BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Crystal Structure of Psb31, a Novel Extrinsic Protein of Photosystem II from a Marine Centric Diatom and Implications for Its Binding and Function

Ryo Nagao,[†] Michihiro Suga,[¶] Ayako Niikura,[¶] Akinori Okumura,[†] Faisal Hammad Mekky Koua,[¶] Takehiro Suzuki,[⊥] Tatsuya Tomo,[‖,§] Isao Enami,[‖] and Jian-Ren Shen*,[¶]

[†]Department of Integrated Sciences in Physics and Biology, College of Humanities and Sciences, Nihon University, Sakurajosui 3-25-40, Setagaya-ku, Tokyo 156-8850, Japan

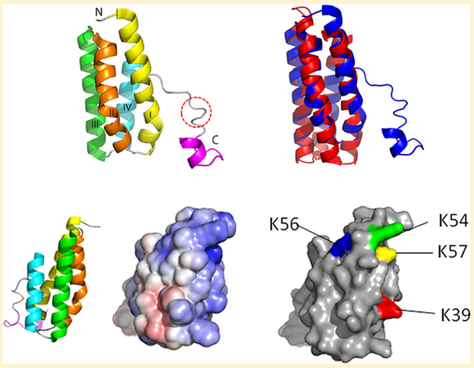
[¶]Graduate School of Natural Science and Technology/Faculty of Science, Okayama University, Tsushima Naka 3-1-1, Kita-ku, Okayama 700-8530, Japan

[⊥]Biomolecular Characterization Team, Discovery Research Institute, RIKEN, Hirosawa 2-1, Wako, Saitama 351-0198, Japan

[‖]Department of Biology, Faculty of Science, Tokyo University of Science, Kagurazaka 1-3, Shinjuku-ku, Tokyo 162-8601, Japan

[§]PRESTO, Japan Science and Technology Agency (JST), 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

ABSTRACT: Psb31 is a fifth extrinsic protein found in photosystem II (PSII) of a centric diatom, *Chaetoceros gracilis*. The protein has been shown to bind directly to PSII in the absence of other extrinsic proteins and serves in part as a substitute for PsbO in supporting oxygen evolution. We report here the crystal structure of Psb31 at a resolution of 1.55 Å. The structure of Psb31 was composed of two domains, one major, N-terminal four helical domain and one minor, flexible C-terminal domain. The four helices in the N-terminal domain were arranged in an up−down−up−down fold, which appeared unexpectedly to be similar to the structure of spinach PsbQ, in spite of their low sequence homology. This suggests that the centric diatom PSII contains another PsbQ-type extrinsic protein in addition to the original PsbQ protein found in the organism. On the other hand, the C-terminal domain of Psb31 has a unique structure composed of one loop and one short helix. Based on these structural analysis and chemical cross-linking experiments, residues responsible for the binding of Psb31 to PSII intrinsic proteins were suggested. The results are discussed in relation to the copy number of extrinsic proteins in higher plant PSII.



Photosystem II (PSII) is a multisubunit, membrane protein complex located in the thylakoid membranes of oxygenic photosynthetic organisms from cyanobacteria to higher plants. PSII catalyzes light-induced electron transfer reactions from water to plastoquinone, leading to the production of molecular oxygen essential for survival of almost all life forms on the earth. PSII from cyanobacteria contains 17 membrane-spanning subunits and 3 peripheral, extrinsic proteins with a total molecular mass of 350 kDa. The 3D structure of oxygen-evolving PSII from a thermophilic cyanobacterium *Thermosynechococcus vulcanus* has been determined at a resolution of 1.9 Å.[1] This analysis revealed the detailed structure and coordination environment of the catalytic center for oxygen evolution, namely, a Mn_4CaO_5-cluster, which was shown to be coordinated by amino acid residues from 2 PSII intrinsic subunits D1 and CP43. In addition, it is also surrounded by a number of other intrinsic and extrinsic proteins as well as several water molecules.

The main PSII intrinsic proteins (D1, D2, CP43, CP47, and Cyt*b*559) are largely conserved from cyanobacteria to higher plants, whereas significant differences are found in the composition of the extrinsic proteins among different species.[2−5] In the cyanobacterial PSII, PsbO (33 kDa), PsbU (12 kDa), and PsbV (17 kDa, cytochrome *c*550) are found as extrinsic proteins. Among these extrinsic proteins, the PsbO protein plays an important role in maintaining the stability and activity of the oxygen-evolving complex (OEC), and is present in all of the oxygenic photosynthetic organisms. The other extrinsic proteins function to optimize the availability of Ca^{2+} and Cl^- cofactors for water oxidation, and are found to be different among different species. While cyanobacterial and red algal PSIIs contain the PsbV and PsbU protein,[4−8] euglena, green algal, and higher plant PSIIs contain the PsbP and PsbQ proteins instead of PsbV and PsbU.[4,5,9−11] In red algal PSII, a fourth extrinsic protein denoted as PsbQ′ is present, which has a low sequence homology to higher plant PsbQ, and is found to be required for the effective binding of PsbV and PsbU.[12,13] Cyanobacterial thylakoid membranes also contain PsbQ-like

Received: June 17, 2013
Revised: August 24, 2013
Published: August 29, 2013

dx.doi.org/10.1021/bi400770d | *Biochemistry* 2013, 52, 6646−6652

(cyanoPsbQ) and PsbP-like (cyanoPsbP) proteins which have been suggested to be necessary for maintaining the optimal activity of PSII in various conditions.[14−16] These two proteins, however, were not found in the crystal structure of cyanobacterial PSII,[1,17−19] and the amount of cyanoPsbP relative to PSII core complexes seemed to be at a substoichiometric level.[14] Thus, these 2 proteins may function to maintain the optimal PSII activity under various unfavorable conditions, whereas whether they are intrinsic components of mature cyanobacterial PSII core complex remains unclear. The 3D structures of higher plant PsbP,[20] PsbQ,[21,22] and cyano PsbQ[23] have been reported.

Recently, we succeeded in preparation of PSII core complex retaining a high oxygen-evolving activity from a marine centric diatom, *Chaetoceros gracilis*.[24,25] The diatom PSII was found to contain, in addition to the four red algal type extrinsic proteins PsbO, PsbQ′, PsbV, and PsbU, a fifth extrinsic protein which was referred to as Psb31 following the nomenclature for PSII subunits.[14] The gene encoding Psb31 was cloned and sequenced,[26] which showed that Psb31 is encoded in the nuclear genome and constitutes one of the extrinsic proteins located in the lumenal side. Genes homologous to *psb31* were found in a red alga and a number of chromophytic algae but not in the other organisms including the green lineage, and its sequence showed no homology to any of the known extrinsic proteins,[26] indicating that it is a novel extrinsic subunit of the diatom PSII. Reconstitution experiments in various combinations with other purified extrinsic proteins showed that Psb31 rebound directly to PSII in the absence of other extrinsic proteins, indicating that Psb31 has its own binding site in the PSII intrinsic proteins.[27] Interestingly, PSII reconstituted with Psb31 alone partially restored the oxygen-evolving activity in the absence of PsbO, indicating that Psb31 serves as a substitute in part for PsbO in supporting oxygen evolution.[27] The 3D structure of Psb31, however, is unknown so far. Because Psb31 has no significant sequence homology to other known proteins, structural modeling does not give reliable results.

In order to gain information regarding the structure and function of Psb31, we expressed the Psb31 protein in *E. coli*, and purified, crystallized it. Here, we report the crystal structure of Psb31 from *C. gracilis* analyzed at a resolution of 1.55 Å; based on this and chemical cross-linking results, we discuss its possible binding site and function in PSII.

■ MATERIALS AND METHODS

Expression, Purification of the Recombinant Psb31, and Crystallization. The gene encoding the mature Psb31 protein from the marine centric diatom *Chaetoceros gracilis* was cloned into the *Kpn*I/*Sal*I site of the pET32a expression vector, resulting in a fusion protein with thioredoxin and $(\mathrm{His})_6$-tag attached at its N-terminus. The recombinant protein was expressed with the host cell BL21 (DE3) and purified by Ni-affinity chromatography. The fusion protein was treated with enterokinase to cleave off the thioredoxin and His-tag and then purified again by the affinity column. Expression of selenomethyonine-labeled (SeMet) Psb31 for phase determination was performed in a modified M9 medium containing 0.1 mg mL^{-1} ampicillin instead of LB medium.[28,29]

The purified Psb31 and SeMet-Psb31 proteins were dialyzed against 10 mM Tris-HCl (pH 8.0), and concentrated to 2.0−8.0 mg mL^{-1}. Crystallization conditions were screened with commercially available screening kits from which microcrystals were obtained under certain conditions. After optimization of the crystallization condition, crystals were obtained at 277 K after 4−5 days by the sitting drop vapor diffusion method, with a reservoir containing 2.0 M ammonium sulfate, 0.1 M citric acid−NaOH (pH 5.0), and 0.2 M ammonium citrate dibasic.

X-ray Data Collection, Processing, Structure Determination, and Refinement. The crystals of Psb31 were flash−frozen in a nitrogen gas stream after being soaked in a cryoprotectant solution containing 35% PEG4000, 5% sucrose, and 5% glucose in 50 mM MES (pH 6.0). X-ray diffraction data was collected at beamlines BL44XU and BL41XU of Spring-8, Japan, using a MarCCD225 detector at 100 K. The crystals diffracted to a resolution of 1.55 Å, and belong to the space group $P3_121$ with unit cell parameters of $a = b = 92.4$, $c = 31.3$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$ (Table 1).

Table 1. Statistics of X-ray Data Collection and Structure Analysis[a]

	native crystal	Se-Met crystal
wavelength (Å)	0.90000	0.97940
space group	$P3_121$	$P3_121$
unit cell (Å)	$a = 92.4$, $b = 92.4$, $c = 31.3$ $\alpha = 90$, $\beta = 90$, $\gamma = 120$	$a = 91.9$, $b = 91.9$, $c = 31.5$ $\alpha = 90$, $\beta = 90$, $\gamma = 120$
resolution (Å)[b]	50−1.55 (1.58−1.55)	50−2.20 (2.28−2.20)
no. of unique reflections	22604	7945
completeness (%)[b]	99.8 (99.5)	99.8 (99.2)
redundancy[b]	6.1 (5.1)	26.2 (24.3)
R_{merge}[b]	0.061 (0.499)	0.110 (0.287)
$I/\sigma(I)$[b]	33.4 (3.3)	48.9 (16.3)
R_{factor}	0.1855	
R_{free}	0.1949	
rmsd for bond lengths (Å)	0.007	
rmsd for bond angles (deg)	1.166	
Ramachandran plot		
favored	98.3	
allowed	1.7	
outliers	0.0	

[a]X-ray diffraction data was collected at beamlines BL44XU and BL41XU of SPring-8, Japan. [b]Numbers in parentheses indicate those of the highest resolution shell.

The diffraction data was processed with the HKL2000 package and CCP4 suite.[30,31] Phase information was obtained from the SeMet labeled Psb31 crystal with the program HKL2MAP,[32] then the phase information was transferred to the native data set. An initial model was built manually with the program COOT.[33] The crystal structure was solved and refined with the CCP4 suite and Phenix.[31,34]

■ RESULTS AND DISCUSSION

Structure of Psb31. The Psb31 protein was successfully expressed and purified from *E. coli*, and its crystals were obtained after screening and optimization of crystallization conditions with commercial screening kits. The diffraction data was collected with synchrotron X-rays, and the crystal structure was solved at a resolution of 1.55 Å with phase information collected from Se-methionine substituted crystals (Table 1). As shown in Figure 1A, 111 residues beginning from Ala 3 to Arg 113 out of 124 total residues of Psb31 were successfully built

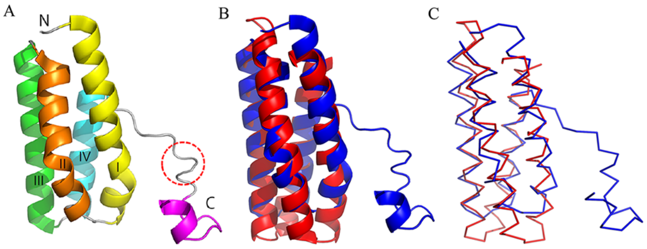


Figure 1. Structure of diatom Psb31 analyzed at 1.55 Å resolution and its comparison with the structure of spinach PsbQ. (A) Structure of diatom Psb31 at 1.55 Å resolution. Helices are numbered I–IV from the N-terminus to the C-terminus. (B) and (C) Comparison between the structure of Psb31 (blue) and the C-terminal helical structure of spinach PsbQ (red) in stick (B) or line (C) representations.

into the structure, although 2 residues in the N-terminal and 11 residues in the C-terminal regions were not visible probably due to their higher mobilities. The structure was composed of a major domain and a minor domain (Figure 1A). The major domain is composed of 4 well-defined helices organized in a typical 4 up−down helix bundle structure in the N-terminal region, whereas the minor domain is located in the C-terminal region and composed of a short, flexible strand and helix. The four α-helices in the major domain are helix-I, residues 6–29; helix-II, residues 32–49; helix-III, residues 54–73; helix-IV, residues 76–89; and they run approximately antiparallel and are connected by short turns. Residues 30–31, 50–53, 74–75 connect α-helices I and II, II and III, III and IV, respectively. Surprisingly, this N-terminal helical domain structure is very similar to that of the C-terminal domain of spinach PsbQ reported previously (Figure 1B and C),[21,22] in spite of a very low sequence identity (around 3.6%) between the N-terminal domain of Psb31 and C-terminal domain of spinach PsbQ. Indeed, the RMSD value between the Cα atoms of Psb31 and PsbQ (PDB ID = 1NZE) is 2.0 Å^2, suggesting a high similarity between the two structures.

The minor, C-terminal domain of Psb31 has a characteristic structure with one short strand (residues 90–107) and one short helix (residues 108–112) (Figure 1A). Four residues in the loop from Glu97 to Ser100 (red circle in Figure 1A) have high temperature factors and their side chains were not clearly observed in the electron density map even at the resolution of 1.55 Å, suggesting that this region is highly flexible. In the crystal structure, these 4 residues (EINS) were modeled as polyalanines. The region from Ser 114 to the C-terminus (Arg 124) is significantly disordered and therefore their structure could not be assigned, suggesting that this C-terminal region is very flexible and mobile, too.

Molecular Surface of Psb31. Figure 2 shows the electrostatic surface potential of Psb31 obtained by the PDB2PQR server.[35] The positive and negative charges are clearly separated on the surface of the Psb31 protein. The positively charged residues (blue area) are mainly composed of Arg 11, Arg 19, Lys 39, Lys 54, Lys 56, Lys 57, Lys 69, Arg 72, and Arg 93, and the negatively charged residues (red area) are mainly composed of Asp 25, Asp 31, Glu 37, Glu 38, Glu 109, and Asp 111. Especially, most of the acidic residues are assembled in the inside region of the helical domain connected with the flexible C-terminal loop. This distribution of surface potential, however, was significantly different from that of the spinach PsbQ protein, in spite of the high similarity in the four helix bundle structure of both proteins (Figure 2).

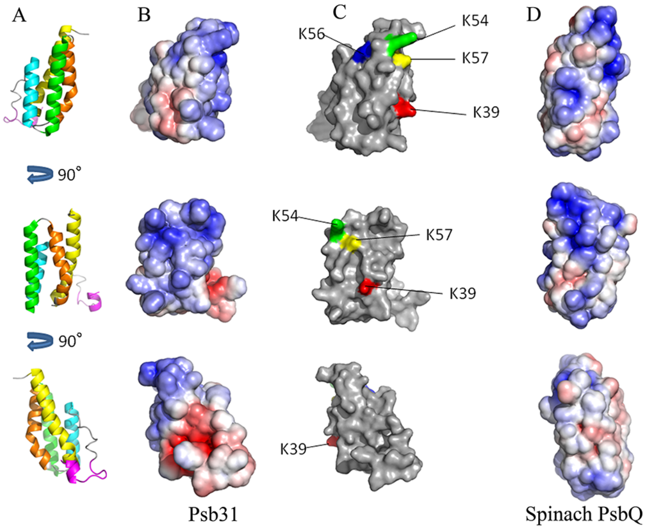


Figure 2. Distribution of electrostatic surface potential of Psb31 obtained by the PDB2PQR server[27] and its comparison with the surface potential of spinach PsbQ. (A) The structure of Psb31 viewed from three angles, with each rotated 90° relative to the one in the uppermost line. (B) The molecular surface of Psb31 colored according to their electrostatic potential, with blue for positive and red for negative. (C) Locations of Lys residues suggested to be responsible for the binding of Psb31 to PSII (Lys 39, Ly54, Lys56, and Lys57). (D) Distribution of molecular surface potential of the spinach PsbQ.

As an extrinsic protein, Psb31 has been shown to bind electrostatically to PSII intrinsic proteins.[24,25] Our preliminary data using chemical modification approaches as reported previously[36] suggest that the positive charges of Psb31 are responsible for its electrostatic interaction with PSII (unpublished data). This is in line with the fact that most of the negative charges on the surface of the Psb31 protein are shielded by the flexible, minor C-terminal domain, so that they would be unable to interact with PSII intrinsic proteins. Among the amino acid residues with positive charges, Lys 39 in helix II, Lys 54, Lys 56, and Lys 57 in helix III are completely conserved within the total 10 Lys residues of Psb31 among different diatom species (Figure 3). The locations of these four Lys residues are shown in Figure 2. One or more of these Lys residues, therefore, is considered to be responsible for the electrostatic interaction between Psb31 and PSII intrinsic proteins

Comparison between the Structures of Psb31 and Higher Plant PsbQ. Although the structure of the major domain of Psb31, namely, the four up−down helix bundle structure, is very similar to that of spinach PsbQ (Figure 1), the structures of the N- and C-terminal regions are largely different between Psb31 and PsbQ. As shown in Figure 4, spinach PsbQ has a flexible N-terminal domain (residues 1–45) that contains two parallel β-strands joining the beginning and end of the N-terminal region and polyproline type II structure,[22] while the diatom Psb31 has only 5 residues in the N-terminal region before the four helix bundle structure. On the other hand, Psb31 has a flexible C-terminal region (residues 90–124) that contains a short strand and short helix, while PsbQ has only one residue in the C-terminal domain after the four helix bundle structure. It thus appears that the long, extra C-terminal region of the Psb31 was moved to the N-terminal region in PsbQ, or vice versa. These structural differences may be related

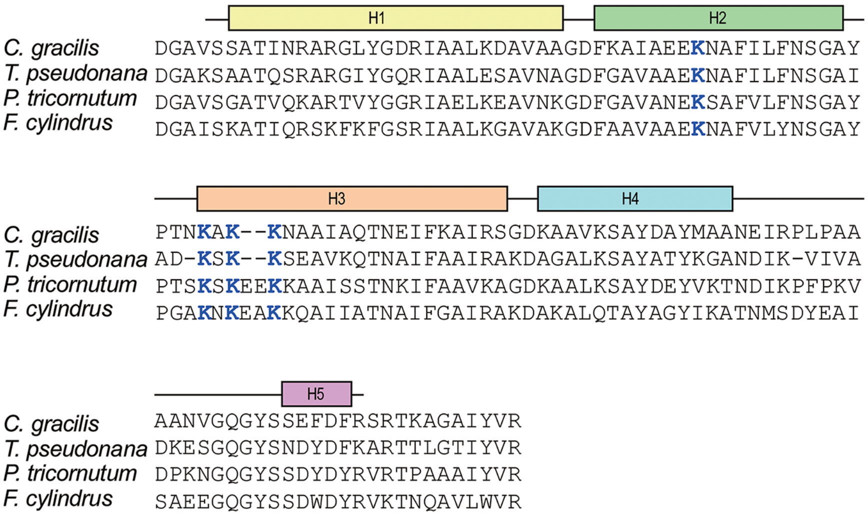


Figure 3. Alignment of the Psb31 sequences from *Chaetoceros gracilis* and some other diatoms, *Thalassiosira pseudonana* (JGI database, protein ID: 24769), *Phaeodactylum tricornutum* (JGI database, protein ID: 46529), and *Fragilariopsis cylindrus* (JGI database, protein ID: 269688). The alignment was performed with the ClustalW software.[37] The Lys residues suggested to be responsible for the Psb31 binding to PSII were highlighted in blue.

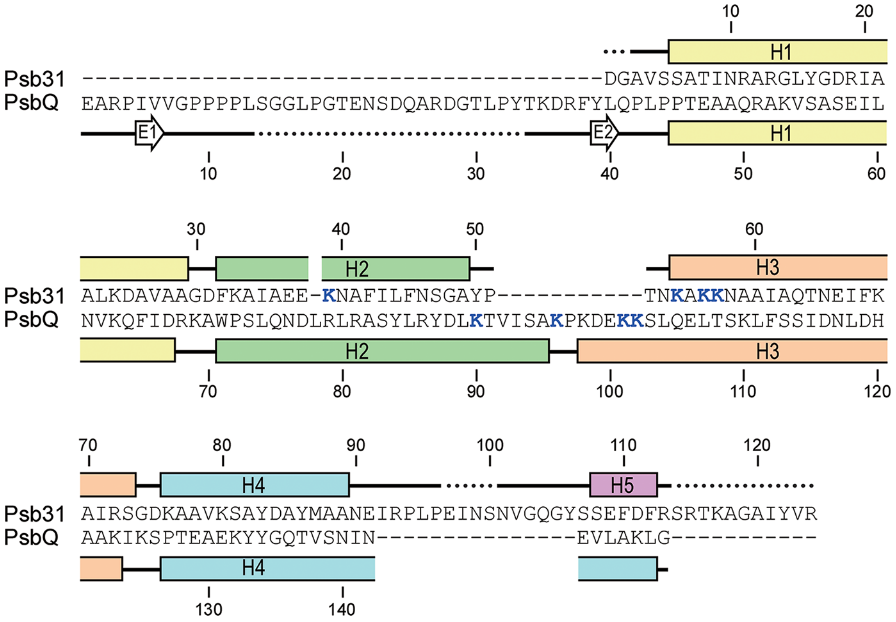


Figure 4. Comparison of the primary and secondary structures between diatom Psb31 and spinach PsbQ based on the Dali server.[34] Lys residues suggested to be responsible for the binding of Psb31 and PsbQ to PSII were colored blue. ″H″ indicates helix, and ″E″ indicates strand.

to differences in their binding properties to PSII and/or their functions.

As described above, the sites responsible for electrostatic binding of Psb31 to PSII is most likely positively charged amino groups of Lys residues on helices II (Lys 39) and III (Lys 54, Lys 56, Lys 57). Our previous results have suggested that the binding partner for Psb31 may be PsbH or Cyt *b*559*α* in PSII intrinsic proteins, because cross-linked products between Psb31 and PsbH or Cyt *b*559*α* were detected by a zero-length cross-linker, EDC, which cross-links carboxyl groups and amino groups that interact electrostatically with each other.[26] Thus, the positive charges of some Lys residues in Psb31 may be in electrostatic interaction with some negative charges of Asp and/or Glu residues in PsbH or Cyt *b*559*α* proteins.

On the other hand, two binding sites of spinach PsbQ with PSII have been reported. One is the N-terminal region of PsbQ, which has been believed for a long time to be necessary for its binding to PSII, because this extrinsic protein cannot rebind to PSII when its N-terminal region is partly degraded.[37] This N-terminal region is absent in the sequence of Psb31, suggesting that the binding of PsbQ to PSII intrinsic proteins is different from that of Psb31.

The other possible site responsible for the binding of higher plant PsbQ to PSII is Lys residues on the C-terminal four helix bundle domain. Meades and co-workers reported that four Lys residues (Lys 90, Lys 96, Lys 101, Lys 102) were labeled with NHS-biotin when PsbQ was modified in solution but they were not labeled when PsbQ is associated with PSII, suggesting that these Lys residues define a site for the interaction between PsbQ and PS II.[38] These Lys residues are located in a cluster that bridges the end of helix II with the beginning of helix III. Surprisingly, the location of these Lys residues in PsbQ is similar to that of four Lys residues Lys 39, Lys 54, Lys 56, Lys 57 of Psb31 that are suggested to be responsible for its binding to PSII (Figures 2 and 3). These results suggest that, in spite of the low sequence similarity and different distribution of electrostatic surface potential between Psb31 and PsbQ, the binding sites on the helical structure may have been conserved. The binding partners for PsbQ are considered to be PsbO and PsbP extrinsic proteins but not PSII intrinsic proteins, because in spinach PSII, PsbQ binds to PSII only through its interaction with both PsbO and PsbP and cannot bind to PSII directly.[39] Furthermore, PsbQ was shown to cross-link with PsbP in spinach PSII using homobifunctional cross-linking reagents.[40] Thus, the partners for the binding of PsbQ appear to be different from that of Psb31.

Origin of Psb31. The 3D structure of cyanoPsbQ was also determined by Jackson et al.,[23] which bore a four helix bundle structure very similar to that of spinach PsbQ. Jackson and co-workers indicated that the α-carbon backbones are clearly related (rmsd = 1.425 Å^2) between cyanoPsbQ and higher plant PsbQ, despite there being only 17% sequence identity between the helical domains of cyanoPsbQ and spinach PsbQ.[23] In addition to cyanoPsbQ and spinach PsbQ, red alga PsbQ′,[12] diatom PsbQ′,[24,25] and green algal PsbQ[10] are known as PsbQ homologous proteins, although their 3D structures have not been determined. Table 2 shows the sequence identities and similarities among these PsbQ homologous proteins together with Psb31. The sequence identities (similarities) of cyanoPsbQ with red algal PsbQ′, diatom PsbQ′, green algal PsbQ, and spinach PsbQ are 16.2% (39.9%), 15.8% (30.9%), 17.1% (34.2%), and 15.1% (35.5%), respectively. The identity is 25.8% even between green algal PsbQ and spinach PsbQ. The identities (similarities) of diatom Psb31 with cyanoPsbQ, red algal PsbQ′, diatom PsbQ′, green algal PsbQ, and spinach PsbQ are 10.6% (25.0%), 12.3% (22.6%), 14.6% (25.5%), 11.7% (22.1%), and 12.1% (25.5%), respectively. It appears that the overall sequence identities between Psb31 and PsbQ are much lower than those among PsbQ proteins from different organisms. Taking into account the overall low sequence identity of PsbQ homologous proteins, however, we may consider that Psb31 belongs to a member of PsbQ family. This is in line with the present structural analysis showing that the diatom Psb31 has a four helix bundle structure very similar to those of cyanoPsbQ and spinach PsbQ. These structural similarities support the notion that cyanoPsbQ was the ancestral version of PsbQ family including Psb31. According to this notion, it may be possible that red algal PsbQ′ and green algal PsbQ arose by divergence of cyanoPsbQ after a primary endosymbiosis event, and diatom PsbQ′ and Psb31 occurred by gene duplication and divergence of the red algal PsbQ′ after a secondary endosymbiosis event. During this process, the diatom Psb31 may have attained the additional C-terminal region not present in the cyanoPsbQ and higher plant PsbQ proteins.

Table 2. Comparison of the Sequence Identities and Similarities between Psb31 and the PsbQ Protein Family[a]

	Psb31	cyano PsbQ	red algal PsbQ′	diatom PsbQ′	green algal PsbQ	spinach PsbQ
Psb31		10.6 (25.0)	12.3 (22.6)	14.6 (25.5)	11.7 (22.1)	12.1 (25.5)
cyano PsbQ	10.6 (25.0)		16.2 (39.9)	15.8 (30.9)	17.1 (34.2)	15.1 (35.5)
red algal PsbQ′	12.3 (22.6)	16.2 (39.9)		24.0 (39.5)	26.0 (42.9)	18.8 (40.9)
diatom PsbQ′	14.6 (25.5)	15.8 (30.9)	24.0 (39.5)		17.3 (34.0)	15.4 (38.3)
green algal PsbQ	11.7 (22.1)	17.1 (34.2)	26.0 (42.9)	17.3 (34.0)		25.8 (43.7)
spinach PsbQ	12.1 (25.5)	15.1 (35.5)	18.8 (40.9)	15.4 (38.3)	25.8 (43.7)	

[a]The sequence identities (%) and similarities (%) (parentheses) were obtained based on ClustalW.[41] CyanoPsbQ: PsbQ of *Synechocystis* sp. PCC 6803;[14] red algal PsbQ′: PsbQ′ of *Cyanidium caldarium*;[13] diatom PsbQ′: PsbQ′ of *Chaetoceros gracilis*;[24] green algal PsbQ: PsbQ of *Chlamydomonas reinhardtii*;[10] spinach PsbQ: PsbQ of *Spinach oleracea*.[9]

However, the possibility remains that cyanoPsbQ is not the ancestral version of Psb31, since the four-helix bundle motif is fairly common in the Protein Data Bank[42] as described by Calderone et al.[21] A search using DALI server[43] showed that the structure of Psb31 is similar to those of spinach PsbQ (*Z* score, 10.2), cyanoPsbQ (*Z* score, 9.9), cytochrome *c* (*Z* score, 8.5), cytochrome *b*562 (*Z* score, 8.1), α-catenin (*Z* score, 7.8), metal-binding protein (*Z* score, 7.7), and adhension kinase (*Z* score, 7.6). A number of other proteins with a similar structural motif were listed by the DALI server. All these structures include a four-helix bundle similar to Psb31, but these proteins have only low sequence similarity and no functional similarity to Psb31. Since Psb31 has high sequence similarities of 68.3−81.0% (51.6−61.9% identities) among different diatom species, and 60.3−65.0% similarities (43.7−51.2% identities) between diatom and brown alga (Table 3), its sequence has been highly conserved during evolution. Thus, the Psb31 protein seems to form a family distinctly different from that of the PsbQ family. If Psb31 was not diverged from cyanoPsbQ, it must have originated from an origin that is not yet well understood.

In any events, the similar structures of Psb31 and PsbQ suggest that there are two copies of PsbQ-like proteins as extrinsic subunits in the diatom PSII. These two copies,

Table 3. Comparison of the Sequence Identities and Similarities of the Psb31 Protein from Different Species[a]

	diatom				brown alga
	C. gracilis	*T. pseudonana*	*P. tricornutum*	*F. cylindrus*	*E. siliculosus*
C. gracilis		58.9 (74.2)	61.9 (81.0)	51.6 (68.3)	49.6 (63.2)
T. pseudonana	58.9 (74.2)		57.9 (68.3)	52.4 (68.3)	51.2 (65.0)
P. tricornutum	61.9 (81.0)	57.9 (68.3)		54.8 (74.6)	43.7 (60.3)
F. cylindrus	51.6 (68.3)	52.4 (68.3)	54.8 (74.6)		45.7 (63.8)
E. siliculosus	49.6 (63.2)	51.2 (65.0)	43.7 (60.3)	45.7 (63.8)	

[a]The sequence identities (%) and similarities (%) (parentheses) were obtained based on analysis with ClustalW.[41]

however, have different binding and functional properties. While the Psb31 protein was shown to be able to bind to PSII independent of other extrinsic subunits, the "real" PsbQ protein binds to PSII only through its interactions with PsbP and PsbO in higher plant PSII. Furthermore, Psb31 functions to replace in part the PsbO protein to support oxygen evolution, whereas PsbQ functions to mitigate the requirement for Cl^- and probably also for Ca^{2+} in the presence of PsbO and PsbP. Thus, the diatom PSII has attained two extrinsic proteins with similar structures to fulfill distinctly different functions. In this respect, it is particularly worth mentioning that the Psb31 protein can function to partly replace that of PsbO. Since higher plant PSII has been reported to contain two copies of PsbO,[44,45] a plausible hypothesis is therefore that during the evolution of diatom to higher plant, Psb31 was replaced by the PsbO protein, resulting in the presence of the two copies of PsbO in higher plant PSII. This explains why cyanobacterial PSII has only one copy of PsbO as revealed by the crystal structural analysis of cyanobacterial PSII.[1,17−19]

■ ASSOCIATED CONTENT

Accession Codes

The structural data reported in this paper has been deposited in the Protein Data Bank with the accession code of 4K7B.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: shen@cc.okayama-u.ac.jp; Tel: +81-86-251-8502; Fax: +81-86-251-8502;.

Present Address

Ryo Nagao, Division of Material Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464−8602, Japan.

Author Contributions

Ryo Nagao and Michihiro Suga contributed equally to this work.

Funding

This work was supported by a Grant-in-Aid for Specially Promoted Research (No. 24000018 to J.R.S.), Grant-in-Aids for Scientific Research (21570038, 22370017 to T.T.), from the Ministry of Education of Japan, and a grant from JST PRESTO (T.T.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Drs. Yasufumi Umena and Keisuke Kawakami who helped in some of the crystallization and data collection experiments. We also thank the staff members at beamlines BL41XU and BL44XU of SPring-8, Japan, for their kind help in the X-ray data collection.

■ ABBREVIATIONS

MES, 2-(*N*-morpholino)ethanesulfonic acid; PSII, Photosystem II

■ REFERENCES

(1) Umena, Y., Kawakami, K., Shen, J.-R., and Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature 473*, 55−60.

(2) Enami, I., Suzuki, T., Tada, O., Nakada, Y., Nakamura, K., Tohri, A., Ohta, H., Inoue, I., and Shen, J.-R. (2005) Distribution of the extrinsic proteins as a potential marker for the evolution of photosynthetic oxygen-evolving photosystem II. *FEBS J. 272*, 5020−5030.

(3) Roose, J. L., Wegener, K. M., and Pakrasi, H. B. (2007) The extrinsic proteins of Photosystem II. *Photosynth. Res. 92*, 369−387.

(4) Enami, I., Okumura, A., Nagao, R., Suzuki, T., Iwai, M., and Shen, J.-R. (2008) Structures and functions of the extrinsic proteins of photosystem II from different species. *Photosynth. Res. 98*, 349−363.

(5) Bricker, T. M., Roose, J. L., Fagerlund, R. D., Frankel, L. K., and Eaton-Rye, J. J. (2012) The extrinsic proteins of Photosystem II. *Biochim. Biophys. Acta 1817*, 121−142.

(6) Shen, J.-R., Ikeuchi, M., and Inoue, Y. (1992) Stoichiometric association of extrinsic cytochrome c_{550} and 12 kDa protein with a highly purified oxygen-evolving photosystem II core complex from *Synechococcus vulcanus*. *FEBS Lett. 301*, 145−149.

(7) Shen, J.-R, and Inoue, Y. (1993) 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.

(8) Enami, I., Murayama, H., Ohta, H., Kamo, M., Nakazato, K., and Shen, J.-R. (1995) Isolation and characterization of a Photosystem II complex from the red alga *Cyanidium caldarium*: association of cytochrome *c*-550 and a 12 kDa protein with the complex. *Biochim. Biophys. Acta 1232*, 208−216.

(9) Berthold, D. A., Babcock, G. T., and Yocum, C. F. (1981) A highly resolved, oxygen-evolving photosystem II preparation from spinach thylakoid membranes: EPR and electron-transport properties. *FEBS Lett. 134*, 231−234.

(10) Suzuki, T., Minagawa, J., Tomo, T., Sonoike, K., Ohta, H., and Enami, I. (2003) Binding and functional properties of the extrinsic proteins in oxygen-evolving photosystem II particle from a green alga, *Chlamydomonas reinhardtii* having his-tagged CP47. *Plant Cell Physiol. 44*, 76−84.

(11) Suzuki, T., Tada, O., Makimura, M., Tohri, A., Ohta, H., Yamamoto, Y., and Enami, I. (2004) Isolation and characterization of oxygen-evolving photosystem II complexes retaining the PsbO, P and Q proteins from *Euglena gracilis*. *Plant Cell Physiol. 45*, 1168−1175.

(12) Enami, I., Kikuchi, S., Fukuda, T., Ohta, H., and Shen, J.-R. (1998) Binding and functional properties of four extrinsic proteins of photosystem II from a red alga, *Cyanidium caldarium*, as studied by release-reconstitution experiments. *Biochemistry 37*, 2787−2793.

(13) Ohta, H., Suzuki, T., Ueno, M., Okumura, A., Yoshihara, S., Shen, J.-R., and Enami, I. (2003) Extrinsic proteins of photosystem II: an intermediate member of PsbQ protein family in red algal PS II. *Eur. J. Biochem. 270*, 4156−4163.

(14) Kashino, Y., Lauber, W. M., Carroll, J. A., Wang, Q., Whitmarsh, J., Satoh, K., and Pakrasi, H. B. (2002) Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. *Biochemistry 41*, 8004−8012.

(15) Thornton, L. E., Ohkawa, H., Roose, J. L., Kashino, Y., Keren, N., and Pakrasi, H. B. (2004) Homologs of plant PsbP and PsbQ proteins are necessary for regulation of photosystem II activity in the cyanobacterium *Synechocystis* 6803. *Plant Cell 16*, 2164−2175.

(16) Summerfield, T. C., Shand, J. A., Bentley, F. K., and Eaton-Rye, J. J. (2005) PsbQ (Sll1638) in *Synechocystis* sp. PCC 6803 is required for photosystem II activity in specific mutants and in nutrient-limiting conditions. *Biochemistry 44*, 805−815.

(17) Kamiya, N., and Shen, J.-R. (2003) Crystal structure of oxygen-evolving photosystem II from *Thermosynechococcus vulcanus* at 3.7-Å resolution. *Proc. Natl. Acad. Sci. U. S. A. 100*, 98−103.

(18) Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J., and Iwata, S. (2004) Architecture of the photosynthetic oxygen-evolving center. *Science 303*, 1831−1838.

(19) Loll, B., Kern, J., Saenger, W., Zouni, A., and Biesiadka, J. (2005) Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II. *Nature 438*, 1040−1044.

(20) Ifuku, K., Nakatsu, T., Kato, H., and Sato, F. (2004) Crystal structure of the PsbP protein of photosystem II from *Nicotiana tabacum*. *EMBO Rep. 5*, 362−367.

(21) Calderone, V., Trabucco, M., Vujičić, A., Battistutta, R., Giacometti, G. M., Andreucci, F., Barbato, R., and Zanotti, G. (2003) Crystal structure of the PsbQ protein of photosystem II from higher plants. *EMBO Rep. 4*, 900−905.

(22) Balsera, M., Arellano, J. B., Revuelta, J. L., de las Rivas, J., and Hermoso, J. A. (2005) The 1.49 Å resolution crystal structure of PsbQ from photosystem II of *Spinacia oleracea* reveals a PPII structure in the N-terminal region. *J. Mol. Biol. 350*, 1051−1060.

(23) Jackson, S. A., Fagerlund, R. D., Wilbanks, S. M., and Eaton-Rye, J. J. (2010) Crystal structure of PsbQ from *Synechocystis* sp. PCC 6803 at 1.8 Å: Implications for binding and function in cyanobacterial photosystem II. *Biochemistry 49*, 2765−2767.

(24) Nagao, R., Ishii, A., Tada, O., Suzuki, T., Dohmae, N., Okumura, A., Iwai, M., Takahashi, T., Kashino, Y., and Enami, I. (2007) Isolation and characterization of oxygen-evolving thylakoid membranes and photosystem II particles from a marine diatom *Chaetoceros gracilis*. *Biochim. Biophys. Acta 1767*, 1353−1362.

(25) Nagao, R., Tomo, T., Noguchi, E., Nakajima, S., Suzuki, T., Okumura, A., Kashino, Y., Mimuro, M., Ikeuchi, M., and Enami, I. (2010) Purification and characterization of a stable oxygen-evolving Photosystem II complex from a marine centric diatom *Chaetoceros gracilis*. *Biochim. Biophys. Acta 1797*, 160−166.

(26) Okumura, A., Nagao, R., Suzuki, T., Yamagoe, S., Iwai, M., Nakazato, K., and Enami, I. (2008) A novel protein in Photosystem II of a diatom *Chaetoceros gracilis* is one of the extrinsic proteins located on lumenal side and directly associates with PSII core components. *Biochim. Biophys. Acta 1777*, 1545−1551.

(27) Nagao, R., Moriguchi, A., Tomo, T., Niikura, A., Nakajima, S., Suzuki, T., Okumura, A., Iwai, M., Shen, J.-R., Ikeuchi, M., and Enami, I. (2010) Binding and functional properties of five extrinsic proteins in oxygen-evolving photosystem II from a marine centric diatom *Chaetoceros gracilis*. *J. Biol. Chem. 285*, 29191−29199.

(28) Doublié, S. (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol. 276*, 523−530.

(29) Ishikawa, K., Higashi, N., Nakamura, T., Matsuura, T., and Nakagawa, A. (2007) The first crystal structure of L-threonine dehydrogenase. *J. Mol. Biol. 366*, 857−867.

(30) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307−326.

(31) Collaborative Computational Project, number 4. The *CCP4* suite: programs for proten crystallography. *Acta Crystallogr., Sect. D* 1994 *50*, 760−763.

(32) Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr., Sect. D 58*, 1772−1779.

(33) Emsley, P., and Cowtan, K. (2004) Coot: model-buiding tools for molecular graphics. *Acta Crystallogr., Sect. D 60*, 2126−2132.

(34) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D 66*, 213−221.

(35) Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res. 32*, W665−W667.

(36) Miura, T., Shen, J.-R., Takahashi, S., Kamo, M., Nakamura, E., Ohta, H., Kamei, A., Inoue, Ya., Domae, N., Takio, K., Nakazato, K., Inoue, Yo., and Enami, I. (1997) Identification of domains on the extrinsic 33-kDa protein possibly involved in electrostatic interaction with photosystem II complex by means of chemical modification. *J. Biol. Chem. 272*, 3788−3798.

(37) Kuwabara, T., Murata, T., Miyao, M., and Murata, N. (1986) Partial degradation of the 18-kDa protein of the photosynthetic oxygen-evolving complex: a study of a binding site. *Biochim. Biophys. Acta 850*, 146−155.

(38) Meades, G. D., Jr., McLachlan, A., Sallans, L., Limbach, P. A., Frankel, L. K., and Bricker, T. M. (2005) Association of the 17-kDa extrinsic protein with photosystem II in higher plants. *Biochemistry 44*, 15216−15221.

(39) Miyao, M., and Murata, N. (1989) The mode of binding of three extrinsic proteins of 33 kDa, 23 kDa and 18 kDa in photosystem II complex of spinach. *Biochim. Biophys. Acta 977*, 315−321.

(40) Enami, I., Mochizuki, Y., Takahashi, S., Kakuno, T., Horio, T., Satoh, K., and Katoh, S. (1990) Evidence from crosslinking for nearest-neighbor relationships among the three extrinsic proteins of spinach photosystem II complexes that are associated with oxygen evolution. *Plant Cell Physiol. 31*, 725−729.

(41) Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics 23*, 2947−2948.

(42) Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D., and Zardecki, C. (2002) The Protein Data Bank. *Acta Crystallogr., Sect. D 58*, 899−907.

(43) Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol. 233*, 123−138.

(44) Popelkova, H., and Yocum, C. F. (2011) PsbO, the manganese-stabilizing protein: analysis of the structure-function relations that provide insights into its role in photosystem II. *J. Photochem. Photobiol. B 104*, 179−190.

(45) Roose, J. L., Yocum, C. F., and Popelkova, H. (2011) Binding stoichiometry and affinity of the manganese-stabilizing protein affects redox reactions on the oxidizing side of photosystem II. *Biochemistry 50*, 5988−5998.