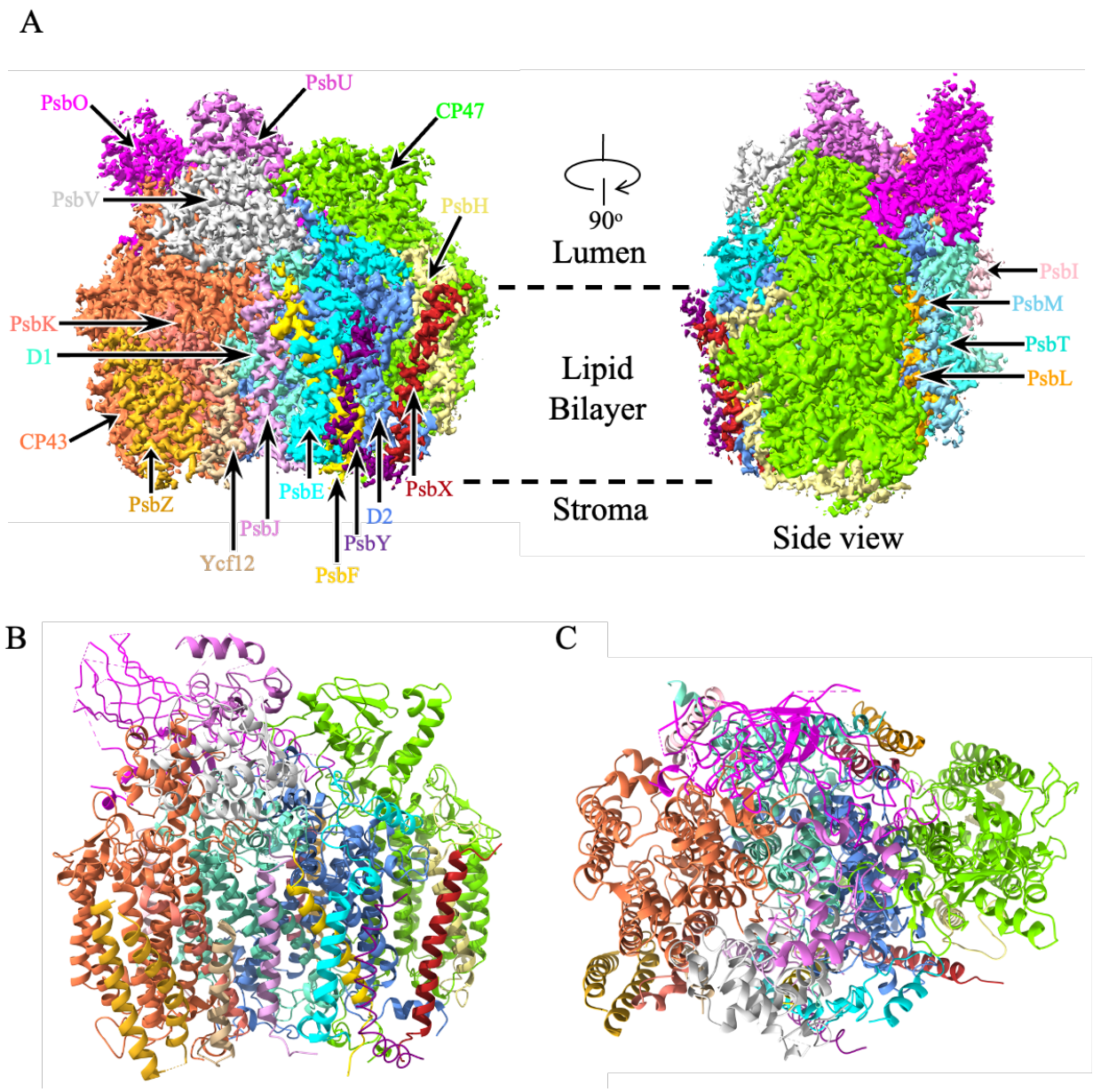


Fig. 1-4. Cryo-EM map (A) and model (B, C) of PSII monomer from *T. vulcanus*.
A, B are side views, and C is the top view from the stromal side. Subunits are indicated by arrows and colored by aquamarine (D1), chartreuse (CP47), coral (CP43), cornflower (D2), cyan (PsbE), gold (PsbF), khaki (PsbH), pink (PsbI), plum (PsbJ), salmon (PsbK), orange (PsbL), skyblue (PsbM), magenta (PsbO), turquoise (PsbT), orchid (PsbU), light gray (PsbV), firebrick (PsbX), purple (PsbY), goldenrod (PsbZ), and tan (Ycf12).

Based on the cryo-EM density map (Figs. 1-5 and 1-6), an overall atomic model of the PSII monomer was built and refined without imposing any symmetry (Fig. 1-4B, C). The resultant PSII

monomer exhibits overall dimensions of ~100 × 80 × 130 Å, and consists of all 20 protein subunits (D1, CP47, CP43, D2, PsbE, PsbF, PsbH, PsbI, PsbJ, PsbK, PsbL, PsbM, PsbO, PsbT, PsbU, PsbV, PsbX, PsbY, PsbZ, and Ycf12) (Table 1-2). The structures of these subunits are very similar to the crystal structure of one monomer of the PSII dimer[18,26] as well as the crystal structure of the PSII monomer (PDB: 3KZI) reported at a lower resolution[27]. However, compared with the high-resolution crystal structure of the PSII dimer, some residues in the monomer are disordered (Table 1-3). On the other hand, the PsbY subunit is retained in the PSII monomer, even though it is slightly disordered owing to the relatively lower resolution in the peripheral regions, making nearly all of the side chains of the amino acid residues in this chain invisible. This subunit is not present in the crystal structure of both PSII dimer[18] and PSII monomer[1] or is only partially present in the crystal structures of Sr-substituted PSII[28] or native PSII[26]. In addition to these protein subunits, there are 35 Chls, two pheophytins, 9 BCRs (one is *β*-cryptoxanthin), 13 lipids, one unknown ligand, two plastoquinones, two haems, and one non-haem iron found in the PSII monomer structure (Table 1-2). Due to the limitations of current resolution, only 16 water molecules are clearly observed in the cryo-EM density map. Among them, seven are coordinated to Chls (Fig. 1-6B), while others are distributed among amino acid residues. The heterocubane cluster Mn_4CaO_5 of the oxygen-evolving center cannot be constructed in the shape of a "distorted chair" as in that of the PSII dimer structure; thus we adopted the structure from the PSII dimer.

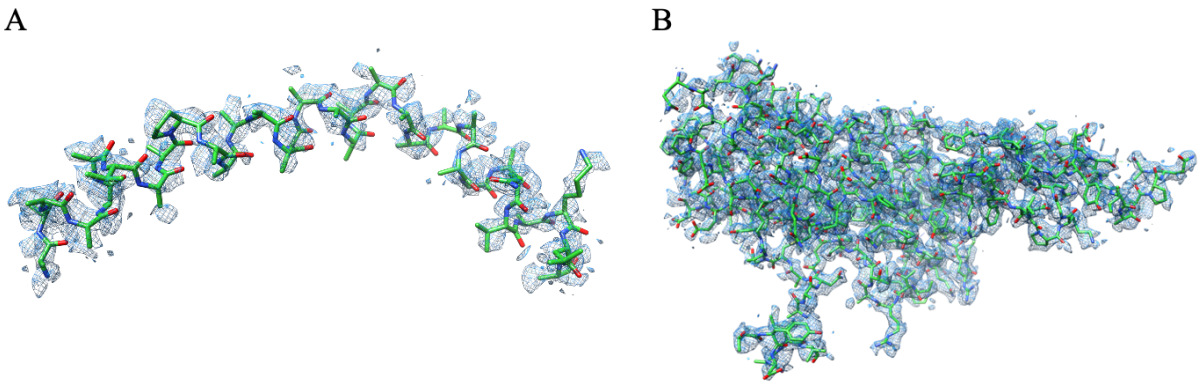


Fig. 1-5. Superposition of the cryo-EM maps and molecular models for some subunits of PSII monomer. A. PsbY. B. PsbO.

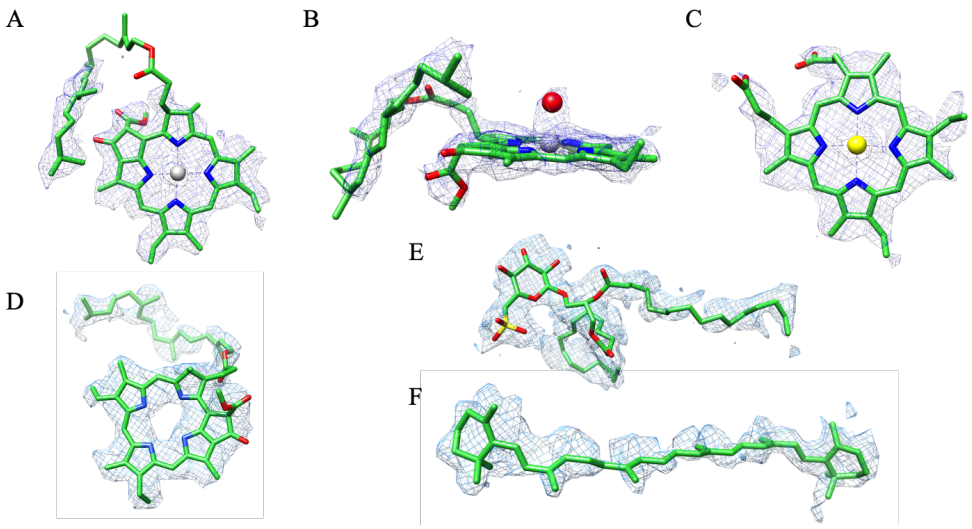


Fig. 1-6. Superposition of cryo-EM maps and molecular models for representative cofactors of PSII monomer.

A. A chlorophyll *a*, B. A chlorophyll *a* coordinated with H_2O molecule, C. Haem, D. A pheophytin *a*, E. An SQDG, F. A *β*-carotene.

Table 1-2. Cofactors in PSII monomer from *T. vulcanus*.

Subunit	Chlorophyll	Carotenoid	Lipid	Others
PsbA	4 Chl *a*	1 BCR	1 SQDG	2 pheophytin
			1 LMG	1 plastoquinone
				1 non-haem iron
				7 water
PsbB	16 Chl *a*	2 BCR	1 LMG	5 water
				1 unknown ligand
PsbC	13 Chl *a*	2 BCR	3 DGD	4 water
PsbD	2 Chl *a*	1 BCR	1 SQDG	1 plastoquinone
			3 LHG	1 bicarbonate ion
			1 LMG	
PsbE				
PsbF				1 HEM
PsbH			1 DGD	1 RRX
PsbI				
PsbJ				
PsbK		2 BCR		
PsbL			1 LHG	
PsbM				
PsbO				
PsbT				
PsbU				
PsbV				1 HEM
PsbX		1 *β*-cryptoxanthin		
PsbY				
PsbZ				
Ycf12				

Table 1-3. Disordered residues (residues those are invisible) of PSII monomer compared with the crystal structure of a single monomer from a PSII dimer.

Chains	Start	End	Chains	Start	End
PsbB	82	87	PsbO	130	135
PsbB	405	413	PsbO	137	141
PsbB	435	436	PsbO	178	181
PsbE	59	62	PsbO	202	203
PsbH	63	64	PsbO	206	209
PsbI	33	36	PsbO	246	246
PsbL	1	7	PsbT	8	8
PsbM	1	3	PsbU	8	8
PsbM	31	33	PsbU	52	57
PsbO	3	5	PsbU	59	60
PsbO	12	15	PsbU	62	70
PsbO	21	28	PsbU	86	86
PsbO	32	37	PsbV	14	16
PsbO	48	49	PsbV	86	87
PsbO	54	68	PsbZ	1	4
PsbO	86	91	PsbZ	30	34
PsbO	110	112	PsbZ	62	62

Comparison between the cryo-EM structure of PSII monomer and the crystal structure of a monomer from the dimer

The cryo-EM structure of PSII monomer was compared with the single monomer of the PSII dimer and shown in Fig. 1-7A. The RMSD value was calculated to be 0.7799 Å, suggesting that the cryo-EM structure of the PSII monomer is very similar to the crystal structure of the monomer within a dimer. However, as mentioned above, the cryo-EM structure of the monomer retained PsbY, a peripheral subunit that is lost or partially retained in the crystal structure of PSII dimer. This suggests that the PsbY is easily detached from PSII, and the cryo-EM structure represents a more physiological state of PSII than that of the crystal structure. A similar situation is seen in the cryo-EM structure of the PSII dimer solved recently at a resolution of 1.95 Å by cryo-EM[27], where the PsbY subunit is also retained.

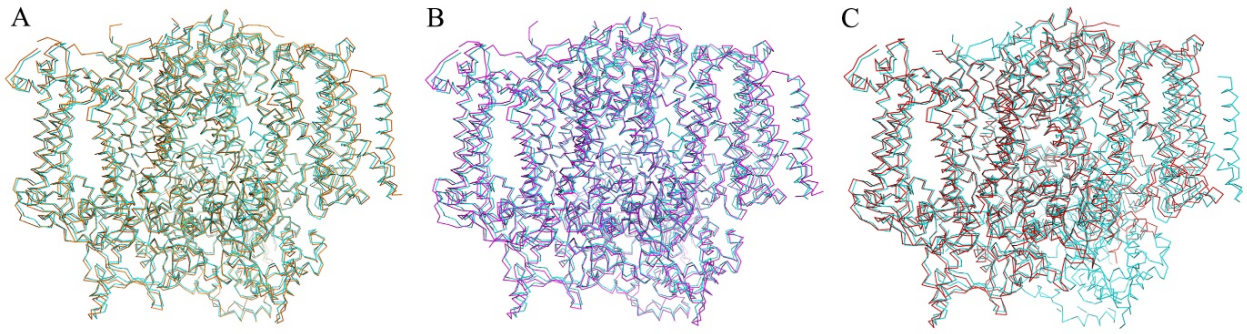


Fig. 1-7. Superposition of the PSII monomer structure in the present study (PDB code: 7EDA, cyan) with a single monomer of the PSII dimer (PDB code: 3WU2, orange) (A), the crystal structure of PSII monomer (PDB code: 3KZI, magenta) (B), and the cryo-EM structure of the apo-PSII monomer (PDB code: 6WJ6, red) (C). The structures are depicted in ribbons.

Differences are found in some regions between the cryo-EM structure of the monomer and the crystal structure of the dimer. There are a number of amino acid residues that are disordered and not visible in the cryo-EM structure, which include 82 to 87, 405 to 413, 435 to 436 of PsbB, 59 to 62 of PsbE, and several regions in the PsbH, PsbI, PsbL, PsbM, PsbT, PsbU, PsbV and PsbZ (Table 1-3). In particular, many residues of PsbO are disordered and not visible in the cryo-EM structure. Most of these residues are located at the periphery regions of PSII, suggesting that they may have more flexibility in solution than those in the crystals. Alternatively, since PsbO of one monomer interacts with CP47 from the adjacent monomer in the dimer structure[18,26], absence of CP47 from the adjacent monomer may destabilize the structure of PsbO in the other monomer, making it flexible and not visible in the cryo-EM structure. This suggests the involvement of PsbO in the dimerization of PSII, as has been reported previously[29].

The PSII monomer contains eight BCRs and one *β*-cryptoxanthin, which is two BCRs less than those in the structure of the single monomer from the PSII dimer (Fig. 1-8A, B, Table 1-1). The BCR (BCR101T) that is present in the monomer structure of the PSII dimer is lost in the cryo-EM structure of the monomer. This BCR is located in the monomer-monomer interface, and participates in the monomer-monomer interactions in the dimer (Fig. 1-8B-D). As can be seen in Fig. 1-8D, this BCR has a number of hydrophobic interactions with a cluster of three BCRs from another monomer with distances in the range of 3.5 - 4.5 Å. The loss of this BCR will therefore weaken the monomer-monomer interactions within the dimer. In addition, a large part of a BCR molecule (BCR619) visible in the dimer structure was not visible in the cryo-EM density, thus

only a small part of this BCR is modelled as an unknown molecule (UNL/619/B) in the structure of the monomer (Fig. 1-8A, B).

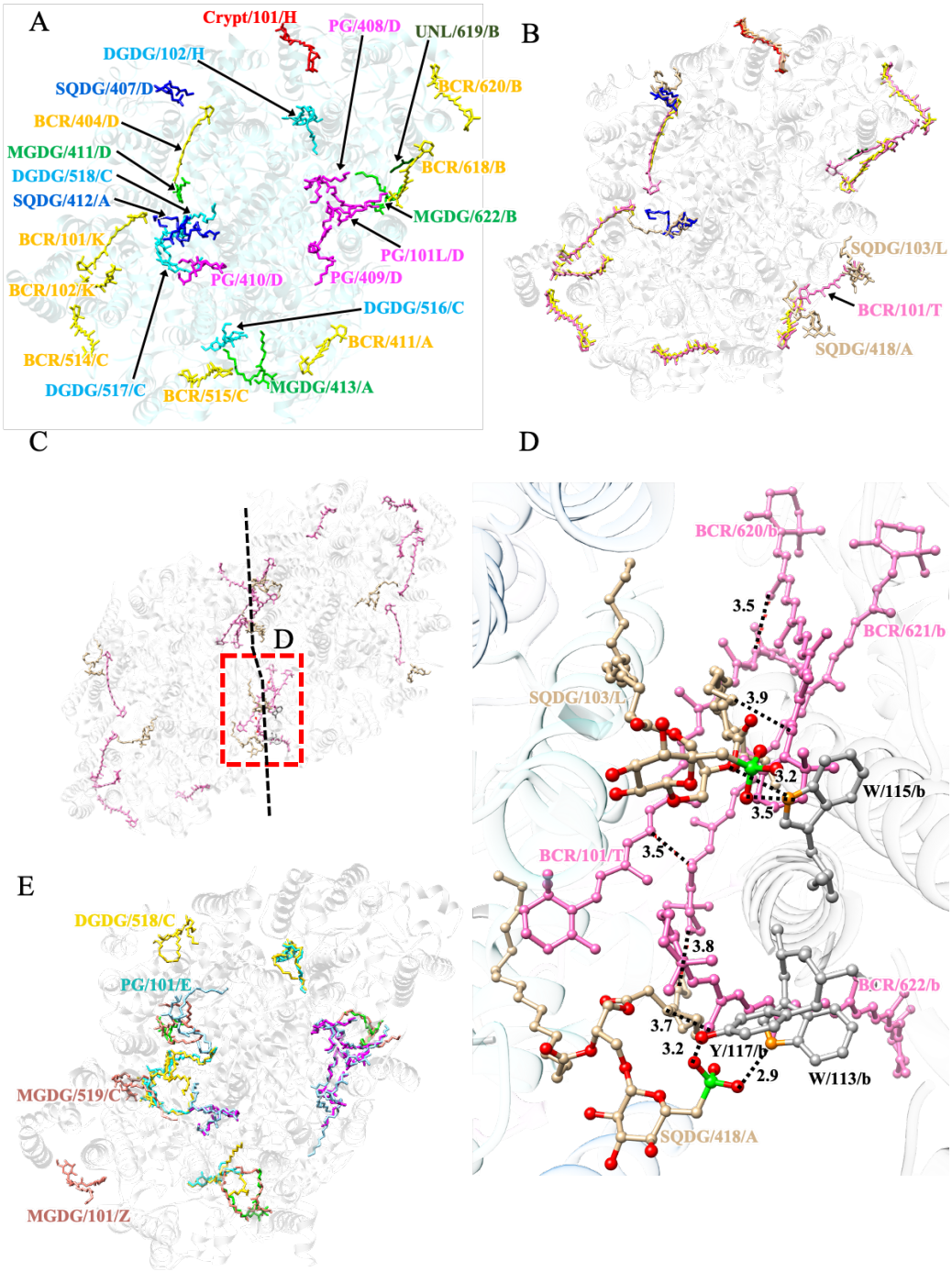


Fig. 1-8. Cofactor distribution in the PSII monomer and comparisons with the single monomer from the PSII dimer (3WU2). A. Distribution of BCR and lipids in the PSII monomer of the present study. Color codes: BCR, yellow; *β*-cryptoxanthin, red; MGDG, green; DGDG, cyan; SQDG, deep

blue; PG, magenta; UNL, dark green. The proteins are represented by cartoons in pale cyan. B. Superposition of BCR, *β*-cryptoxanthin and SQDG between the structures of PSII monomer and the single monomer of the crystal structure of PSII dimer (3WU2). BCRs, SQDGs from the PSII dimer structure are colored in hot pink and tan, respectively. One BCR and two SQDGs that are present in the monomer-monomer interface of the dimer but absent in the present monomer structure are labeled. C. Structure of the PSII dimer, showing the monomer-monomer interface within the dimer (dashed black line). D. An enlarged view of the boxed area in panel C, showing interactions of BCR and SQDG in one monomer with the BCR and the residues from another monomer. BCRs, SQDGs and residues of the adjacent monomer that interact with the two SQDGs are labeled and depicted in hot pink, tan and gray, respectively. Distances between two atoms that have major interactions are labeled. E. Superposition of MGDG, DGDG and PG in the present monomer structure with the single monomer structure from the PSII dimer (3WU2). MGDG, DGDG and PG of the present monomer structure are colored as in panel A, and those from the PSII dimer structure are colored in salmon, gold and light teal, respectively. Molecules that are found only in the dimer structure are labeled.

Table 1-4. Differences in the number of cofactors among the structures of monomers.

PDB ID	7EDA	3WU2	3KZI	6WJ6
Species	*T. vulcanus*	*T. vulcanus*	*T. elongatus*	*S.* sp. PCC6803
Method	Cryo-EM	X-ray	X-ray	Cryo-EM
β-carotenes	8 + 1*	10 + 1*	11	9
SQDGs	2	4	4	7
MGDGs	3	5	8	3
DGDGs	4	5	8	2
PGs	4	5	2	4

*One is *β*-cryptoxanthin.

Another notable difference between the current monomer structure and the crystal structure of the PSII dimer is that, the current PSII monomer contains two SQDGs, whereas each monomer of the PSII dimer contains four SQDGs (Fig. 1-8A and B, Table 1-4). The two SQDGs that are lost in the cryo-EM structure of the monomer are also located in the monomer-monomer interface and participate in the monomer-monomer interactions in the PSII dimer (Fig. 1-8C, D). Among

them, the head of SQDG103L is hydrogen-bonded to Trp115 of PsbB from the adjacent monomer, whereas the head of SQDG418A is hydrogen-bonded to Trp113 and Tyr117 of PsbB of the adjacent monomer. The lack of these two SQDGs, together with the lack of the BCR molecule mentioned above, will weaken the monomer-monomer interactions significantly, thus making the monomer unable to form the dimer. In addition, there are some β-DDM and unknown molecules located in the monomer-monomer interface in the crystal structure of the PSII dimer (not shown), but these molecules are lost in the cryo-EM structure of the PSII monomer. Some of these molecules may also represent lipids in the native PSII dimer, and their loss will also weaken the monomer-monomer interactions.

The current PSII monomer structure contains three distearoylmonogalactosyl diglyceride (MGDG), four digalactosyldiacyl glycerol (DGDG) and four PG molecules (Fig. 1-8A, Table 1-4 and 1-2). Compared with the single monomer of the dimer, the current monomer lost two MDGD, one DGDG and one PG molecules (Fig. 1-8E, Table 1-4). These molecules are either located in the periphery regions of the monomer or sites far from the monomer-monomer interface in the dimer; thus, they are not related with the dimer formation.

Comparisons with other monomer structures

For comparisons with other monomer structures, superpositions of the current cryo-EM structure of monomer with the crystal structure of the PSII monomer (3KZI)[1] and the cryo-EM structure of the apo-PSII monomer (6WJ6)[30] were shown in Fig. 1-7B and C, respectively. The RMSD values were calculated to be 1.0308 Å, and 1.4948 Å, respectively, between the chains that are present in the two structures compared. These values suggest that the overall structure among the cryo-EM structures of different monomers and the crystal structure of the PSII monomer are largely similar, although the cryo-EM structure of the apo-PSII monomer lacks three extrinsic proteins and three small membrane-spanning subunits, which are omitted from its comparison with the current cryo-EM structure.

The differences concerning the position of BCRs and SQDGs between the cryo-EM structure and the crystal structure of the PSII monomer are summarized in Fig. 1-9A and Table 1-4. One more BCR (BCR115) is found in the crystal structure of the monomer; however, this BCR is located near the stromal side in the vicinity of PsbJ instead of the BCR located near the monomer-monomer interface in the crystal structure of the PSII dimer. This BCR may account for the

photoprotection against overexcitation of the photosynthetic system[31], or structural stability of phycobilisomes[6], but is not related with the dimer formation due to its far distance.

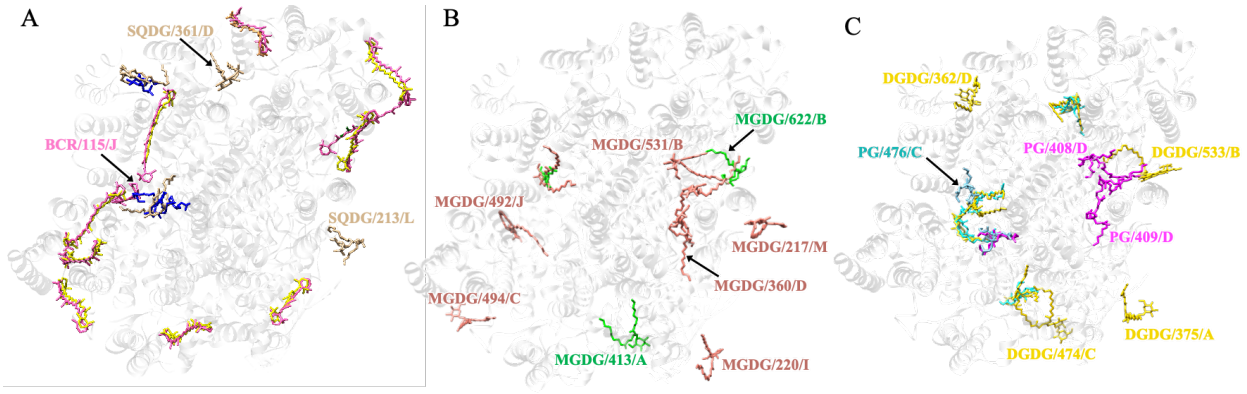


Fig. 1-9. Comparison of the present cryo-EM structure of PSII monomer with the crystal structure of the monomer (3KZI).

A. Superposition of BCR, SQDG and the unknown molecule in the present monomer structure with those of the crystal structure (3KZI). BCRs, SQDGs and the unknown molecule of the present monomer structure are colored the same as in Fig. 4A, and BCRs, SQDGs from the PSII monomer of the crystal structure (3KZI) are colored in hot pink and tan, respectively. The BCR and SQDG molecules that are present in the crystal structure but absent in the cryo-EM structure of the monomer are labeled. B. Superposition of MGDG in the present monomer structure with those of the crystal structure (3KZI). MGDG of the present monomer structure are colored in green, and those from the PSII monomer of the crystal structure (3KZI) are colored in salmon, respectively. The MGDG molecules that are located in the monomer-monomer interface in the crystal structure of the monomer, as well as some typical molecules that are different between the present cryo-EM structure (7EDA) and the crystal structure (3KZI), are labeled. C. Superposition of DGDG and PG in the present monomer structure with those of the crystal structure (3KZI). DGDG and PG of the present monomer structure are colored in cyan and magenta, and those from the PSII monomer of the crystal structure (3KZI) are colored in gold and light teal, respectively. The DGDG and PG molecules that are located in the monomer-monomer interface in the crystal structure of the monomer, as well as some typical molecules that are different between the present cryo-EM structure (7EDA) and the crystal structure (3KZI), are labeled.

Two more SQDG (SQDG361D and SQDG213L) are also found in the crystal structure of the PSII monomer (3KZI), whose negatively charged headgroups are facing the stromal surface (Fig. 1-9A, Table 1-4). Among these additional SQDGs, one (SQDG213L) is located near the monomer-monomer interface similar to one of the two SQDGs seen in the crystal structure of the PSII dimer (Fig. 1-8B and C), and thus may be involved in the dimer formation. The other SQDG (SQDG361D) is located far from the monomer-monomer interface, thus it is not related with the dimer formation. Instead, this SQDG may assist the attraction of protons from the stroma and affect the exchange process of free plastoquinones[3,32], as it is rather close to Q_B.

There are eight MGDGs, eight DGDGs and two PGs, in the crystal structure of the PSII monomer (3KZI) (Fig. 1-9B and C, Table 1-4). The numbers of MGDG and DGDG are much more than those in the cryo-EM structure of the monomer or the crystal structure of the PSII dimer, whereas the number of PG is two less than that in the cryo-EM structure of the PSII monomer or three less than the crystal structure of the dimer. Among these molecules, two MGDGs (MGDG492 and MGDG494) are located in the far periphery of the PSII monomer, and one of them (MGDG217) is located in the adjacent monomer structure of the PSII dimer. Among the two PGs found, one is the same as that of the cryo-EM structure whereas the other one (PG476) is found in a new position. The DGDG molecules only found in the crystal structure of the PSII monomer are also not located in the monomer-monomer interface (Fig. 1-9C). Whether these molecules are indeed present, and if yes, what functions they may have, need to be examined further, as the crystal structure of the PSII monomer is solved at a relatively lower resolution so that there may be ambiguities regarding the presence and locations of these molecules. Some of these additional lipids have much shorter acyl chains, suggesting that they may not be independent lipid molecules but rather detergent molecules or a part of the Chl phytol tails. Indeed, the number of lipid molecules assigned in the crystal structure of the monomer is much larger than those found in the high-resolution crystal structure[18] or cryo-EM structure of the PSII dimer[27], suggesting that at least some of the lipid molecules assigned in the monomer may not represent real lipids.

Comparison of the cryo-EM structure of the native PSII monomer from *T. vulcanus* with the cryo-EM structure of apo-PSII monomer from *Synechocystis* sp. PCC 6803 revealed another different situation. There are nine BCRs and seven SQDGs in the Synechocystis apo-PSII monomer (Fig. 1-10A and B, Table 1-4); among them, one BCR (BCR101K) that is present in the current monomer structure but lost in the apo-PSII monomer structure is located close to PsbK and

not related with the dimer formation. The BCR molecule that is found in the PSII dimer but lost in the cryo-EM structure of the current monomer was also absent in the cryo-EM structure of the apo-PSII monomer, indicating its importance for the formation of the PSII dimer. However, three SQDGs (101 T, 103 T and 409A) are newly appeared in the monomer-monomer interface (Fig. 1-10B), among which two (SQDG103T and SQDG409A) are in the similar positions with the two SQDGs of the dimer structure (Fig. 1-10B). These suggest that these SQDGs may not be involved in the dimer formation, as PSII forms a monomer even these SQDGs are present.

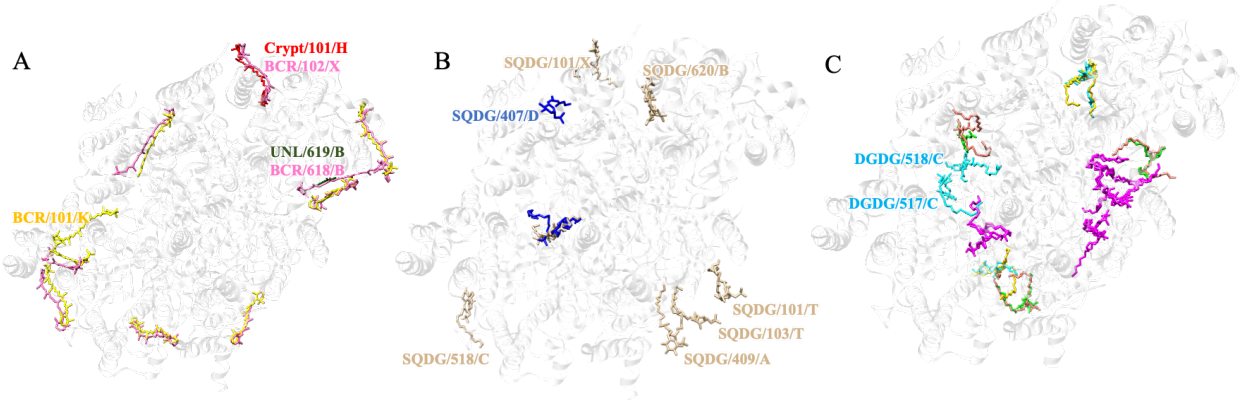


Fig. 1-10. Comparison of the present monomer structure with the cryo-EM structure of the apo-PSII monomer (6WJ6).

A. Superposition of BCR (including a *β*-cryptoxanthin and an unknown molecule) in the present monomer structure with the cryo-EM structure of the apo-PSII monomer. BCRs, *β*-cryptoxanthin and the unknown molecule of the present monomer structure are colored the same as in Fig. 4A, and BCRs from the cryo-EM structure of the apo-PSII monomer (6WJ6) are colored in hotpink. B. Superposition of SQDG in the present monomer structure with those of the cryo-EM structure of the apo-PSII monomer. SQDGs of the present monomer structure are colored in deep blue, and those from the cryo-EM structure of the apo-PSII monomer are colored in tan. One SQDG molecule only present in the current monomer structure and six SQDG molecules only present in the apo-PSII monomer structure are labeled. C. Superposition of MGDG, DGDG and PG in the present monomer structure with those of the cryo-EM structure of the apo-PSII monomer. MGDG, DGDG and PG of the present monomer structure are colored in green, cyan and magenta, and those from the cryo-EM structure of the apo-PSII monomer are colored in salmon, gold and light

teal, respectively. Two DGDG molecules found only in the present monomer structure (7EDA) are labeled.

The number of MGDG, PG in the apo-PSII monomer is the same as that in the cryo-EM structure of the native PSII monomer, whereas the number of DGDG is two in the apo-PSII monomer, which is two less than that of the native PSII monomer (Fig. 1-10C, Table 1-4). The two additional DGDGs in the native PSII monomer are far away from the monomer-monomer interface, thus they are not related with the dimer formation.

Discussions

PSII functions as a dimer in vivo, however, PSII monomer also exists in vivo in various amounts[10,33-36], and it was even reported that the native form of PSII may be a monomer[36]. One role of the PSII monomer is to serve as an intermediate or precursor during the assembly and damage/repair processes of the dimer, which may proceed from monomer to dimer and vice versa. Alternatively, the PSII monomer may be present in the regions of thylakoid membranes different from that of the dimer in vivo. However, it is not clear how dimer forms and what factors contribute to the dimer formation.

Our structural analysis of the PSII monomer revealed important differences in the monomer-monomer interface between the structures of the monomer and dimer. One BCR (BCR101T) and two SQDGs (SQDG103L and SQDG418A) are present in the monomer-monomer interface of the crystal structure of the dimer solved at a high resolution[18,26], which are also conserved in the cryo-EM structure of the PSII dimer reported recently at a high resolution[27]. However, these molecules are lost in the present monomer structure. The BCR molecule forms extensive hydrophobic interactions with a cluster of three BCR molecules from the adjacent monomer. Indeed, BCR has been reported to play important roles in mediating formation of the dimer[5-7] as well as in the assembly of PSII[37-39]. This indicates that the BCR molecule is required for the formation of the PSII dimer, and suggests that BCR is not only important for energy transfer/dissipation but also plays a role in mediating molecular interactions.

The two head groups and acyl tails of the two SQDGs (SQDG103L and SQDG418A) that are present in the dimer but lost in the monomer form multiple hydrogen-bonds and hydrophobic interactions with the amino acid residues and BCR101 from the adjacent monomer. These suggest

important roles that these SQDGs may play in the formation of the dimer either directly or indirectly through the BCR molecule. The absence of these SQDGs in the monomer structure thus may account for the failure of dimer formation. This also agrees with reports that deletion of an SQDG synthesis gene resulted in the depletion of SQDG in the photosynthetic membrane and an increase of PSII monomer[3,40].

Comparison of the present monomer structure with the crystal structure of the native PSII monomer[1] or the cryo-EM structure of the apo-PSII monomer[30] revealed some different situations. The BCR molecule is also absent in these structures, illustrating its importance in the formation of the dimer. However, one additional SQDG (SQDG213L), two additional MGDG (MGDG217M and MGDG220I) and one additional DGDG (DGDG375A) molecules are found in the monomer-monomer interface of the crystal structure of the PSII monomer, and three additional SQDGs (SQDG101T, SQDG103T, SQDG409A) are found in the monomer-monomer interface of the cryo-EM structure of the apo-PSII monomer. These suggest that these lipid molecules may not play important roles in the formation of the dimer, as PSII forms a monomer even in their presence. However, cautions are required in explaining these results. The crystal structure of PSII monomer is solved from a closely related cyanobacterium *T. elongatus*, thus its lipid distribution should be the same as the cryo-EM structure of the monomer solved in the present study. However, the resolution of the crystal structure is 3.6 Å, which is not enough to distinguish among different lipid molecules, and some of the lipid molecules may be assigned based on the dimer structure reported previously. Thus, it needs to be confirmed if these additional lipid molecules are indeed present in the monomer-monomer interface of the monomer. Indeed, the crystal structure of the monomer contained even more lipid molecules than each monomer of the dimer (Table 1-4), suggesting that some of the lipid molecules assigned may not represent true lipids. On the other hand, the cryo-EM structure of the apo-PSII monomer is solved from another cyanobacterium *Synechocystis* sp. PCC 6803, which lacks the three extrinsic proteins and the oxygen-evolving activity. Whether the differences found in the distribution of lipids in the two cryo-EM structures of the monomers are due to species differences and/or differences in the PSII composition need to be examined in the future.

In addition to the above differences, there are a number of *β*-DDM molecules and some unknown molecules located in the monomer-monomer interface, especially in the high-resolution crystal structure of the PSII dimer[18]. Some of these molecules should represent native lipids in

vivo, and may therefore contribute to the formation of the dimer. Most of these molecules are not found in the cryo-EM structure of the PSII monomer, which may account partly for the failure of dimer formation.

There are some additional differences in the distributions of BCRs and lipids in the regions not related to the dimer formation among the four types of structures compared. The functional significances of these differences will require further examinations in the future.

Another factor contributing to the dimerization is PsbO, whose function has been elaborated with a PsbO knockout mutant[29]. PsbO has some interactions with the CP47 subunit of the other monomer, illustrating its importance in stabilizing the dimer[18,26,29,41]. The structure of PsbO exhibits a higher flexibility compared with the other subunits in the present monomer (Table 1-5), suggesting that it may be due to the absence of the other monomer. The structure of the apo-PSII monomer also supports this suggestion[30], where some *β*-sheet secondary structure of the soluble domain of CP47 was lacking due to the loss of extrinsic subunits, especially PsbO.

References

1. Broser, M. *et al.* Crystal structure of monomeric photosystem II from *Thermosynechococcus elongatus* at 3.6-Å resolution. *J Biol Chem* **285**, 26255-26262 (2010).
2. Huang, G. *et al.* Structural insights into a dimeric Psb27-photosystem II complex from a cyanobacterium *Thermosynechococcus vulcanus*. *Proc Natl Acad Sci USA* **118** (2021).
3. Nakajima, Y. *et al.* Thylakoid membrane lipid sulfoquinovosyl-diacylglycerol (SQDG) is required for full functioning of photosystem II in *Thermosynechococcus elongatus*. *J Biol Chem* **293**, 14786-14797 (2018).
4. Guskov, A. *et al.* Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. *Nat Struct Mol Biol* **16**, 334-342 (2009).
5. Zakar, T., Laczko-Dobos, H., Toth, T. N. & Gombos, Z. Carotenoids assist in cyanobacterial photosystem II assembly and function. *Front Plant Sci* **7** (2016).
6. Toth, T. N. *et al.* Carotenoids are essential for the assembly of cyanobacterial photosynthetic complexes. *Biochim Biophys Acta Bioenerg* **1847**, 1153-1165 (2015).
7. Sozer, O. *et al.* Involvement of carotenoids in the synthesis and assembly of protein subunits of photosynthetic reaction centers of *Synechocystis* sp. PCC 6803. *Plant Cell Physiol* **51**, 823-835 (2010).

8. Shen, J. R., Ikeuchi, M. & Inoue, Y. Stoichiometric association of extrinsic cytochrome c550 and 12 kDa protein with a highly purified oxygen-evolving photosystem II core complex from *Synechococcus vulcanus*. *FEBS Lett* **301**, 145-149 (1992).
9. Koike, H., Hanssum, B., Inoue, Y. & Renger, G. Temperature-dependence of S-state transition in a thermophilic cyanobacterium, *Synechococcus-vulcanus* copeland measured by absorption changes in the ultraviolet region. *Biochim Biophys Acta* **893**, 524-533 (1987).
10. Shen, J. R. & Kamiya, N. Crystallization and the crystal properties of the oxygen-evolving photosystem II from *Synechococcus vulcanus*. *Biochemistry* **39**, 14739-14744 (2000).
11. Shen, J. R. & Inoue, Y. Binding and functional properties of two new extrinsic components, cytochrome c-550 and a 12-kDa protein, in cyanobacterial photosystem II. *Biochemistry* **32**, 1825-1832 (1993).
12. Ikeuchi, M. & Inoue, Y. A new 4.8-kDa polypeptide intrinsic to the PS II reaction center, as revealed by modified SDS-PAGE with improved resolution of low-molecular-weight proteins. *Plant Cell Physiol* **29**, 1233-1239 (1988).
13. Wittig, I., Braun, H. P. & Schagger, H. Blue native PAGE. *Nat Protoc* **1**, 418-428 (2006).
14. Maki-Yonekura, S. *et al.* Hexameric and pentameric complexes of the ExbBD energizer in the Ton system. *Elife* **7** (2018).
15. Zhang, J. J. *et al.* JADAS: A customizable automated data acquisition system and its application to ice-embedded single particles. *J Struct Biol* **165**, 1-9 (2009).
16. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7** (2018).
17. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221 (2015).
18. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55-60 (2011).
19. Scheres, S. H. W. & Chen, S. X. Prevention of overfitting in cryo-EM structure determination. *Nat Methods* **9**, 853-854 (2012).
20. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J comput chem* **25**, 1605-1612 (2004).
21. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).

22. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
23. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
24. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.
25. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).
26. Suga, M. *et al.* Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99-U265 (2015).
27. Kato, K. *et al.* High-resolution cryo-EM structure of photosystem II reveals damage from high-dose electron beams. *Commun Biol* **4** (2021).
28. Koua, F. H. M., Umena, Y., Kawakami, K. & Shen, J.-R. Structure of Sr-substituted photosystem II at 2.1 Å resolution and its implications in the mechanism of water oxidation. *Proc Natl Acad of Sci USA* **110**, 3889-3894 (2013).
29. Lundin, B. *et al.* Towards understanding the functional difference between the two PsbO isoforms in Arabidopsis thaliana--insights from phenotypic analyses of psbo knockout mutants. *Photosynth Res* **98**, 405-414 (2008).
30. Gisriel, C. J. *et al.* Cryo-EM structure of monomeric photosystem II from *synechocystis* sp. pcc 6803 lacking the water-oxidation complex. *Joule* (2020).
31. Latowski, D., Kuczynska, P. & Strzalka, K. Xanthophyll cycle - a mechanism protecting plants against oxidative stress. *Redox Rep* **16**, 78-90 (2011).
32. Van Eerden, F. J., Melo, M. N., Frederix, P. W. J. M., Periole, X. & Marrink, S. J. Exchange pathways of plastoquinone and plastoquinol in the photosystem II complex. *Nat Commun* **8** (2017).
33. Kuhl, H. *et al.* Towards structural determination of the water-splitting enzyme - Purification, crystallization, and preliminary crystallographic studies of photosystem II from a thermophilic cyanobacterium. *J Biol Chem* **275**, 20652-20659 (2000).
34. Watanabe, M., Iwai, M., Narikawa, R. & Ikeuchi, M. Is the photosystem II complex a monomer or a dimer? *Plant Cell Physiol* **50**, 1674-1680 (2009).

35. Kern, J. *et al.* Purification, characterisation and crystallisation of photosystem II from Thermosynechococcus elongatus cultivated in a new type of photobioreactor. *Biochim Biophys Acta Bioenerg* **1706**, 147-157 (2005).
36. Takahashi, T. *et al.* Photosystem II complex in vivo is a monomer. *J Biol Chem* **284**, 15598-15606 (2009).
37. Telfer, A. What is beta-carotene doing in the photosystem II reaction centre? *Philos Trans R Soc Lond B Biol Sci* **357**, 1431-1439 (2002).
38. Nixon, P. J., Michoux, F., Yu, J. F., Boehm, M. & Komenda, J. Recent advances in understanding the assembly and repair of photosystem II. *Ann Bot* **106**, 1-16 (2010).
39. Gruszecki, W. I. & Strzalka, K. Carotenoids as modulators of lipid membrane physical properties. *Biochim Biophys Acta* **1740**, 108-115 (2005).
40. Mizusawa, N. & Wada, H. The role of lipids in photosystem II. *Biochim Biophys Acta Bioenerg* **1817**, 194-208 (2012).
41. Zhu, Q. J. *et al.* Function of PsbO-Asp158 in photosystem II: effects of mutation of this residue on the binding of PsbO and function of PSII in *Thermosynechococcus vulcanus*. *Photosynth Res* **146**, 29-40 (2020).

CHAPTER 2

Purification and analysis of D1-V185T mutant photosystem II dimer from the thermophilic cyanobacterium *Thermosynechococcus elongatus*

Abstract

Photosystem II (PSII) uses light energy to drive a series of electron transfer reactions, leading to the splitting of water and production of molecular oxygen. Mn_4CaO_5 cluster is the active site of PSII in water oxidation, and proceeds through five sequential oxidation states (S_0 to S_4) to catalyze the water-oxidizing reaction. The D1-V185 residue is a residue of the D1 subunit of PSII reaction center, and has been shown to be important in the PSII water oxidation reaction. Directed mutation of this residue to threonine (V185T) has been shown to perturb the oxidation state transition of the Mn_4CaO_5 cluster, resulting in a decrease in oxygen evolution. In this work, we purified and characterized PSII dimers from a D1-V185T mutant of the thermophilic cyanobacterium *Thermosynechococcus elongatus* (*T. elongatus*). The resulting PSII dimers were found to have an oxygen-evolving activity significantly lower than that of the wild-type strain. Particle images of the purified sample were examined using negative staining electron microscopy, which showed that they are indeed dimers. We will use cryo-electron microscopy to analyze its structure in order to reveal the structural changes caused by the mutation that causes the significant decrease in oxygen evolution.

Introduction

Oxygen-evolving complex (OEC), also known as the water-splitting complex, is the catalytic center of photosystem II (PSII) for water oxidation, which is consisted of a Mn_4CaO_5 cluster. The accurate positions of metal atoms in the dark stable state are well defined in a "distorted chair" shape by a 1.9-Å resolution PSII dimer crystal structure published in 2011[1]. A "damage-free" structure was analyzed by serial femtosecond rotational crystallography (SF-ROX) in 2015[2], in which the atomic distances in OEC are further refined.

The water oxidation reaction catalyzed by OEC proceeds according to the Kok cycle[3] (Fig. 2-1), which shows that OEC is sequentially oxidized following four single flashes so that it cycles through five redox states denoted S_n (n = 0 - 4). Among these S-states, S_1-state is the dark-stable state and S_4-state is a transient state which simultaneously transits to the S_0-state with the concomitant release of molecular oxygen. EPR measurements indicated that the S_2-state exhibits two different ground state configurations[4,5], a low spin (S_2^{LS}) S = 1/2 configuration and a high spin (S_2^{HS}) S = 5/2 configuration. In the S_2^{LS} configuration, the net oxidation states of the 4 Mn ions are considered to be: $Mn4^{IV}$, $Mn3^{IV}$, $Mn2^{IV}$, $Mn1^{III}$. In this structure, Mn4 and Mn3 are connected by a di-μ-oxo bridge involving O4 and O5, whereas Mn1, Mn2, Mn3, Ca and O1, O2, O3 and O5 form an open cubane structure, in which the distance between O5 and Mn1 is longer than that in high spin state[6,7], see Fig. 2-2.

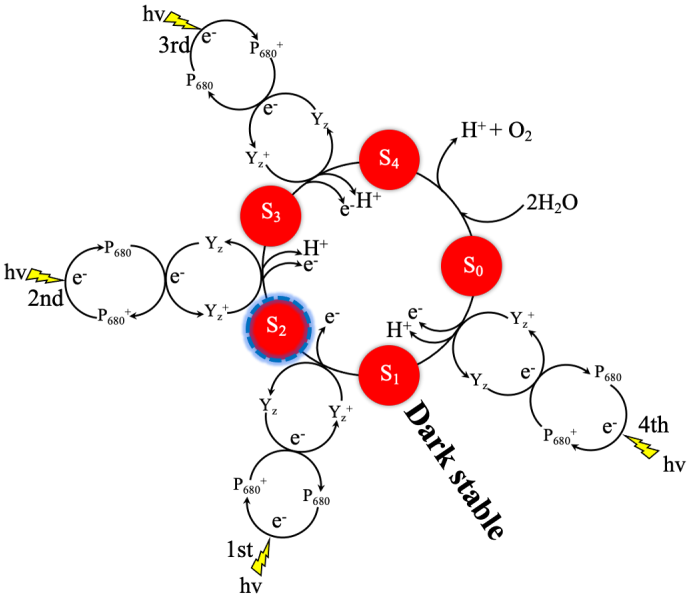


Fig. 2-1. S_n-state in the Kok cycle of the water oxidation reaction, and electron and proton transfer reactions accompanying the S_n-state transitions.

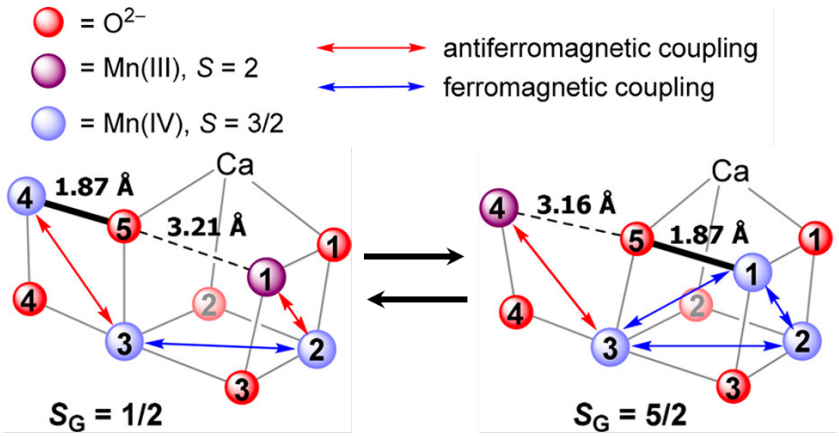


Fig. 2-2. Proposed Mn_4CaO_5 structures of two ground states in S_2[7,8], S_2^{LS} (left) shows an open cubane configuration, and S_2^{HS} (right) shows a closed cubane configuration. The suggested distances of O5-Mn4 and O5-Mn1 are labelled, respectively.

It has been demonstrated that the mutations of valine 185 (V185) in the D1 subunit of PSII (the position of this residue relative to the Mn_4CaO_5 cluster is shown in Fig. 2-3) have a strong effect on the water oxidation of PSII[9-11]. The side chain of V185, which is only 3.8 Å from the Mn_4CaO_5 cluster, faces the water channel from Mn_4CaO_5 cluster to the luminal surface of PSII. The hydrophobic V185 occupies a position contacting water molecules that are located between Y_Z and the putative proton gate residue, D1-Asp61, and at a position close to the oxo bridged oxygen (O5) of the cluster. Substitution of V185 by a hydrophilic residue threonine would perturb the water cavity containing H_2O molecules that interact with the Mn_4CaO_5 cluster, amino acid ligands of the Mn_4CaO_5 cluster, Y_Z and the chloride ion of PSII. The V185 was also reported to contribute to stabilize the S_2^{LS} state, and a site-directed mutant from valine to threonine, V185T, was suggested to accumulate a high proportion of the S_2^{HS} state by Sugiura et al[10]. They proposed that the proximity of the chloride environment with V185 was alternated by the mutation V185T, together with the perturbations of the hydrogen bond network, and the water molecules close to O5 which is proposed to be one of the substrate water molecules. However, the above findings are observed based on the functional measurements, and the structure of the mutant PSII is lacking. With the goal of analyzing the structure of the mutant PSII by cryo-EM, we collected the mutant cells, purified PSII from them, and analyzed their oxygen-evolving activity. We further imaged

the purified samples by negative staining EM, which showed that the purified samples are largely PSII dimers that may be used for structural analysis by cryo-EM.

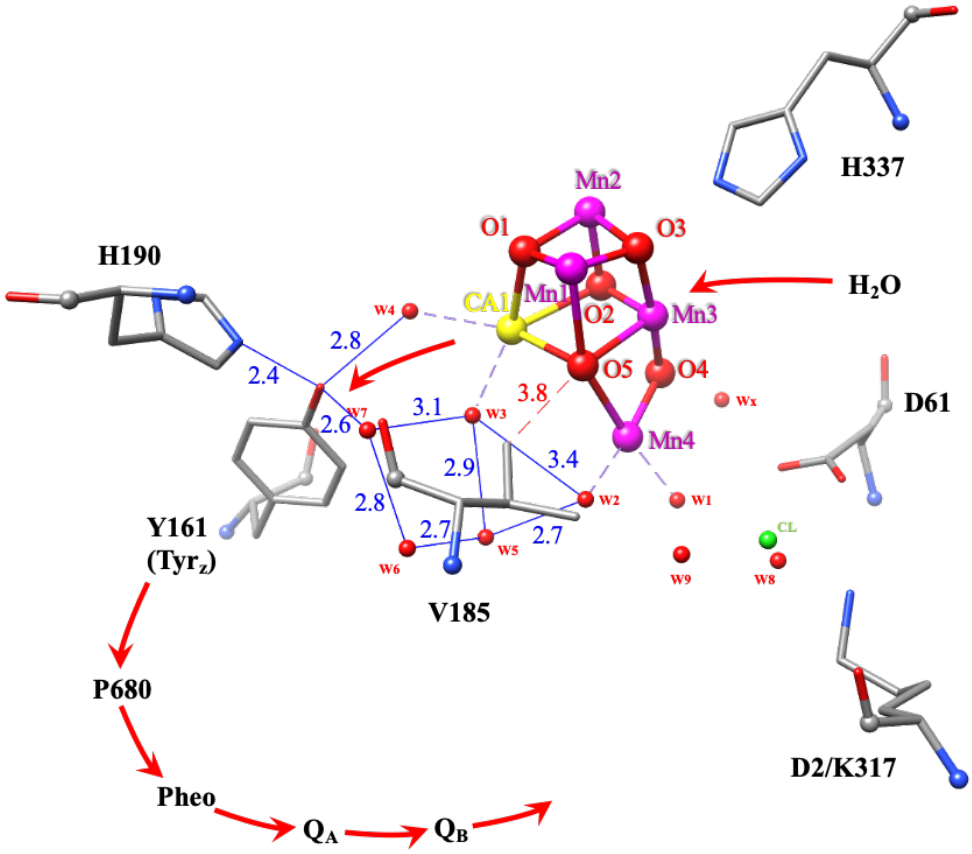


Fig. 2-3. An overview of the electron transfer pathway (depicted in red arrow), some related residues within PSII, and hydrogen bond network around Y_Z (PDB code: 3WU2).

The hydrogen bonds (in Å) around Y_Z are depicted in solid blue lines, the bonds between metal atoms of the Mn_4CaO_5 cluster and their coordinated water ligands are depicted in dashed purple lines, the distance (in Å) between the side chain of V185 and Mn_4CaO_5 cluster is depicted in dashed red line. Mn: magenta; Ca: yellow; Cl: green; O red; N: blue; carbon: gray.

Materials and Methods

Construction of the *T. elongatus* D1-V185T mutant strain

The D1 subunit of PSII is encoded by three *psbA* genes, $psbA_1$, $psbA_2$ and $psbA_3$ in the thermophilic cyanobacterium *T. elongatus*. The strain used in this study was a Δ$psbA_1$, Δ$psbA_2$ deletion mutant, called WT*3, constructed from *T. elongatus*. For the construction of the V185T site-directed mutant, the nucleotides at positions +552, +553 and +554 in the $psbA_3$ gene were

substituted to introduce a V185T substitution and to create a restriction site of Acl I on a plasmid DNA as shown in Fig. 1-4. The mutant strain was constructed by Sugiura, et al[10].

Fig. 2-4. Nucleotide sequence of *psbA*$_3$ and the amino acid sequences around V185 in the WT*3 strain and the *psbA*$_3$-V185T mutant. Position of V185 is indicated with red letters. Numbers correspond to the position from the initial start codon. Substituted nucleotides in the V185T mutant are indicated in small letters. For selection of the mutant, an Acl I restriction site (letters in italic) was created in the *psbA*$_3$-Thr185 construct at the same time.

Cell culture of mutant strain

Liquid culture of the *T. elongatus* D1-V185T mutant strain was grown at 50 °C with bubbling of air containing 3 % CO_2, at a photon density of 13.3 µmol photons m^{-2} s^{-1} with LED light dominated in wavelengths for plant growth. Cells are cultured in 1 x DTN medium with 4 types of antibiotics added in 20,000 times dilution of the concentration summarized in Table 2-1. For maintaining the strain, the OD_{730} of the culture solution was kept between 0.6 and 1.5. Different volumes of Erlenmeyer flasks were used, and the cultures were gradually transferred from 50 mL to 500 mL, 1 L, and 2 L until the cells were harvested. However, it is difficult to keep cells alive in 5 L, so the maximum culture volume is kept below 2 L.

Table 2-1. Composition of culture medium for the mutant strain.

A. 1 x DTN medium. B. 40 x DTN medium. C. 200 x Sulfate Mix. D. Micronutrients. E. Antibiotics.

A

Stock	For 5 L
40 x DTN medium (B)	125 mL
14.4 mM $FeCl_3 \cdot 6H_2O$	1.3 mL
200 x Sulfate Mix (C)	25 mL
1 M Tricine/NaOH	80 mL
1 M NH_4Cl	0.5 mL
Micronutrients (D)	2.5 mL

*Adjust the pH to 7.60 ~ 7.65.

B

Stock	For 1 L
EDTA-2Na	7.6 g
$MgSO_4 \cdot 7H_2O$	4.0 g
KNO_3	4.0 g
$NaNO_3$	28.0 g
Na_2HPO_4	11.2 g
$CaCl_2$	1.6 g
NaCl	320 mg

*Adjust the pH to 7.60 ~ 7.65.

C

Stock	For 500 mL
$Na_2S_2O_3 \cdot 5H_2O$	20.0 g
Na_2SO_3	10.0 g
$NaHSO_3$	16.8 g

D

Stock	For 1 L
H_2SO_4 (conc)	0.5 mL
H_3BO_3	0.5 g
$MnSO_4 \cdot 5H_2O$	3.6 g
$ZnSO_4 \cdot 7H_2O$	0.5 g
$CuSO_4 \cdot 5H_2O$	25.0 mg
$Na_2MoO_4 \cdot 2H_2O$	25.0 mg
$CoCl_2 \cdot 6H_2O$	45.0 mg
Na_2SO_4	4.0 mg

E

Stock	Concentration
Chloramphenicol (Cm)	40 μg mL^{-1}
Spectinomycin (Sp)	25 μg mL^{-1}
Streptomycin (St)	10 μg mL^{-1}
Gentamycin (Gm)	24 μg mL^{-1}

Verification of the *psbA₃* gene in mutant cells

After growth of the mutant strain for several generations, the genotype of the strain is confirmed by the restriction enzyme Acl I. Because the Acl I digesting site is designed at the nucleotide position +552 in the mutant strain, two primers, forward (Fw) and reverse (Rv) primers at positions +318 and +872, respectively, are synthesized (Fig. 2-5), and the genomic DNA is amplified using these two primers by PCR. This generates a 554-bp DNA fragment. After digestion with Acl I, it can be seen whether the Acl I still exists or not.

An aliquot of 1 mL cells in culture medium of the mutant and wild strain were harvested by centrifugation at 15,000 rpm for 1 min. The precipitation was freeze-thawed once for cell disruption after suspended in 50 μL volume, and used as the DNA template for PCR amplification. The DNA amplification by PCR was performed as Table 2-2A, B.

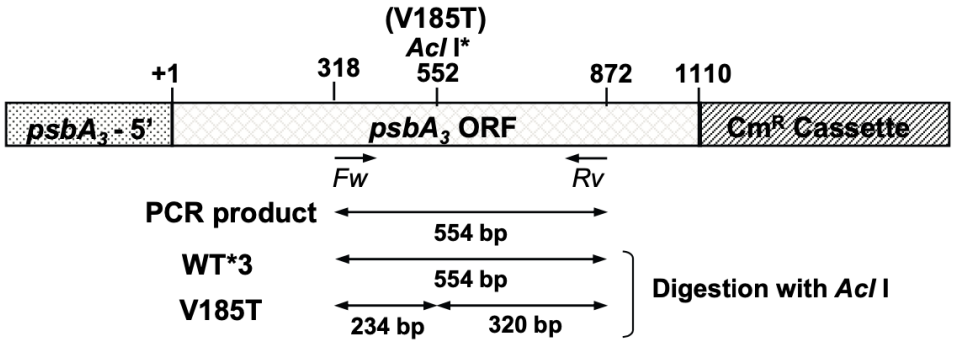


Fig. 2-5. Physical map around *psbA₃*.

The PCR product should give rise to 234 bp and 320 bp DNA fragments after digestion with Acl I, as the Acl I site is introduced into the mutant strain but not in the wild-type strain.

For purification of the PCR product, 1 volume of PCR product was mixed with 5 volumes of Buffer PB (5 M guanidine hydrochloride, 30 % (v/v) isopropanol) and loaded into a QIAquick column that was placed in a 2 mL collection tube provided by the manufacturer. The mixture was centrifuged for 1 min in the tabletop microcentrifuge to discard the flow-through and allow binding of the PCR products to the column. Excess salt was removed by adding 750 μL Buffer PE (2 mM Tris-HCl (pH 7.5), 80 % (v/v) ethanol) and centrifugation for 1 min. The flow-through was discarded and QIAquick column was put back to the collection tube, then another one-minute centrifugation was performed to remove residual ethanol from Buffer PE. DNA was eluted by

adding 50 µl milliQ water to the QIAquick column and centrifugation for 1 min, and collected in a new tube.

The purified DNA was cleaved by Acl I (Table 1-2C) at 37 ˚C for 2 h, and the product amplified from the wild strain cells was used as control. DNA digestion (3 µL) was analyzed by agarose gel (2.0 %) electrophoresis at 125 mV, 30 min after mixing with 10 µL loading buffer.

Table 1-2. DNA amplification and Acl I digestion.

A. DNA amplification reaction components. B. PCR program. C. Reaction components for restriction digestion.

A

Template DNA	10.0 µL
10 x reaction buffer	5.0 µL
$MgCl_2$	2.0 µL
dNTP	5.0 µL
Primer	2 x 0.5 µL
Biotag (last)	0.5 µL
Distilled water	26.5 µL
Total	50.0 µL

B

Step	Temperature	Time	
Initial denature	96 ˚C	2 min	
Denature	96 ˚C	30 s	35 cycles
Annealing	50 ˚C	30 s	
extension	72 ˚C	50 s	
Final extension	72 ˚C	2 min	
	4 ˚C	∞	

C

Amplified DNA	5.0 µL
CutSmart buffer	1.0 µL
Acl I (last)	1.0 µL
Distilled water	3.0 µL
Total	10.0 µL

PSII dimer purification

The mutant cells were centrifuged at 13,800 g, 25 ˚C for 20 min (R9A rotor, HITACHI KOKI), and the precipitation was suspended with Buffer-A (40 mM KH_2PO_4-KOH (pH 6.8)). After centrifugation at 13,800 g, 25 ˚C for 20 min again, the precipitated cells were suspended with Buffer-B (0.4 M mannitol, 40 mM KH_2PO_4-KOH (pH 6.8)). The suspension was incubated at 38 ˚C, 90 rpm for 2.5 h in dark after addition of 1 mM EDTA-2Na and 1.43 g L^{-1} lysozyme in their final concentrations. The incubated cells were centrifuged again at 13,800 g, 4 ˚C for 25 min, and

the precipitation was suspended with Buffer-D (25 % Glycerol, 20 mM Hepes-NaOH (pH 7.0), 10 mM $MgCl_2$), and stored at -80 ˚C until use.

After thawing, the cells were diluted with Buffer-C (30 mM Hepes-NaOH (pH 7.0), 10 mM $MgCl_2$) in 10 times to completely disrupt the cells. The suspension was then centrifuged at 13,800 g, 4 ˚C for 15 min, and the precipitation obtained was resuspended with Buffer-D and shaken with 1 % (w/v) DNase at 4 ˚C for 30 min in the dark. After the treatment of DNase, the suspension was centrifuged at 13,800 g, 4 ˚C for 15 min. The precipitation collected was used as thylakoid membranes, and they were suspended with Buffer-C to reach a concentration of ~ 1 mg Chl mL^{-1}. The thylakoid membranes obtained were stirred with 0.5 % lauryldimethylamine oxide (LDAO) at 27 ˚C for 20 min, and the supernatant collected after centrifugation (22,000 g, 4 ˚C, 15 min, T15A41) was mixed with a final concentration of 13 % (w/v) PEG1,500 and stirred for 10 min on ice. Subsequently, the mixture was centrifuged at 22,000 g, 4 ˚C for 12 min, and the precipitation was suspended with Buffer-D and stored at -80 ˚C until use, which was crude PSII and denoted LDAO-PSII.

The LDAO-PSII was thawed, treated with 1 % (w/v) n-dodecyl-*β*-D-maltoside (*β*-DDM) at 4 ˚C for 1 h in the dark, and centrifuged at 22,000 g, 4 ˚C for 12 min. The supernatant was injected into a Q-Sepharose High Performance Anion Exchange Column (Cytiva) with a bed volume of 30 mL after filtration with a 0.45 μm filter. The column was washed with 170 mM NaCl in A-buffer (5 % (w/v) Glycerol, 30 mM Mes-NaOH (pH 6.0), 3 mM $CaCl_2$, 0.03 % (w/v) *β*-DDM) for 5.5 column volumes, then eluted with a linear gradient of NaCl from 170 mM to 300 mM for 12.5 column volumes to separate PSII dimer, monomer, PSI, and other fractions. The PSII dimer fraction was collected and exchanged with storage buffer (20 mM Mes-NaOH (pH 6.0), 20 mM NaCl, 3 mM $CaCl_2$) using 15 mL Amicon Ultra centrifugal filters (Millipore) with a 100 kDa cutoff, and the concentrated sample was suspended in storage buffer and quick-frozen in liquid nitrogen (-180 ˚C) until analysis. All the purification procedures were performed under dim green light.

Biochemical analysis

Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was performed according to previous protocols[12-14], with a 20 % acrylamide gel containing 6.75 M urea. An aliquot of 2.5 μg Chl of thylakoid membrane and 5 μg Chl of purified sample from the mutant

strain, 5 μg Chl of thylakoid membrane, PSI trimer, PSII dimer and monomer from the wild strain were adjusted to 10 μL with water, respectively. The samples were solubilized in 9.3 % lithium lauryl sulfate, 0.33 M dithiothreitol, and 0.33 M Tris-HCl (pH 8.5) at 60 °C for 40 min, and loaded onto the gel.

Blue native PAGE (BN-PAGE) was performed according to previous protocols[14,15] with a 7.5 % acrylamide gel. The samples were first treated with 1 % n-dodecyl-*β*-D-maltoside (*β*-DDM) and centrifuged at 22,000 g for 10 min to remove unsolubilized materials. The supernatant was further treated with 2 % *β*-DDM at 25 °C for 10 min, and mixed with Coomassie Brilliant Blue G-250 before being applied onto the gel.

Negative staining EM and reference-free 2D classification using RELION

To confirm the overall conformation of the purified mutant PSII dimer, negative staining EM and reference-free 2D classification were performed. The purified sample was diluted 10 times with the elution buffer without NaCl. An aliquot of 4 μL of the sample was dropped on the surface of a carbon coated EM grid which was glow-discharged to be hydrophilic. After standing for 30 s, the excessive solution was blotted off using a filter paper, then a 4 μL aliquot of 16 % $MoNH_4$ was applied onto the grid. Another blot off was performed after waiting for 30 s. The prepared grid was then loaded into the HITACHI-7650 transmission EM operated at 80 keV. Negative staining images were collected with a CCD detector, and the candidate particles were auto-picked and classified to check the overall conformations using RELION 3.1[16].

Measurements of oxygen-evolving activity

The oxygen-evolving activity was measured under continuous, saturating illumination at 25 °C using a Clark-type oxygen electrode (Hansatech Instrument). Wild-type PSII dimer and the purified mutant dimer were diluted at 10 μg Chl mL^{-1} in the storage buffer, and 0.3 mM phenyl-p-benzoquinone (PPBQ) and 1 mM potassium ferricyanide ($K_3[Fe(CN)_6]$) were used as electron acceptors. The activities were measured twice to calculate an average value.

Results and Discussions

Confirmation of the *psbA*$_3$ gene in mutant cells

As Fig. 2-6A shows, only one band at 554 bp appeared after DNA amplification from cultured cells of both wild-type strain and mutant strain. After treatment with Acl I, the band of the mutant strain was digested into two bands, one at 320 bp and the other one at 234 bp, whereas the band of the wild-type strain remains single at 554 bp (Fig. 1-6B). This result indicates that the Acl I digesting site existed in the mutant strain, and it was not contaminated by the wild-type stain, ensuring our latter studies possible using the purified mutant strain.

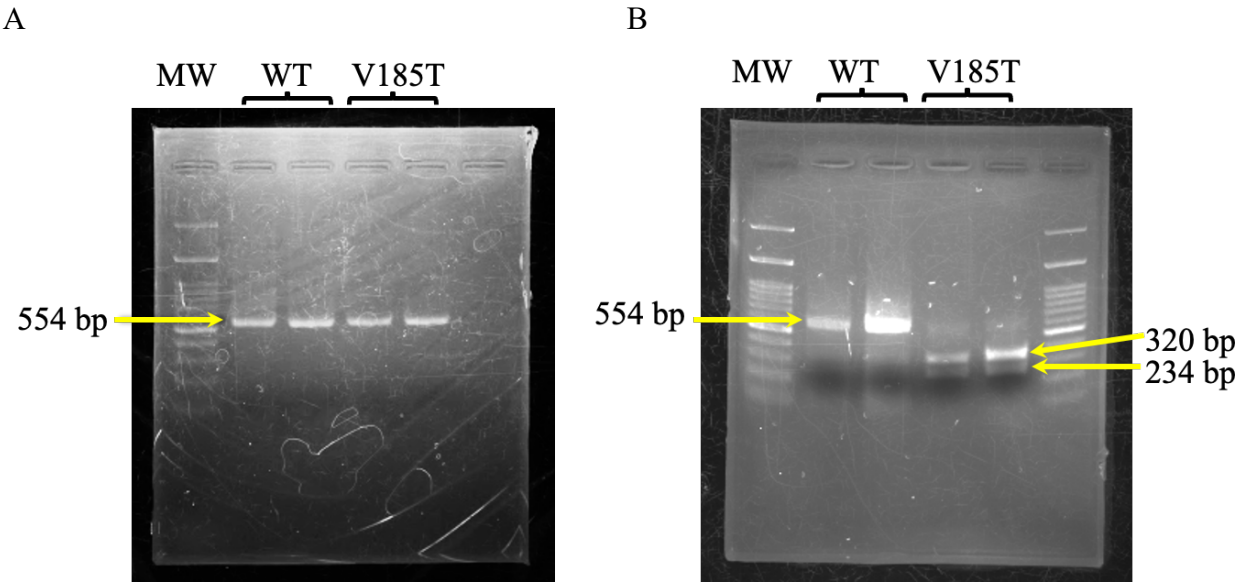


Fig. 2-6. Analysis of *psbA*$_3$ gene using agarose electrophoresis.
A. Products after DNA amplification using PCR from cultured cells of wild-type strain and the mutant D1-V185T strain. B. Agarose gel electrophoresis of PCR products after treatment with Acl I.

Purification of the PSII dimer

PSII dimer was purified by a combination of LDAO and *β*-DDM solubilization, and the final elution profile of *β*-DDM solubilized crude PSII from the mutant strain is shown in Fig. 2-7. One of the peaks marked in dashed red rectangle was demonstrated to be PSII dimer by the following analysis.

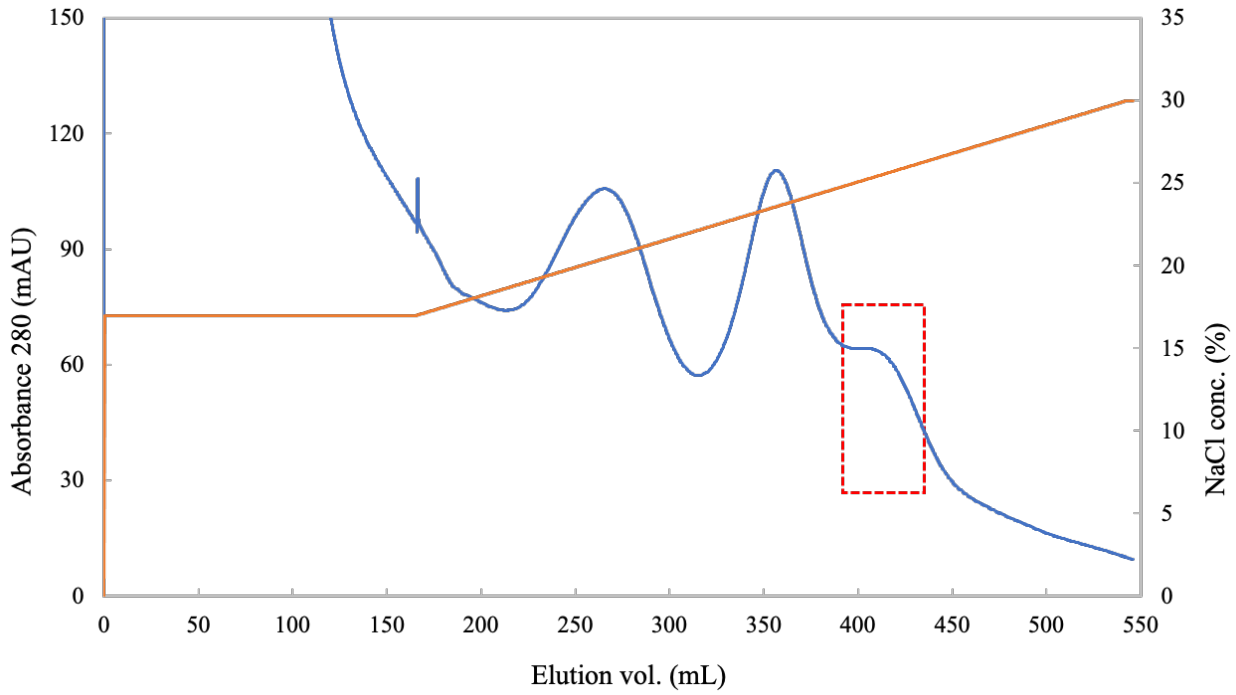


Fig. 2-7. FPLC elution profile of anion exchange chromatography of the mutant crude PSII sample. Blue curve: absorbance at 280 nm; orange line: linear gradient from 170 mM to 300 mM NaCl. The target peak is marked with dashed red rectangle.

Results of SDS-PAGE and BN-PAGE

The result of SDS-PAGE is shown in Fig. 2-8A. The mutant sample contains all the known PSII subunits, which are almost the same as the PSII dimer and monomer from the wild-type strain. Among them, some bands of the extrinsic subunits which are located in the peripheral region of PSII luminal side and function to maintain the activity and stability of OEC, such as Cyt c550 (PsbV) and 12 kDa (PsbU), are relatively weaker than those of the wild-type strain. The samples are further analyzed by BN-PAGE (Fig. 2-8B), which showed that the mutant PSII is mainly in a dimeric state, but a weak band of monomer is also contained (marked with the red dashed rectangle). The amount of the monomeric particles is relatively small and can be excluded in SPA, so they will not affect the final 3D reconstruction by negative staining EM and cryo-EM.

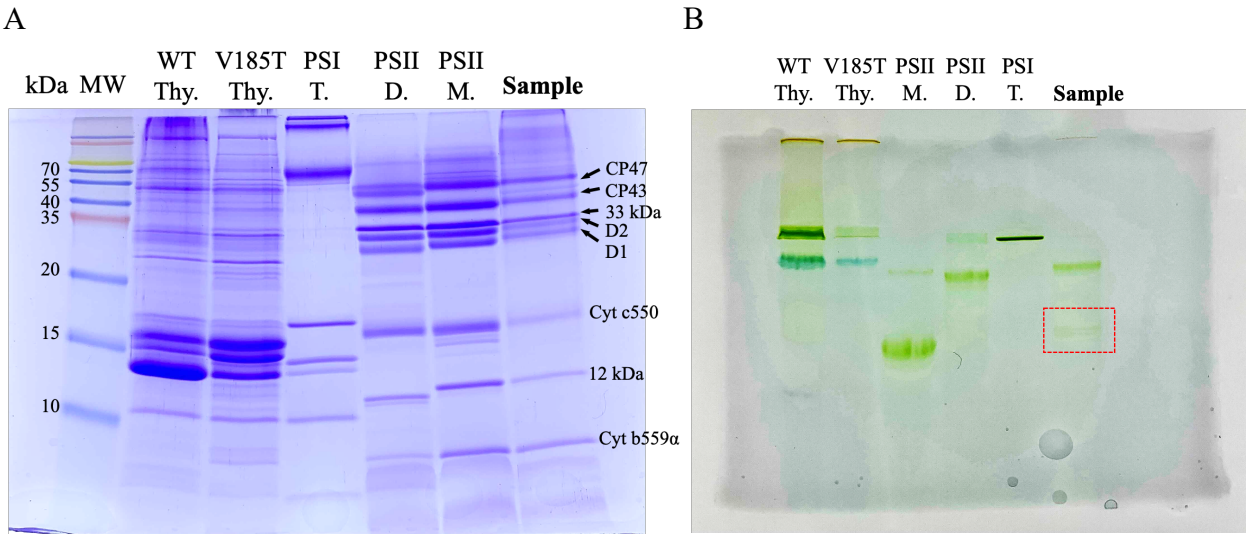


Fig. 2-8. Biochemical analysis of the purified samples.

A. Analysis of subunit components using SDS-PAGE. PSI T. PSII D. and PSII M. represent PSI trimer, PSII dimer and PSII monomer purified from the wild-type strain, and subunits contained in the mutant PSII sample are labelled. B. Analysis of oligomeric states using BN-PAGE. The abbreviations of sample for each lane are the same as (A). The small band of monomer is marked in a dashed red rectangle.

Low oxygen-evolving activity in the mutant strain

The oxygen-evolving activities of PSII dimers of the wild-type and mutant strains are shown in Fig. 2-9 and Table 2-3. The activity of the mutant PSII dimer was 59.0 μmol O_2 mg Chl^{-1} h^{-1}, which is less than 5 % of that of the wild-type strain. This low activity may partly be due to the loss of part of the extrinsic subunits during the purification, as shown by the weaker bands in SDS-PAGE (Fig. 2-8A). However, the extrinsic subunits are still present in the purified PSII dimers, so the loss of the activity cannot be solely caused by loss of the extrinsic proteins. Therefore, change of Val185 to Thr may have affected the structure of the Mn_4CaO_5 cluster itself or its ligand environment or its immediate environment, thereby resulting in the large loss of the activity.

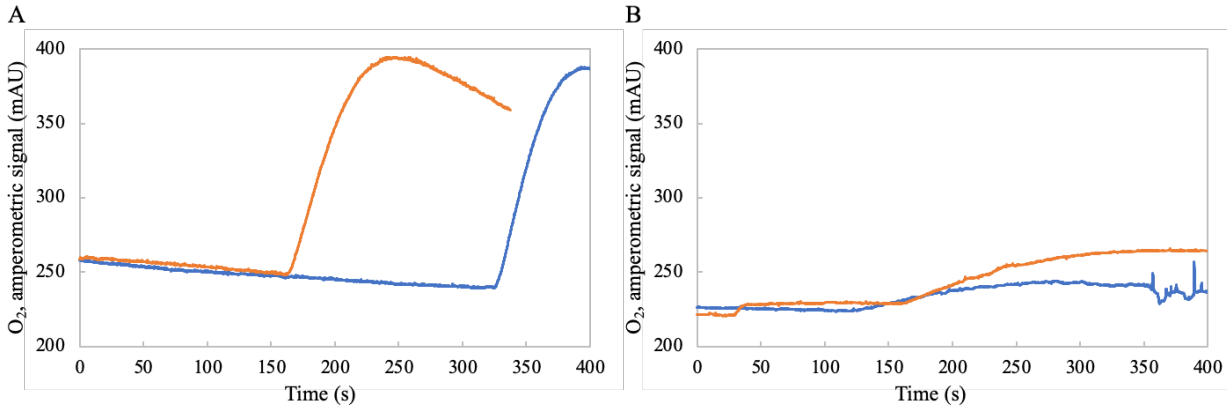


Fig. 2-9. The plots of oxygen-evolving activity against time.

Oxygen amperometric signals of PSII dimers purified from the wild-type (A) and the mutant D1-V185T strains (B). Each sample was measured twice which are depicted in orange and blue curves, respectively.

Table 2-3. Average values of oxygen-evolving activity of PSII dimers from the wild-type and the mutant D1-V185T strains.

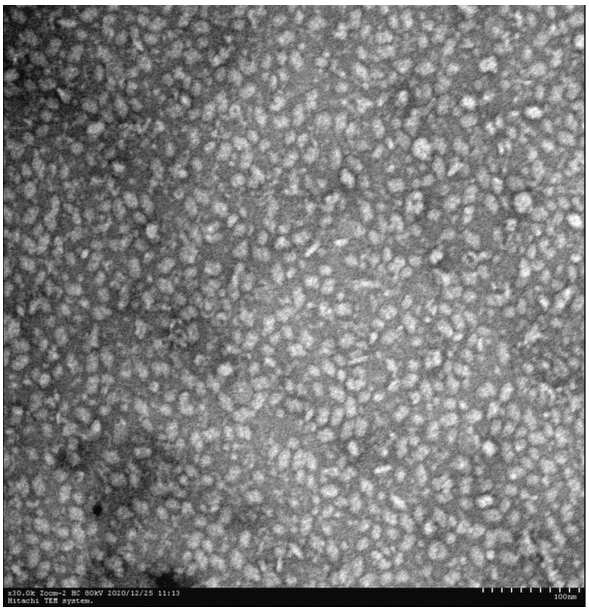
PSII Type	Oxygen-evolving activity (μmol mg Chl^{-1} h^{-1})	Ratio to WT
WT	1208.4	100.00 %
D1-V185T	58.98	4.88 %

Dimeric conformations observed using negative staining EM

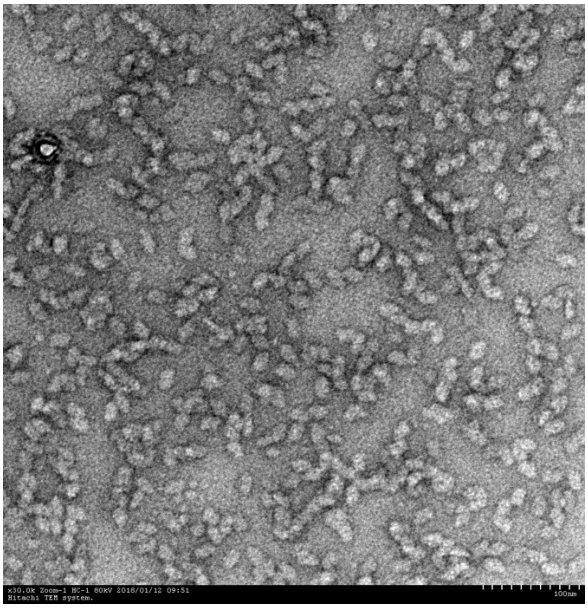
In total, 33 negative staining images were collected from the purified mutant PSII dimer using transmission electron microscope (TEM) operated at 120 keV at a nominal magnification of 30,000. An example of the typical images is shown in Fig. 2-10A, and an image collected from PSII dimers of wild-type strain is shown in Fig. 2-10B. In general, the particles in the mutant PSII image are distinct from each other, which are different from some particles present in aggregated states from the wild-type strain. A small proportion of some PSII monomers and possibly PSI trimers are also visible in our sample, suggesting their partial contaminations. A total of 29,755 particles (averagely 902 particles per image) were auto-picked and used for the subsequent classifications. Finally, 5,635 particles were found to be present in dimeric conformations from 5 classes (Fig. 2-10C). These dimeric particles have an overall conformation very similar to those of

the wild-type strain, suggesting that the V185T mutation does not affect the overall conformation of PSII dimer significantly. Detailed analysis of the structure of the PSII dimer will be achieved by cryo-EM at a later time when the machine time is available.

A



B



C

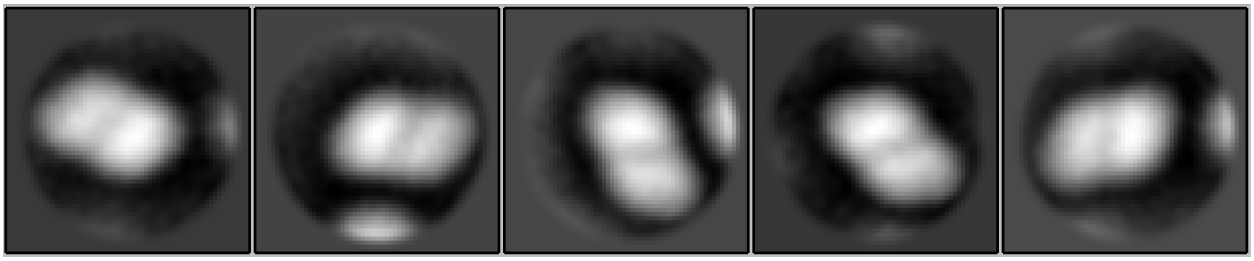


Fig. 2-10. Representative negative staining images and the results of 2D classification. Negative staining image of the purified PSII sample from the mutant D1-V185T strain (A) and the PSII dimer from wild-type strain (B). C. Five typical PSII particle sets in dimeric conformations screened by 2D classifications.

Conclusions

In summary, we successfully purified PSII dimer from the *T. elongatus* D1-V185T mutant strain. We observed an extremely low oxygen-evolving activity from the purified sample, which may indicate the importance of the V185 residue in oxygen evolution. Negative staining EM images showed that a large amount of particles from the purified sample are present in the dimeric conformation. The detailed structure of the PSII dimer from the mutant strain in the S_1-state will be analyzed using cryo-EM to reveal the effect of the mutation on the structure of the Mn_4CaO_5 cluster and its nearby environment.

References

1. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55-60 (2011).
2. Suga, M. *et al.* Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99-103 (2015).
3. Kok, B., Forbush, B. & McGloin, M. Cooperation of charges in photosynthetic O2 evolution–I. A linear four step mechanism. *Photochem Photobiol* **11**, 457-475 (1970).
4. Boussac, A. & Rutherford, A. W. Comparative study of the g=4.1 EPR signals in the S_2 state of photosystem II. *Biochim Biophys Acta Bioenerg* **1457**, 145-156 (2000).
5. Pokhrel, R. & Brudvig, G. W. Oxygen-evolving complex of photosystem II: correlating structure with spectroscopy. *Phys Chem Chem Phys* **16**, 11812-11821 (2014).
6. Boussac, A. *et al.* The low spin - high spin equilibrium in the S_2-state of the water oxidizing enzyme. *Biochim Biophys Acta Bioenerg* **1859**, 342-356 (2018).
7. Lee, H. B. *et al.* Tetranuclear [$(Mn^{III}Mn_3^{IV}O_4$] complexes as spectroscopic models of the S_2 state of the oxygen evolving complex in photosystem II. *J Am Chem Soc* **140**, 17175-17187 (2018).
8. Pantazis, D. A., Ames, W., Cox, N., Lubitz, W. & Neese, F. Two interconvertible structures that explain the spectroscopic properties of the oxygen-evolving complex of photosystem II in the S_2 state. *Angew Chem Int Ed Engl* **51**, 9935-9940 (2012).

9. Dilbeck, P. L., Bao, H., Neveu, C. L. & Burnap, R. L. Perturbing the water cavity surrounding the manganese cluster by mutating the residue D1-Valine 185 has a strong effect on the water oxidation mechanism of photosystem II. *Biochemistry* **52**, 6824-6833 (2013).
10. Sugiura, M. *et al.* Probing the role of Valine 185 of the D1 protein in the Photosystem II oxygen evolution. *Biochim Biophys Acta Bioenerg* **1859**, 1259-1273 (2018).
11. de Lichtenberg, C. *et al.* The D1-V185N mutation alters substrate water exchange by stabilizing alternative structures of the Mn_4Ca-cluster in photosystem II. *Biochim Biophys Acta Bioenerg*, 148319 (2020).
12. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).
13. Schagger, H. & Vonjagow, G. Tricine sodium dodecyl-sulfate polyacrylamide-gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**, 368-379 (1987).
14. Kawakami, K., Iwai, M., Ikeuchi, M., Kamiya, N. & Shen, J. R. Location of PsbY in oxygen-evolving photosystem II revealed by mutagenesis and X-ray crystallography. *FEBS Lett* **581**, 4983-4987 (2007).
15. Wittig, I., Braun, H. P. & Schagger, H. Blue native PAGE. *Nat Protoc* **1**, 418-428 (2006).
16. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7** (2018).

CHAPTER 3

Comparison of two cryo-EM image processing softwares, cryoSPARC and RELION, based on a high-resolution photosystem II dataset

Abstract

Structural analysis by cryo-electron microscopy (cryo-EM) uses a single particle analysis (SPA) technique to reconstruct biological macromolecular 3D structure from numerous 2D images recorded by the microscope. Several software packages are designed for SPA, such as RELION, cryoSPARC, EMAN2, FREALIGN, cisTEM, and THUNDER, etc. Nowadays, RELION and cryoSPARC stand out as the mainly used tools in cryo-EM structural analysis owing to their excellent performances, but there are still some differences between them that are required to be compared and evaluated. In this study, we compared these two software packages based on a photosystem II (PSII) high-resolution cryo-EM data set in many aspects, which include the resolution of the data, B-factor of the structure, quality of the map obtained, and the calculation speed. The results showed that while the electron potential map generated by RELION showed more features that may represent the structure in more details, cryoSPARC has a higher processing speed and uses less cost of computing resources. Therefore, one way to process the cryo-EM data may adopt a mixed procedure with processing of the data first with cryoSPARC, followed by the 3D refinement with RELION, which may produce a reasonably good electron potential map at a high speed.

Introduction

PSII is a large membrane-embedded protein complex that exist in higher plants, algae and cyanobacteria, which functions to oxidize water and generate oxygen. Its high-resolution structure has long been a critical subject to investigate the mechanisms of photosynthesis. As the catalytic center of water splitting and oxygen evolution, the oxygen-evolving complex (OEC), a Mn_4CaO_5 cluster is usually under the spotlight to evaluate structural quality.

Some traditional techniques are used to determine the 3D structure of PSII. The crystal structure of PSII dimer has been resolved at a resolution of 1.9 Å using SPring-8 synchrotron radiation X-rays in 2011[1], in which all the metal atoms of the Mn_4CaO_5 cluster, together with the coordinated ligands have been identified and located. The damage-free structure was further determined at 1.95 Å resolution using an X-ray free electron laser (XFEL) in 2015[2]. However, these structures are solved in crystals where the PSII molecules may suffer from constraints of crystal packing, and one of such effects is seen in the PsbY subunit. The PsbY subunit is a peripheral, trans-membrane subunit of PSII, but it is not observed in the high-resolution crystal structure[1] and observed only in one of the two monomers in the XFEL structure due to the crystal packing. Thus, a structure of PSII in solution would help us to better understand the native conformation of PSII.

A turning point appeared along with the breakthrough of cryo-electron microscopic structural analysis technique, which brings us into the era of studying structures of protein in solutions without crystallization. Similar to other structural analysis techniques, structural studies using cryo-electron microscopy (cryo-EM) involve protein purification, cryo-EM grid preparation and image acquisition, image processing, and structural refinement, in order to obtain a final structural model of the protein. Among these processes, image processing of the original pictures obtained by cryo-EM provides a bridge between macromolecular images and microstructures. The cryo-EM technique uses single particle analysis (SPA) to obtain the structure, which allows us to unveil the macromolecules at near-atomic resolution, and thereby provides us an opportunity to investigate the biological mechanisms based on the structure obtained, as well as by theoretical calculation methods. In SPA, purified proteins or protein complexes are embedded in a thin layer of vitreous ice film through rapid plunge freezing after blotting, which preserves the physiological states of the sample. Transmission electron microscopy (TEM) is then used to collect numerous 2D snapshots of the samples. As the proteins are oriented randomly within the ice, these images

can represent the molecules from various angles, and can be recombined into a high-resolution 3D reconstruction. However, different image-processing software packages may produce different data. For example, the overall resolution of a photosynthetic reaction center complex from a green sulfur bacterium was solved at 2.73 Å by RELION and 2.71 Å by cryoSPARC[3]. This indicates that differences of structural features may exist from the same dataset processed differently.

Electron beam damage is another factor related with the performance of the data processing softwares. Protein samples are exposed under consecutive incident electron beams, and movie frames are recorded and stacked in slices. In each collected movie, the frames from early exposure suffer no or less radiation damages and hence contain more high-resolution structural information, whereas the latter frames suffer more accumulated electron doses which may cause changes in the valences of the metal ions, inter-atom distances, and even breakages of chemical bonds. Generally, this loss of information can be accounted for weighted averaging exposure frames in image-processing software, but the details of their algorithms differ from each other in different softwares, so it is worthy comparing them in practice to find a best way of data processing. In this study, we compared the data processing results by RELION and cryoSPARC based on a PSII high-resolution dataset in terms of resolution of the overall maps, B-factor of the structures, quality of maps, calculation speed, etc. We found that the electron potential map generated by RELION showed more features that may represent the structure in more details, whereas cryoSPARC has a higher processing speed and uses less cost of computing resources. Therefore, cryoSPARC and RELION should be used in combination, with processing of the data first with cryoSPARC, followed by the 3D refinement with RELION, to obtain a reasonably good electron potential map at a high speed.

Materials and Methods

Purification of PSII and cryo-EM data acquisition

The details of PSII sample purification are described in previous publications[4-6], and the cryo-EM grid preparation and data collection in this study are described in Kato et al[7]. The total particles used in this study are 444,729, which are collected from 2,160 images.

Image processing with RELION

Movie frames were aligned and summed using the MotionCor2 software[8] to obtain a final dose weighted image. Estimation of the contrast transfer function (CTF) was performed using the

CTFFIND4 program[9]. An electron potential map at an overall resolution of 1.95 Å was processed using RELION 3.0[10], and these processes are described in Kato et al[7]. The data collection and processing parameters are summarized in Table 3-1.

Table 3-1. Parameters of data collection in this study and data processing in RELION.

Data set ID	ARM-60K
Final sample conc.	1.94 mg Chl mL^{-1}
Grid	Quantifoil R1.2/1.3, Mo 300 mesh
EM	JEOL CRYO ARM 300 RIKEN
Voltage	300 kV
Detector	Gatan K2 summit in counting mode
Nominal mag. (pixel size)	60,000 (0.822 Å)
Nominal defocus range	-0.8 ~ -1.6 μm
Total electron dose	82.85 $e^{-}Å^{-2}$
Exposure time	10.0 sec
Number of frames	50
Number of Micrographs	2,160
Imposed symmetry	C2
Initial particle images	444,729
Final particle images	174,099
Map sharpening B-factor	-34.2 $Å^2$ (-31.5 $Å^2$)*
Map resolution	1.98 Å (1.95 Å)*

*After micelle density subtraction

Workflow of data processing with cryoSPARC

The same cryo-EM dataset was processed using cryoSPARC, and the processing workflow is shown briefly in Fig. 3-1. A total of 174,099 PSII dimer particles in good quality were imported from RELION. To generate an initial model for the following 3D refinement, *ab-initio* reconstruction was first run without any symmetry. For yielding a higher-resolution map, two rounds of homogeneous refinement and non-uniform refinement with a *C*2 symmetry are applied in which per-particle defocus and per-group CTF refinement were additionally imposed along with the second round, where the non-uniform refinement is a unique algorithm developed by

cryoSPARC. It is demonstrated that non-uniform refinement has the ability to drastically improve the results of membrane proteins and smaller proteins[11]. Finally, a local refinement with a *C*2 symmetry was performed to obtain the final map (Fig. 3-1).

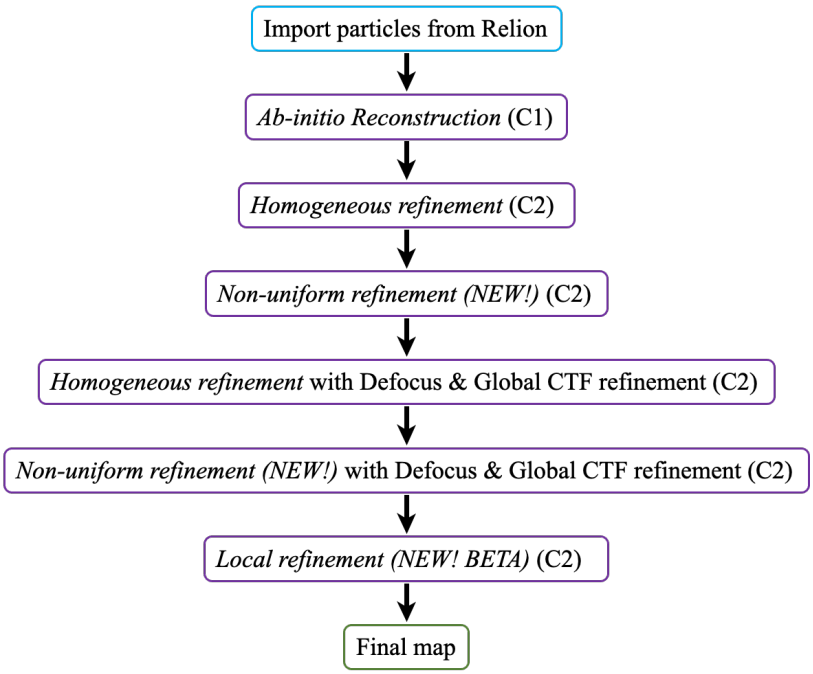


Fig. 3-1. Processing workflow of cryoSPARC in this study.

Results and Discussions

Comparisons at full electron dose

In transmission electron microscopy (TEM), a sequence of movie frames is obtained under the irradiation of the incident electron beam within ten to several tens of seconds, the time elapsed between two adjacent acquisition frames lasts only less than a second. A PSII cryo-EM map at full electron dose (82.85 $e^{-}Å^{-2}$) (which contains 50 frames per movie) was resolved at 1.95 Å resolution using RELION[7]. For comparison, a cryo-EM map at full dose was also processed using cryoSPARC, and the processing results are shown in Fig. 3-2. The best resolution calculated using the tight mask with correction by noise substitution[12] (measurement of overfitting and resolution validation) was 1.84 Å, and the average B-factor is 34.5 $Å^2$ for all of the PSII atoms modelled.

The processed data given by two software packages, cryoSPARC and RELION, are compared in Table 3-2. It is shown that cryoSPARC gives a higher nominal resolution than RELION (1.95 Å), whereas the average B-factor, another evaluation parameter that reflects the flexibility of atoms,

side chains, or even whole regions, is lower by cryoSPARC than that given by RELION which is 31.9 Å^2. This may suggest an electron potential map with more noise or ambiguities generated by cryoSPARC than that generated by RELION.

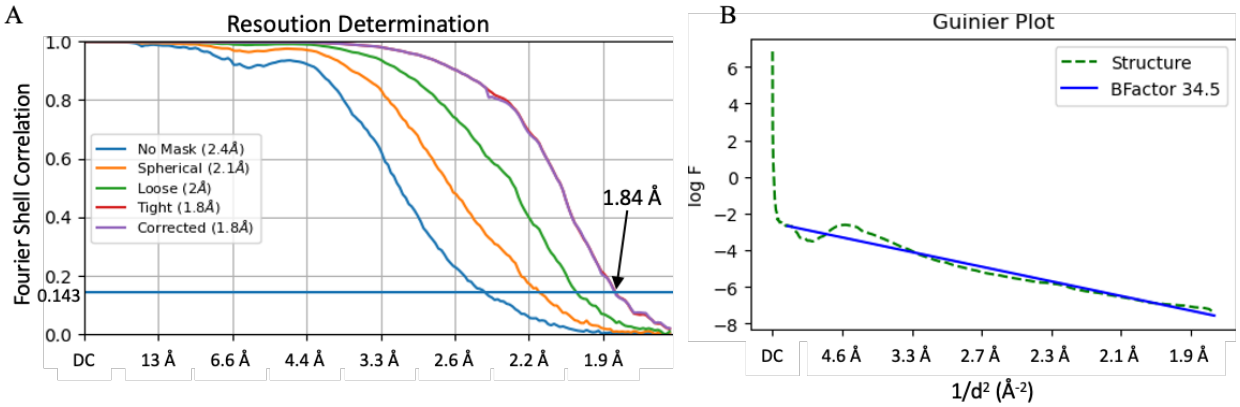


Fig. 3-2. Plots of resolutions (A) and Guinier plot (B) obtained from the cryoSPARC GUI using full frames of the PSII dimer dataset.

Resolution of a 3D structure reconstruction was determined using "Gold Standard" Fourier Shell Correlation (GSFSC) criterion (cut off = 0.143) where the data had been split into two randomly chosen half sets earlier. Guinier plot shows the natural logarithm of the spherically averaged structure factor amplitude (F) for a protein against $1/d^2$, where d is the resolution (Å).

Table 3-2. Comparison of data processed by RELION and cryoSPARC using full frames of the PSII dimer dataset.

Data	RELION (82.85 e$^-$Å^{-2})	cryoSPARC (82.85 e$^-$Å^{-2})
Res. (Å)	1.95	1.84
B-factor (Å^2)	-31.9	-34.5
No. of frames used per movie	50	50

The above results suggested that although the nominal resolution obtained by cryoSPARC is higher than that obtained with RELION with the same dataset, the electron potential map may be worse by cryoSPARC than that by RELION. However, these data are not enough to conclude that RELION is better than cryoSPARC or vice versa, and map densities obtained from the two

softwares need to be compared in order to determine which one is better than the other. For comparison, we choose the Mn_4CaO_5 cluster because it is a compound isolated from the bulk proteins and contains both Mn, Ca ions which have strong densities and O atoms which have relatively low densities[1]. The map qualities on the metal and oxygen atoms of the Mn_4CaO_5 cluster should be a useful indicator to evaluate the different image processing softwares. The atomic model of the Mn_4CaO_5 cluster extracted from the crystal structure of *T. vulcanus* PSII[1] is superimposed with the map solved using cryoSPARC and RELION in Fig. 3-3. The intensities of metal atoms are almost identical between the two softwares, but the densities of O atoms are different. The maps surrounding O1, O3, and O5 solved by cryoSPARC are weaker than those by RELION. Therefore, we conclude that the resolution given by cryoSPARC is overestimated, it is just a nominal value, and the output map given by RELION represents more features than that by cryoSPARC.

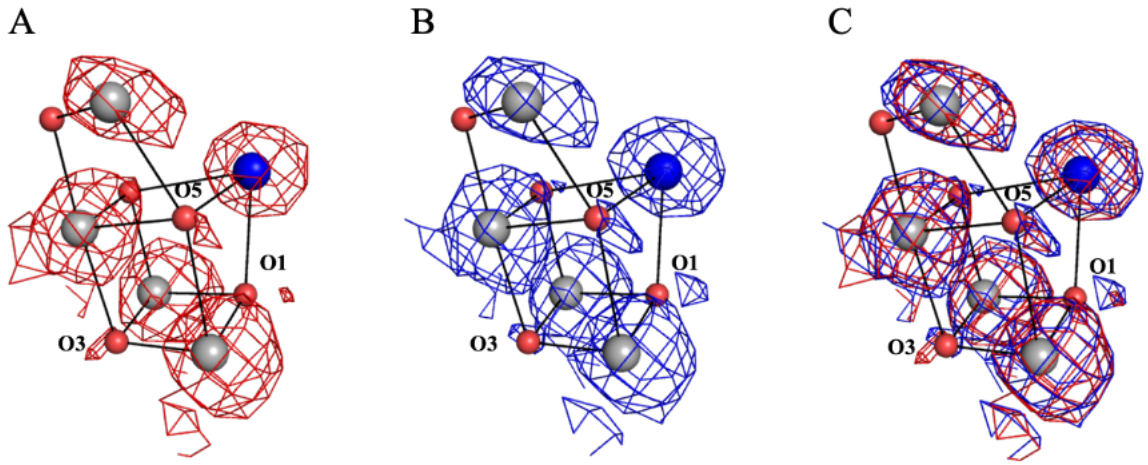


Fig. 3-3. Comparisons of maps surrounding the Mn_4CaO_5 cluster generated by cryoSPARC and RELION.

A. The Mn_4CaO_5 cluster in the 1.84-Å cryo-EM maps (red) solved by cryoSPARC. B. The Mn_4CaO_5 cluster in the 1.95-Å cryo-EM map (blue) solved by RELION. C. Superposition of the two maps. The atomic model is extracted from *T. vulcanus* PSII (PDB code: 3WU2), Mn: gray; Ca: blue; O: red. The densities are depicted at 4 σ.

Evaluations of beam damage and electron dosages

The conventional TEM uses only two elastic signals, namely, the direct beam and the diffracted beams. Beam damage is caused by inelastic scattering while electrons traversing the specimen.

The deposition of energy on the specimen by inelastic scattering events would cause breakage of chemical bonds and ultimately structural collapse, which is an energy-loss process. Generally, this damage can be alleviated by energy filter equipped with the microscope, but in this case some energy loss is inevitable. In order to obtain a high-contrast image, high-dose electron beams are usually imposed which may therefore cause more severe electron beam damage and change the native states of molecules. One of the solutions to mitigate this damage is to observe an image at a much decreased dose, and it has been demonstrated in the RELION-solved PSII structure with the high-resolution dataset that the damage could be greatly reduced at a much decreased dose while keeping a relatively high resolution. It is therefore necessary to compare the ability of two software packages in recovering the structural features from the damage at much reduced electron doses. A series of electron potential maps have been produced with RELION from the high-resolution PSII dataset using gradually reduced movie frames as shown in Table 3-3[7], which showed that the electron beam damage can be largely reduced by using the initial two frames of each movie. A same processing strategy was therefore performed using cryoSPARC with the PSII dataset as that used in RELION, and the results are summarized in Table 3-3 and Fig. 3-4.

Table 3-3. Processing data given by RELION and cryoSPARC against the electron dosages.

Dose ($e^{-}Å^{-2}$)	No. of frames used per movie	RELION		cryoSPARC	
		Res. (Å)	B-factor ($Å^2$)	Res. (Å)	B-factor ($Å^2$)
3.3	2	2.12	-41.7	*	*
5	3	2.09	-41.0	2.23	-49.8
10	6	2.00	-40.0	1.96	-41.8
20	12	1.98	-38.4	1.87	-39.3
40	24	1.98	-35.9	1.84	-35.4
82.85	50	1.95	-31.9	1.84	-34.5

*: In the case of two frames, cryoSPARC failed to produce an electron potential map.

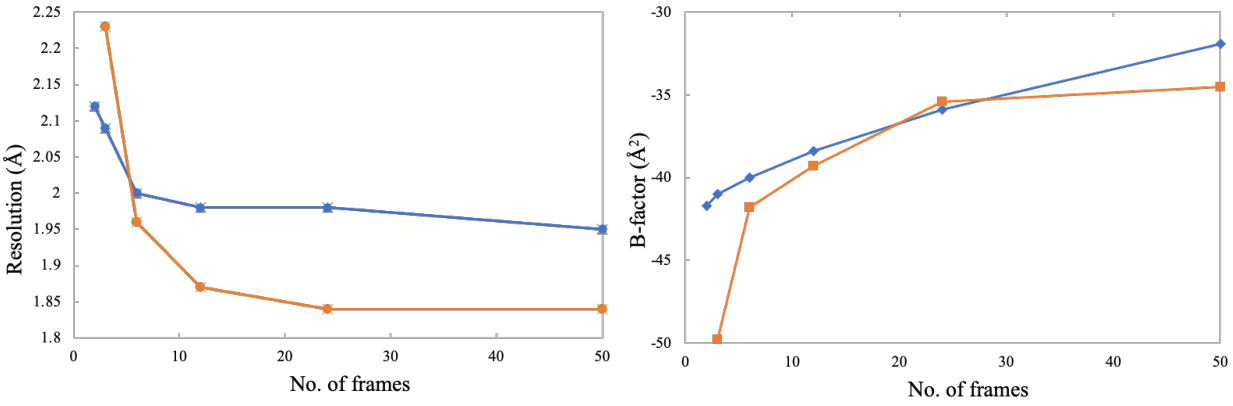


Fig. 3-4. Plots of resolutions (left) and B-factors (right) given by RELION and cryoSPARC against electron dosages.

Blue curve: RELION; Orange curve: cryoSPARC. Note that, in the case of two frames, cryoSPARC failed to produce an electron potential map.

The results showed that, although cryoSPARC failed to produce an electron potential map at the lowest dose with the initial two frames of each movie, it tends to give higher resolutions with the use of more movie frames. However, the resolution is dramatically decreased by using the number of movie frames less than 10. In the case of only the initial two frames used, construction of a reasonable map was failed by cryoSPARC presumably due to the utilization of too few information, and the molecular particles embedded in ice drifted too much with the initial strong beam. In general TEM, strong beam-induced movement during the first 1 - 2 $e^{-}Å^{-2}$ of exposure typically makes the first few movie frames unusable, so these frames are best to be discarded[13] (Fig. 3-5). Therefore, if only the initial two or three frames are fed to the calculations, it would induce a low signal-to-noise ratio. Motion Correction is designed to correct beam-induced sample motion recorded on dose fractionated movie stacks, but with the support of such low number of frames, it may be not effective to track particle motions due to the initial great movement. This will further affect the global CTF refinement and defocus refinement in the processing with cryoSPARC. RELION performs better at low electron dose/fewer movie frames, because it may refer to the full-frame motion correction to calibrate the motions of the initial two frames, which constitutes one of the advantages of RELION. Regarding the B-factors, the results achieved by cryoSPARC are higher than that given by RELION, and not so good as the resolutions given, even in the cases in which nominal higher resolutions are produced.

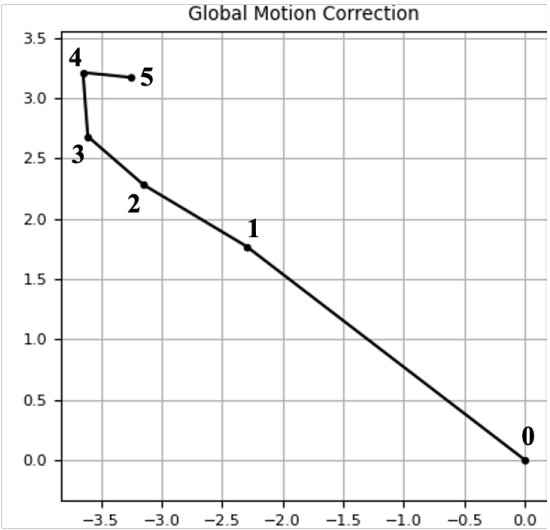


Fig. 3-5. A particle trajectory calibrated by global motion correction in cryoSPARC.

Disulfide bond is one of the sensitive bonds to the damage by electron beams, and is easily broken if exposed to excessive dosages. Thus, the degree of breakage and recovery of the disulfide bonds in the protein structures is a useful criterion to evaluate the degrees of damage and recovery from beam damage of two software packages. We compared the densities surrounding a disulfide bond between Cys 19 and Cys 44 in PsbO at 10 e^{-}Å^{-2} solved by RELION and cryoSPARC (Fig. 3-6), a dose rate at which the breakage just occurred in the solved map by RELION. The result showed very tiny differences between them, indicating that the two software packages perform almost equally in respect to the radiation damage and repair.

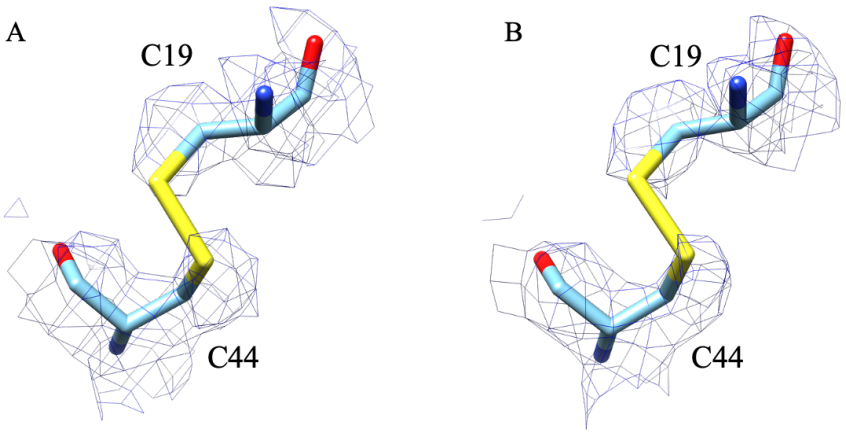


Fig. 3-6. The disulfide bond between Cys 19 and Cys 44 in PsbO at an electron beam dose of 10 e^{-}Å^{-2} solved by RELION (A) and cryoSPARC (B). The densities are depicted at 4 σ.

Higher calculation speed achieved by cryoSPARC

Cryo-EM map refinement is a computationally expensive process, the primary burden is to search and optimize orientation parameters that align each 2D single particle image to a 3D density map best. CryoSPARC is reported to accelerate this search nine times faster than the GPU enabled RELION on a ribosome complex cryo-EM dataset[14,15]. This significant speed-up is attributed to their branch-and-bound maximum likelihood optimization algorithm, which can quickly and inexpensively discard large regions of the search space in which the optimal pose is not included, and the next candidate is then calculated that approximates the optimal pose more tightly. This process is repeated to approach the best alignment[15]. Even though cryoSPARC might produce a poorer map than RELION from above comparisons, we may obtain at least a quick-check map to evaluate our future processing strategy of the cryo-EM data by using cryoSPARC first.

Another algorithm employed in cryoSPARC is stochastic gradient descent (SGD), which is implemented in the *Ab-initio* reconstruction function. Map calculation is a nonconvex optimization problem, which is sensitive to the random initialization of local optima, but their SGD algorithm is sufficiently robust to enable reconstruction of multiple 3D classes from independent arbitrary initializations[15]. Therefore, a combined strategy is usually adopted to calculate a cryo-EM map, in which cryoSPARC is utilized to generate an initial low-resolution model using *Ab-initio* reconstruction first, which is fed to RELION to guide the 3D refinement to obtain the final map[16-18]. This approach may provide us with a reasonably good electron potential map at a faster computational speed with a lower computational cost.

Conclusions

In summary, we compared performances of two cryo-EM image processing software packages, RELION and cryoSPARC, based on a high-resolution PSII dataset. RELION is still a more powerful tool for the cryo-EM map refinement, even though cryoSPARC can give a nominal higher resolution data, the map output from cryoSPARC is not as good as that from RELION. Their abilities of recovery from electron beam damage are almost same. The advantage of cryoSPARC is its higher calculation speed, which saves time and allows us to examine the data quality we collected quickly. Also, the *Ab-initio* reconstruction in cryoSPARC is often used to provide a reference model for subsequent 3D reconstruction in RELION, and this mixed procedure

may help us to achieve a reasonably good electron potential map with a higher speed and lower computational cost.

References

1. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55-60 (2011).
2. Suga, M. *et al.* Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99-U265 (2015).
3. Chen, J. H. *et al.* Architecture of the photosynthetic complex from a green sulfur bacterium. *Science* **370** (2020).
4. Koike, H., Hanssum, B., Inoue, Y. & Renger, G. Temperature-dependence of S-state transition in a thermophilic cyanobacterium, *Synechococcus-vulcanus* copeland measured by absorption changes in the ultraviolet region. *Biochim Biophys Acta* **893**, 524-533 (1987).
5. Shen, J. R., Ikeuchi, M. & Inoue, Y. Stoichiometric association of extrinsic cytochrome c550 and 12 kDa protein with a highly purified oxygen-evolving photosystem II core complex from *Synechococcus vulcanus*. *FEBS Lett* **301**, 145-149 (1992).
6. Shen, J. R. & Inoue, Y. Binding and functional properties of two new extrinsic components, cytochrome c-550 and a 12-kDa protein, in cyanobacterial photosystem II. *Biochemistry* **32**, 1825-1832 (1993).
7. Kato, K. *et al.* High-resolution cryo-EM structure of photosystem II reveals damage from high-dose electron beams. *Commun Biol* **4**, 382 (2021).
8. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
9. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221 (2015).
10. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7** (2018).
11. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat Methods* **17**, 1214-1221 (2020).

12. Chen, S. X. *et al.* High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicro* **135**, 24-35 (2013).
13. Sigworth, F. J. Principles of cryo-EM single-particle image processing. *Microscopy (Oxf)* **65**, 57-67 (2016).
14. Wong, W. *et al.* Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. *Elife* **3** (2014).
15. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).
16. Caffalette, C. A. & Zimmer, J. Cryo-EM structure of the full-length WzmWzt ABC transporter required for lipid-linked O antigen transport. *Proc Natl Acad Sci USA* **118** (2021).
17. Zabret, J. *et al.* Structural insights into photosystem II assembly. *Nat Plants* **7**, 524-538 (2021).
18. Xie, J. F. *et al.* Structure of the human sodium leak channel NALCN in complex with FAM155A. *Nat Commun* **11** (2020).

GENERAL CONCLUSIONS

PSII usually exists in dimeric form, and its high-resolution structure is also solved for the dimer. However, a part of PSII exists in monomer in vivo, and its structure is helpful for us to understand the PSII assembly/repair process. In this study, a cryo-EM structure of PSII monomer from *T. vulcanus* was solved at 2.78 Å resolution, which was compared with the 1.9-Å resolution crystal structure of dimeric PSII from the same organism[1], 3.6-Å resolution crystal structure of monomeric PSII from *T. elongatus*[2], and 2.58-Å resolution cryo-EM structure of an apo-PSII monomer lacking the water oxidation complex from *Synechocystis* sp. PCC 6803[3]. The results showed that while the overall structures are very similar except some extrinsic subunits, apparent differences in pigments and lipids in the monomer-monomer interface are found between the PSII monomer and dimer. One *β*-carotene and two SQDG molecules are found in the monomer-monomer interface of the dimer structure but absent in the monomer structure. Additionally, a large part of the PsbO subunit was disordered in the monomer structure, which is not in the dimer structure and has been reported to interact with CP47 of the other monomer[1,4]. These results indicated the importance of the *β*-carotene, SQDGs and PsbO in the formation of PSII dimer.

D1-V185 is located close to the Mn_4CaO_5 cluster and suggested to be important for maintaining the structure and activity of the Mn_4CaO_5 cluster. In order to examine the function of D1-V185 from structural point of view, we cultured bulk volumes of D1-V185T site-directed mutant strain from *T. elongatus* and confirmed the mutation of this residue in the *psbA₃* gene of the mutant cells. PSII dimer was successfully purified from the mutant strain by anion exchange chromatography using Q-Sepharose column. A significantly lower oxygen-evolving activity was detected in the purified sample, which is in agreement with the report from Sugiura, et al[5]. The lower activity may indicate the important role of V185 in the oxygen evolution process. Negative staining EM images were collected from the purified sample, which showed that a large amount of particles are present in dimeric conformations. Cryo-EM images will be collected from the sample to analyze the PSII dimer structure of the dimer in a high resolution.

For the cryo-EM single particle analysis, RELION and cryoSPARC are two popular software packages that are widely used to process the cryo-EM images. Their performances are compared based on a PSII high-resolution cryo-EM dataset[6] in terms of resolution, B-factors, quality of the electron potential maps, etc. The results showed that the resolution values given by cryoSPARC

were higher than that by RELION, but the B-factors given by cryoSPARC were generally higher than that given by RELION, and the electron potential map generated by RELION showed more features that may represent the structure in more details. This suggests that the nominal resolution values given by cryoSPARC is not reliable and the electron potential maps should be checked to confirm the quality of the model. One of the advantages of cryoSPARC is its higher calculation speed and less cost of computing resources. Therefore, one way to process the cryo-EM data may adopt a mixed procedure with the generation of an initial model first with cryoSPARC, followed by the 3D refinement with RELION, which may provide us with a reasonably good electron potential map at a fast speed.

References

1. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55-60 (2011).
2. Broser, M. *et al.* Crystal structure of monomeric photosystem II from *Thermosynechococcus elongatus* at 3.6-Å resolution. *J Biol Chem* **285**, 26255-26262 (2010).
3. Gisriel, C. J. *et al.* Cryo-EM structure of monomeric photosystem II from *Synechocystis* sp. pcc 6803 lacking the water-oxidation complex. *Joule* (2020).
4. Suga, M. *et al.* Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99-U265 (2015).
5. Sugiura, M. *et al.* Probing the role of Valine 185 of the D1 protein in the photosystem II oxygen evolution. *Biochim Biophys Acta Bioenerg* **1859**, 1259-1273 (2018).
6. Kato, K. *et al.* High-resolution cryo-EM structure of photosystem II reveals damage from high-dose electron beams. *Commun Biol* **4**, 382 (2021).

ACKNOWLEDGEMENTS

There are many people whom I want to thank for their contributions to my Ph.D. course.

I would like to first express my deepest appreciation to Prof. Jian-Ren Shen for providing such a valuable chance to complete my Ph.D. course and the far-sighted guidance all along.

I would like to appreciate Dr. Fusamichi Akita for his patient instructions, and kind encouragement in my daily work. I enjoyed very comfortable collaborations with him and became a more experienced researcher under his influence.

I would like to thank Prof. Takashi Ogura, Prof. Takashi Osumi, Prof. Hiderou Yoshida, Prof. Atsuno Miyazawa from University of Hyogo, and Prof. Zhengqiang Li, Prof. Xueqi Fu from Jilin University for providing me a chance to study abroad.

I would like to thank Prof. Yoshitsugu Shiro, Prof. Yoshiki Higuchi, Prof. Hiderou Yoshida, and Prof. Masaru Tateno from University of Hyogo for their logical suggestions in my Ph.D. course. Also, I would like to thank all the students and staff in Leading Program for their sincere cares about my daily life and academic career.

I would like to thank Dr. Yoshiki Nakajima for the suggestions in sample preparation and other daily affairs, Dr. Koji Kato for the guidance in data processing, Dr. Tianyi Jiang and Dr. Hongjie Li for their suggestions for my research and helps in my daily life.

I would like to thank Prof. Koji Yonekura and Prof. Hideki Shigematsu and their lab members for the help of cryo-EM data acquisition.

I would like to thank all the members of Shen Lab for creating such an impressive academic atmosphere.

I would like to thank my family for their ongoing unconditional support and understanding, and the companions of my friends in Japan and China.