Winter 12-15-2016

Investigating the Life-cycle of Photosystem II Using Mass Spectrometry

Daniel Avram Weisz
Washington University in St. Louis

WASHINGTON UNIVERSITY

Department of Chemistry

Dissertation Examination Committee:
Michael L. Gross, Co-chair
Himadri B. Pakrasi, Co-chair
Robert E. Blankenship
Dewey Holten
Joseph Jez

Investigating the Life-cycle of Photosystem II Using Mass Spectrometry

by

Daniel Avram Weisz

A dissertation presented to
The Graduate School
of Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

December, 2016

Saint Louis, Missouri

Table of Contents

List of figures

Chapter 4: Structural localization of PsbQ in Photosystem II using chemical cross-linking and mass spectrometry

Chapter 5: Conclusions, future directions, and additional work

List of tables

List of Abbreviations

A. thaliana, *Arabidopsis thaliana*

BS^3, (bis(sulfosuccinimidyl)suberate)

C. reinhardtii, *Chlamydomonas reinhardtii*

BS^3, bis(sulfosuccinimidyl)suberate

DSP, dithiobis(succinimidyl propionate)

DTSSP, 3,3'-dithiobis(sulfosuccinimidyl propionate)

EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

ESI, electrospray ionization

FAB, fast atom bombardment

FTICR, Fourier transform ion cyclotron resonance

GEE, glycine ethyl ester

LC, liquid chromatography

LHCII, Light-harvesting complex II

LMM, low-molecular-mass

LTQ, linear ion trap quadrupole

MALDI, matrix-assisted laser desorption ionization

MS, mass spectrometry

N. tabacum, *Nicotiana tabacum*

NHS, N-hydroxysuccinimide

OCP, orange carotenoid protein

PBS, phycobilisome

PSII, Photosystem II

PC, plastocyanin

PTM, post-translational modification

QqQ, triple-quadrupole

SCP, small CAB-like protein

Synechococcus 7002, *Synechococcus* sp. PCC 7002

Synechocystis 6803, *Synechocystis* sp. PCC 6803

T. elongatus, Thermosynechococcus elongatus

T. vulcanus, Thermosynechococcus vulcanus

TOF, time-of-flight

WOC, water-oxidizing complex

Acknowledgements

I would like to thank my advisors, Dr. Michael Gross and Dr. Himadri Pakrasi for their wise, caring guidance during my graduate studies. I am lucky to have worked in the labs of two such brilliant scientists and excellent mentors. I also thank the other members of my thesis committee, Drs. Bob Blankenship, Dewey Holten, and Joe Jez, for helping to steer my research in the right direction with their insightful suggestions during annual committee meetings.

I was trained primarily by Dr. Haijun Liu in the Pakrasi lab and by Dr. Hao Zhang in the Gross lab; I am so grateful to them for their patience and expertise, from which I benefitted greatly. I would also like to thank Drs. Kim Wegener, Michelle Liberton, and Anindita Bandyopadhyay in the Pakrasi lab, and Drs. Henry Rohrs and Manolo Plasencia in the Gross lab, for training and helpful advice throughout.

I thank Dr. Phil Jackson (University of Sheffield) and Dr. Hao Zhang for helpful discussions while writing Chapter 1. Dr. Haijun Liu purified Photosystem II complexes from the Δ*ctp*A-His47, Δ*ctp*A:Δ*psb27*-His47, and His27 strains of *Synechocystis* 6803 analyzed by immunoblotting in Chater 2, Figure 1A. I thank members of Dr. Christoph Borchers' laboratory (University of Victoria), namely Dr. Evgeniy Petrotchenko, Karl Makepeace, and Jason Serpa, for helpful advice while implementing their isotope-encoded cross-linking platform in our lab. The protein-protein docking computations in Chapter 2 were performed by Dr. Sundarapandian Thangapandian in the lab of Dr. Emad Tajkhorshid (University of Illinois, Urbana-Champaign). I would like to thank Sundar for an enjoyable, productive collaboration. I thank Dr. Terry Bricker for the HT3 strain of *Synechocystis* 6803 and other members of the Gross and Pakrasi labs for collegial discussions that have contributed to the work described in this dissertation.

The majority of the experimental work in Chapter 4 was conducted by Dr. Haijun Liu; Figures 1, 10C, 11, and S1 represent my work. I co-wrote the first paper below which describes

part of the work in Chapter 4, and wrote the second paper below which describes the rest of the work in Chapter 4:

Liu, H., Zhang, H., Weisz, D.A., Vidavsky, I., Gross, M.L., Pakrasi, H.B. (2014). MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. *Proc. Natl. Acad. Sci. USA* 111, 4638-4643.

Liu, H*., Weisz, D.A*., and Pakrasi, H.B. (2015). Multiple copies of the PsbQ protein in a cyanobacterial photosystem II assembly intermediate complex. *Photosynth. Res.* 126, 375-383. *co-first authors

The work in each chapter was supported by the Photosynthetic Antenna Research Center, an Energy Frontier Research Center funded by the U.S. Department of Energy (DOE), Office of Basic Energy Sciences (Grant DE-SC 0001035 to Drs. Gross and Pakrasi). Additionally, work in Chapter 1 was supported by the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy (grant no. DE-FG02-99ER20350 to Dr. Pakrasi); work in Chapter 2 was supported by the National Institutes of Health (Grants U54-GM087519, and P41-GM104601 to Dr. Tajkhorshid.), and we acknowledge computing resources provided by Blue Waters at National Center for Supercomputing Applications (NCSA), and XSEDE (grant TG-MCA06N060 to Dr. Tajkhorshid.). The mass spectrometry research in Chapter 2 was supported in part by the National Institutes of Health (Grant 2P41GM103422 to Dr. Gross). Funding for the work in Chapter 4 was provided by National Science Foundation Grant MCB0745611 (to Dr. Pakrasi) and the National Institute of General Medical Sciences (8 P41 GM103422-35) of the NIH (to Dr. Gross). We also thank the mass spectrometry facility of the Photosynthetic Antenna Research Center (PARC), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award# DE-SC0001035, for providing access to the mass spectrometer (Synapt G2) and data analysis (MassMatrix).

I would also like to thank my undergraduate research advisor, Dr. Jeffrey Agar, and the members of his lab at Brandeis University, especially Drs. Jared Auclair, Jennifer Cobb, and Kristin Boggio, for initial training in cyanobacterial culture and protein mass spectrometry; the late Charlie Ingersoll, for the good times we shared building a cyanobacterial growth chamber in the Rosenstiel machine shop at Brandeis; and to Dr. Greg Petsko, for sparking my interest in biochemistry and proteins while at Brandeis.

Ellie, thank you for your constant love and support, which has been especially meaningful during the busy period of writing this dissertation.

I would like to express my appreciation to my parents for everything, always; in this context, especially for their helpful guidance and faith in me.

Dedication

This dissertation is dedicated in loving memory of my grandfather, George Weisz. He kept his completed chemistry doctoral dissertation in a drawer unsubmitted, rather than compromise his values in the Communist regime under which he lived. His unwavering principles are a constant inspiration.

Abstract

Photosystem II (PSII) is a protein complex found embedded in the thylakoid membranes of all organisms that perform oxygenic photosynthesis. PSII converts sunlight into chemical energy, filling our atmosphere with molecular oxygen in the process and supporting nearly all life on Earth. PSII undergoes frequent light-induced damage as an unavoidable result of the electron transfer reactions it catalyzes. When damaged, PSII is disassembled, repaired, and reassembled in an intricate, tightly regulated process. The structure and mechanism of function of active PSII are relatively well-understood, due to the available crystal structures of the active complex and many years of biochemical and biophysical investigation. However, many aspects of the broader PSII life-cycle are less clear.

In this work, several structural aspects of the PSII life-cycle are investigated, with an approach that emphasizes mass spectrometry (MS)-based tools. The field of protein MS is developing rapidly, and, especially in the last several years, has become a key tool for addressing a variety of questions in the field of photosynthesis. Chapter 1 provides an in-depth review of the ways in which MS has been, and can be, applied to PSII life-cycle research. This chapter presents the relevant MS-based techniques, as well as the knowledge that has been gained about the PSII life-cycle through their application.

The cyanobacterial Psb28 protein binds transiently to a PSII assembly intermediate complex, exerting a protective effect on this complex. However, since Psb28 dissociates before assembly is complete, it is not found in the crystal structure and its structural location within the complex has remained unknown. We used isotope-encoded chemical cross-linking followed by mass spectrometry to identify the binding partners of Psb28 in the model cyanobacterium *Synechocystis* sp. PCC 6803, the organism used throughout this work. We identified three cross-links between Psb28 and the α- and β- subunits of cytochrome b_{559} (PsbE and PsbF), pinpointing

the structural location of Psb28 on the cytosolic surface of PSII in close association with these subunits. Our results allow us to propose several mechanisms by which Psb28 could exert its protective effect.

In Chapter 3, we used high-resolution tandem MS to identify oxidative modifications in PSII. We found that the total number of modified residues increased by over 50% following light incubation, with the D1 protein showing the most marked increase (3.3-fold) of the proteins we monitored. These results strongly support the idea that ROS are generated as a byproduct of PSII photochemistry and that they damage PSII subunits, especially D1, which has the fastest turnover rate of all the subunits. By mapping the modified residues onto the PSII crystal structure, we found that the lumen-side residues form two nearly continuous, roughly linear "arms" starting at the Mn_4Ca cluster and radiating outward all the way to the surface of PSII. We propose that these two "arms" are oxygen/ROS exit channels that protect PSII by removing ROS from the complex after they are generated at the Mn_4Ca cluster. It has long been believed that PSII must contain such channels, and this study provides the most complete and descriptive molecular-level evidence yet for their existence and location.

Chapter 4 describes a study that used cross-linking and MS to identify the binding location of PsbQ on the lumenal surface of PSII. Though PsbQ is a necessary component of PSII complexes with highest oxygen-evolving activity, it is not found in the available cyanobacterial crystal structures. Our results show that PsbQ helps stabilize the PSII dimer, providing a structural basis to explain our biochemical data and previous findings. A novel PSII subcomplex with multiple copies of the PsbQ protein was also discovered, and its characterization is described in this chapter. Based on our results, we propose it is a late PSII assembly intermediate that stabilizes the active PSII dimer just before association of the other lumenal extrinsic proteins PsbU and PsbV.

Chapter 5 summarizes the results of this work and provides an outlook for future MS contributions to PSII research. In addition, the preliminary results of an ongoing study to characterize a novel PSII subcomplex are presented.

Chapter 1: Introduction

The use of advanced mass spectrometry to dissect the life-cycle of Photosystem II

This chapter was adapted from:

Weisz, D.A., Gross, M.L., and Pakrasi, H.B. (2016). The use of advanced mass spectrometry to dissect the life-cycle of Photosystem II. *Front. Plant Sci.* 7, 617.

Summary

Photosystem II (PSII) is a photosynthetic membrane-protein complex that undergoes an intricate, tightly regulated cycle of assembly, damage, and repair. The available crystal structures of cyanobacterial PSII are an essential foundation for understanding PSII function, but nonetheless provide a snapshot only of the active complex. To study aspects of the entire PSII life-cycle, mass spectrometry (MS) has emerged as a powerful tool that can be used in conjunction with biochemical techniques. This chapter presents the MS-based approaches that are used to study PSII composition, dynamics, and structure, and reviews the information about the PSII life-cycle that has been gained by these methods. This information includes the composition of PSII subcomplexes, discovery of accessory PSII proteins, identification of post-translational modifications and quantification of their changes under various conditions, determination of the binding site of proteins not observed in PSII crystal structures, conformational changes that underlie PSII functions, and identification of water and oxygen channels within PSII. The conclusion of this introduction outlines the work presented in the subsequent chapters.

Introduction

Since the late 1990s, mass spectrometry (MS) has become a central tool for the study of proteins and their role in biology. The advent of electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) permits the ionization of peptides and proteins and their introduction into the gas phase, enabling their analysis by MS. The typical "bottom-up" workflow that emerged in the wake of these breakthroughs involves: 1) enzymatic digestion (often by trypsin) of a protein to produce peptides of small enough size (typically 1-3 kDa) to be ionized and fragmented efficiently in a mass spectrometer; 2) liquid chromatographic (LC) separation of the peptides; and 3) online (or offline) injection of the separated peptides into a mass spectrometer.

The "top-down" approach is an attractive alternative that eliminates the protein digestion step, but the subsequent steps are generally more difficult for intact proteins than peptides, and this approach is currently best-suited for small, soluble proteins. After injection of the peptides, a typical tandem MS analysis consists of: 1) ionization of the peptide sample by ESI or MALDI and introduction into the gas phase; 2) measurement of the mass-to-charge (m/z) ratio of the intact peptide (also referred to as "MS^1" analysis); and 3) fragmentation of the precursor ion and measurement of its "product-ion" spectrum ("MS/MS" or "MS^2" analysis), which provides information about the peptide's amino acid sequence. When genomic information is available to predict the sequence of all proteins in the organism, computer analysis of the peptide masses and product-ion spectra can determine the highest-scoring match for each peptide from the protein database. This highest-scoring match is taken as the identity of the peptide, assuming data quality meets certain statistical criteria. A given protein is then determined to have been present in the sample if the quality and number of its peptide hits meet an additional set of statistical criteria. The ability to identify many proteins in a sample at once by MS has become the cornerstone of the field of proteomics.

Protein identification is only the most basic application of MS-based proteomics, and it has traditionally been described as the first "pillar" of the field. The second pillar is characterization of the many proteoforms that exist for each protein, arising, e.g., from splice variants and post-translational modifications (PTMs). These two pillars address questions about the *composition* of a protein sample. The third pillar is quantification—either absolute or relative—of proteins using isotopic labeling or label-free approaches. This pillar is typically used to address questions about the *dynamics* of a system- how composition of proteins or proteoforms changes over time, space, or under different environmental conditions or perturbations. A proposed fourth pillar focuses on

the emerging area of structural proteomics that uses MS-based techniques to address questions about the three-dimensional *structure* of proteins and protein complexes in a cell.

These four pillars of proteomics have each become indispensable tools for gleaning information about photosynthesis (Bricker et al., 2015; Battchikova et al., 2015; Heinz et al., 2016) and, in particular for this work, the life-cycle of PSII. A search for publications containing

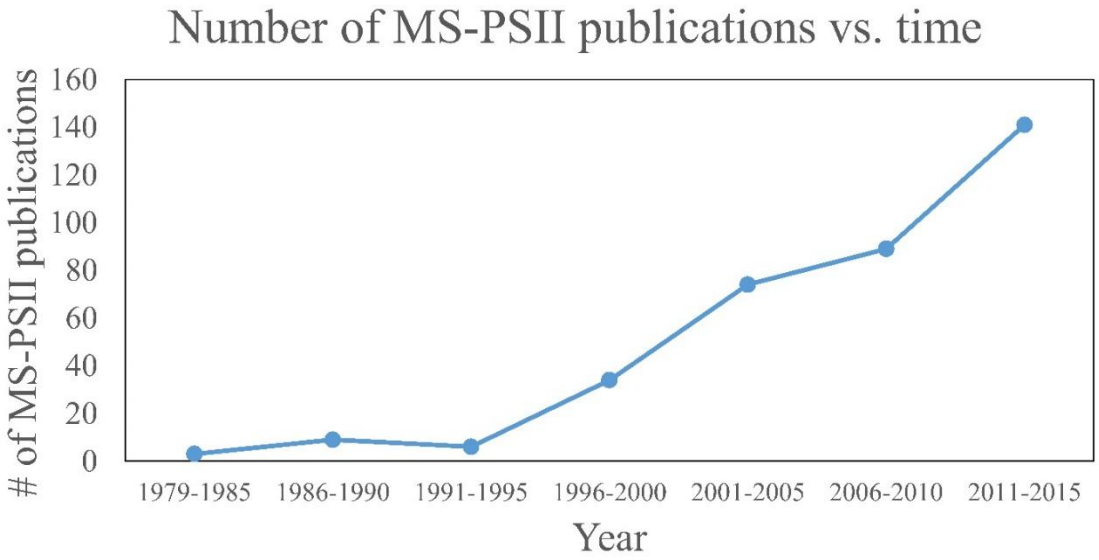


Figure 1. Publications that use MS for PSII research, 1979-2015.

Publications that contain "Photosystem II" and "mass spectrometry" in their article title, abstract, or keywords were searched on the Scopus database. Each data point represents the total number of publications for that range of years.

both "Photosystem II" and "mass spectrometry" in the article title, abstract, and/or keywords was performed on the Scopus database. The results, displayed in Fig. 1, show that prior to the advent of ESI and MALDI in the late 1980s, publications were nearly zero per year. Starting in the early 1990s and continuing through 2015, publications have risen steadily, with around 20-30 publications per year in the last several years. The rise can be attributed to method and instrument

development, and to increasing accessibility of MS instrumentation to biology researchers. An overview of how MS-based tools are typically applied to PSII life-cycle research is given in Table 1. This chapter focuses on MS of proteins. However, it should be noted that another widely used application of MS in PSII research is the analysis of the isotopic composition of evolved oxygen by membrane-inlet mass spectrometry. This technique has yielded significant insight into the mechanistic aspects of water oxidation by PSII (reviewed in Shevela and Messinger, 2013).

In the sections that follow, we consider questions of PSII composition, dynamics, and structure separately. For each area, a brief overview of the relevant MS-based tools is given, followed by examples of several PSII life-cycle research areas that have benefitted from these techniques. In the final section, the outlook for future contributions of MS techniques to PSII life-cycle research is discussed.

Table 1. Overview of the role of MS in PSII life-cycle research

Kind of information	Information desired	MS-based technique
Composition	PSII subunits present in a complex	Bottom-up MS (intact or top-down MS for LMM subunits)
	Accessory proteins that associate with PSII	Bottom-up MS
	PTMs	Bottom-up MS
Dynamics	Protein and PTM changes between samples	Label-free or isotopic-label-based relative quantification
	PSII subunit lifetime	Rate of unlabeled protein disappearance after isotopic label exposure
	Relative position of subcomplexes in PSII life-cycle	Relative isotopic label incorporation after pulse
Structure	Binding site of proteins not found in PSII crystal structures	Cross-linking, footprinting
	Conformational changes	Footprinting, quantify changes in modification extent
	Water and oxygen channel detection	Footprinting; mapping oxidative modifications

Composition of PSII complexes

MS-based methods to study the composition of PSII complexes

PSII subunits with soluble domains

The bottom-up MS workflow is highly effective at identifying soluble proteins or proteins with soluble domains. It is, therefore, the main MS strategy that has been used to detect the core PSII proteins D1, D2, CP43, and CP47 (which are transmembrane proteins but have multiple soluble domains), the extrinsic (soluble) PSII proteins, or unknown PSII-bound proteins. Bottom-up MS analysis can be preceded by either in-gel or in-solution digestion of the protein, each with advantages. Gel electrophoresis serves as a one- or two-dimensional fractionation step, simplifying the mixture to be analyzed by MS. Using this approach to remove interferences can improve instrument sensitivity towards proteins in the band of interest. Native PAGE, either alone or followed by denaturing SDS-PAGE (2D-BN-PAGE), is a common choice for resolving multiple protein complexes in a thylakoid membrane or purified PSII preparation; unknown bands can be excised and analyzed by MS to identify components of specific complexes (Granvogl et al., 2008; Pagliano et al., 2014; Gao et al., 2015). However, targeted band excision can miss potentially important proteins that migrated at positions not selected for in-gel digestion. In theory, native PAGE can remove unbound proteins from complexes, simplifying MS analysis; however, disruption of certain relevant protein-protein interactions in complexes cannot ever be fully excluded. Alternatively, in-solution digestion allows a more comprehensive analysis of the protein components in a sample, but without the sample simplification or complex-specific resolution provided by prior SDS-PAGE or native PAGE.

MS instrumentation, as well as membrane-protein sample preparation (Whitelegge, 2013; Battchikova et al., 2015; Heinz et al., 2016) and bioinformatics capabilities, has improved over the last two decades to facilitate PSII life-cycle research (Table 2 summarizes the kinds of experiments

that have been performed and the main MS instrumentation and features that enable them). Early mass spectrometers that were applied to PSII research, especially triple-quadrupole (QqQ) and MALDI-time-of-flight (MALDI-TOF) instruments, had relatively low sensitivity, resolving power, and mass accuracy (on the order of 100-several hundred ppm) (Michel et al., 1988; Sharma et al., 1997a,b,c; Frankel et al., 1999). Scarcity of genomic sequence data combined with low instrument sensitivity, mass accuracy, and fragmentation efficiency meant that sample analysis was mainly restricted to highly purified PSII complexes or individual subunits, with poor capability for novel protein identification. The mid-2000s saw the appearance of higher-performing instruments, especially the hybrid quadrupole-TOF (Q-TOF) and increasing availability of genomic sequence data for commonly studied photosynthetic organisms. These enabled routine bottom-up identification of the main subunits of PSII complexes (those with soluble domains) from more complex starting mixtures and identification of novel PSII-associated proteins (Kashino et al., 2002; Heinemeyer et al., 2004; Komenda et al., 2005). The fragmentation efficiency of the Q-TOF, however, still limited sequence coverage of proteins. The development and distribution of Fourier transform instruments (ion cyclotron resonance and orbitraps) sometimes interfaced with ion traps provided improved fragmentation efficiency and enabled analysis of highly complex mixtures with higher sequence coverage than ever before. These instruments allow proteome-wide experiments, enable routine confident PTM site identification, and have opened the door for bottom-up MS experiments on photosynthetic systems not before feasible (see Table 2 and sections below).

Table 2. MS instruments and instrument features for PSII life-cycle research applications
(footnotes on next page)

Biological application	High mass accuracy (MS^1)	High mass accuracy (MS^2)	High sensitivity/ Good sequence coverage	QqQ	TOF	Q-TOF	LTQ-Orbitrap	Q-Exactive[a]	LTQ-FT-ICR
ID proteins- purified PSII complex/ simple mixture/ gel band	Med	Low	Low	+	+	++	++	++	++
ID LMM subunits- purified PSII complex- intact/ top-down	Med	Low	Low	+	+	++	++	++	++
ID proteins- membranes/ complex mixture/ unknown protein scarch	High	Med	Med	-	-	+	++	++	++
ID modifications- targeted search	High	Med	Med	+	+	+	++	++	++
ID modifications- non-targeted search (PTMs, footprinting)	High	Med	High	-	-	-	+	++	+
Quantification of proteins/ modifications[b]- targeted search	High	Med	Med	+	+	+	++	++	++
Quantification of proteins/ modifications[b]- non-targeted scarch	High	Med	High	-	-[c]	-	+	++	+
Cross-linking	High	Med-High	High	-	-	-	+	++	+

High = high priority; Med = medium priority; Low = low priority. "-" = undesirable instrument choice; "+" = acceptable instrument choice; "++" = desirable instrument choice.

a- The Q-Exactive is the most sensitive instrument listed. For experiments where it is given an equal rating as other instruments, high sensitivity was not deemed absolutely critical to the experiment. However, if a Q-Exactive is readily accessible, it is generally the preferred choice of the instruments listed. Other high-performing instruments have been released recently and are expected to be highly useful for PSII research as well.

b- Ratings are assuming precursor-ion-based quantification, as has been used in the large majority of studies focused on the PSII life-cycle. Product-ion-based quantification is relevant for studies that use iTRAQ and some forms of spectral counting.

c- As an exception, rough quantification of relative LMM subunit stoichiometry between samples has been performed by intact-mass measurement on a MALDI-TOF (Sugiura et al., 2010a)

The low-molecular-mass (LMM) subunits

Fully assembled PSII contains around 13 low-molecular-mass (LMM) proteins (<10 kDa) whose transmembrane domains account for around 40-85% of the sequence. Identification of these very hydrophobic proteins by bottom-up LC-MS/MS is challenging, with typically four or fewer LMM proteins detected (Granvogl et al., 2008; Haniewicz et al., 2013; Pagliano et al., 2014). Difficulties are associated with the proteins' hydrophobicity and lack of soluble domains, which lead to sample losses during preparation, poor tryptic digestion due to infrequent arginines and lysines, slow elution during chromatography, and poor ionization efficiency due to lack of abundant proton-accepting residues. Fractionation by gel electrophoresis carries the additional challenge of extracting the protein from the gel, made more difficult because tryptic digestion sites are infrequent (Granvogl et al., 2008).

To circumvent these difficulties, intact-mass measurement (no MS/MS fragmentation of the protein) and more recently top-down MS strategies have been employed, both of which avoid protein digestion and are able to identify nearly all the LMM subunits in a purified complex (summarized in Table 3). Intact-mass measurement of the LMM subunits was demonstrated by both ESI and MALDI methods, using QqQ and MALDI-TOF instruments (see references cited in Table 3). Both methods achieve roughly 50-200 ppm mass accuracy; especially without fragmentation data, this would typically not be enough for confident identification of an unknown protein. However, because there are only approximately 13 LMM subunits, predicted masses, which are available from genomic sequences in many organisms, are distinctive, and because the

starting sample is typically a purified PSII complex, these intact-mass measurements are routinely accepted as confident identifications.

MS/MS fragmentation of intact LMM subunits, however, can be induced using both ESI and MALDI, although ESI has been more successful (see Table 3 and references cited therein). Whitelegge and co-workers (Thangaraj et al., 2010) identified eleven LMM proteins in purified PSII from *G. sulphuraria* with a linear ion trap quadrupole-Fourier transform ion cyclotron resonance (LTQ-FTICR) instrument after offline LC and confirmed several modifications. They employed both collisional-activated dissociation (CAD) and electron-capture dissociation (ECD) to fragment the proteins, but CAD gave better results for all LMM proteins. Eichacker and co-workers (Granvogl et al., 2008) demonstrated top-down analysis on a Q-TOF with sequence coverage ranging from 14-82%. This method has been used in several other recent studies (Plöscher et al., 2009; Boehm et al., 2011; Boehm et al., 2012). Notably, Eichacker and co-workers (Granvogl et al., 2008) developed a protocol to perform in-gel extraction of intact LMM proteins prior to top-down analysis (capable of extracting all but the PsbZ protein from the gel matrix). This technique can be used to analyze individual BN-PAGE bands and, thus, identify the LMM components specific to individual types of PSII complexes in heterogeneous mixtures such as a thylakoid membrane proteome or affinity-tagged PSII complexes.

Table 3. Identification of LMM subunits by MS

MS technique	Ionization method	# LMM subunits identified	Instrument	Mass accuracy (MS^1)	References
Bottom-up	ESI, MALDI	0-4	Variety	~1-100 ppm (peptides)	Many, e.g., Pagliano et al., 2014; Haniewicz et al., 2013; Plöscher et al., 2011; Kereïche et al., 2008
Intact	ESI	9-11	QqQ	~50-200 ppm	Thangaraj et al., 2010; Laganowsky et al., 2009; Gómez et al., 2002; Sharma et al., 1997a
	MALDI	9-13	MALDI-TOF	~50-200 ppm	Pagliano et al., 2014; Nakamori et al., 2014; Nowaczyk et al., 2012; Pagliano et al., 2011; Sugiura et al., 2010
Top-down	ESI	8-13	Q-TOF, LTQ-FTICR	3-30 ppm (Q-TOF), <1-5 ppm (LTQ-FTICR)	Boehm et al., 2012; Boehm et al., 2011; Thangaraj et al., 2010; Plöscher et al., 2009; Granvogl et al., 2008
	MALDI	5	MALDI-TOF/TOF	~50-200 ppm	Pagliano et al., 2014

PSII life-cycle application: Composition of subcomplexes

Many subcomplexes form during the PSII life-cycle, and MS has played a critical role, in combination with gel electrophoresis, immunoblotting, crystallography, electron microscopy and other biochemical techniques, in identifying their components (Heinz et al., 2016). A schematic of the life-cycle is shown in Fig. 2 (for reviews of the life-cycle and the subcomplexes that form, see Baena-González and Aro, 2002; Aro et al., 2005; Nixon et al., 2010; Shi et al., 2012; Komenda et al., 2012b; Nickelsen and Rengstl, 2013; Järvi et al., 2015; Heinz et al., 2016). A summary of the main subcomplexes whose composition has been studied by MS is found in Table 4 (for completeness, several other subcomplexes are also included). MS analysis generally allows more rapid, comprehensive, and definitive profiling of PSII subunits than other methods, and is especially useful for the LMM subunits that tend to stain poorly on gels. However, owing to the high sensitivity of MS and because relative quantification by MS is not straightforward, it can be difficult to distinguish a trace component of a complex from one that is stoichiometric. Immunoblotting, therefore, complements MS for characterizing composition of subcomplexes.

At the start of *de novo* PSII assembly, each of the four core subunits D1, D2, CP47, and CP43, forms a pre-complex with specific LMM components. Using a ΔD1 mutant in *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis* 6803) and top-down ESI-MS on a Q-TOF, Nixon and co-workers (Boehm et al., 2011) showed that the CP47 pre-complex contains the LMM subunits PsbH, PsbL, and PsbT, whereas the CP43 pre-complex contains the LMM subunits PsbK and Psb30. In this study, it was not possible by MS alone to demonstrate fully stoichiometric binding, just co-purification, of those LMM subunits to CP47 and CP43. However, these results are consistent with the PSII crystal structures and other non-MS-based results (Boehm et al., 2011 and references cited therein). Previous evidence implies PsbZ could also associate with the CP43

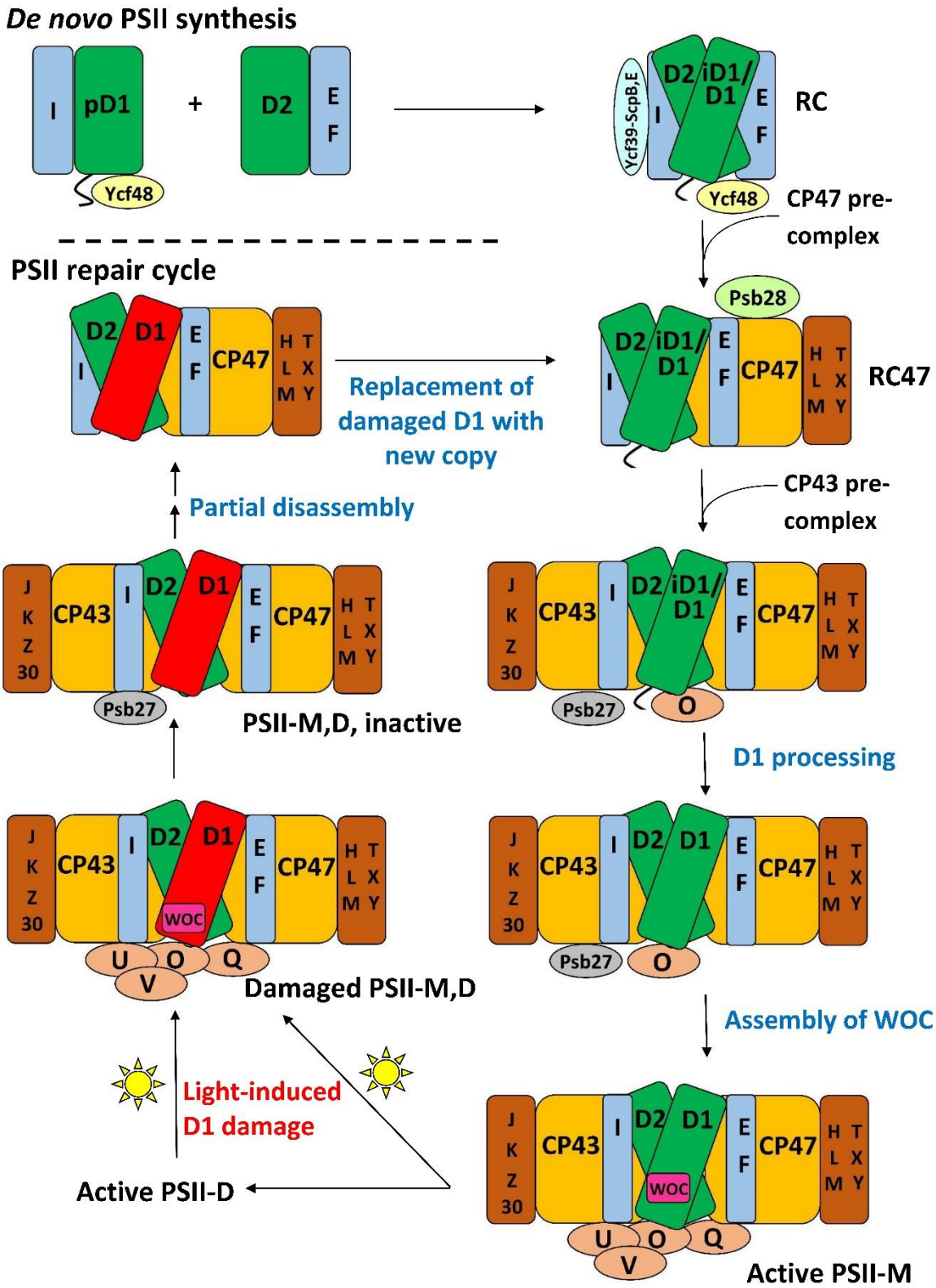

De novo PSII synthesis
I
pD1
Ycf48
+
D2
E
F
Ycf39-ScpB,E
D2
iD1/
D1
RC
CP47 pre-
complex
PSII repair cycle
D2
D1
CP47
H T
L X
M Y
Replacement of
damaged D1 with
new copy
Psb28
RC47
Partial disassembly
CP43 pre-
complex
J
K
Z
30
CP43
PSII-M,D, inactive
Psb27
O
D1 processing
WOC
U
O
Q
V
Damaged PSII-M,D
Assembly of WOC
Light-induced
D1 damage
Active PSII-D
Active PSII-M

Figure 2. A schematic of the PSII life-cycle. Refer to the text for description of each step. This schematic represents the cyanobacterial PSII life-cycle. The subcomplex progression is similar in algae and higher plants, though several homologous subunits are named differently in these species than in cyanobacteria, and certain subunits are unique to each group (see Rokka et al., 2005; Shi et al., 2012; Nickelsen and Rengstl, 2013;Järvi et al., 2015; Heinz et al., 2016). In algae and higher plants, damaged complexes migrate from thylakoid grana to stromal lamellae for repair and the first steps of reassembly (Tikkanen and Aro, 2014; Järvi et al., 2015). In cyanobacteria, chloroplasts and such inter-thylakoid structure are absent, and repair is not believed to require spatial migration of damaged complexes. *De novo* PSII synthesis through RC formation appears to begin in specialized membrane subfractions in cyanobacteria, algae, and higher plants before PSII migration to the general thylakoid membrane space, though the details of this process in the various species classes remains to be resolved (Zak et al., 2001; Nickelsen et al., 2011; Nickelsen and Rengstl, 2013; Rast et al., 2015).

Table 4. Composition of complexes in the PSII life-cycle by MS and other methods

Sub-complex	Composition by MS	Species[a]	Reference	Composition by other methods	Species[a]	Reference
D1 pre-complex	ND	-	-	pD1, I, (Ycf48[c])	*Syn.* 6803	Dobáková et al., 2007
D2 pre-complex	ND	-	-	D2, E, F	*Syn.* 6803	Komenda et al., 2008
CP47 pre-complex	CP47, H, L, T	*Syn.*6803	Boehm et al., 2011	CP47, H	*Syn.* 6803	Komenda et al., 2005
CP43 pre-complex	CP43, K, Psb30	*Syn.* 6803	Boehm et al., 2011	CP43, K, Z, Psb30[c]	*C. reinhardtii*, *T. elongatus*	Iwai et al., 2007; Sugimoto and Takahashi, 2003
RC	D1, D2, E, F, I	Pea	Sharma et al., 1997a,b	D1, D2, E, F, I, (W[d])	Spinach	Irrgang et al., 1995; Ikeuchi and Inoue, 1988; Nanba and Satoh, 1987
RCII*	D1/iD1, D2, E, F, I, Ycf48, Ycf39, ScpB, ScpE	*Syn.* 6803	Knoppová et al., 2014	D1/iD1, D2, E, F, I, Ycf48, Ycf39, ScpE	*Syn.* 6803	Knoppová et al., 2014; Komenda et al., 2008; Dobáková et al., 2007; Komenda et al., 2004
RCIIa	ND	-	-	D1/iD1, D2, E, F, I, Ycf48	*Syn.* 6803	Knoppová et al., 2014; Komenda et al., 2008; Dobáková et al., 2007
RC47 monomer	D1, D2, CP47, E, F, I, L, M, T, X, Y, Psb28, Psb28-2	*Syn.* 6803	Boehm et al., 2012	D1, D2, CP47, E, F, H, Psb28, Psb28-2	*Syn.* 6803	Boehm et al., 2012; Dobáková et al., 2009; Komenda et al., 2004
RC47 monomer	D1, D2, CP47, E, F, I, T_c, W	Spinach	Zheleva et al., 1998	D1, D2, CP47, E, F, H, I, M, R, T_c	Spinach	Rokka et al., 2005
RC47 dimer	D1, D2, CP47, E, F, I, K, L, T_c, W	Spinach	Zheleva et al., 1998	ND	-	-
PSII monomer, inactive	Psb27	*Syn.* 6803, *T. elongatus*	Liu et al., 2013a; Grasse et al., 2011	RC47 components + CP43, Psb27	*Syn.* 6803, *T. elongatus*	Liu et al., 2013a; Grasse et al., 2011; Roose and Pakrasi, 2008
PSII monomer/dimer, active	D1, D2, CP47, CP43, E, F, H, I, J, K, L, M, O, Q, U, V, T, X, Y, Z, Psb30	*Syn.* 6803, *T. elongatus*	Nowaczyk et al., 2012; Kashino et al., 2002	Crystal structure shows all components as by MS except PsbQ	*Syn.* 6803, *T. elongatus*	Umena et al., 2011; Roose et al., 2007
PSII monomer/dimer, active	D1, D2, CP47, CP43, E, F, H, I, K, L, M, O, R, T_C, X	*N. tabacum*	Haniewciz et al., 2015; Granvogl et al., 2008	D1, D2, CP47, CP43, O, P, S; R, W (pea)	*A. thaliana*	Pagliano et al., 2011; Caffarri et al., 2009
PSII-LHCII supercomplexes[b]	D1, D2, CP47, CP43, E, F, H, I, K, L, M, O,R, T_C, W, X; Lhcb1-4, 6	*N. tabacum*	Haniewciz et al., 2015; Granvogl et al., 2008	D1, D2, CP47, CP43, O, P, Q, S; Lhcb1-6; W (pea)	*A. thaliana*	Kouřil et al., 2012; Caffarri et al., 2009; Thidholm et al., 2002
PSII-PSI-PBS megacomplex	D1, D2, CP47, CP43, O, U, V; PsaA, B, C, D, E, F, L, Slr0172, Ycf4, PC; ApcA, B, C, D, E, F, CpcA, B, D, G1, G2	*Syn.* 6803	Liu et al., 2013b	ND	-	-
PSII-PSI-LHCII megacomplex	D1, D2, CP47, CP43, E; PsaA,B,L; LhcbM1, LhcbM10, Lhca2	*Ulva* sp.	Gao et al., 2015	D1, D2, CP47, CP43; PsaB,D,F,G, K,L; unspecified LHCII subunits	*A. thaliana*	Järvi et al., 2011

ND, not determined

a- When two species are listed in the same subcomplex entry, the protein components are the union of those found in the individual studies.

b-characterization of specific PSII-LHCII supercomplexes.

c-Uncertain; evidence is suggestive

d-Subsequent studies indicate PsbW presence in this complex may be an artifact of solubilization conditions (discussed in the text).

pre-complex (Iwai et al., 2007; Guskov et al., 2009; Takasaka et al., 2010), but it was not detected by MS in this study. As determined by affinity purification and immunoblotting, the D1 pre-complex contains PsbI and possibly Ycf48 (Dobáková et al., 2007). It was suggested that a Ycf39-ScpB-ScpE complex may also associate as early as this stage to insert chlorophyll into D1 (Knoppová et al., 2014). The D2 pre-complex contains PsbE and PsbF (Müller and Eichacker, 1999; Komenda et al., 2008).

The D1 and D2 pre-complexes merge to form the reaction center (RC) complex, the earliest subcomplex capable of charge separation (Baena-González and Aro, 2002; Dobáková et al., 2007). The RC complex, initially isolated from spinach and wheat by detergent solubilization of thylakoid membranes, was characterized by gel electrophoresis and immunoblotting to contain D1, D2, PsbE, PsbF, and PsbI (Nanba and Satoh, 1987; Ikeuchi and Inoue, 1988). Intact-mass and bottom-up MS studies later confirmed this composition (Sharma et al., 1997a,b,c). Several biochemical studies detected the 10-kDa PsbW subunit, which is found in green algae and higher plants but not cyanobacteria, as an additional component (Irrgang et al., 1995; Lorković et al., 1995; Shi and Schröder, 1997). Subsequently, more specific studies (including an MS-based one (Granvogl et al., 2008)) showed that PsbW associates later, to dimers during formation of PSII-Light-harvesting complex II (LHCII) supercomplexes (see below) (Shi et al., 2000; Thidholm et al., 2002; Rokka et al., 2005; Granvogl et al., 2008). Despite attaching to PSII at a late stage of assembly, PsbW may bind tightly to the D1/D2 surface and, thus, remain partially attached to the RC complex during solubilization, while other peripheral PSII subunits are removed, explaining the controversy (Rokka et al., 2005). This case highlights that subcomplexes obtained from detergent solubilization, a technique used especially in early PSII subcomplex studies, do not necessarily represent subcomplexes that form *in vivo*. An alternative major method for isolating PSII

subcomplexes is purifying them from mutant strains that are "blocked" at a particular stage of assembly. Such complexes are indeed formed *in vivo*, but it is possible that the altered relative quantity of PSII subunits in the thylakoid membrane arising from the mutation may lead to artefactual binding of certain subunits to some subcomplexes (Thidholm et al., 2002). In cyanobacteria, two slightly different forms of the RC complex were observed, labeled RCII* and RCIIa, which differ slightly in accessory protein content (see Table 3 and the section below). MS was critical in RCII* component characterization, and was indirectly used for RCIIa characterization as well by gel and immunoblot comparison (Knoppová et al., 2014).

The next complex formed during PSII assembly is the RC47 intermediate, also called the CP43-less core monomer in plants, formed by attachment of the CP47 pre-complex to the RC complex. In 1998, Barber and co-workers (Zheleva et al., 1998) showed by MS that the monomeric RC47 complex from spinach contains the D1, D2, CP47, PsbE, PsbF, PsbI, $PsbT_c$, and PsbW proteins, and the dimeric form contains, in addition, PsbK and PsbL. From the later studies on PsbW cited above, PsbW presence may arise from a tight binding to the D1/D2 surface, not *in vivo* presence in the RC47 complex during assembly. Based on Nixon and co-workers' study (Boehm et al., 2011) on the CP47 pre-complex in *Synechocystis* 6803, it would be expected that RC47 also contains PsbH. Indeed, a more recent MS-based study of the RC47 complex from *Synechocystis* 6803 identified all the proteins found by Barber and co-workers (Zheleva et al., 1998) in their monomeric RC47 complex (except PsbW which is not found in cyanobacteria), plus PsbH, PsbM, PsbX, PsbY, and Psb28 (Boehm et al., 2012).

Attachment of the CP43 pre-complex to RC47 forms the inactive PSII monomer (Nickelsen and Rengstl, 2013). Active monomeric PSII is formed upon D1 processing (Liu et al., 2013a), dissociation of Psb27 (Liu et al., 2013a), assembly of the water-oxidizing manganese-

calcium cluster and photoactivation (Dasgupta et al., 2008), and binding of PsbO, PsbU, and PsbV (cyanobacteria) or PsbO, PsbP, and PsbQ (algae and higher plants) (Bricker et al., 2012). Active monomers dimerize and can attach to the phycobilisome antenna complex (cyanobacteria) (Mullineaux, 2008) or various oligomeric states of LHCII complexes (algae and higher plants) (Kouřil et al., 2012).

Although crystal structures of active PSII dimers from cyanobacteria are available, several MS studies of fully-assembled cyanobacterial PSII have provided independent confirmation of the subunits present in purified complexes under more native conditions (Sugiura et al., 2010a; Nowaczyk et al., 2012). Using native conditions has even helped discover a component (PsbQ) that was lost during crystallization (Kashino et al., 2002; Roose et al., 2007). The majority of PSII from algae and higher plants is found in several PSII dimer-LHCII supercomplexes (for a review see Kouřil et al., 2012). MS studies (in concert with other techniques) have identified their subunit compositions, even in the absence of crystal structures of the complexes from these organisms. Eichacker and co-workers (Granvogl et al., 2008) showed that the four PSII-LHCII supercomplexes in *Nicotiana tabacum* contain identical PSII core and LMM subunits (of the eight LMM subunits identified), and that only PSII-LHCII supercomplexes contain the PsbW protein. These results support previous studies that suggest that PsbW may facilitate linkage of LHCII trimers to PSII (Shi et al., 2000; Thidholm et al., 2002; Rokka et al., 2005). Using both bottom-up and top-down MS techniques, Pagliano and co-workers (Pagliano et al., 2014) found that the various supercomplexes in pea contain identical core and LMM subunits, but that the $C_2S_2M_2$ supercomplex contains the PsbQ, PsbR, PsbP, Lhcb3, and Lhcb6 proteins whereas the C_2S_2 supercomplex does not. In light of the stabilizing effect of the PsbQ and PsbP proteins on oxygen evolution, this finding raises interesting questions about the role of the C_2S_2 supercomplex.

Another recent study used MS to characterize PSII-LHCII supercomplexes in *N. tabacum* and found a few differences in subunit composition; in particular, the C_2S_2 supercomplex contained Lhcb1 isoform CB25, while the $C_2S_2M_2$ supercomplex did not (Haniewicz et al., 2015).

Several studies indicate that PSII-PSI-antenna megacomplexes can form in both cyanobacteria and higher plants. Using *in vivo* cross-linking, Blankenship and co-workers (Liu et al., 2013b) captured a PSII-PSI-phycobilisome megacomplex in *Synechocystis* 6803**.** The authors used MS to demonstrate presence of subunits from each complex in the preparation (Tables 3 and 5), and identified cross-links revealing specific inter-complex subunit interactions. Aro and co-workers (Tikkanen et al., 2008b; Tikkanen et al., 2010) showed that LHCII can transfer excitation energy to PSI in grana margins of higher plants as a means of balancing energy flux under varying light conditions. In support of this hypothesis, two PSII-PSI-LHCII megacomplexes from *Arabidopsis thaliana* were observed by a novel large-pore BN-PAGE system (Järvi et al., 2011), and more recently, a PSII-PSI-LHCII megacomplex was identified by MS from the macroalga *Ulva* sp. under drought stress conditions (Gao et al., 2015).

PSII life-cycle application: Identification of accessory proteins

Many accessory proteins bind transiently to PSII subcomplexes during the PSII life-cycle, serving key regulatory roles, but are not present in the crystal structure owing to their absence in fully assembled PSII. For reviews of the accessory proteins of PSII, see Shi et al. (2012); Komenda et al. (2012b); Nickelsen and Rengstl (2013); Mabbitt et al. (2014); Järvi et al. (2015); Heinz et al. (2016). Bottom-up MS has played a key role in identifying some of the known ones, and others likely remain to be identified. Identifying a previously unknown PSII-associated protein in this manner, however, is not straightforward because the mass spectrometers used for bottom-up analysis are so sensitive that dozens of contaminant proteins are often detected even in "purified"

complexes. Low signal intensity of a peptide compared to those of known PSII peptides does not necessarily indicate a contaminant at low abundance because different peptides have different intrinsic ionization efficiencies, and many accessory proteins bind sub-stoichiometrically to PSII. Certain contaminant proteins such as NDH-1 complex subunits (Nowaczyk et al., 2012), ATP synthase subunits (Komenda et al., 2005), phycobilisome subunits (Kufryk et al., 2008), certain ribosomal proteins (Liu et al., 2011b), and several carbon dioxide-concentrating mechanism proteins (Kufryk et al., 2008; Liu et al., 2011b) are frequently observed. Careful examination of the full list and consideration of the experimental conditions are needed to distinguish plausible PSII-interaction candidates from contaminant proteins (Kashino et al., 2002). Although different MS search software packages use different algorithms for scoring protein hits, a strict statistical confidence threshold should be employed and reported. Overall, although a simple bottom-up experiment is a powerful tool to suggest new candidate proteins that associate with PSII, subsequent targeted experiments on each one are needed to confirm the interaction.

This strategy has proven successful many times for identifying new PSII interaction partners. An early example (Kashino et al., 2002) analyzed SDS-PAGE bands by MALDI-TOF MS from a highly purified PSII preparation and identified several novel proteins, Sll1638 (PsbQ), Sll1252, and Sll1398 (Psb32), that appeared to be plausible PSII interaction partners. Follow-up biochemical studies targeting these proteins confirmed their role in the PSII life-cycle and elucidated functional aspects of each (Wegener et al., 2011; Inoue-Kashino et al., 2011; Bricker et al., 2012). A later proteomic study of purified PSII complexes revealed that the Slr0144-Slr-0152 proteins, all part of one operon, associate with PSII, leading to further characterization of their role in PSII assembly (Wegener et al., 2008). In other cases, specific subcomplexes were isolated before MS analysis and identification of accessory proteins. For example, analysis of a gel band

from Δ*ctpA*-HT3-PSII revealed that the Psb27 protein binds specifically to a PSII subcomplex that accumulates before D1 processing (Roose and Pakrasi, 2004), initiating the studies that ultimately elucidated its role in PSII assembly (Nowaczyk et al., 2006; Roose and Pakrasi, 2008; Grasse et al., 2011; Liu et al., 2011a,b; Komenda et al., 2012a). MS analysis showed that the Ycf39, ScpB (HliC), and ScpE (HliD) proteins bind specifically to the RCII* form of the reaction center complex, but not the related RCIIa form (Knoppová et al., 2014). The specific binding of the accessory proteins Psb28 (Dobáková et al., 2009; Boehm et al., 2012) and Psb28-2 (Boehm et al., 2012) to the RC47 complex, and of Ycf48 to RCII* and RCIIa (Knoppová et al., 2014), was initially discovered by immunoblotting, but the proteins' presence was confirmed by MS, strengthening the finding.

PSII life-cycle application: Identification of PTMs

Identification of processing events to form mature PSII proteins

The D1 protein is synthesized as a precursor protein (pD1) with a C-terminal extension that gets cleaved during PSII assembly (Takahashi et al., 1988). An early study using peptide sequencing showed that in spinach, cleavage occurs after Ala-344, removing nine C-terminal residues (Takahashi et al., 1988). Several years later, it was found that, in *Synechocystis* 6803, cleavage also occurs after Ala-344, removing 16 C-terminal residues (Nixon et al., 1992). In this study, peptide sequencing as well as fast atom bombardment (FAB)-MS (a predecessor for ESI and MALDI) were used to pinpoint this cleavage site. Ala-344 serves as a ligand for a Mn ion in the water oxidation cluster (Umena et al., 2011) so that without cleavage, PSII remains incapable of oxygen evolution (Roose and Pakrasi, 2004). The extension, thus, protects early assembly intermediates from harmful premature water oxidation activity. Interestingly, although D1 in higher plants is cleaved in a single step, cyanobacterial D1 is cleaved in two steps, and an

intermediate D1 (iD1) is formed transiently (Inagaki et al., 2001). Although the iD1 cleavage site remained unknown for two decades, in 2007, MS and biochemical evidence demonstrated that the CtpA protease cleaves after Ala-352 to form iD1, which is then cleaved again after Ala-344 to form mature D1 (Komenda et al., 2007). The significance of the two-step cleavage remains unknown, although iD1 may serve as a signal for transferring an early PSII assembly intermediate from the cytoplasmic to the thylakoid membrane (Komenda et al., 2007).

The CP43 protein also appears to be cleaved before, or during an early stage of, PSII assembly. Tandem MS analysis identified a CP43 peptide in spinach starting with a modified form of Thr-15 (Michel et al., 1988). Based on the genomic sequence, the preceding residue is a leucine, so this peptide would not be a predicted trypsin cleavage product. It was also found that the N-terminus of CP43 is blocked from analysis by Edman degradation, likely owing to N-terminal modification. Taken together, these results show that the first 14 residues of CP43 are cleaved, leaving Thr-15 as the mature protein's N-terminus (Michel et al., 1988). Subsequent studies identified the corresponding CP43 peptide in *A. thaliana* (Vener et al., 2001) and *Synechocystis* 6803 (Wegener et al., 2008), suggesting that this cleavage is conserved. Crystal structures of cyanobacterial PSII were not able to resolve the most N-terminal portion of CP43, so those structures do not address this question of CP43 cleavage (Loll et al., 2005; Umena et al., 2011).

Cyanobacterial Psb27, PsbQ, and PsbP have unusually hydrophobic properties for soluble lumen-localized proteins and contain a lipoprotein signal motif and conserved cysteine in their N-terminal regions (Thornton et al., 2004; Nowaczyk et al., 2006; Fagerlund and Eaton-Rye, 2011). This led to the suggestion that they are N-terminally lipid-modified and, thus, anchored to the lumenal surface of the thylakoid membrane. Using lipase treatment and MALDI-TOF MS, Rögner and co-workers (Nowaczyk et al., 2006) showed that Psb27 from *Thermosynechococcus elongatus*

does indeed contain such a modification. Also using MALDI-TOF MS, Wada and co-workers (Ujihara et al., 2008) confirmed this finding with Psb27 from *Synechocystis* 6803 and also found that *Synechocystis* 6803 PsbQ, recombinantly expressed in *E. coli*, is also N-terminally lipid modified. Notably, this group developed a method to extract lipid-modified peptides from a gel matrix after in-gel digestion, enabling downstream MS analysis (Ujihara et al., 2008). During PSII assembly, it is important that Psb27 binds to the lumenal surface before the other extrinsic proteins (Liu et al., 2013a), and the lipid anchor may facilitate this sequence by keeping Psb27 in close proximity at all times. A similar role for the lipid anchor of PsbQ was proposed recently (Liu et al., 2015). A lipid modification on PsbP has not yet been demonstrated although strong suggestive evidence indicates its presence (Fagerlund and Eaton-Rye, 2011).

2.4.2. Identification of phosphorylation sites

In the early 1980s phosphorylation of the four PSII subunits that later came to be known as D1, D2, CP43, and PsbH, was observed. These studies were conducted *in vivo* and *in vitro* using ^{32}P labeling of whole cells and thylakoid membranes from *Chlamydomonas reinhardtii* and pea, with detection of phosphoproteins by autoradiography (Steinback et al., 1982; Owens and Ohad, 1982; Owens and Ohad, 1983). Immunoblotting with antibodies that recognize phosphorylated residues was introduced later and became another popular detection method (Rintamäki et al., 1997). Neither of these methods, however, reveal the modified residue. This information was first obtained by gas-phase sequencing using Edman degradation, which demonstrated that the PsbH phosphorylation site is Thr-2, its N-terminus, in spinach (Michel and Bennett, 1987) and *C. reinhardtii* (Dedner et al., 1988). Since then, MS analysis has replaced Edman degradation and become the dominant method for phosphorylation-site determination, as it is higher-throughput, more definitive, and not limited by N-terminal blockage (e.g., acetylation).

The main sites identified are presented below (for reviews, see Vener, 2007; Pesaresi et al., 2011; Puthiyaveetil and Kirchhoff, 2013).

Tandem MS demonstrated phosphorylation of D1-Thr-2, D2-Thr-2, and CP43-Thr-15, the mature proteins' N-termini, in spinach (Michel et al., 1988), *A. thaliana* (Vener et al., 2001), and *C. reinhardtii* (Turkina et al., 2006). Phosphorylation of CP43 was also observed at Thr-20, Thr-22, and Thr-346 in spinach (Rinalducci et al., 2006), and at Thr-346 and Ser-468 in *A. thaliana* (Sugiyama et al., 2008; Reiland et al., 2009). MS analysis showed that PsbH is phosphorylated at its N-terminus in *A. thaliana*, supporting the Edman degradation data from spinach and *C. reinhardtii*, and additionally demonstrated phosphorylation of Thr-4 (Vener et al., 2001). Intact-mass MS evidence also indicates double PsbH phosphorylation in spinach and pea (Gómez et al., 1998; Gómez et al., 2002). More recently, phosphorylation of the extrinsic proteins PsbP, PsbQ, and PsbR was observed in phosphoproteomic studies of *A. thaliana* (Sugiyama et al., 2008; Lohrig et al., 2009; Reiland et al., 2009). Although not discussed here, phosphorylation of LHCII is well-documented, and it regulates state transitions in green algae and higher plants (for reviews see Lemeille and Rochaix, 2010; Minagawa, 2011; Schönberg and Baginsky, 2012; Tikkanen and Aro, 2014; Tikhonov, 2015).

Phosphorylation of PSII subunits is not absolutely required for PSII repair (Bonardi et al., 2005) but assists in transferring damaged PSII complexes from the stacked thylakoid grana to stromal lamellae, where repair occurs. Phosphorylation appears to induce architectural changes in the stacked grana and increase membrane fluidity in such a way as to promote mobility of damaged PSII centers to the stromal lamellae for repair (Tikkanen et al., 2008a; Fristedt et al., 2009; Fristedt et al., 2010; Herbstová et al., 2012; Järvi et al., 2015). For many of the PSII phosphorylation sites, light intensity and/or other environmental conditions affect the phosphorylation extent, with

implications for the functional significance of these modifications. MS analysis has played a critical role in these quantitative studies, and methodology for such measurements is discussed in the dynamics section below. For reviews that discuss the role of PSII phosphorylation, see Pesaresi et al. (2011); Mulo et al. (2012); Schönberg and Baginsky (2012); Järvi et al. (2015).

PSII phosphorylation may not be needed in cyanobacteria owing to the lack of spatial organization of thylakoids (Mulo et al., 2012). However, a recent global proteomics study of the cyanobacterium *Synechococcus* sp. PCC 7002 (hereafter *Synechococcus* 7002) found that a portion of D1 copies are phosphorylated at their N-terminus, Thr-2 (Yang et al., 2014), as in higher plants. This finding opens the possibility for a role of phosphorylation in PSII turnover in cyanobacteria.

2.4.3. Identification of oxidative and other modifications

Light is necessary for PSII function, but even low light intensities can lead to PSII damage, particularly of the D1 protein. Damage triggers partial PSII disassembly, D1 degradation, insertion of a new D1 copy, and PSII re-assembly (Nickelsen and Rengstl, 2013). When the rate of damage exceeds that of repair, photosynthesis is inhibited, referred to as photoinhibition. Photodamage can be initiated in several ways, but a common result of each mechanism is production of highly oxidizing species (e.g., singlet O_2, other reactive oxygen species (ROS), or radical PSII cofactors). These species rapidly oxidize PSII subunits, ultimately rendering the complex non-functional. For reviews of the photoinhibition process, see Barber and Andersson (1992); Adir et al. (2003); Pospíšil (2009); Allahverdiyeva and Aro (2012); Tyystjärvi (2013).

Though oxidative damage of PSII was long believed to be responsible for photoinhibition (Telfer et al., 1994), MS studies provided the first concrete evidence for specific oxidative modifications of PSII. Bottom-up MS analysis of the D1 and D2 subunits from pea PSII found up to three +16 oxidative modifications (each representing incorporation of an oxygen atom) on

certain peptides (Sharma et al., 1997c). Interestingly, not all peptides were oxidized, but the oxidized ones were all located near the predicted D1 and D2 redox cofactor sites, supporting the idea that radical redox cofactors themselves, or ROS produced by reaction with them, cause oxidative damage to PSII. More recently, Bricker and co-workers (Frankel et al., 2012; Frankel et al., 2013b) used tandem MS to identify oxidized residues on spinach D1, D2, and CP43 that are located near the Q_A, $Pheo_{D1}$, and manganese cluster sites, all reasonable sources of oxidizing species. Additionally, tryptophan oxidation products in spinach were identified on CP43-Trp-365 and D1-Trp-317, which are located near the manganese cluster (17 and 14 Å, respectively, in the crystal structure from *T. elongatus*) (Anderson et al., 2002; Dreaden et al., 2011; Kasson et al., 2012). By monitoring the digested peptides' absorption at 350 nm, the authors found that these tryptophan oxidations are correlated with increased light intensity and decreased oxygen evolution. Other modifications to PSII subunits were also detected by MS (Gómez et al., 2002; Gómez et al., 2003; Anderson et al., 2004; Rexroth et al., 2007; Sugiura et al., 2013). Notably, a recent global proteomics study of *Synechococcus* 7002 identified many new PSII PTMs (Yang et al., 2014), but the functional significance of these modifications remains to be determined.

Dynamics: Quantitative or semi-quantitative changes in PSII proteins and PTMs

MS-based methods to study PSII dynamics

Most MS-based quantification experiments seek the relative, not absolute quantity of a protein or PTM in one sample compared to another. We focus here on relative quantification methods because nearly all the work on PSII dynamics fell into that category.

Gel-based quantification

Perhaps the most basic MS-based semi-quantitative method is in-gel digestion at the same band in two different sample lanes, prompted by a significant staining-intensity difference between

the two bands. This approach was used frequently when analyzing different purified PSII complexes (Liu et al., 2011b; Knoppová et al., 2014), yielding information about accessory proteins that bind specifically to certain subcomplexes. A proper loading control (typically equal chlorophyll) must be used to ensure a meaningful comparison. Multiple proteins are typically identified by MS in both bands, however, so it may not be immediately apparent which protein is the main component (Liu et al., 2011b). Confirmation may be necessary by western blotting or one of the more quantitative MS-based techniques described below.

The accuracy of gel-based quantification can be improved by introducing a second electrophoretic separation dimension before in-gel digestion and LC-MS/MS. Semi-quantitative two-dimensional denaturing gel electrophoresis (2DE) (distinct from 2D BN-PAGE described above), a popular technique especially in early proteomics studies, usually first separates proteins by size and then on the basis of pI (Rabilloud et al., 2010). The difference in staining intensity indicates the relative content of that protein in each sample. Because two proteins migrate less often together in two dimensions than in one, separation and quantification accuracy are improved. 2DE is useful for large-scale studies such as whole-cell or whole-organelle proteome profiling that require higher-resolution separation than a 1D gel provides. However, in recent years, 2DE has declined in popularity owing to its numerous drawbacks (reviewed in Rabilloud et al., 2010) and the improvements in other more versatile quantitative MS methods. Such large-scale proteomics studies have detected expression-level changes in several PSII proteins in response to a variety of stress conditions (e.g., Ingle et al., 2007; Aryal et al., 2011; Li et al., 2011; Guerreiro et al., 2014). However, insights into the PSII life-cycle have mainly emerged from more focused studies on purified PSII complexes.

Label-free quantification

Some MS-based relative quantification methods use a so-called label-free approach, but the better approach, when feasible, is to introduce a stable isotope into the sample. For label-free quantification, the samples to be compared are analyzed by LC-MS/MS separately. A variety of software tools can then be used to obtain an extracted ion chromatogram (EIC) of any peptide. The EIC displays the total intensity (peak area) of that peptide. Comparing the intensities of the same peptide from two different samples indicates the relative content of that peptide in those samples. Although the concept is simple, accurate label-free quantification depends on a number of factors: equal sample loading (on a relevant basis, e.g., chlorophyll concentration), reproducible LC runs, and appropriate normalization during data analysis. For quantification of proteins, data from component peptides must be merged in a statistically sound way (Bantscheff et al., 2012; Nahnsen et al., 2013). Thorough mass spectral sampling of possible precursors—not as crucial in non-quantitative experiments—is necessary for accurate peak definition, but that typically diverts instrument time from obtaining product-ion spectra that give information for peptide identification and sequence coverage (Bantscheff et al., 2012). Various strategies have been designed to address this challenge (e.g., data-independent acquisition approaches such as MS^E (Silva et al., 2006; Grossmann et al., 2010) and "all-ion fragmentation" (Geiger et al., 2010) especially when combined with Ultra-Performance LC (UPLC) (Bantscheff et al., 2012). Label-free quantification by spectral counting, which involves comparing the total number of product-ion (MS/MS) spectra obtained for a given peptide or protein, is a common approach (Lundgren et al, 2010), although that has been used in fewer PSII-related studies (Fristedt and Vener, 2011; Stöckel et al., 2011). Label-free quantification of intact proteins is more direct than comparing peptides, but best applied for small proteins. Intact-mass spectra (MALDI and ESI) of the LMM PSII proteins indeed have

been used in a number of instances for label-free quantification between states (Laganowsky et al., 2009; Sugiura et al., 2010a).

Isotope label-based quantification

The alternative to label-free quantification is introduction of a stable isotope label into one of the two samples being compared (certain methods also allow greater multiplexing, see below). In contrast to the label-free approach, the labeled and unlabeled samples (often called "heavy" and "light") are mixed and analyzed in a single LC-MS/MS run. The mass spectra of the light and heavy peptide show two peaks shifted slightly in mass. Comparison of their peak areas, just as in label-free quantification, indicates the relative amount of that peptide in each sample (Bantscheff et al., 2012). Although comparing peak areas from a single LC-MS/MS run eliminates the concerns of label-free LC reproducibility and ion suppression, labeling introduces additional sample preparation steps and often involves costly reagents.

Isotopic labeling (with ^{2}H, ^{13}C, ^{15}N, or ^{18}O) of all proteins can be accomplished during cell growth (metabolic labeling), or by labeling a subset of proteins or peptides at various stages after cell lysis (chemical or enzymatic labeling). In the SILAC method ("stable isotope labeling by amino acids in cell culture") (reviewed in Chen et al., 2015), addition of labeled arginine or lysine to the growth medium results in incorporation of only the labeled form of that amino acid into all proteins. Hippler and co-workers (Naumann et al., 2007) used a SILAC-based method to measure changes in expression of PSII subunits and other proteins in *C. reinhardtii* under iron deficiency, and Jacobs and co-workers (Aryal et al., 2011) used this method to measure light-dark diurnal cycles in *Cyanothece* sp. ATCC 51142. A more common approach in PSII life-cycle research, however, has been ^{15}N metabolic labeling (see "Measuring the temporal dynamics of life-cycle events using isotopic labeling*"* below), in which the growth medium is modified so that the only

nitrogen source is a labeled salt such as potassium nitrate or ammonium chloride (Gouw et al., 2010).

Isotopic labeling at the peptide or protein level during downstream processing after cell lysis is an alternative to metabolic labeling. Tandem mass tags (TMT) (Thompson et al., 2003), isotope tags for relative and absolute quantification (iTRAQ) (Ross et al., 2004), enzymatic ^{18}O labeling, and isotope-coded affinity tags (ICAT) can be used in proteomics experiments in photosynthetic organisms (Thelen and Peck, 2007). TMT and iTRAQ are related approaches that have become popular recently (Bantscheff et al., 2012). Both modify peptides with one of several possible isobaric tags that produce reporter ions during MS/MS fragmentation. Each sample is labeled with a different tag, but owing to the tags' isobaric nature, identical peptides from each sample are observed together chromatographically and as a single peak in a low-resolving power mass spectrum. Each tag, however, contains a unique reporter ion that appears as a distinct peak in the product-ion (MS/MS) spectra, and the ratio of these ions reveals the relative amounts of that peptide in each sample. The iTRAQ reagent modifies primary amines, and TMT tags are available that modify primary amines, thiols, or carbonyl groups. Advantages of these labeling approaches include the ability to multiplex up to eight or ten samples, greater than with metabolic and other chemical labeling methods, and the isobaric nature of the same peptide across all samples reduces both LC separation demands and MS data complexity (Bantscheff et al., 2012). Although many proteomics studies on photosynthetic organisms have used these chemical labeling methods, most have not focused on PSII life-cycle issues (Thelen and Peck, 2007; Battchikova et al., 2015). Two relevant examples include the detection of elevated PsbO cysteine oxidation under DCMU and dark conditions (Guo et al., 2014), and intriguing evidence that PSII thermotolerance in

Synechocystis 6803 may arise in part from antenna trimming and an increased rate of electron transfer to the cytochrome b_6/f complex (Rowland, et al. 2010).

PSII life-cycle application: Measuring changes in phosphorylation levels

As mentioned in the composition section above, phosphorylation of PSII subunits affects membrane fluidity and inter-thylakoid dynamics, thus playing a role in facilitating PSII turnover in green algae and higher plants (see the reviews cited in that section for in-depth treatment of this topic). Many of the studies that have contributed to our current understanding of this process used MS quantification techniques to compare phosphorylation levels between samples and under different environmental conditions.

When using peak-area-based label-free quantification to determine the change in a modified peptide between samples, it is crucial that the peak area of the unmodified peptide be taken into account as well, to distinguish a true change in modification *extent* from simply an increased level of protein expression in one of the states. This method is demonstrated in a study of phosphorylation and nitration in *A. thaliana* grown under low and high light regimes (Galetskiy et al, 2011b). The authors first normalized each modified-peptide peak area in each sample to that of its unmodified counterpart and then compared the modified peptides' normalized peak area to each other. This method can reveal *fold-changes in modification extent* between the two states, but not the absolute percentage of that peptide that contains the modification (the "*modification stoichiometry*").

To find the modification stoichiometry, it is necessary to know in addition the relative "flyability" (ionization efficiency) of the modified and unmodified peptides. Vierstra and co-workers (Vener et al., 2001) showed that the relative flyabilities of six synthetic phosphopeptides and their non-phosphorylated counterparts are nearly identical. Suggesting this as a general

phenomenon for phosphorylated peptides, they estimated the modification stoichiometry for the phosphorylated peptides of D1, D2, CP43, PsbH, and an LHCII protein. In 2010, Vener and co-workers (Fristedt et al., 2010) calculated the actual relative flyability ratio for these PSII peptide pairs, and reported reliable modification stoichiometry for these proteins for the first time under the various conditions in their study. Interestingly, the flyability ratios were indeed close to 1 for each pair (ranging from 0.89-1.23), supporting the earlier suggestion that this may be the case for most phosphorylated/non-phosphorylated peptide pairs (Vener et al., 2001). Other studies have since used those flyability ratios to determine changes in the modification stoichiometry of those same phosphorylation sites under other growth conditions (Fristedt and Vener, 2011; Romanowska et al., 2012; Samol et al., 2012). Knowledge of modification stoichiometry under different conditions is quite valuable; it enabled, for example, a greater level of confidence and detail in the model proposed for how phosphorylation affects thylakoid membrane stacking than would have been possible with fold-change data alone (Fristedt et al., 2010).

Chemical isotopic labeling of peptides has also been applied fruitfully to the study of PSII phosphorylation. Immobilized metal-ion affnity chromatography (IMAC) is a standard protocol for enrichment of phosphopeptides, taking advantage of the interaction between phosphoryl groups and a Fe^{3+}-agarose matrix (Andersson and Porath, 1986). Given that free carboxyl groups can also interact with the resin, it has become common to convert free carboxylates to methyl esters after digestion and prior to IMAC, to avoid this interaction (Ficarro et al., 2002). Vener and co-workers (Vainonen et al., 2005) modified this approach by using deuterated methanol (CD_3) as the esterification reagent for one sample, and unlabeled methanol for a second sample to quantify by "isotope encoding". After mixing the samples and analyzing by LC-MS/MS, the relative amount of each phosphorylated peptide in the two samples is quantified by comparison of their mass

spectral peak areas. It should be noted that this approach does not reveal the modification stoichiometry of any phosphorylation site; rather the techniques described above still need to be performed to gain that information. Instead, as with other isotope-labelling strategies, it enables more confident and straightforward comparisons of the level of any given peptide between samples. This labeling method was used to study phosphorylation of PSII under a variety of conditions and genetic backgrounds (Vainonen et al., 2005; Lemeille et al., 2010; Fristedt and Vener, 2011; Samol et al., 2012).

PSII life-cycle application: Measuring changes in oxidation levels

As discussed above, oxidation of PSII subunits is a well-documented phenomenon, and occurs, at least partially, from oxidizing species generated during the electron transfer reactions of PSII, especially under stress. However, relatively few studies have quantified changes in PSII subunit oxidation under different controlled conditions. Adamska and co-workers (Galetskiy et al., 2011a) used label-free quantification to compare oxidation and nitration (also associated with oxidative stress) levels of thylakoid membrane protein complexes from *A. thaliana* grown under low and high light. They found significantly more modified sites in PSII than in the PSI, cytochrome b_6/f, and ATP synthase complexes. Interestingly, the modified D1, D2, and PsbO sites increased around 2-5-fold, whereas CP47, CP43, PsbE, and PsbR oxidation levels remained roughly constant. D1 and D2 bind most of the redox-active cofactors of PSII, so the increased oxidation especially of these two proteins is not surprising. Similarly, by measuring the increase in 350 nm absorption, Barry and co-workers (Dreaden et al., 2011; Kasson et al., 2012) found that two tryptophan oxidation products increase after exposure to high light, with a corresponding decrease in oxygen evolution activity. Adamska and co-workers (Galetskiy et al., 2011b) found that nitration levels in assembled PSII complexes decrease after exposure to high light, but increase

in PSII subcomplexes. This may imply that once nitrated, PSII complexes are damaged and targeted for disassembly and repair.

PSII life-cycle application: Measuring the temporal dynamics of life-cycle events using isotopic labeling

Measurement of PSII subunit lifetimes has focused mainly on D1, using immunodetection following addition of a protein-synthesis inhibitor or by radioisotope pulse-chase labeling with detection by autoradiography or phosphorimaging (Aro et al., 1993; Mullet and Christopher, 1994; Ohnishi and Murata, 2006). Recently, several studies used ^{15}N labeling pulses and quantified the disappearance of unlabeled PSII subunits using MS. This method enables simultaneous detection of a larger number of PSII subunits and eliminates any concern of overlapping signal from proteins with similar electrophoretic mobility (Yao et al., 2012b). From surveying nine PSII subunits from *Synechocystis* 6803, Vermaas and co-workers (Yao et al., 2012a) found that protein half-lives range from 1.5-33 hours in a PSI-less mutant grown under low light (4 μmol $m^{-2}s^{-1}$ photon flux). In WT *Synechocystis* 6803 grown under 75 μmol $m^{-2}s^{-1}$ photon flux, half-lives of D1, D2, CP47, and CP43 ranged from <1-11 hours (Yao et al., 2012b). In both studies, D1 exhibited the shortest half-life. These studies highlight the wide range in PSII subunit lifetime and the tight regulation of protein synthesis and PSII assembly that must occur to ensure constant proper stoichiometric availability of all subunits. Interestingly, the chlorophyll half-life was several times longer than that of the core chlorophyll-binding proteins, but the half-life was reduced in the absence of the small CAB-like proteins (SCPs), implying that SCPs play a role in chlorophyll recycling during PSII turnover (Yao et al., 2012a).

Rögner, Nowaczyk, and co-workers demonstrated an elegant application of ^{15}N labeling by purifying several subcomplexes in the PSII life-cycle after a pulse with ^{15}N (from $^{15}NH_4Cl$).

Comparing extents of incorporation of ^{15}N in different subcomplexes (e.g., monitoring D1 and D2 peptides) reveals the subcomplexes' position in the PSII life-cycle. Using this method, the authors demonstrated that in *T. elongatus*, Psb27 binds to a monomeric subcomplex early in the PSII assembly process (Nowaczyk et al., 2006), and that Psb27 binds again during disassembly to inactive dimers (Grasse et al., 2011). This information fits well with the current understanding of Psb27 as a gatekeeper preventing manganese cluster assembly in immature complexes (Liu et al., 2013a; Mabbitt et al., 2014).

Cyanobacteria contain multiple versions of the *psbA* gene, and the resulting versions of the D1 protein have some different properties and are expressed preferentially under different environmental conditions (for reviews see Mulo et al., 2009; Sugiura and Boussac, 2014). For example, the *psbA1* gene product in *T. elongatus* is dominant under standard growth conditions, but expression of the *psbA3* gene product, which differs from the PsbA1 copy by ~21 residues, increases under high light conditions (Clarke et al., 1993; Kós et al., 2008; Mulo et al., 2009). Characterization of PSII from mutants that express only specific versions of the gene has shown differences in electron-transfer properties, with the implication that PsbA3 assists in photoprotection of PSII under light stress conditions (Sugiura et al., 2010b; Sander et al., 2010). D1-copy expression was mainly monitored on the transcript level (Golden et al., 1986; Komenda et al., 2000; Kós et al., 2008; Sugiura et al., 2010b). However, using ^{15}N labeling and MS-based quantification, Rögner and co-workers showed that PsbA3 incorporation on the protein level could be monitored unambiguously in *T. elongatus* under high light conditions (Sander et al., 2010) and in the Δ*psbJ* mutant (Nowaczyk et al., 2012). Those studies used ^{15}N-labeled PSII from a strain that only expresses the PsbA3 copy as a standard for 100% incorporation; relative peak area of the

unlabeled PsbA3 peptides compared to this standard is a measure of the incorporation. Such definitive monitoring should allow further detailed studies of *psbA* gene incorporation dynamics.

Progress has also recently been made on the role of the PsbA4 D1 copy; an iTRAQ labeling study found elevated expression of PsbA4 in *Cyanothece* sp. PCC 7822 in the dark (Welkie et al., 2014), providing complementary evidence to that of Pakrasi and co-workers (Wegener et al., 2015) who found that PsbA4 incorporation into PSII renders the complex non-functional. PsbA4 replaces PsbA1 at night in cyanobacterial species that fix nitrogen during this time, protecting against even the trace levels of oxygen evolution that could occur and damage the nitrogenase enzyme (Wegener et al., 2015).

Structure: Determining protein-protein interactions in PSII complexes

MS-based methods to study PSII structure

X-ray crystallography remains the benchmark for determining the structure of protein complexes, but besides fully-assembled active PSII, many complexes that form during the PSII life-cycle are too transient and low in abundance to be easily amenable to crystallography. Valuable information about protein-protein interactions within PSII was obtained from immunogold labeling (Tsiotis et al., 1996; Promnares et al., 2006) and yeast two-hybrid assays (Schottkowski et al., 2009; Komenda et al., 2012a; Rengstl et al., 2013), but the former is primarily suitable for large PSII complexes (Dobáková et al., 2009), and the latter is time-consuming and low-throughput. Both provide relatively low-resolution structural information. Recently, advanced structural proteomics techniques bypass the limitations of the above techniques and offer higher-resolution structural data (although still lower than X-ray crystallography). Either chemical cross-linking or protein footprinting followed by MS detection of these modifications are enabled by MS instruments with high sensitivity, resolving power, and <1-5 ppm mass accuracy on

orbitrap- and FTICR-based instruments (Table 2). These methods allow not only identification of the binding partners of a specific protein but also a low-resolution mapping of the binding site.

Chemical cross-linking

Briefly, the chemical cross-linking technique (reviewed in Sinz, 2014) uses a small molecule with two functional groups on either end that can react with protein residues, separated by a spacer arm (typically less than 14 Å). Many types of cross-linkers are available (Paramelle et al., 2013). The ones most commonly used in PSII research (Bricker et al., 2015) can react with either the primary amine of a lysine and protein N-terminus (and under certain conditions, to a lesser extent with the hydroxyl group of a serine, threonine, or tyrosine (Mädler et al., 2009)), or with the carboxylate of aspartate and glutamate side chains and protein C-termini. After both sides of the cross-linker react with neighboring proteins, digestion, LC-MS/MS, and specialized data analysis can identify cross-linked peptides. Inter-protein cross-linked peptides provide structural information about the complex because the two linked residues are constrained to the spacer arm-length distance from each other.

Cross-linking has been used for decades to study protein-protein interactions (Clegg and Hayes, 1974; Wetz and Habermehl, 1978; Walleczek et al., 1989; Back et al., 2003; Sinz, 2014), but its power was limited until modern MS instrumentation and the proteomics platform enabled high-throughput analysis and confident identification of linked peptides (Rappsilber, 2011). Identification of cross-linked peptides by MS is more challenging than for a typical protein digest, especially for large complexes, because the candidate peptide database increases roughly with the square of the number of peptides. As a result, false positives based on the mass spectrum are common even with high mass accuracy instruments, making high-quality product-ion spectra critical for a confident assignment. Despite powerful and constantly improving cross-link search

algorithms (Rinner et al., 2008; Xu and Freitas, 2009; Petrotchenko and Borchers, 2010; Yang et al., 2012; Götze et al., 2012; Götze et al., 2015; Hoopmann et al., 2015), manual verification of the product-ion spectra of hits is highly recommended. Successful cross-linking requires high sequence coverage and high mass accuracy as is now practical with orbitrap- and FTICR-based instruments (Table 2).

Because cross-linked peptides give typically low-intensity signals compared to those of unlinked peptides, they are often not selected for fragmentation by the instrument's traditional "highest-abundance ion" selection criteria. Several strategies have been developed to improve cross-link selection and/or reduce false positives. They include various methods to enrich for cross-linked peptides before LC-MS/MS (Chu et al., 2006; Kang et al., 2009; Leitner et al., 2012; Fritzsche et al., 2012); use of isotope-coded linkers whose "fingerprint" increases confidence in an identification and can enable real-time guided selection of cross-links for fragmentation (Müller et al., 2001; Pearson et al., 2002; Seebacher et al., 2006; Petrotchenko et al., 2014); and MS-cleavable linkers that simplify data analysis by cleaving a cross-linked peptide into its component peptides before fragmentation (Kao et al., 2011; Petrotchenko et al., 2011; Weisbrod et al., 2013; Buncherd et al., 2014).

Protein footprinting

Protein footprinting is another MS-based structural technique that has been used to study PSII. Its principle is that a protein residue's solvent accessibility determines its susceptibility to modification by a reagent in the solution; residues buried in a protein-protein interface are less susceptible to modification than surface-exposed residues. These modifications are then detected by MS. Instruments with high sensitivity, resulting in high sequence coverage, are critical so that footprinting experiments yield maximal information (Table 2). A common approach is hydroxyl

radical footprinting using the well-established technique of synchrotron radiolysis of water to generate the radicals (Takamoto and Chance, 2006; Wang and Chance, 2011). Fast photochemical oxidation of proteins (FPOP) is a more recent hydroxyl radical fooptrinting technique that uses a laser pulse to generate the radicals and can probe protein dynamics that occur on a faster timescale, down to microseconds (Gau et al., 2011). Hydroxyl radical footprinting can modify 14 of the 20 amino acid side chains (Wang and Chance, 2011). Another technique, glycine ethyl ester (GEE) labeling, adapts a long-standing method for modifying and cross-linking carboxylate groups in proteins (Hoare and Koshland, 1967; Swaisgood and Natake, 1973) for protein footprinting (Wen et al., 2009; Gau et al., 2011). It is easier to implement than hydroxyl radical footprinting, and data interpretation is simpler, but it can only probe changes on aspartate, glutamate, and protein C-termini.

PSII life-cycle application: Cross-linking and footprinting to determine interactions among PSII subunits

Early cross-linking studies on PSII provided information about subunit connectivity before PSII crystal structures were available. Many studies focused on the lumenal extrinsic proteins (Enami et al., 1987; Bricker et al., 1988; Odom and Bricker, 1992; Han et al., 1994), which are more easily accessible to soluble cross-linkers, but interactions involving the transmembrane subunits can also be detected (Tomo et al., 1993; Seidler, 1996; Harrer et al., 1998). In the absence of the MS-based platforms currently available, gel electrophoresis and immunoblotting identify cross-linked products and their likely component proteins. Those methods are still helpful today as confirmation and when cross-linked peptides are not detected by MS (Hansson et al., 2007; Nagao et al., 2010; Liu et al., 2011a; Liu et al., 2014b), but MS provides much greater confidence in the identification and pinpoints the exact cross-linked residues. Notably, Satoh and co-workers

(Enami et al., 1998) used FAB-MS to identify intramolecular cross-linked peptides in PsbO, and deduced the linked residues even without MS/MS capability.

Since these early studies, crystal structures have elucidated the connectivity between the components of active cyanobacterial PSII. As a result, more recent cross-linking studies have focused on accessory proteins that bind only to subcomplexes and/or that are not found in the crystal structures, though work has continued on the lumenal extrinsic PSII subunits from algae and higher plants, PsbP and PsbQ, which differ significantly from their cyanobacterial counterparts (Bricker et al., 2012) (results are summarized in Table 5). Cross-linking-MS has also been recently applied to study interactions within the phycobilisome (Tal et al., 2014) and between the phycobilisome and the photoprotective orange carotenoid protein (OCP) (Zhang et al., 2014; Liu et al., 2016), reviewed in Bricker et al., 2015.

With complementary use of the cross-linkers EDC and DTSSP, Pakrasi and co-workers (Liu et al., 2011a) demonstrated that the accessory protein Psb27 binds on the lumenal surface of CP43. Because this interaction is transient and occurs in only a small fraction of PSII centers in the cell at a given time, the authors purified PSII complexes from the Δ*ctp*A mutant strain of *Synechocystis* 6803 that accumulates such complexes (Liu et al., 2011b), maximizing chances of capturing and observing Psb27 inter-protein cross-links. The two cross-linked species detected were used to map Psb27 onto the PSII crystal structure, showing how Psb27 accomplishes its role as a gatekeeper, protecting partially assembled PSII complexes from gaining premature harmful water oxidation activity (Roose and Pakrasi, 2008). Recently, Nowaczyk and co-workers (Cormann et al., 2016) identified a different cross-link between Psb27 and CP43 in *T. elongatus* using an isotope-encoded version of the BS^3 cross-linker. Despite the different cyanobacterial

Table 5. Summary of MS-based PSII cross-linking studies[a]

Species	Cross-linked subunit 1	Cross-linked subunit 2	Cross-linker	Method notes	Reference
Syn. 6803	Psb27	CP43	EDC, DTSSP	• Cross-linked species enriched on gel • In-gel digestion; trypsin or chymotrypsin • LTQ-Orbitrap XL • MassMatrix search software	Liu et al., 2011a
Syn. 6803	PsbQ	CP47, PsbO	EDC, DTSSP	• No cross-link enrichment • In-solution digestion; trypsin • LTQ-Orbitrap XL • MassMatrix search software	Liu et al., 2014b
Syn. 6803	D2, CP43, CP47	ApcE	DSP	• No pre-MS cross-link enrichment • In-solution digestion; trypsin + LysC • LTQ-Orbitrap XL • ≥+3 charge states selected for MS^2 to maximize cross-link selection • MassMatrix search software	Liu et al., 2013b
T. elongatus	Psb27	CP43	BS^3 d0/d12	• Isotope-encoded cross-linker • No cross-link enrichment • In-solution digestions; trypsin • Orbitrap Elite Velos Pro • StavroX search software	Cormann et al., 2016
C. reinhardtii	PsbP	PsbQ	EDC	• Wash step isolates extrinsic proteins after cross-linking • Cross-linked species enriched on gel • In-gel digestion; trypsin or Asp-N • Ultraflex MALDI-TOF • MS^1 only; trypsin and Asp-N samples independently indicate the same cross-linked residues	Nagao et al., 2010
Spinach	PsbP	PsbQ	BS^3	• Wash step isolates extrinsic proteins after cross-linking • Cross-linked species enriched on gel • In-gel digestion; trypsin +/- LysC • LTQ-FTICR • MassMatrix search software	Mummadi-setti et al., 2014
Spinach	PsbP	PsbE	EDC	• Biotin-tagged PsbP isolates the free protein + its cross-linked partners • Cross-linked species enriched on gel • In-gel digestion; trypsin • LTQ-Orbitrap XL • MassMatrix search software	Ido et al., 2012
Spinach	PsbP	PsbR, CP26	EDC	• Biotin-tagged PsbP isolates the free protein + its cross-linked partners • Cross-linked species enriched on gel • In-gel digestion; trypsin • LTQ-Orbitrap XL • MassMatrix search software	Ido et al., 2014

[a] Only inter-protein cross-links that reveal interactions not detectable in the available PSII crystal structures are shown here.

species used in the two studies, and the different Psb27 residues that were cross-linked, both cross-links localize Psb27 to the same domain on CP43 (Liu et al., 2011; Cormann et al., 2016).

Cyanobacterial PsbQ is a component of active PSII (Roose et al., 2007), but is not found in any of the crystal structures, presumably because it is destabilized under crystallization conditions. Pakrasi and co-workers (Liu et al., 2014b) again used EDC and DTSSP in parallel and detected a PsbQ-CP47 and two PsbQ-PsbO cross-links by MS. A PsbQ-PsbQ cross-link that appears to arise from two different copies of the protein was also detected. Taken together, these results position PsbQ along the lumenal PSII dimer interface, consistent with evidence that PsbQ stabilizes the PSII dimer (Liu et al., 2014b). In this study, in-solution digestion was used instead of in-gel digestion to avoid losses of large cross-linked peptides that are difficult to extract from the gel matrix.

Several recent studies have probed the binding sites of the higher plant lumenal extrinsic proteins PsbP and PsbQ, which help optimize Ca^{2+} and Cl^- binding properties at the oxygen-evolving center (Bricker et al., 2012). Ifuku and co-workers (Ido et al., 2012; Ido et al., 2014) identified cross-links in spinach PSII between PsbP and PsbE, PsbR, and CP26 by MS and provided MS-based evidence for PsbP-CP43, PsbQ-CP43 and PsbQ-CP26 cross-links. The suggestive evidence arose from MS identification of CP43 or CP26 in individual cross-linked gel bands after affinity pull-downs using biotin-tagged PsbP or PsbQ (Ido et al., 2014). Their binding model for PsbP is different than that proposed by Bricker and co-workers (Mummadisetti et al., 2014), who identified nine intra-protein cross-links between the N-terminal and C-terminal regions of spinach PsbP that constrain significantly its binding conformation. The authors also identified a PsbP-PsbQ cross-link, consistent with that observed in *C. reinhardtii* by Enami and co-workers (Nagao et al., 2010).

The PsbQ-CP43 interaction in spinach PSII suggested by Ifuku and co-workers (Ido et al., 2014) contrasts with the PsbQ-CP47 cross-link identified in *Synechocystis* 6803 by Pakrasi and co-workers (Liu et al., 2014b) and their evidence for a PsbQ-PsbQ interaction at the PSII dimer interface. Significant sequence differences between cyanobacterial and plant PsbQ may explain this discrepancy. Bricker and co-workers (Mummadisetti et al., 2014) also found cross-linking evidence for a PsbQ-PsbQ interaction in spinach that may require a position at the dimer interface, consistent with the Pakrasi group's results in *Synechocystis* 6803. However, they suggest that that interaction could in theory arise from an inter-PSII-dimer interaction, and, thus, the results could alternatively be consistent with the Ifuku group's positioning of spinach PsbQ near CP43. Interestingly, the recently published crystal structure of PSII from the eukaryotic red alga *Cyanidium caldarium* indeed shows PsbQ′ binding to the lumenal surface of CP43 (Ago et al., 2016). PsbQ′ shares relatively low sequence homology to green algal or higher plant PsbQ; and though PsbQ′ can functionally replace PsbQ at least partially in *C. reinhardtii*, it cannot bind to spinach PSII (Ohta et al., 2003). Therefore, the red algal PsbQ′-CP43 interaction supports Ifuku and co-workers' (Ido et al., 2014) similar conclusion in spinach**,** but at the same time it does not necessarily contradict the alternate PsbQ-CP47 interaction observed by the other groups in spinach and *Synechocystis* 6803. The recent characterization of an active PSII complex from *Synechocystis* 6803 with multiple copies of the PsbQ protein (Liu et al., 2015) hints at one possible reconciliation of these findings, if such a complex is present in other species as well. Despite some discrepancies, these results begin to elucidate the binding orientation of the higher plant lumenal extrinsic proteins, suggesting a mechanism for stabilization of PSII-LHCII supercomplexes (Ido et al., 2014), and paving the road for further structural studies.

Though the advanced techniques for improving cross-link identification described in the methods section above have largely not yet been applied to PSII studies (with the exception of the recent use of isotope-encoded BS^3 by Nowaczyk and co-workers (Cormann et al., 2016)), several other creative approaches have been used. Enami and co-workers (Nagao et al., 2010) improved identification confidence by detecting the same cross-linked residues in peptides from two separate digestion experiments, one with trypsin and one with Asp-N. Pakrasi and co-workers (Liu et al., 2011a) provided strong evidence, using the thiol-cleavable cross-linker DTSSP and 2D gel electrophoresis, that Psb27 and CP43 cross-link to each other, allowing targeted data analysis and providing higher confidence in the subsequent MS cross-link identification. Ifuku and co-workers (Ido et al., 2012; Ido et al., 2014) used a biotin-tagged PsbP or PsbQ to purify only those cross-linked proteins. Although this method is not as efficient as purifying only cross-linked peptides by means of a tagged linker, because following digestion many non-linked peptides from the tagged protein will be present, it does simplify sample complexity and focuses on cross-links containing a particular protein of interest. Notably, Blankenship and co-workers (Liu et al., 2013b) demonstrated that *in-vivo* cross-linking of thylakoid membrane complexes is possible and can capture interactions between protein complexes that are otherwise difficult to preserve after cell lysis. Using the membrane-permeable cross-linker DSP, they captured a PSII-PSI-phycobilisome megacomplex and identified five cross-links between PSII subunits and the PBS, and five between PSI subunits and the PBS, providing the first molecular-level description of the interface of these complexes.

Like cross-linking, protein footprinting is a technique that has long been used in PSII structural studies but that has become significantly more powerful in combination with modern MS. Early studies using N-hydroxysuccinimidobiotin (NHS-biotin) and other modification

reagents investigated the binding site of higher plant PsbO to PSII. In the absence of MS detection, specific modification sites could either not be identified (Bricker et al., 1988) or were localized to particular protein domains by N-terminal sequencing of peptides (Frankel and Bricker, 1992). With the rise of protein MS in the mid-1990s, MALDI-TOF and FAB-MS were used to identify modified peptides; lack of MS/MS capability, however, produced lower-confidence peptide identification than is achievable today, and meant that specific modified residues could only be pinpointed in favorable cases (Frankel and Bricker, 1995; Miura et al., 1997; Frankel et al., 1999). Nonetheless, these pioneering footprinting studies demonstrated, e.g., that PsbO interacts with Loop E of CP47 (Frankel and Bricker, 1992), and that charged residues on the surface of PsbO are involved in its interaction with PSII (Miura et al., 1997; Frankel et al., 1999).

Recently, hydroxyl radical footprinting using synchrotron radiolysis of water was used to study the binding surfaces of spinach PsbP and PsbQ to PSII, with detection of modified residues by MS (Mummadisetti et al., 2014). The results reveal buried regions on the surface of these proteins that complement the authors' cross-linking data and suggest these proteins' binding interfaces to other PSII subunits. The data also confirm and elaborate on the binding region identified by this group in a previous study using NHS-biotin as footprinting reagent (Meades et al., 2005).

Although the above footprinting studies detected whether or not a residue was modified in a given state, it is also possible to analyze footprinting data quantitatively to detect a conformational change in a complex in two different states. The label-free approaches described above can be used to monitor the relative change in modification, normalized to the unmodified peptide, in different PSII complexes. The utility of this approach was demonstrated in a study of the role of Psb27 in PSII assembly (Liu et al., 2013a) using GEE labeling. The authors monitored

the relative changes in aspartate and glutamate modification of three PSII complexes representing different stages of PSII assembly, not only extending previous information about the Psb27 binding site (Liu et al., 2011a; Komenda et al., 2012a), but also demonstrating a conformational change upon D1 processing that prompts Psb27 dissociation and permits assembly of the oxygen evolving complex (Liu et al., 2013a). Blankenship and co-workers have also used quantitative GEE labeling to detect a light-dependent conformational change in the OCP protein that appears to underlie its photoprotective function (Liu et al., 2014a). The recent implementation of isotopically-labeled GEE (iGEE) footprinting (Zhang et al., 2016) will streamline, and increase confidence in, quantitative comparisons of modification extent between states.

Hydroxyl radical footprinting has also been used to identify putative water and oxygen channels in PSII (Frankel et al., 2013a), a topic that has been explored previously through computational studies (Murray and Barber, 2007; Ho et al., 2008; Gabdulkhakov et al., 2009; Vassiliev et al., 2012). This study provides general experimental support for the existence of such channels, confirms specific channel identifications from computational work (Ho et al., 2008; Vassiliev et al., 2012), and proposes a previously unidentified putative oxygen/ROS exit channel (see Bricker et al., 2015 for a discussion of the MS-based and computational results).

This work

As discussed in the introduction above, the available crystal structures of PSII provide an extremely valuable, detailed map of the protein subunits and cofactors in active PSII. This structural information has been critical in elucidating the sequence and energetics of the electron transfer chain in PSII. The 1.9 Å crystal structure (PDB 3WU2) (Umena et al., 2011) has even contributed insights into the mechanism of water oxidation with its high-resolution depiction of the geometry of the Mn_4Ca cluster. Nonetheless, as the crystal structures reflect the fully-

assembled, active complex, they can provide little structural insight into the several dozen accessory proteins that are now known to interact with PSII during various other stages of its life-cycle (Nickelsen and Rengstl, 2013; Järvi et al., 2015; Heinz et al., 2016). It has even become clear that an additional protein (PsbQ), is present in the active complex of cyanobacterial PSII, yet is missing from the crystal structures (Thornton et al., 2004; Roose et al., 2007).

The identification and characterization of increasing numbers of accessory PSII proteins in recent years has coincided with recent dramatic improvements in method development for chemical cross-linking combined with MS (Rappsilber, 2011; Sinz, 2014). In this context, we sought to identify the binding site of the elusive Psb28 protein using an isotopically-encoded chemical cross-linker followed by MS analysis. Psb28 binds in an unknown location primarily to a low-abundance PSII assembly intermediate complex, RC47. Psb28 exerts a protective effect on RC47, especially under stress conditions of high light or high temperature (Sakata et al., 2013). Our results, presented in Chapter 2, show that Psb28 binds on the cytosolic surface of PSII in close association with PsbE and PsbF, the α- and β-subunits of cytochrome b_{559}, an essential component of the PSII reaction center. The structural location of Psb28 that we have determined allows us to propose several mechanisms by which it could exert its protective effect on the RC47 intermediate.

In the work presented in Chapter 3, we used high-resolution tandem MS to identify oxidative modifications in PSII. PSII residues are believed to be modified by reactive oxygen species (ROS) unavoidably produced by redox-active PSII cofactors. These modifications have long been suspected as a major reason for the frequent damage and turnover of PSII proteins. The sensitivity and speed of the latest mass spectrometers affords the possibility to characterize these modifications definitively and with high throughput, a challenging or impossible task by

any other method. Our results show that oxidative modifications increase in response to light, and map two putative paths of travel for oxygen/ROS from the Mn_4Ca cluster all the way to the surface of PSII.

Chemical cross-linking followed by mass spectrometry was implemented again in the work presented in Chapter 4, in this case to identify the binding site of the cyanobacterial PsbQ protein. As mentioned above, this protein is a component of active PSII, yet was apparently lost during crystallization so its binding site remained unknown. In this study, it was determined that PsbQ associates closely with CP47 and PsbO, and binds near the dimer interface of PSII. In addition, a novel PSII complex that contains four copies of PsbQ was identified and characterized. An updated model of the PSII life-cycle is presented that includes this complex.

Chapter 5 summarizes the results of this work as well as the role of MS in PSII life-cycle research, and future directions in this area. The current status of an ongoing study is also presented, which involves characterization of a novel PSII subcomplex lacking the reaction center proteins. Table 6 below is a modified version of Table 1, and it highlights the PSII life-cycle topics addressed, and the MS tools used to study them, in this work.

Table 6. Modified version of Table 1 that provides an overview of the kind of work conducted in this study

Kind of information	Information desired	MS-based technique	Applied in chapter
Composition	PSII subunits present in a complex	Bottom-up MS (intact or top-down MS for LMM subunits)	2-4
	Accessory proteins that associate with PSII	Bottom-up MS	
	PTMs	Bottom-up MS	3
Dynamics	Protein and PTM changes between samples	Label-free or isotopic-label-based relative quantification	3*
	PSII subunit lifetime	Rate of unlabeled protein disappearance after isotopic label exposure	
	Relative position of subcomplexes in PSII life-cycle	Relative isotopic label incorporation after pulse	
Structure	Binding site of proteins not found in PSII crystal structures	Cross-linking, footprinting	2, 4
	Conformational changes	Footprinting, quantify changes in modification extent	
	Water and oxygen channel detection	Footprinting; mapping oxidative modifications	3

Blue text indicates topics addressed, and methods used, in the subsequent chapters.

*A label-free quantitative analysis of the results in Chapter 3 is ongoing; preliminary results of the label-free analysis are included.

References

Adir, N., Zer, H., Shochat, S., and Ohad, I. (2003). Photoinhibition - a historical perspective. *Photosynth. Res.* 76, 343-370.

Ago, H., Adachi, H., Umena, Y., Tashiro, T., Kawakami, K., Kamiya, N., et al. (2016). Novel features of eukaryotic Photosystem II revealed by its crystal structure analysis from a red alga. *J. Biol. Chem.* 291, 5676-5687.

Anderson, L.B., Maderia, M., Ouellette, A.J.A., Putnam-Evans, C., Higgins, L., Krick, T., et al. (2002). Posttranslational modifications in the CP43 subunit of photosystem II. *Proc. Natl. Acad. Sci. USA* 99, 14676-14681.

Anderson, L.B., Ouellette, A.J.A., Eaton-Rye, J., Maderia, M., MacCoss, M.J., Yates, J.R., et al. (2004). Evidence for a post-translational modification, aspartyl aldehyde, in a photosynthetic membrane protein. *J. Am. Chem. Soc.* 126, 8399-8405.

Andersson, L., and Porath, J. (1986). Isolation of phosphoproteins by immobilized metal (Fe^{3+}) affinity chromatography. *Anal. Biochem.* 154, 250-254.

Aro, E.-M., and Allahverdiyeva, Y. (2012). "Photosynthetic responses of plants to excess light: Mechanisms and conditions for photoinhibition, excess energy dissipation, and repair," in *Photosynthesis: Plastid biology, energy conversion and carbon assimilation*. *Advances in photosynthesis and respiration*, eds. J.J. Eaton-Rye, B.C. Tripathy, T.D. Sharkey (Dordrecht, the Netherlands: Springer), 275-297.

Aro, E.M., McCaffery, S., and Anderson, J.M. (1993). Photoinhibition and D1 protein degradation in peas acclimated to different growth irradiances. *Plant Physiol.* 103, 835-843.

Aro, E.-M., Suorsa, M., Rokka, A., Allahverdiyeva, Y., Paakkarinen, V., Saleem, A., et al. (2005). Dynamics of photosystem II: a proteomic approach to thylakoid protein complexes. *J. Exp. Bot.* 56, 347-356.

Aryal, U.K., Stöckel, J., Krovvidi, R.K., Gritsenko, M.A., Monroe, M.E., Moore, R.J., et al., (2011). Dynamic proteomic profiling of a unicellular cyanobacterium *Cyanothece* ATCC51142 across light-dark diurnal cycles. *BMC Syst. Biol.* 5, 194.

Back, J.W., de Jong, L., Muijsers, A.O., and de Koster, C.G. (2003). Chemical cross-linking and mass spectrometry for protein structural modeling. *J. Mol. Biol.* 331, 303-313.

Baena-González, E., and Aro, E.-M. (2002). Biogenesis, assembly and turnover of photosystem II units. *Phil. Trans. R. Soc. Lond. B.* 357, 1451-1459.

Bantscheff, M., Lemeer, S., Savitski, M.M., and Kuster, B. (2012). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal. Bioanal. Chem.* 404, 939-965.

Barber, J., and Andersson, B. (1992). Too much of a good thing—light can be bad for photosynthesis. *Trends Biochem. Sci.* 17, 61-66.

Battchikova, N., Angeleri, M., and Aro, E.-M. (2015). Proteomic approaches in research of cyanobacterial photosynthesis. *Photosynth. Res.* 126, 47-70.

Boehm, M., Romero, E., Reisinger, V., Yu, J., Komenda, J., Eichacker, L.A., et al. (2011). Investigating the early stages of Photosystem II assembly in *Synechocystis* sp. PCC 6803: Isolation of CP47 and CP43 complexes. *J. Biol. Chem.* 286, 14812-14819.

Boehm, M., Yu, J., Reisinger, V., Beckova, M., Eichacker, L. A., Schlodder, E., et al., (2012). Subunit composition of CP43-less photosystem II complexes of *Synechocystis* sp. PCC 6803: implications for the assembly and repair of photosystem II. *Phil. Trans. R. Soc. B* 367, 3444-3454.

Bonardi, V., Pesaresi, P., Becker, T., Schleiff, E., Wagner, R., Pfannschmidt, T., et al. (2005). Photosystem II core phosphorylation and photosynthetic acclimation require two different protein kinases. *Nature* 437, 1179-1182.

Bricker, T.M., Mummadisetti, M.P., and Frankel, L.K. (2015). Recent advances in the use of mass spectrometry to examine structure/function relationships in photosystem II. *J. Photochem. Photobiol. B* 152, 227-246.

Bricker, T.M., Odom, W.R., and Queirolo, C.B. (1988). Close association of the 33-kDa extrinsic protein with the apoprotein of CPa1 in photosystem II. *FEBS Lett.* 231, 111-117.

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.

Brunner, A.M., Lossl, P., Liu, F., Huguet, R., Mullen, C., Yamashita, M., et al. (2015). Benchmarking multiple fragmentation methods on an Orbitrap Fusion for top-down phospho-proteoform characterization. *Anal. Chem.* 87, 4152-4158.

Buncherd H., Roseboom W., de Koning L.J., de Koster C.G., de Jong L. (2014). A gas phase cleavage reaction of cross-linked peptides for protein complex topology studies by peptide fragment fingerprinting from large sequence database. *J. Proteomics* 108, 65-77.

Caffarri, S., Kouřil, R., Kereiche, S., Boekema, E.J., and Croce, R. (2009). Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.* 28, 3052-3064.

Catherman, A.D., Skinner, O.S., Kelleher, N.L. (2014). Top Down proteomics: Facts and perspectives. *Biochem. Biophys. Res. Commun.* 445, 683-693.

Chen, X., Wei, S., Ji, Y., Guo, X., and Yang, F. (2015). Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics* 15, 3175-3192.

Chu F., Mahrus S., Craik C.S., and Burlingame A.L. (2006). Isotope-coded and affinity-tagged cross-linking (ICATXL): An efficient strategy to probe protein interaction surfaces. *J. Am. Chem. Soc.* 128, 10362-10363.

Clarke, A.K., Soitamo, A., Gustafsson, P., and Öquist, G. (1993). Rapid interchange between two distinct forms of cyanobacterial photosystem II reaction center protein D1 in response to photoinhibition. *Proc. Natl. Acad. Sci. USA* 90, 9973-9977.

Clegg, C., and Hayes, D. (1974). Identification of neighboring proteins in ribosomes of *Escherichia coli*- Topographical study with crosslinking reagent dimethyl suberimidate. *Eur. J. Biochem.* 42, 21-28.

Cormann, K.U., Möller, M., and Nowaczyk, M.M. (2016). Critical assessment of protein cross-linking and molecular docking: An updated model for the interaction between Photosystem II and Psb27. *Front. Plant Sci.* 7, 157.

Dasgupta, J., Ananyev, G., and Dismukes, G.C. (2008). Photoassembly of the water-oxidizing complex in photosystem II. *Coord. Chem. Rev.* 252, 347-360.

Dedner, N., Meyer, H.E., Ashton, C., and Wildner, G.F. (1988). N-terminal sequence analysis of the 8 kDa protein in *Chlamydomonas reinhardii*: Localization of the phosphothreonine. *FEBS Lett*. 236, 77-82.

Dobáková, M., Sobotka, R., Tichý, M., and Komenda, J. (2009). 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.

Dobáková, M., Tichý, M., and Komenda, J. (2007). Role of the PsbI protein in photosystem II assembly and repair in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol.* 145, 1681-1691.

Dreaden, T.M., Chen, J., Rexroth, S., and Barry, B.A. (2011). N-formylkynurenine as a marker of high light stress in photosynthesis. *J. Biol. Chem.* 286, 22632-22641.

Enami, I., Kamo, M., Ohta, H., Takahashi, S., Miura, T., Kusayanagi, M., et al. (1998). Intramolecular cross-linking of the extrinsic 33-kDa protein leads to loss of oxygen evolution but not its ability of binding to Photosystem II and stabilization of the manganese cluster. *J. Biol. Chem.* 273, 4629-463.

Enami, I., Satoh, K., and Katoh, S. (1987). Crosslinking between the 33 kDa extrinsic protein and the 47 kDa chlorophyll-carrying protein of the PSII reaction center core complex. *FEBS Lett*. 226, 161-165.

Fagerlund, R.D., and Eaton-Rye, J.J. (2011). The lipoproteins of cyanobacterial photosystem II. *J. Photochem. Photobiol. B* 104, 191-203.

Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., et al. (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 20, 301-305.

Frankel, L.K., and Bricker, T.M. (1992). Interaction of CPa-1 with the manganese-stabilizing protein of Photosystem II: Identification of domains on CPa-1 which are shielded from N-

hydroxysuccinimide biotinylation by the manganese-stabilizing protein. *Biochemistry* 31, 11059-11064.

Frankel, L.K., and Bricker, T.M. (1995). Interaction of the 33-kDa extrinsic protein with Photosystem II: Identification of domains on the 33-kDa protein that are shielded from NHS-biotinylation by Photosystem II. *Biochemistry* 34, 7492-7497.

Frankel, L.K., Cruz, J.A., and Bricker, T.M. (1999). Carboxylate groups on the manganese-stabilizing protein are required for its efficient binding to Photosystem II. *Biochemistry* 38, 14271-14278.

Frankel, L.K., Sallans, L., Bellamy, H., Goettert, J.S., Limbach, P.A., and Bricker, T.M. (2013a). Radiolytic mapping of solvent-contact surfaces in Photosystem II of higher plants: Experimental identification of putative water channels within the photosystem. *J. Biol. Chem.* 288, 23565-23572.

Frankel, L.K., Sallans, L., Limbach, P.A., and Bricker, T.M. (2013b). Oxidized amino acid residues in the vicinity of Q_A and $Pheo_{D1}$ of the Photosystem II reaction center: Putative generation sites of reducing-side reactive oxygen species. *PLoS ONE* 8, e58042.

Frankel, L.K., Sallans, L., Limbach, P., and Bricker, T.M. (2012). Identification of oxidized amino acid residues in the vicinity of the Mn_4CaO_5 cluster of Photosystem II: Implications for the identification of oxygen channels within the photosystem. *Biochemistry* 51, 6371-6377.

Frese, C.K., Altelaar, A.F.M., van den Toorn, H., Nolting, D., Griep-Raming, J., Heck, A.J.R., et al. (2012). Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal. Chem.* 84, 9668-9673.

Fristedt, R., Granath, P., and Vener, A.V. (2010). A protein phosphorylation threshold for functional stacking of plant photosynthetic membranes. *PLoS ONE* 5, e10963.

Fristedt, R., Willig, A., Granath, P., Crevecoeur, M., Rochaix, J.-D., and Vener, A.V. (2009). Phosphorylation of Photosystem II controls functional macroscopic folding of photosynthetic membranes in *Arabidopsis*. *Plant Cell* 21, 3950-3964.

Fristedt, R., and Vener, A.V. (2011). High light induced disassembly of Photosystem II supercomplexes in *Arabidopsis* requires STN7-dependent phosphorylation of CP29. *PLoS ONE* 6, e24565.

Fritzsche, R., Ihling, C.H., Götze, M., and Sinz, A. (2012). Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun. Mass Spectrom.* 26, 653-658.

Gabdulkhakov, A., Guskov, A., Broser, M., Kern, J., Mueh, F., Saenger, W., et al., (2009). Probing the accessibility of the Mn_4Ca cluster in Photosystem II: Channels calculation, noble gas derivatization, and cocrystallization with DMSO. *Structure* 17, 1223-1234.

Galetskiy, D., Lohscheider, J.N., Kononikhin, A.S., Popov, I.A., Nikolaev, E.N., and Adamska, I. (2011a). Mass spectrometric characterization of photooxidative protein modifications in *Arabidopsis thaliana* thylakoid membranes. *Rapid Commun. Mass Spectrom.* 25, 184-190.

Galetskiy, D., Lohscheider, J.N., Kononikhin, A.S., Popov, I.A., Nikolaev, E.N., and Adamska, I. (2011b). Phosphorylation and nitration levels of photosynthetic proteins are conversely regulated by light stress. *Plant Mol. Biol.* 77, 461-473.

Gao, S., Gu, W., Xiong, Q., Ge, F., Xie, X., Li, J., et al. (2015). Desiccation enhances phosphorylation of PSII and affects the distribution of protein complexes in the thylakoid membrane. *Physiol. Plantarum* 153, 492-502.

Gau, B., Garai, K., Frieden, C., and Gross, M.L. (2011). Mass spectrometry-based protein footprinting characterizes the structures of oligomeric apolipoprotein E2, E3, and E4. *Biochemistry* 50, 8117-8126.

Geiger, T., Cox, J., and Mann, M. (2010). Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol. Cell. Proteomics* 9, 2252-2261.

Golden, S.S., Brusslan, J., and Haselkorn, R. (1986). Expression of a family of *psbA* genes encoding a photosystem II polypeptide in the cyanobacterium *Anacystis nidulans* R2. *EMBO J.* 5, 2789-2798.

Gómez, S.M., Bil', K.Y., Aguilera, R., Nishio, J.N., Faull, K.F., and Whitelegge, J.P. (2003). Transit peptide cleavage sites of integral thylakoid membrane proteins. *Mol. Cell. Proteomics* 2, 1068-1085.

Gómez, S.M., Nishio, J.N., Faull, K.F., and Whitelegge, J.P. (2002). The chloroplast grana proteome defined by intact mass measurements from liquid chromatography mass spectrometry. *Mol. Cell. Proteomics* 1, 46-59.

Gómez, S.M., Park, J.J., Zhu, J., Whitelegge, J.P., and Thornber, J.P. (1998). "Isolation and characterization of a novel xanthophyll-rich pigment-protein complex from spinach," in *Photosynthesis: Mechanisms and Effects Vol 1,"* ed. G. Garab (Dordrecht, the Netherlands: Kluwer Academic Publishers), 353-356.

Götze, M., Pettelkau, J., Fritzsche, R., Ihling, C.H., Schaefer, M., and Sinz, A. (2015). Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J. Am. Soc. Mass. Spectrom.* 26, 83-97.

Götze, M., Pettelkau, J., Schaks, S., Bosse, K., Ihling, C., Krauth, F., et al. (2012). StavroX-A Software for analyzing crosslinked products in protein interaction studies. *J. Am. Soc. Mass. Spectrom.* 23, 76-87.

Gouw, J.W., Krijgsveld, J., and Heck, A.J.R. (2010). Quantitative proteomics by metabolic labeling of model organisms. *Mol. Cell. Proteomics* 9, 11-24.

Granvogl, B., Zoryan, M., Plöscher, and Eichacker, L.A. (2008). Localization of 13 one-helix integral membrane proteins in Photosystem II subcomplexes. *Anal. Biochem.* 383, 279-288.

Grasse, N., Mamedov, F., Becker, K., Styring, S., Rögner, M., and Nowaczyk, M.M. (2011). Role of novel dimeric Photosystem II (PSII)-Psb27 protein complex in PSII repair. *J. Biol. Chem.* 286, 29548-29555.

Grossmann, J., Roschitzki, B., Panse, C., Fortes, C., Barkow-Oesterreicher, S., Rutishauser, D., et al. (2010). Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. *J. Proteomics* 73, 1740-1746.

Guerreiro, A.C.L., Benevento, M., Lehmann, R., van Breukelen, B., Post, H., Giansanti, P., et al. (2014). Daily rhythms in the cyanobacterium *Synechococcus elongatus* probed by high-resolution mass spectrometry-based proteomics reveals a small defined set of cyclic proteins. *Mol. Cell. Proteomics* 13, 2042-2055.

Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A., and Saenger, W. (2009). Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. *Nat Struct Mol Biol*. 16, 334-342.

Guo, J., Nguyen, A.Y., Dai, Z., Su, D., Gaffrey, M.J., Moore, R.J., et al. (2014). Proteome-wide light/dark modulation of thiol oxidation in cyanobacteria revealed by quantitative site-specific redox proteomics. *Mol. Cell. Proteomics* 13, 3270-3285.

Han, K.C., Shen, J.R., Ikeuchi, M., and Inoue, Y. (1994). Chemical cross-linking studies of extrinsic proteins in cyanobacterial photosystem II. *FEBS Lett.* 355, 121-124.

Haniewicz, P., De Sanctis, D., Büchel, C., Schröder, W.P., Loi, M.C., Kieselbach, T., et al. (2013). Isolation of monomeric photosystem II that retains the subunit PsbS. *Photosynth. Res.* 118, 199-207.

Haniewicz, P., Floris, D., Farci, D., Kirkpatrick, J., Loi, M.C., Büchel, C. (2015). Isolation of plant Photosystem II complexes by fractional solubilization. *Front. Plant Sci.* 6, 1100.

Harrer, R., Bassi, R., Testi, M.G., and Schäfer, C. (1998). Nearest-neighbor analysis of a photosystem II complex from *Marchantia polymorpha* L. (liverwort), which contains reaction center and antenna proteins. *Eur. J. Biochem.* 255, 196-205.

Hansson, M., Dupuis, T., Stromquist, R., Andersson, B., Vener, A.V., and Carlberg, I. (2007). The mobile thylakoid phosphoprotein TSP9 interacts with the light-harvesting complex II and the peripheries of both photosystems. *J. Biol. Chem.* 282, 16214-16222.

Heinemeyer, J., Eubel, H., Wehmhoner, D., Jansch, L., and Braun, H.P. (2004). Proteomic approach to characterize the supramolecular organization of photosystems in higher plants. *Phytochemistry* 65, 1683-1692.

Heinz, S., Liauw, P., Nickelsen, J., and Nowaczyk, M. (2016). Analysis of photosystem II biogenesis in cyanobacteria. *Biochim. Biophys. Acta*, 1857, 274-287.

Herbstová, M., Tietz, S., Kinzel, C., Turkina, M., and Kirchhoff, H. (2012). Architectural switch in plant photosynthetic membranes induced by light stress. *Proc. Natl. Acad. Sci. USA* 109, 20130-20135.

Ho, F.M., and Styring, S. (2008). Access channels and methanol binding site to the $CaMn_4$ cluster in Photosystem II based on solvent accessibility simulations, with implications for substrate water access. *Biochim. Biophys. Acta* 1777, 140-153.

Hoare, D.G., and Koshland, D.E. (1967). A method for quantitative modification and estimation of carboxylic acid groups in proteins. *J. Biol. Chem.* 242, 2447-2453.

Hoopmann, M.R., Zelter, A., Johnson, R.S., Riffle, M., MacCoss, M.J., Davis, T.N., et al. (2015). Kojak: Efficient analysis of chemically cross-linked protein complexes. *J. Proteome Res.* 14, 2190-2198.

Howery, A.E., Elvington, S., Abraham, S.J., Choi, K.H., Dworschak-Simpson, S., Phillips, S., et al. (2012). A designed inhibitor of a CLC antiporter blocks function through a unique binding mode. *Chem. Biol.* 19, 1460-1470.

Ido, K., Kakiuchi, S., Uno, C., Nishimura, T., Fukao, Y., Noguchi, T., et al. (2012). The conserved His-144 in the PsbP protein is important for the interaction between the PsbP N-terminus and the Cyt b_{559} subunit of Photosystem II. *J. Biol. Chem.* 287, 26377-26387.

Ido, K., Nield, J., Fukao, Y., Nishimura, T., Sato, F., and Ifuku, K. (2014). Cross-linking evidence for multiple interactions of the PsbP and PsbQ proteins in a higher plant Photosystem II supercomplex. *J. Biol. Chem.* 289, 20150-20157.

Ikeuchi, M., and Inoue, Y. (1988). A new 4.8-kDa polypeptide instrinsic 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.

Inagaki, N., Yamamoto, Y., and Satoh, K. (2001). A sequential two-step proteolytic process in the carboxyl-terminal truncation of precursor D1 protein in *Synechocystis* sp. PCC6803. *FEBS Lett*. 509, 197-201.

Ingle, R.A., Schmidt, U.G., Farrant, J.M., Thomson, J.A., and Mundree, S.G. (2007). Proteomic analysis of leaf proteins during dehydration of the resurrection plant *Xerophyta viscosa*. *Plant Cell Environ.* 30, 435-446.

Inoue-Kashino, N., Kashino, Y., Orii, H., Satoh, K., Terashima, I., and Pakrasi, H.B. (2011). S4 protein Sll1252 is necessary for energy balancing in photosynthetic electron transport in *Synechocystis* sp. PCC 6803. *Biochemistry* 50, 329-339.

Irrgang, K.D., Shi, L.X., Funk, C., and Schröder, W.P. (1995). A nuclear-encoded subunit of the Photosystem II reaction center. *J. Biol. Chem.* 270, 17588-17593.

Iwai, M., Suzuki, T., Dohmae, N., Inoue, Y., and Ikeuchi, M. (2007). Absence of the PsbZ subunit prevents association of PsbK and Ycf12 with the PSII complex in the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. *Plant Cell Physiol*. 48, 1758-1763.

Järvi, S., Suorsa, M., and Aro, E.-M. (2015). Photosystem II repair in plant chloroplasts - Regulation, assisting proteins and shared components with photosystem II biogenesis. *Biochim. Biophys. Acta* 1847, 900-909.

Järvi, S., Suorsa, M., Paakkarinen, V., Aro, E.-M. (2011). Optimized native gel systems for separation of thylakoid protein complexes: novel super- and mega-complexes. *Biochem J.* 439, 207-214.

Kang, S., Mou, L., Lanman, J., Velu, S., Brouillette, W., Prevelige, and P.E., Jr. (2009). Synthesis of biotin-tagged chemical cross-linkers and their applications for mass spectrometry. *Rapid Commun. Mass Spectrom.* 23, 1719-1726.

Kao, A., Chiu, C.-l., Vellucci, D., Yang, Y., Patel, V.R., Guan, S., et al. (2011). Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol. Cell. Proteomics* 10, (M110):002212.

Kashino, Y., Lauber, W.M., Carroll, J.A., Wang, Q.J., Whitmarsh, J., Satoh, K., et al. (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.

Kasson, T.M.D., Rexroth, S., and Barry, B.A. (2012). Light-induced oxidative stress, N-formylkynurenine, and oxygenic photosynthesis. *PLoS ONE* 7, 42220.

Knoppová, J., Sobotka, R., Tichý, M., Yu, J., Konik, P., Halada, P., et al. (2014). Discovery of a chlorophyll binding protein complex involved in the early steps of Photosystem II assembly in *Synechocystis*. *Plant Cell* 26, 1200-1212.

Komenda, J., Hassan, H.A.G., Diner, B.A., Debus, R.J., Barber, J., and Nixon, P.J. (2000). Degradation of the Photosystem II D1 and D2 proteins in different strains of the cyanobacterium *Synechocystis* PCC 6803 varying with respect to the type and level of *psbA* transcript. *Plant Mol. Biol.* 42, 635-645.

Komenda, J., Knoppová, J., Kopečná, J., Sobotka, R., Halada, P., Yu, J., et al. (2012a). The Psb27 assembly factor binds to the CP43 complex of Photosystem II in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol*. 158, 476-486.

Komenda, J., Kuviková, S., Granvogl, B., Eichacker, L.A., Diner, B.A., and Nixon, P.J. (2007). 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* 1767, 829-837.

Komenda, J., Nickelsen, J., Tichý, M., Prášil, O., Eichacker, L., and Nixon, P.J. (2008). 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.

Komenda, J., Reisinger, V., Müller, B.C., Dobáková, M., Granvogl, B., and Eichacker, L.A. (2004). Accumulation of the D2 protein is a key regulatory step for assembly of the Photosystem II reaction center complex in *Synechocystis* PCC 6803. *J. Biol. Chem.* 279, 48620-48629.

Komenda, J., Sobotka, R., and Nixon, P.J. (2012b). Assembling and maintaining the Photosystem II complex in chloroplasts and cyanobacteria. *Curr. Opin. Plant Biol.* 15, 245-251.

Komenda, J., Tichý, M., and Eichacker, L.A. (2005). The PsbH protein is associated with the inner antenna CP47 and facilitates D1 processing and incorporation into PSII in the cyanobacterium *Synechocystis* PCC 6803. *Plant Cell Physiol.* 46, 1477-1483.

Kós, P.B., Deák, Z., Cheregi, O., and Vass, I. (2008). Differential regulation of *psbA* and *psbD* gene expression, and the role of the different D1 protein copies in the cyanobacterium *Thermosynechococcus elongatus* BP-1. *Biochim. Biophys. Acta* 1777, 74-83.

Kouřil, R., Dekker, J.P., and Boekema, E.J. (2012). Supramolecular organization of photosystem II in green plants. *Biochim. Biophys. Acta* 1817, 2-12.

Kufryk, G., Hernandez-Prieto, M.A., Kieselbach, T., Miranda, H., Vermaas, W., and Funk, C. (2008). Association of small CAB-like proteins (SCPs) of *Synechocystis* sp. PCC 6803 with Photosystem II. *Photosynth. Res.* 95, 135-145.

Laganowsky, A., Gómez, S.M., Whitelegge, J.P., and Nishio, J.N. (2009). Hydroponics on a chip: Analysis of the Fe deficient *Arabidopsis* thylakoid membrane proteome. *J. Proteomics* 72, 397-415.

Leitner, A., Reischl, R., Walzthoeni, T., Herzog, F., Bohn, S., Förster, F., et al. (2012). Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol. Cell. Proteomics* 11, M111.014126.

Lemeille, S. and Rochaix, J.-D. (2010). State transitions at the crossroad of thylakoid signalling pathways. *Photosynth. Res.* 106, 33-46.

Lemeille, S., Turkina, M., Vener, A.V., and Rochaix, J.-D. (2009). Stt7-dependent phosphorylation during state transitions in the green alga *Chlamydomonas reinhardtii*. *Mol. Cell. Proteomics* 9, 1281-1295.

Li, X.-J., Yang, M.-F., Zhu, Y., Liang, Y., and Shen, S.-H. (2011). Proteomic analysis of salt stress responses in rice shoot. *J. Plant Biol.* 54, 384-395.

Liu, H., Chen, J., Huang, R.-C., Weisz, D., Gross, M.L., and Pakrasi, H.B. (2013a). Mass spectrometry-based footprinting reveals structural dynamics of Loop E of the chlorophyll-binding protein CP43 during Photosystem II assembly in the cyanobacterium *Synechocystis* 6803. *J. Biol. Chem.* 288, 14212-14220.

Liu, H., Huang, R.Y.-C., Chen, J., Gross, M.L., and Pakrasi, H.B. (2011a). Psb27, a transiently associated protein, binds to the chlorophyll binding protein CP43 in photosystem II assembly intermediates. *Proc. Natl. Acad. Sci. USA* 108, 18536-18541.

Liu, H., Roose, J.L., Cameron, J.C., and Pakrasi, H.B. (2011b). A genetically tagged Psb27 protein allows purification of two consecutive Photosystem II (PSII) assembly intermediates in *Synechocystis* 6803, a cyanobacterium. *J. Biol. Chem.* 286, 24865-24871.

Liu, H., Weisz, D.A., and Pakrasi, H.B. (2015). Multiple copies of the PsbQ protein in a cyanobacterial photosystem II assembly intermediate complex. *Photosynth. Res.* 126, 375-383.

Liu, H., Zhang, H., King, J., Wolf, N., Prado, M., Gross, M.L., et al. (2014a). Mass spectrometry footprinting reveals the structural rearrangements of cyanobacterial orange carotenoid protein upon light activation. *Biochim. Biophys. Acta* 1837, 1955-1963.

Liu, H., Zhang, H., Niedzwiedzki, D., Prado, M., He, G., Gross, M.L., et al. (2013b). Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. *Science* 342, 1104-1107.

Liu, H., Zhang, H., Orf, G.S., Lu, Y., Jiang, J., King, J.D., et al. (2016). Dramatic domain rearrangements of the cyanobacterial orange carotenoid protein upon photoactivation. *Biochemistry* 55, 1003-1009.

Liu, H., Zhang, H., Weisz, D.A., Vidavsky, I., Gross, M.L., Pakrasi, H.B. (2014b). MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. *Proc. Natl. Acad. Sci. USA* 111, 4638-4643.

Lohrig, K., Mueller, B., Davydova, J., Leister, D., and Wolters, D.A. (2009). Phosphorylation site mapping of soluble proteins: bioinformatical filtering reveals potential plastidic phosphoproteins in *Arabidopsis thaliana*. *Planta* 229, 1123-1134.

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.

Lorković, Z.J., Schröder, W.P., Pakrasi, H.B., Irrgang, K.D., Herrmann, R.G., and Oelmüller, R. (1995). Molecular characterization of PsbW, a nuclear-encoded component of the Photosystem II reaction center complex in spinach. *Proc. Natl. Acad. Sci. USA* 92, 8930-8934.

Lundgren, D.H., Hwang, S.-I., Wu, L., and Han, D.K. (2010). Role of spectral counting in quantitative proteomics. *Expert Rev. Proteomics* 7, 39-53.

Mabbitt, P.D., Wilbanks, S.M., and Eaton-Rye, J.J. (2014). Structure and function of the hydrophilic Photosystem II assembly proteins: Psb27, Psb28 and Ycf48. *Plant Physiol. Bioch.* 81, 96-107.

Mädler, S., Bich, C., Touboul, D., and Zenobi, R. (2009). Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J. Mass Spectrom.* 44, 694-706.

Meades, G.D., 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.

Mehmood, S., Allison, T.M., and Robinson, C.V. (2015). Mass spectrometry of protein complexes: From origins to applications. *Annu. Rev. Phys. Chem.* 66, 453-474.

Michel, H.P., and Bennett, J. (1987). Identification of the phosphorylation site of an 8.3 kDa protein from photosystem II of spinach. *FEBS Lett*. 212, 103-108.

Michel, H., Hunt, D.F., Shabanowitz, J., and Bennett, J. (1988). Tandem mass spectrometry reveals that three Photosystem II proteins of spinach chloroplasts contain *N*-acetyl-*O*-phosphothreonine at their NH_2 termini. *J. Biol. Chem.* 263, 1123-1130.

Minagawa, J. (2011). State transitions-The molecular remodeling of photosynthetic supercomplexes that controls energy flow in the chloroplast. *Biochim. Biophys. Acta* 1807, 897-905.

Miura, T., Shen, J.R., Takahashi, S., Kamo, M., Nakamura, E., Ohta, H., et al. (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.

Müller, B., and Eichacker, L.A. (1999). Assembly of the D1 precursor in monomeric Photosystem II reaction center precomplexes precedes Chlorophyll *a*-triggered accumulation of Reaction Center II in barley etioplasts. 11, 2365-2377.

Müller, D.R., Schindler, P., Towbin, H., Wirth, U., Voshol, H., Hoving, S., et al. (2001). Isotope tagged cross linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal. Chem.* 73, 1927-1934.

Mullet, J.E., and Christopher, D.A. (1994). Separate photosensory pathways coregulate blue-light/ultraviolet-A-activated *psb*D-*psb*C transcription and light-induced D2 and CP43 degradation in barley (*Hordeum vulgare*) chloroplasts. *Plant Physiol*. 104, 1119-1129.

Mullineaux, C.W. (2008). Phycobilisome-reaction centre interaction in cyanobacteria. *Photosynth. Res.* 95, 175-182.

Mulo, P., Sakurai, I., and Aro, E.-M. (2012). Strategies for *psbA* gene expression in cyanobacteria, green algae and higher plants: From transcription to PSII repair. *Biochim. Biophys. Acta* 1817, 247-257.

Mulo, P., Sicora, C., and Aro, E.-M. (2009). Cyanobacterial *psbA* gene family: optimization of oxygenic photosynthesis. *Cell. Mol. Life Sci.* 66, 3697-3710.

Mummadisetti, M., Frankel, L.K., Bellamy, H.D., Sallans, L., Goettert, J.S., Brylinski, M., et al. (2014). Use of protein cross-linking and radiolytic footprinting to elucidate PsbP and PsbQ interactions within higher plant Photosystem II. *Proc. Natl. Acad. Sci. USA* 111, 16178-16183.

Murray, J.W., and Barber, J. (2007). Structural characteristics of channels and pathways in photosystem II including the identification of an oxygen channel. *J. Struct. Biol.* 159, 228-237.

Nagao, R., Suzuki, T., Okumura, A., Niikura, A., Iwai, M., and Dohmae, N. (2010). Topological analysis of the extrinsic PsbO, PsbP and PsbQ proteins in a green algal PSII complex by cross-linking with a water-soluble carbodiimide. *Plant Cell Physiol.* 51, 718-727.

Nahnsen, S., Bielow, C., Reinert, K., and Kohlbacher, O. (2013). Tools for label-free peptide quantification. *Mol. Cell. Proteomics* 12, 549-556.

Nakamori, H., Yatabe, T., Yoon, K.-S., and Ogo, S. (2014). Purification and characterization of an oxygen-evolving photosystem II from *Leptolyngbya* sp. strain O-77. *J. Biosci. Bioeng.* 118, 119-124.

Nanba, O., and Satoh, K. (1987). Isolation of a photosystem II reaction center consisting of D-1 and D-2 polypeptides and cytochrome *b*-559. *Proc. Natl. Acad. Sci. USA* 84, 109-112.

Naumann, B., Busch, A., Allmer, J., Ostendorf, E., Zeller, M., Kirchhoff, H., et al. (2007). Comparative quantitative proteomics to investigate the remodeling of bioenergetic pathways under iron deficiency in *Chlamydomonas reinhardtii*. *Proteomics* 7, 3964-3979.

Nickelsen, J., and Rengstl, B. (2013). Photosystem II assembly: From cyanobacteria to plants. *Annu. Rev. Plant Biol.* 64, 609-635.

Nickelsen, J., Rengstl, B., Stengel, A., Schottkowski, M., Soll, J., and Ankele, E. (2011). Biogenesis of the cyanobacterial thylakoid membrane system - an update. *FEMS Microbiol. Lett.* 315, 1-5.

Nixon, P.J., Trost, J.T., and Diner, B.A. (1992). Role of the carboxy terminus of polypeptide D1 in the assembly of a functional water-oxidizing manganese cluster in Photosystem II of the cyanobacterium *Synechocystis* sp. PCC 6803: Assembly requires a free carboxyl group at C-terminal position 344. *Biochemistry* 31, 10859-10871.

Nixon, P.J., Michoux, F.,Yu, J., Boehm, M., and Komenda, J. (2010) Recent advances in understanding the assembly and repair of photosystem II. *Ann. Bot.* 106, 1-16.

Nowaczyk, M.M., Hebeler, R., Schlodder, E., Meyer, H.E., Warscheid, B., and Rögner, M. (2006). Psb27, a cyanobacterial lipoprotein, is involved in the repair cycle of Photosystem II. *Plant Cell* 18, 3121-3131.

Nowaczyk, M.M., Krause, K., Mieseler, M., Sczibilanski, A., Ikeuchi, M., and Rögner, M. (2012). Deletion of *psbJ* leads to accumulation of Psb27-Psb28 photosystem II complexes in *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta* 1817, 1339-1345.

Odom, W.R., and Bricker, T.M. (1992). Interaction of CPa-1 with the manganese-stabilizing protein of Photosystem II: Identification of domains cross-linked by 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide. *Biochemistry* 31, 5616-5620.

Ohnishi, N., and Murata, N. (2006). Glycinebetaine counteracts the inhibitory effects of salt stress on the degradation and synthesis of D1 protein during photoinhibition in *Synechococcus* sp. PCC 7942. *Plant Physiol.* 141, 758-765.

Ohta, H., Suzuki, T., Ueno, M., Okumura, A., Yoshihara, S., Shen, J.R., et al. (2003). Extrinsic proteins of photosystem II - An intermediate member of the PsbQ protein family in red algal PSII. *Eur. J. Biochem.* 270, 4156-4163.

Owens, G.C., and Ohad, I. (1983). Changes in thylakoid polypeptide phosphorylation during membrane biogenesis in *Chlamydomonas reinhardii* y-1. *Biochim. Biophys. Acta* 722, 234-241.

Owens, G.C., and Ohad, I. (1982). Phosphorylation of *Chlamydomonas reinhardi* chloroplast membrane proteins in vivo and in vitro. *J. Cell. Biol.* 93, 712-718.

Pagliano, C., Chimirri, F., Saracco, G., Marsano, F., and Barber, J. (2011). One-step isolation and biochemical characterization of a highly active plant PSII monomeric core. *Photosynth. Res.* 108, 33-46.

Pagliano, C., Nield, J., Marsano, F., Pape, T., Barera, S., Saracco, G., et al. (2014). Proteomic characterization and three-dimensional electron microscopy study of PSII-LHCII supercomplexes from higher plants. *Biochim. Biophys. Acta* 1837, 1454-1462.

Paramelle, D., Miralles, G., Subra, G., and Martinez, J. (2013). Chemical cross-linkers for protein structure studies by mass spectrometry. *Proteomics* 13, 438-456.

Pearson K.M., Pannell L.K., and Fales H.M. (2002). Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Commun. Mass Spectrom.* 16, 149–159.

Pesaresi, P., Pribil, M., Wunder, T., and Leister, D. (2011). Dynamics of reversible protein phosphorylation in thylakoids of flowering plants: The roles of STN7, STN8 and TAP38. *Biochim. Biophys. Acta* 1807, 887-896.

Petrotchenko E.V., and Borchers C.H. (2010). ICC-CLASS: isotopically-coded cleavable crosslinking analysis software suite. *BMC Bioinformatics*

Petrotchenko E.V., Makepeace K.A.T., Serpa J.J., Borchers C.H. (2014). Analysis of protein structure by cross-linking combined with mass spectrometry. *Methods Mol. Biol.* 1156, 447-463.

Petrotchenko E.V., Serpa J.J., Borchers, C.H. (2011). An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol. Cell. Proteomics* 10, 10.1074/mcp.M110.001420.

Plöscher, M., Granvogl, B., Zoryan, M., Reisinger, V., and Eichacker, L.A. (2009). Mass spectrometric characterization of membrane integral low molecular weight proteins from photosystem II in barley etioplasts. *Proteomics* 9, 625-635.

Plöscher, M., Reisinger, V., Eichacker, L.A. (2011). Proteomic comparison of etioplast and chloroplast protein complexes. *J. Proteomics* 74, 1256-1265.

Pospíšil, P. (2009). Production of reactive oxygen species by photosystem II. *Biochim. Biophys. Acta* 1787, 1151-1160.

Promnares, K., Komenda, J., Bumba, L., Nebesarova, J., Vacha, F., and Tichy, M. (2006). Cyanobacterial small chlorophyll-binding protein ScpD (HliB) is located on the periphery of Photosystem II in the vicinity of PsbH and CP47 subunits. *J. Biol. Chem.* 281, 32705-32713.

Puthiyaveetil, S., and Kirchhoff, H. (2013). A phosphorylation map of the photosystem II supercomplex C2S2M2. *Front. Plant Sci.* 4, 459.

Rabilloud, T., Chevallet, M., Luche, S., and Lelong, C. (2010). Two-dimensional gel electrophoresis in proteomics: Past, present and future. *J. Proteomics* 73, 2064-2077.

Rappsilber, J. (2011). The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* 173, 530-540.

Rast, A., Heinz, S., and Nickelsen, J. (2015). Biogenesis of thylakoid membranes. *Biochim. Biophys. Acta* 1847, 821-830.

Reiland, S., Messerli, G., Baerenfaller, K., Gerrits, B., Endler, A., Grossmann, J., et al. (2009). Large-scale Arabidopsis phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. *Plant Physiol.* 150, 889-903.

Rengstl, B., Knoppová, J., Komenda, J., and Nickelsen, J. (2013). Characterization of a *Synechocystis* double mutant lacking the photosystem II assembly factors YCF48 and Sll0933. *Planta* 237, 471-480.

Rexroth, S., Wong, C.C.L., Park, J.H., Yates, J.R., III, Barry, B.A. (2007). An activated glutamate residue identified in Photosystem II at the interface between the manganese-stabilizing subunit and the D2 polypeptide. *J. Biol. Chem.* 282, 27802-27809.

Rinalducci, S., Larsen, M.R., Mohammed, S., and Zolla, L. (2006). Novel protein phosphorylation site identification in spinach stroma membranes by titanium dioxide microcolumns and tandem mass spectrometry. *J. Proteome Res.* 5, 973-982.

Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L., Beck, M., Schmidt, A., et al. (2008). Identification of cross-linked peptides from large sequence databases. *Nat. Methods* 5, 315-318.

Rintamäki, E., Salonen, M., Suoranta, U.M., Carlberg, I., Andersson, B., and Aro, E.-M. (1997). Phosphorylation of Light-harvesting Complex II and Photosystem II core proteins shows different irradiance-dependent regulation *in vivo*: Application of phosphothreonine antibodies to analysis of thylakoid phosphoproteins. *J. Biol. Chem.* 272, 30476-30482.

Rokka, A., Suorsa, M., Saleem, A., Battchikova, N., and Aro, E.M. (2005). Synthesis and assembly of thylakoid protein complexes: multiple assembly steps of photosystem II. *Biochem. J.* 388, 159-168.

Romanowska, E., Wasilewska, W., Fristedt, R., Vener, A.V., and Zienkiewicz, M. (2012). Phosphorylation of PSII proteins in maize thylakoids in the presence of Pb ions. *J. Plant Physiol.* 169, 345-352.

Roose, J.L., Kashino, Y., and Pakrasi, H.B. (2007). The PsbQ protein defines cyanobacterial Photosystem II complexes with highest activity and stability. *Proc. Natl. Acad. Sci. USA* 104, 2548-2553.

Roose, J.L., and Pakrasi, H.B. (2004). Evidence that D1 processing is required for manganese binding and extrinsic protein assembly into Photosystem II. *J. Biol. Chem.* 279, 45417-45422.

Roose, J.L., and Pakrasi, H.B. (2008). The Psb27 protein facilitates manganese cluster assembly in Photosystem II. *J. Biol. Chem.* 283, 4044-4050.

Ross, P.L., Huang, Y.L.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., et al. (2004). Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 3, 1154-1169.

Rowland, J.G., Simon, W.J., Nishiyama, Y., and Slabas, A.R. (2010). Differential proteomic analysis using iTRAQ reveals changes in thylakoids associated with Photosystem II-acquired thermotolerance in *Synechocystis* sp. PCC 6803. *Proteomics* 10, 1917-1929.

Ryan, C.M., Souda, P., Bassilian, S., Ujwal, R., Zhang, J., Abramson, J., et al. (2010). Post-translational modifications of integral membrane proteins resolved by top-down Fourier transform mass spectrometry with collisionally activated dissociation. *Mol. Cell. Proteomics* 9, 791-803.

Samol, I., Shapiguzov, A., Ingelsson, B., Fucile, G., Crevecoeur, M., Vener, A.V., et al. (2012). Identification of a Photosystem II phosphatase involved in light acclimation in *Arabidopsis*. *Plant Cell* 24, 2596-2609.

Sander, J., Nowaczyk, M., Buchta, J., Dau, H., Vass, I., Deák, Z., et al. (2010). Functional characterization and quantification of the alternative PsbA copies in *Thermosynechococcus elongatus* and their role in photoprotection. *J. Biol. Chem.* 29851-29856.

Schönberg, A., and Baginsky, S. (2012). Signal integration by chloroplast phosphorylation networks: an update. *Front. Plant Sci.* 3, 256.

Schottkowski, M., Gkalympoudis, S., Tzekova, N., Stelljes, C., Schuenemann, D., Ankele, E., et al. (2009). Interaction of the periplasmic PratA factor and the PsbA (D1) protein during biogenesis of Photosystem II in *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 284, 1813-1819.

Seebacher, J., Mallick, P., Zhang, N., Eddes, J., Aebersold, R., and Gelb, M.H. (2006). Protein cross-linking analysis using mass spectrometry, isotope-coded cross-linkers, and integrated computational data processing. *J. Proteome Res.* 5, 2270-2282.

Seidler, A. (1996). Intermolecular and intramolecular interactions of the 33-kDa protein in photosystem II. *Eur. J. Biochem.* 242, 485-490.

Sharma, J., Panico, M., Barber, J., and Morris, H.R. (1997a). Characterization of the low molecular weight photosystem II reaction center subunits and their light-induced modifications by mass spectrometry. *J. Biol. Chem.* 272, 3935-3943.

Sharma, J., Panico, M., Barber, J., and Morris, H.R. (1997b). Purification and determination of intact molecular mass by electrospray ionization mass spectrometry of the photosystem II reaction center subunits. *J. Biol. Chem.* 272, 33153-33157.

Sharma, J., Panico, M., Shipton, C.A., Nilsson, F., Morris, H.R., and Barber, J. (1997c). Primary structure characterization of the photosystem II D1 and D2 subunits. *J. Biol. Chem.* 272, 33158-33166.

Shaw, J.B., Li, W., Holden, D.D., Zhang, Y., Griep-Raming, J., Fellers, R.T., et al. (2013). Complete protein characterization using top-down mass spectrometry and ultraviolet photodissociation. *J. Am. Chem. Soc.* 135, 12646-12651.

Shevela, D., and Messinger, J. (2013). Studying the oxidation of water to molecular oxygen in photosynthetic and artificial systems by time-resolved membrane-inlet mass spectrometry. *Front. Plant Sci.* 4, 473.

Shi, L.-X., Hall, M., Funk, C., and Schröder, W. (2012). Photosystem II, a growing complex: Updates on newly discovered components and low molecular mass proteins. *Biochim. Biophys. Acta* 1817, 13-25.

Shi, L.-X., Lorković, Z.J., Oelmüller, R., and Schröder, W.P. (2000). The low molecular mass PsbW protein is involved in the stabilization of the dimeric Photosystem II complex in *Arabidopsis thaliana*. *J. Biol. Chem.* 275, 37945-37950.

Shi, L.-X., and Schröder, W.P. (1997). Compositional and topological studies of the PsbW protein in spinach thylakoid membrane. *Photosynth. Res.* 53, 45-53.

Silva, J.C., Gorenstein, M.V., Li, G.Z., Vissers, J.P.C., and Geromanos, S.J. (2006). Absolute quantification of proteins by LCMS^{E} - A virtue of parallel MS acquisition. *Mol. Cell. Proteomics* 5, 144-156.

Sinz, A. (2014). The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert Rev. Proteomics* 11, 733-743.

Steinback, K.E., Bose, S., and Kyle, D.J. (1982). Phosphorylation of the light-harvesting chlorophyll-protein regulates excitation energy distribution between Photosystem II and Photosystem I. *Arch. Biochem. Biophys.* 216, 356-361.

Stöckel, J., Jacobs, J.M., Elvitigala, T.R., Liberton, M., Welsh, E.A., Polpitiya, A.D., et al. (2011). Diurnal rhythms result in significant changes in the cellular protein complement in the cyanobacterium *Cyanothece* 51142. *PLoS ONE* 6, e16680.

Sugimoto, I., and Takahashi, Y. (2003). Evidence that the PsbK polypeptide is associated with the Photosystem II core antenna complex CP43. *J. Biol. Chem.* 278, 45004-45010.

Sugiura, M., Iwai, E., Hayashi, H., and Boussac, A. (2010a). Differences in the interactions between the subunits of Photosystem II dependent on D1 protein variants in the thermophilic cyanobacterium *Thermosynechococcus elongatus*. *J. Biol. Chem.* 285, 30008-30018.

Sugiura, M., Kato, Y., Takahashi, R., Suzuki, H., Watanabe, T., Noguchi, T., et al. (2010b). Energetics in Photosystem II from *Thermosynechococcus elongatus* with a D1 protein encoded by either the psbA$_1$ or psbA$_3$ gene. *Biochim. Biophys. Acta* 1797, 1491-1499.

Sugiura, M., Koyama, K., Umena, Y., Kawakami, K., Shen, J.-R., Kamiya, N., et al. (2013). Evidence for an unprecedented histidine hydroxyl modification on D2-His336 in Photosystem II of *Thermosynechoccocus vulcanus* and *Thermosynechoccocus elongatus*. *Biochemistry* 52, 9426-9431.

Sugiura, M., and Boussac, A. (2014). Some Photosystem II properties depending on the D1 protein variants in Thermosynechococcus elongatus. *Biochim. Biophys. Acta* 1837, 1427-1434.

Sugiyama, N., Nakagami, H., Mochida, K., Daudi, A., Tomita, M., Shirasu, K., et al. (2008). Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in *Arabidopsis*. *Mol. Syst. Biol.* 4, 193.

Swaisgood, H., and Natake, M. (1973). Effect of carboxyl group modification on some of the enzymatic properties of L-glutamate dehydrogenase. *J. Biochem.* 74, 77-86.

Takahashi, M., Shiraishi, T., and Asada, K. (1988). COOH-terminal residues of D1 and the 44 kDa CPa-2 at spinach photosystem II core complex. *FEBS Lett*. 240, 6-8.

Takamoto, K., and Chance, M.R. (2006). Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. *Annu. Rev. Biophys. Biomol. Struct.* 35, 251-276.

Takasaka, K., Iwai, M., Umena, Y., Kawakami, K., Ohmori, Y., Ikeuchi, M., et al. (2010) Structural and functional studies on Ycf12 (Psb30) and PsbZ-deletion mutants from a thermophilic cyanobacterium. *Biochim. Biophys. Acta* 1797, 278-284.

Tal, O., Trabelcy, B., Gerchman, Y., and Adir, N. (2014). Investigation of phycobilisome subunit interaction interfaces by coupled cross-linking and mass spectrometry. *J. Biol. Chem.* 289, 33084-33097.

Telfer, A., Bishop, S.M., Phillips, D., and Barber, J. (1994). Isolated photosynthetic reaction center of Photosystem II as a sensitizer for the formation of singlet oxygen: Detection and quantum yield determination using a chemical trapping technique. *J. Biol. Chem.* 269, 13244-13253.

Thangaraj, B., Ryan, C.M., Souda, P., Krause, K., Faull, K.F., Weber, A.P.M., et al. (2010). Data-directed top-down Fourier-transform mass spectrometry of a large integral membrane protein complex: Photosystem II from *Galdieria sulphuraria*. *Proteomics* 10, 3644-3656.

Thelen, J.J., and Peck, S.C. (2007). Quantitative proteomics in plants: Choices in abundance. *Plant Cell* 19, 3339-3346.

Thidholm, E., Lindström, V., Tissier, C., Robinson, C., Schröder, W.P., and Funk, C. (2002). Novel approach reveals localisation and assembly pathway of the PsbS and PsbW proteins into the photosystem II dimer. *FEBS Lett*. 513, 217-222.

Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., et al. (2003). Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* 75, 1895-1904.

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 *Synechopystis* 6803. *Plant Cell* 16, 2164-2175.

Tikhonov, A.N. (2015). Induction events and short-term regulation of electron transport in chloroplasts: an overview. *Photosynth. Res.* 125, 65-94. doi: 10.1007/s11120-015-0094-0

Tikkanen, M., and Aro, E.-M. (2014). Integrative regulatory network of plant thylakoid energy transduction. *Trends Plant Sci.* 19, 10-17.

Tikkanen, M., Grieco, M., Kangasjarvi, S., and Aro, E.-M. (2010). Thylakoid protein phosphorylation in higher plant chloroplasts optimizes electron transfer under fluctuating light. *Plant Physiol.* 152, 723-735.

Tikkanen, M., Nurmi, M., Kangasjarvi, S., and Aro, E.-M. (2008a). Core protein phosphorylation facilitates the repair of photodamaged photosystem II at high light. *Biochim. Biophys. Acta* 1777, 1432-1437.

Tikkanen, M., Nurmi, M., Suorsa, M., Danielsson, R., Mamedov, F., Styring, S., et al. (2008b). Phosphorylation-dependent regulation of excitation energy distribution between the two photosystems in higher plants. *Biochim. Biophys. Acta* 1777, 425-432.

Tomo, T., Enami, I., and Satoh, K. (1993). Orientation and nearest-neighbor analysis of *psb*I gene product in the photosystem II reaction center complex using bifunctional cross-linkers. *FEBS Lett*. 323, 15-18.

Tsiotis, G., Walz, T., Spyridaki, A., Lustig, A., Engel, A., and Ghanotakis, D. (1996). Tubular crystals of a Photosystem II core complex. *J. Mol. Biol.* 259, 241-248.

Turkina, M.V., Kargul, J., Blanco-Rivero, A., Villarejo, A., Barber, J., and Vener, A.V. (2006). Environmentally modulated phosphoproteome of photosynthetic membranes in the green alga *Chlamydomonas reinhardtii*. *Mol. Cell Proteomics* 5, 1412-1425.

Tyystjärvi, E. (2013). Photoinhibition of Photosystem II. *Int. Rev. Cell. Mol. Biol.* 300, 243-303.

Ujihara, T., Sakurai, I., Mizusawa, N., and Wada, H. (2008). A method for analyzing lipid-modified proteins with mass spectrometry. *Anal. Biochem.* 374, 429-431.

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

Vainonen, J.P., Hansson, M., and Vener, A.V. (2005). STN8 protein kinase in *Arabidopsis thaliana* is specific in phosphorylation of Photosystem II core proteins. *J. Biol. Chem.* 280, 33679-33686.

Vassiliev, S., Zaraiskaya, T., and Bruce, D. (2012). Exploring the energetics of water permeation in photosystem II by multiple steered molecular dynamics simulations. *Biochim. Biophys. Acta* 1817, 1671-1678.

Vener, A.V. (2007). Environmentally modulated phosphorylation and dynamics of proteins in photosynthetic membranes. *Biochim. Biophys. Acta* 1767, 449-457.

Vener, A.V., Harms, A., Sussman, M.R., and Vierstra, R.D. (2001). Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of *Arabidopsis thaliana*. *J. Biol. Chem.* 276, 6959-6966.

Walleczek, J., Martin, T., Redl, B., Stofflermeilicke, M., and Stoffler, G. (1989). Comparative cross-linking study on the 50S ribosomal subunit from *Escherichia coli*. *Biochemistry* 28, 4099-4105.

Wang, L.W., and Chance, M.R. (2011). Structural mass spectrometry of proteins using hydroxyl radical based protein footprinting. *Anal. Chem.* 83, 7234-7241.

Wegener, K.M., Bennewitz, S., Oelmüller, R., and Pakrasi, H.B. (2011). The Psb32 protein aids in repairing photodamaged Photosystem II in the cyanobacterium *Synechocystis* 6803. *Mol. Plant* 4, 1052-1061.

Wegener, K.M., Nagarajan, A., and Pakrasi, H.B. (2015). An atypical *psbA* gene encodes a sentinel D1 protein to form a physiologically relevant inactive Photosystem II complex in cyanobacteria. *J. Biol. Chem.* 290, 3764-3774.

Wegener, K.M., Welsh, E.A., Thornton, L.E., Keren, N., Jacobs, J.M., Hixson, K.K., et al. (2008). High sensitivity proteomics assisted discovery of a novel operon involved in the assembly of Photosystem II, a membrane protein complex. *J. Biol. Chem.* 283, 27829-27837.

Weisbrod, C.R., Chavez, J.D., Eng, J.K., Yang, L., Zheng, C., and Bruce, J.E. (2013). *In vivo* protein interaction network identified with a novel real-time cross-linked peptide identification strategy. *J. Proteome Res.* 12, 1569-1579.

Welkie, D., Zhang, X., Markillie, M.L., Taylor, R., Orr, G., Jacobs, J., et al. (2014). Transcriptomic and proteomic dynamics in the metabolism of a diazotrophic cyanobacterium, *Cyanothece* sp. PCC 7822 during a diurnal light-dark cycle. *BMC Genomics* 15, 1185.

Wen, J.Z., Zhang, H., Gross, M.L., and Blankenship, R.E. (2009). Membrane orientation of the FMO antenna protein from *Chlorobaculum tepidum* as determined by mass spectrometry-based footprinting. *Proc. Natl. Acad. Sci. USA* 106, 6134-6139.

Wetz, K., and Habermehl, K.-O. (1979). Topographical studies on poliovirus capsid proteins by chemical modification and cross-linking with bifunctional reagents. *J. Gen. Virol.* 44, 525-534.

Whitelegge, J.P. (2013). Integral membrane proteins and bilayer proteomics. *Anal. Chem.* 85, 2558-2568.

Xu H., Freitas, M.A. (2009). MassMatrix: A database search program for rapid characterization of proteins and peptides from tandem mass spectrometry data. *Proteomics* 9, 1548–1555.

Yang, B., Wu, Y.-J., Zhu, M., Fan, S.-B., Lin, J., Zhang, K., et al. (2012). Identification of cross-linked peptides from complex samples. *Nat. Methods* 9, 904-909.

Yang, M.-k., Yang, Y.-h., Chen, Z., Zhang, J., Lin, Y., Wang, Y., et al. (2014). Proteogenomic analysis and global discovery of posttranslational modifications in prokaryotes. *Proc. Natl. Acad. Sci. USA.* 111, E5633-E5642.

Yao, Danny C. I., Brune, D.C., Vavilin, D., and Vermaas, W.F.J. (2012a). Photosystem II component lifetimes in the cyanobacterium *Synechocystis* sp. strain PCC 6803: Small Cab-like proteins stabilize biosynthesis intermediates and affect early steps in chlorophyll synthesis. *J. Biol. Chem.* 287, 682-692.

Yao, D.C.I., Brune, D.C., and Vermaas, W.F.J. (2012b). Lifetimes of photosystem I and II proteins in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett*. 586, 169-173.

Zak, E., Norling, B., Maitra, R., Huang, F., Andersson, B., and Pakrasi, H.B. (2001). The initial steps of biogenesis of cyanobacterial photosystems occur in plasma membranes. *Proc. Natl. Acad. Sci. USA* 98, 13443-13448.

Zhang, H., Liu, H., Blankenship, R.E., and Gross, M.L. (2016). Isotope-encoded carboxyl group footprinting for mass spectrometry-based protein conformational studies. *J. Am. Soc. Mass Spectrom.* 27, 178-181.

Zhang, H., Liu, H., Niedzwiedzki, D.M., Prado, M., Jiang, J., Gross, M.L., et al. (2014). Molecular mechanism of photoactivation and structural location of the cyanobacterial orange carotenoid protein. *Biochemistry* 53, 13-19.

Zheleva, D., Sharma, J., Panico, M., Morris, H.R., and Barber, J. (1998). Isolation and characterization of monomeric and dimeric CP47-reaction center Photosystem II complexes. *J. Biol. Chem.* 273, 16122-16127.

Chapter 2

Structural analysis of Photosystem II: Mass spectrometry-based cross-linking study shows that the Psb28 protein binds to cytochrome b_{559}

This chapter was adapted from:

Weisz, D.A., Liu, H., Zhang, H., Thangapandian, S., Tajkhorshid, E., Gross, M.L., Pakrasi, H.B. Structural analysis of Photosystem II: Mass spectrometry-based cross-linking study shows that the Psb28 protein binds to cytochrome b_{559}. *Manuscript under review.*

Summary

Photosystem II (PSII), a large pigment protein complex, undergoes rapid turnover under natural conditions. The assembly of PSII, with its numerous subunit proteins and cofactors, must occur without oxidative damage prior to becoming fully functional. Psb28, the only cytoplasmic extrinsic protein in PSII, protects the RC47 assembly intermediate of PSII and assists its efficient conversion into functional PSII. Its role is particularly important under stress conditions when PSII damage occurs frequently. Psb28 is not found, however, in any PSII crystal structure, and its structural location has remained unknown. In this study, we used chemical cross-linking combined with mass spectrometry to capture the transient interaction of Psb28 with PSII. We detected three cross-links between Psb28 and the α- and β-subunits of cytochrome b_{559}, an essential component of the PSII reaction-center complex. These distance restraints enable us to position Psb28 on the cytosolic surface of PSII directly above cytochrome *b*559 in close proximity to the Q_B site. Protein-protein docking results also support Psb28 binding in this region. Determination of the Psb28 binding site and other biochemical evidence allow us to propose several mechanisms by which Psb28 exerts its protective effect on the RC47 intermediate. This study also shows that isotopically-labeled cross-linking with the "mass tags" selection criteria allows confident identification of more cross-linked peptides in PSII than has been previously reported. This approach thus holds promise to identify other transient protein-protein interactions in membrane protein complexes.

Introduction

Photosystem II (PSII) is a multi-subunit pigment-protein complex embedded in the thylakoid membranes of cyanobacteria, algae, and plants. PSII uses light energy to oxidize water to molecular oxygen, simultaneously reducing plastoquinone. Active PSII consists of approximately 20 protein subunits and multiple light-harvesting and redox-active cofactors (Bricker et al., 2012; Suga et al., 2015).

Owing to demanding electron-transfer chemistry, PSII undergoes frequent oxidative damage, necessitating a complex cycle of repair and re-assembly (reviewed in Weisz et al., 2016). The assembly occurs stepwise via multiple transient intermediate complexes that are difficult to study owing to their low abundance, relatively short lifetimes, and heterogeneity. Crystal structures of the active complex from thermophilic cyanobacteria are available (Umena et al., 2011; Hellmich et al., 2014; Suga et al., 2015), but they do not capture the transient interactions of the various accessory proteins that bind at other stages of the life-cycle. Nevertheless, significant progress has been made in characterizing these intermediates through complementary use of genetic modification, biochemical purification and analysis, and mass spectrometry (Nickelsen and Rengstl, 2013; Bricker et al., 2015; Heinz et al., 2016). Many accessory proteins bind exclusively to inactive subcomplexes at various stages of the PSII life-cycle, and these proteins are of interest for their regulatory roles in PSII assembly and/or repair (Nickelsen and Rengstl., 2013; Järvi et al., 2015; Heinz et al., 2016). Because the crystal structures of assembly intermediate complexes have not been determined, the binding sites of these accessory proteins are largely unknown. An exception is the Psb27 accessory protein, whose binding site on the lumenal surface of PSII was identified using chemical cross-linking and mass spectrometry (Liu et al., 2011a; Cormann et al., 2016). This knowledge complements functional studies of Psb27 and provides mechanistic insight

into how Psb27 protects PSII assembly intermediates from prematurely acquiring oxygen evolution capability (Roose and Pakrasi, 2004; Liu et al., 2011b; Komenda et al., 2012; Liu et al., 2013). The lack of comparable binding-site information for other PSII accessory proteins limits our understanding of their functions.

The Psb28 accessory protein was first identified in a proteomic analysis of PSII from *Synechocystis* sp. PCC 6803 (hereafter referred to as *Synechocystis* 6803) (Kashino et al., 2002a), and is likely the unidentified 12-kDa protein observed earlier in a cyanobacterial PSII preparation (Ikeuchi et al., 1995). Psb28 is found across a wide range of cyanobacteria, and has homologues in higher plants as well, but little is known about its binding site and function (Mabbitt et al., 2014). Dobáková et al. (2009) found that Psb28 binds mainly to the monomeric CP43-less PSII assembly intermediate, referred to as "RC47", a finding confirmed subsequently (Boehm et al., 2012; Sakata et al., 2013). Although deletion of the *psb28* gene had little physiological effect under typical growth conditions (Dobáková et al., 2009; Sakata et al., 2013), its absence impairs PSII recovery after photodamage under high light conditions, especially at high temperature (Sakata et al., 2013), suggesting a critical role of this protein during increased PSII turnover. Psb28 is the only known extrinsic cyanobacterial PSII protein found on the cytoplasmic surface of PSII (Dobáková et al., 2009; Nickelsen and Rengstl, 2013; Järvi et al., 2015; Heinz et al., 2016), and it may interact with the cytoplasmic surface of CP47 based on the co-migration of these two proteins during 2D-Blue-Native-SDS-PAGE (Dobáková et al., 2009).

To identify the Psb28 binding site in PSII, we used an isotopically labeled chemical cross-linker and mass spectrometry (MS) to capture covalently the protein-protein interactions in PSII (Rappsilber, 2011; Sinz, 2014; Leitner et al., 2016). The isotope encoding allowed us to use an on-the-fly precursor-ion selection mechanism that takes advantage of the isotopic "fingerprint" of

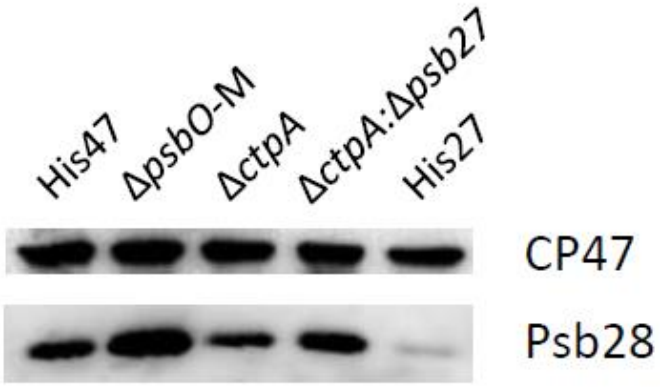
cross-linked peptides to facilitate identification (Petrotchenko et al., 2014). Remarkably, Psb28 cross-linked to PsbE and PsbF, the α- and β-subunits, respectively, of cytochrome b_{559} (Cyt b_{559}), two of the five polypeptides that comprise the core "reaction center" of PSII (Nanba and Satoh, 1987; Ikeuchi and Inoue, 1988). Using these cross-links as distance restraints, and supported by protein-protein docking results, we generated a Psb28 binding model that shows Psb28 binding on the cytoplasmic surface of PSII directly above the heme of Cyt b_{559} and in close proximity to the Q_B site. On the basis of the binding model, we propose several mechanisms for the protective effect of Psb28 on the RC47 assembly intermediate of PSII.

Results

Quantification of Psb28 levels in PSII from several mutant strains

Only approximately one in seven PSII complexes purified from the CP47-His-tagged strain ("His47") contain Psb28 (Fig. 1C, D). This result is reasonable given that Psb28 is found mainly in the RC47 intermediate, and RC47 accounts for approximately 10% of PSII complexes present in the wild type cell (Dobáková et al., 2009; Boehm et al., 2012). To increase the chances of detecting Psb28 cross-linking products, we screened PSII complexes for elevated Psb28 content from several mutant strains that accumulate assembly intermediates (Fig. 1A). PSII from the Δ*psbO*-His47 strain contains the most Psb28 and, therefore, was used in the subsequent experiments. The deletion of *psbO* in this strain prevents PSII dimerization (Liu et al., 2014), allowing the monomeric PSII fraction (Δ*O*-M) to be isolated by glycerol-gradient ultracentrifugation (see "Materials and Methods") and used in cross-linking. PsbO binds on the lumenal surface of PSII and stabilizes the manganese cluster, the site of water oxidation (Bricker et al., 2012). The Δ*psbO* strain still assembles active water-splitting PSII and grows photoautotrophically, but the complexes are more susceptible to photodamage and require more

A.



B.

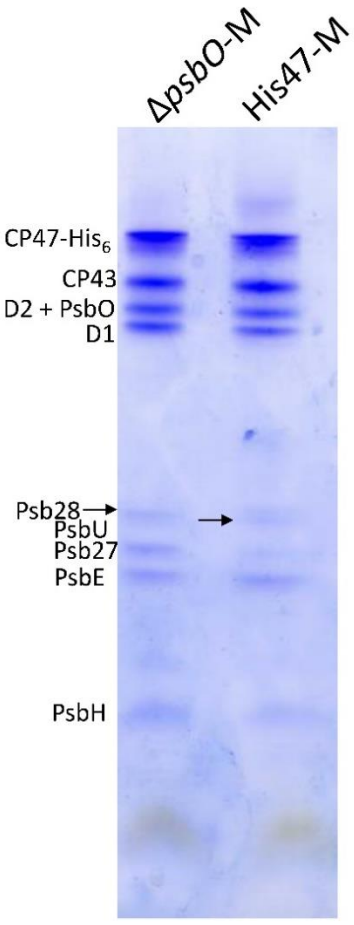
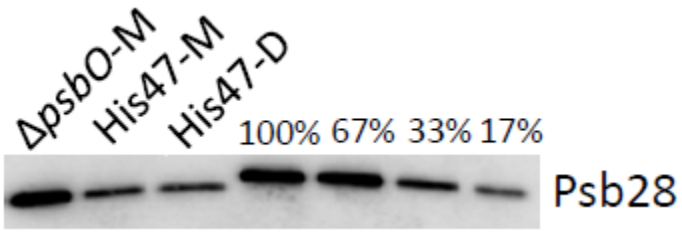


Figure 1. (legend follows)

C.



D.

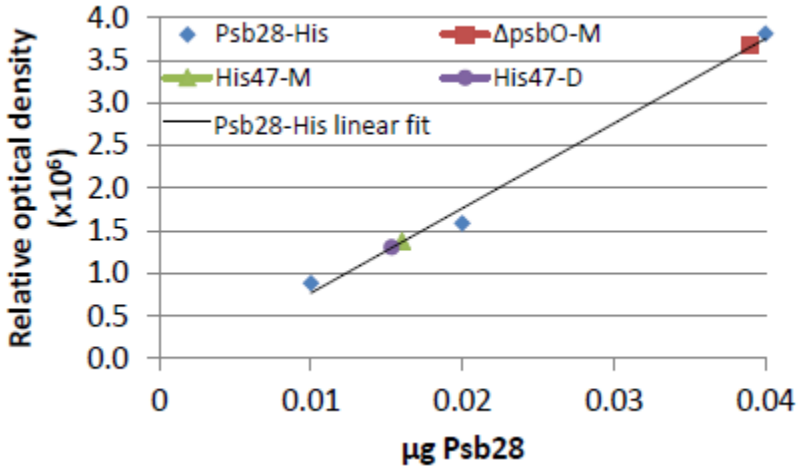


Figure 1. Screening of mutant strains for elevated Psb28 content, characterization of the Δ*psbO* PSII monomer, and quantification of Psb28 content in Δ*psbO* PSII.

A. Comparison of Psb28 content in PSII complexes from the His47 and several different mutant strains.

B. Protein gel comparing PSII subunit composition of the His47 and Δ*psbO* PSII monomers.

C. Immunoblot comparing Psb28 content in the Δ*psbO* PSII monomer, the His47 PSII monomer and dimer, and a dilution series of Psb28-His purified from *E. coli*, respectively.

D. Calibration curve for quantification of Psb28 content in Δ*psbO* PSII and His47 PSII using the *E. coli* –purified Psb28 as a standard.

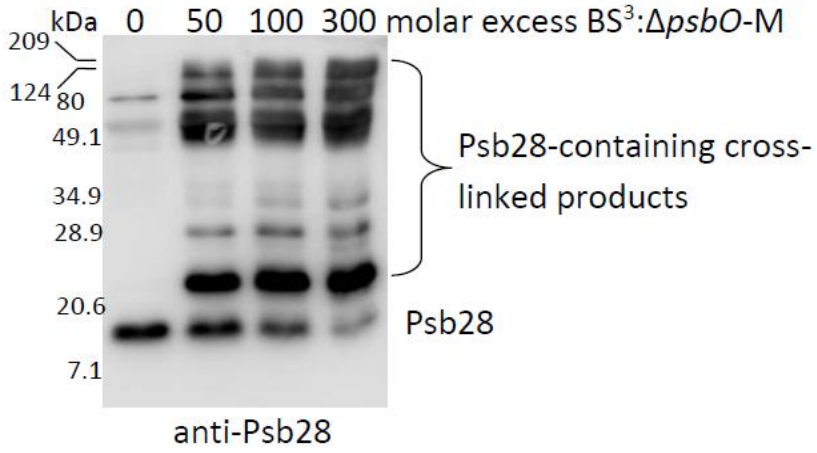


Figure 2. Immunoblot comparing Δ*psbO*-PSII before and after cross-linking with 50, 100, and 300 molar excess BS^3:PSII.

frequent repair than wild-type cells (Burnap et al., 1992). The elevated level of Psb28 implies an increased steady-state concentration of the RC47 complex in this mutant (Dobáková et al., 2009; Boehm et al., 2012; Sakata et al., 2013), which is expected as RC47 is a key intermediate in the PSII repair cycle (Nickelsen and Rengstl, 2013).

We quantified the Psb28 content in ΔO-M as well as the His47 PSII monomer and dimer by calibrating with known quantities of recombinantly expressed and purified Psb28-His (Fig. 1C, D). Approximately one in three Δ*O*-M PSII complexes contained Psb28 whereas approximately one in seven His47-M and His47-D PSII complexes contained Psb28. Other studies found Psb28 mainly in the monomeric RC47 complex (Dobáková et al., 2009; Boehm et al., 2012; Sakata et al., 2013), whereas we detected approximately equal levels in monomeric and dimeric His47-PSII. A dimeric RC47 complex containing Psb28 was observed, however, in two previous studies (Boehm et al., 2012; Sakata et al., 2013) but at lower levels than the monomeric RC47. This complex may form during disassembly, like the dimeric Psb27-containing species (Grasse et al., 2011), implying

a role for Psb28 during disassembly as previously proposed (Nowaczyk et al., 2012; Sakata et al., 2013). The increased level of dimeric Psb28-containing complexes in our His47-PSII preparation compared to that in two previous studies might have arisen from differences in the purification conditions.

Cross-link data analysis

We subjected the Δ*O*-M sample to cross-linking with a 1:1 mixture of BS^3 cross-linker labeled with 12 deuteriums, and its unlabeled analog. This cross-linker can modify primary amines and thus targets lysine residues and protein N-termini (Paramelle et al., 2013), and it gives an isotopic doublet in the mass spectrometer to facilitate identification of cross links. We evaluated the yield over a relative molar concentration of cross-linker:PSII between 50 and 300. We analyzed the cross-linked products by SDS-PAGE and immunoblotting using anti-Psb28 antibodies (Fig. 2) and found multiple Psb28-containing bands after cross-linking, but not in the non-linked control sample, indicating successful cross linking.

We chose the sample that was cross-linked with 300-molar excess linker:PSII for analysis by MS because it showed the highest number and intensity of Psb28-containing cross-linked bands. We identified 18 cross-links and 22 mono-linked sites (for which only one side of the cross-linker reacted with a protein) (Tables S1 and S2). The isotopic doublet "fingerprint" of this cross-linker (Petrotchenko et al., 2014) enabled highly confident identification of cross-links, removing much of the ambiguity that often hinders cross-link analysis (Sinz, 2014) (described further in Supplementary Results and Discussion).

Using the PSII crystal structure data, we could measure the C_α-C_α distance between the linked residues for thirteen of the linked peptides (Table S2). These distances served as a control to assess the quality of the cross-linking reaction and data analysis. We used 30 Å as an upper C_α-C_α

distance threshold for cross-links to be considered consistent with the crystal structure (Hellmich et al., 2014), a typical value used for this cross-linker (Walzthoeni et al., 2011; Herzog et al., 2012; Fischer et al., 2013; Merkley et al., 2014; Shi et al., 2015; Zelter et al., 2015). The distance distribution of identified cross-links (Fig. S1) was similar to the distributions typically found for this cross-linker (Rappsilber, 2011; Herzog et al., 2012; Leitner et al., 2012; Zheng et al., 2013; Merkley et al., 2014) and indicates that our cross-linking data are of high quality and not the result of perturbation of the native structure of the complex (described further in Supplementary Results and Discussion).

Identification of Psb28 cross-links

We identified three inter-protein cross-links containing Psb28: Psb28-K8—PsbE-S2 (Fig. 3A), Psb28-K8—PsbF-A2 (Fig. 3B), and Psb28-A2—PsbF-A2 (Fig. S2). PsbE-S2, PsbF-A2, and Psb28-A2 are susceptible to modification by BS^3 because they are the N-terminal residues of the mature form of these proteins following *in vivo* cleavage of the N-terminal methionine residues (confirmed by intact-mass measurement of each of these proteins, see Fig. S3). We identified a cross-link between PsbE-S2 and PsbF-A2 as well, indicative of the proximity of these two residues and their ability to each form a cross-link with Psb28-K8. The mass spectra for these cross-links—which were not detected in a control, non-linked sample—display the doublet feature characteristic of the isotopic mixture of cross-linker used. Nearly all major fragments in the product-ion spectra match predicted peptide fragments, resulting in highly confident identification of these cross-links (see Fig. 3 and Supplementary Results and Discussion). We also identified the mono-linked peptide corresponding to each of these cross-linked residues (see Table S1), a valuable cross-check since for a given modifiable residue, a mono-link should form more readily than a cross-link. We selected the Δ*psb*O-His47 strain for our cross-linking-MS experiments because it accumulates

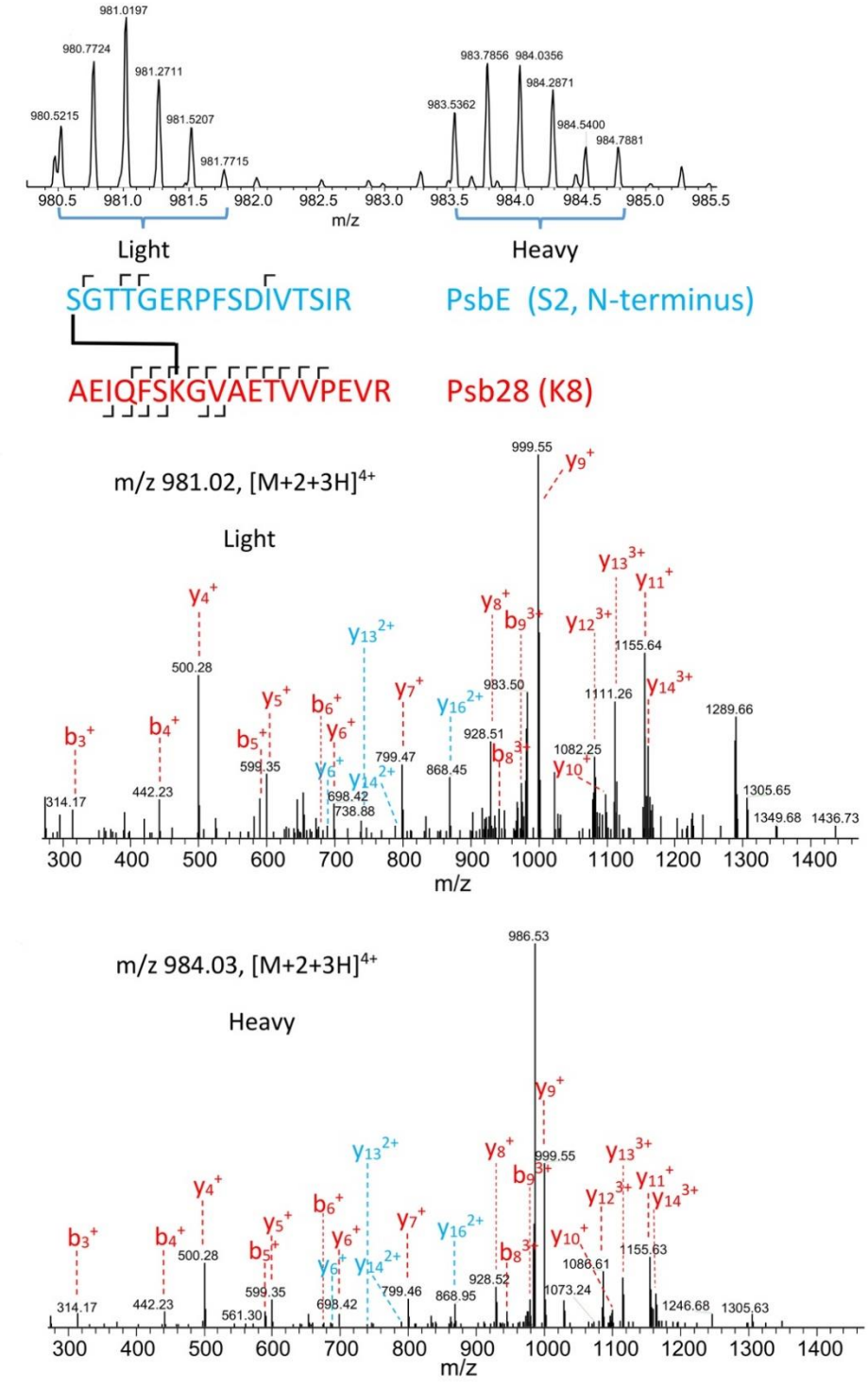


Figure 3A. (legend follows)

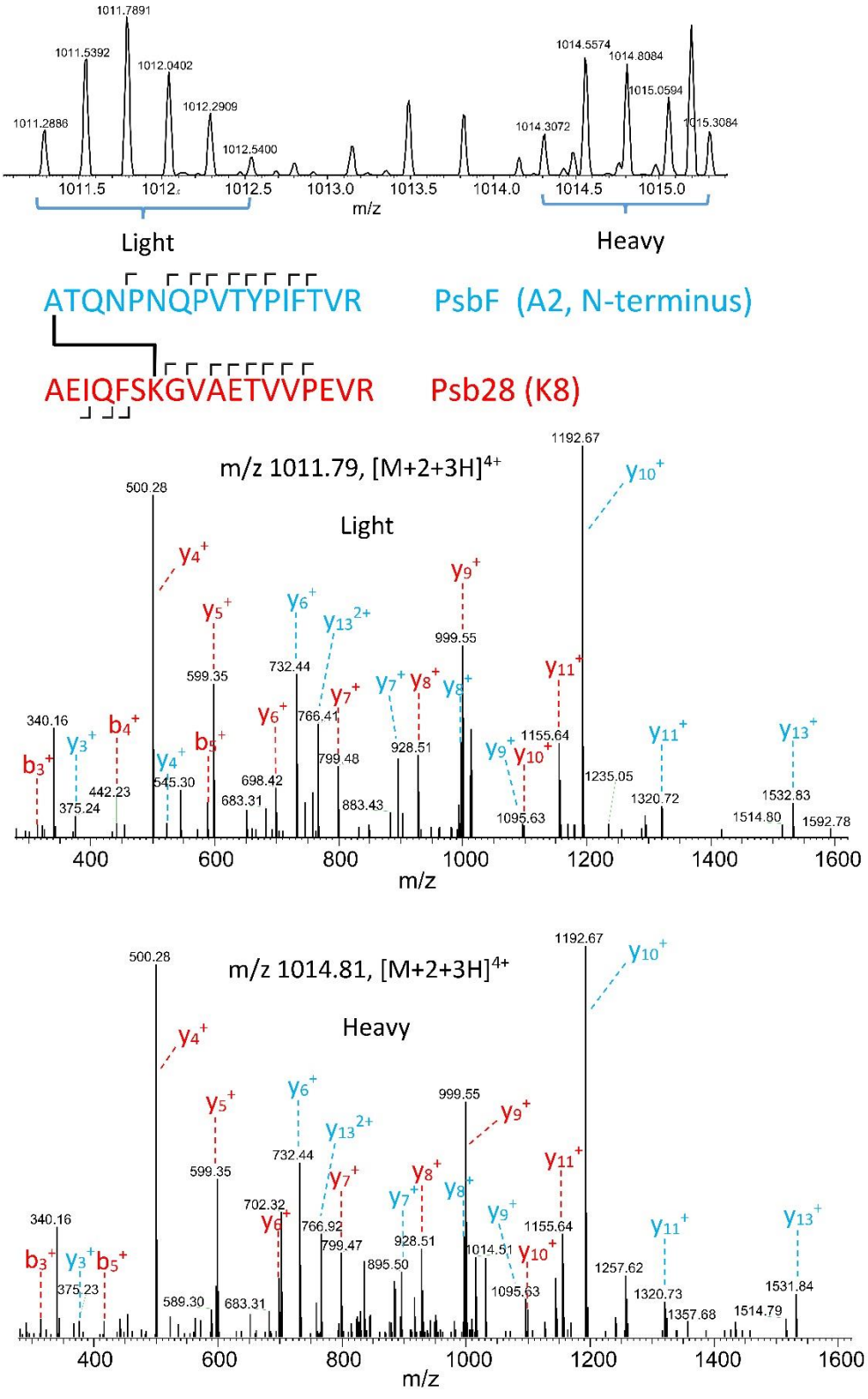


Figure 3B. (legend follows)

Figure 3. Mass spectrometric data showing a cross-link between (A) Psb28-K8 and PsbE-S2 and (B) Psb28-K8 and PsbF-A2. The MS (precursor-ion) spectra for each cross-link is shown at the top of each figure; it displays the isotopic "fingerprint" of a real cross-link: a doublet of peaks of equal intensity, shifted from each other by 12 Da. The two lower spectra in both figures are the MS/MS (product-ion) spectra of the light and heavy form of the cross-linked peptide, generated after fragmentation of that peptide by higher energy collisional dissociation (HCD). The MS/MS spectra also display the "fingerprint" of a real cross-link: they are essentially identical despite originating from different precursor ions, and some cross-linker-containing peaks are shifted by +12 Da in the spectrum of the heavy form. The black notches in between residues in the peptide sequence correspond to fragment ions that were observed in the MS/MS spectra of both the light and heavy forms (also labeled on the spectra themselves), which taken together, permit confident assignment of the spectra to that particular cross-linked peptide. For additional discussion see Supplementary Results and Discussion.

more Psb28 than the other strains we screened. After identifying the Psb28 cross-links in this strain, we searched the MS-1 spectra of cross-linked His47 samples to check for their presence there as well, to rule out the possibility that the Δ*psb*O mutation led to a non-natural interaction between Psb28 and PsbE and PsbF. We did detect the cross-links in the MS-1 spectra of His47 samples (Figs. S4-S6), demonstrating that the association of Psb28 to PsbE and PsbF is not an artifact of *psb*O deletion in the Δ*psb*O-His47 strain we used.

Psb28-PSII docking

To examine possible binding modes of Psb28 to PSII by a method complementary to cross-linking, we performed protein docking with DOT 2.0 (for details see "Materials and Methods"). The top 4000 docked conformations clustered into four spatial groups. One of these clusters ("Cluster 1"), which contains 1282 conformations, is on the cytosolic surface of PSII. Because Psb28 binds to the cytosolic surface of PSII (Dobáková et al., 2009), we restricted our analysis to the Cluster 1 conformations. Further examination of the top conformations showed that they are all localized near the surface above the N-termini of PsbE and PsbF (Fig. S7 and Table S3), supporting the cross-linking results by showing this area is a favorable spot for Psb28 binding.

Discussion

Psb28 dissociates from PSII before Psb27 attachment

In contrast to the other screened PSII mutants, His27-PSII contained almost no Psb28 (Fig. 1A). The His27 strain is the only one in which the polyhistidine tag is not on the CP47 protein, but on Psb27 instead. The absence of Psb28 in His27-PSII indicates that Psb28 dissociates from PSII before, or concomitant with, binding of Psb27. This observation is consistent with the presence of Psb28 mainly in the RC47 complex (Dobáková et al., 2009; Boehm et al., 2012; Sakata et al., 2013), to which CP43, and, therefore, Psb27, have not yet bound (Liu et al., 2011b; Komenda et al., 2012; Liu et al., 2013). Nowaczyk *et al*. (Nowaczyk et al., 2012) detected both Psb27 and Psb28 in a subcomplex purified from the Δ*psbJ* mutant of *T. elongatus*. This finding implies that the absence of PsbJ delays Psb28 dissociation, or it may simply be a result of subcomplex heterogeneity in the preparation in which those two proteins were detected.

The Psb28 binding site

We found cross-links between Psb28-K8 and the N-termini of PsbE and PsbF, and between Psb28-A2 and the N-terminus of PsbF. Protein-protein docking between Psb28 and the RC47 subunits of PSII provided independent support that the cytosolic surface in the region above PsbE and PsbF is a favorable site for Psb28 binding. The first eleven residues of PsbF are not resolved in the PSII crystal structure, and two of the three Psb28 cross-links are located at the PsbF N-terminus. Therefore, we used the earliest resolved residue, PsbF-Y13, as a proxy for PsbF-A2 in positioning Psb28 above the cytosolic surface of PsbE and PsbF in a manner consistent with the cross-links (Fig. 4). We positioned Psb28 such that the distance between cross-linked residues is within the 30 Å threshold mentioned above. In our model, the cross-link distances are in fact considerably lower (15, 10, and 14 Å for the PsbE—Psb28, PsbF—Psb28-K8, and PsbF—Psb28-

A2 cross-links, respectively), indicating that these cross-links are self-consistent and represent a close interaction between Psb28, PsbE, and PsbF. In addition to PsbE and PsbF, the model shows Psb28 sharing a binding interface with CP47, PsbX, and PsbY, and binding directly above the heme of cytochrome b_{559} and in close proximity to Q_B (discussed below). The close association of Psb28 and CP47 in our model is consistent with the observation that the two proteins form a

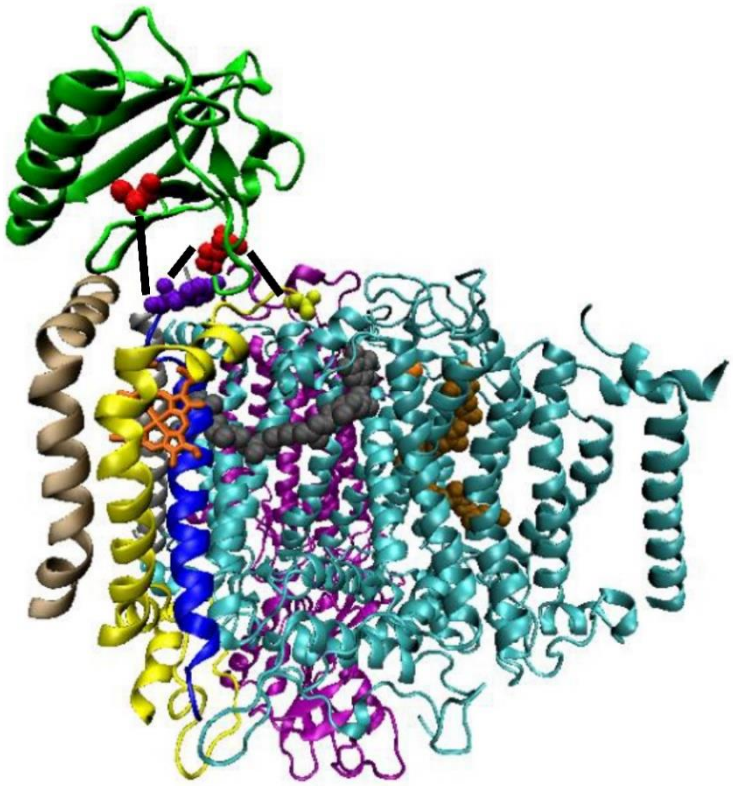


Figure 4. Binding model of Psb28 to RC47 based on cross-linking results. Psb28 binds on the cytosolic surface of PSII sharing an interface with PsbE, PsbF, CP47, PsbX and PsbY. Black lines are shown connecting cross-linked residues (PsbF-Y13 is shown as a proxy for the unresolved PsbF-A2). Cyan- RC47 PSII proteins; yellow, PsbE; blue, PsbF; green, Psb28; purple, CP47; silver, PsbX; tan, PsbY; red, Psb28-A2 (upper red residue) and Psb28-K8 (lower red residue); yellow, PsbE-A2; violet, PsbF-Y13; orange, heme, non-heme Fe; ochre, Q_A; gray, Q_B.

stable complex with each other during native-gel electrophoresis (Dobáková et al., 2009; Boehm et al., 2012).

The energetics of the top docked Psb28 conformations (Table S3) show that electrostatic interactions dominate hydrophobic interactions as the major force stabilizing the binding of Psb28 to RC47. This is consistent with the fact that Psb28 can be removed from PSII membranes by washing with 1 M $CaCl_2$, 0.1 M Na_2CO_3, or 0.1 M NaOH (Dobáková et al., 2009). Previous work identified two Psb28 surface cavities that are mostly formed by highly conserved residues, and it was speculated that these cavities may be important for binding of Psb28 to PSII (Yang et al., 2011). In our model, one of these cavities is present at the interface of Psb28 and RC47 (Cavity 4 in Yang et al., 2011).

Functional implications

Deletion of the *psb28* gene in *Synechocystis* results in little or no phenotypic difference under typical growth conditions (Dobáková et al., 2009; Sakata et al., 2013). Under high-light conditions, however, cells lacking Psb28 reassemble functional PSII after damage at a reduced rate compared to wild-type (Sakata et al., 2013). For the *dgdA* mutant, in which the RC47 intermediate is longer-lived owing to impaired CP43 attachment, absence of Psb28 results in a further decrease in reassembly of functional PSII after damage (Sakata et al., 2013). Combined with the knowledge that Psb28 binds primarily to the RC47 intermediate (Dobáková et al., 2009; Boehm et al., 2012; Sakata et al., 2013), these results show that Psb28 protects the RC47 intermediate, helping ensure it does not undergo damage before conversion into functional PSII.

Under typical growth conditions, the relatively low light levels and short-lived presence of RC47 make the protective role of Psb28 not critical (Dobáková et al., 2009; Sakata et al., 2013). However, under stress conditions such as high light and/or high temperature—when damage occurs more frequently—or under conditions when the RC47 intermediate is longer-lived and is more likely to incur damage, the protective role of Psb28 becomes critical (Sakata et al., 2013), and in its absence, optimal conversion of RC47 into functional PSII is prevented.

Cross-linking shows the Psb28 binding site is on the cytosolic surface above the Cyt b_{559} subunits PsbE and PsbF. Our model positions Psb28 approximately 9-12 Å directly above the heme of Cyt b_{559} (16-20 Å from the heme Fe) and 23-27 Å from the redox-active aromatic ring of the quinone Q_B (Fig. 4). With this information, we propose three ways in which Psb28 might exert its protective effect on RC47:

1) Cyt b_{559} plays a photo-protective role as part of a secondary electron-transfer pathway within PSII. Cyt b_{559} can serve as an electron acceptor from Q_B or the plastoquinone pool under acceptor-side stress conditions (*e.g.*, high light), and as an electron donor ultimately for P_{680}^+ under donor-side stress conditions (*e.g.*, before manganese cluster assembly) (Shinopoulos and Brudvig, 2012; Chu and Chiu, 2016). Both of these roles limit oxidative damage to PSII by shortening the lifetime of reactive electron-transfer intermediates during stress conditions in which the primary redox pathway is disrupted. Cyt b_{559} exists in different redox forms that permit multiple redox roles under different conditions (Shinopoulos and Brudvig, 2012; Chu and Chiu, 2016). Interconversion between these forms is presumably governed by the surrounding protein environment. Despite extensive investigation, the specific conditions that give rise to each form remain unclear. Given the proximity of Psb28 to the heme of Cyt b_{559}, Psb28 may tune the heme's redox potential in RC47, which suffers from both donor-side stress (unassembled manganese cluster) and acceptor-

side stress (altered Q_B site) (Carpenter et al., 1990; Diner et al., 1991; Boehm et al., 2012). Studying the redox properties of Cyt b_{559} after purification of the RC47 complex (Boehm et al., 2012) from a strain containing and a strain lacking Psb28 could determine its effect on Cyt b_{559}.

2) Experimental evidence suggests that one mechanism of protection of RC47 is an alteration in its acceptor-side electron-transfer dynamics—specifically impaired electron transfer from Q_A to Q_B. This impaired transfer is part of an established mechanism that protects partially assembled PSII complexes from damage (Johnson et al., 1995; Komenda and Masojidek, 1998; Shinopoulos and Brudvig, 2012; Cardona et al., 2012). Boehm and co-workers (2012) purified the RC47 complex from a strain of *Synechocystis* 6803 that lacks the CP43 protein, and, therefore, assembles PSII complexes only up to the RC47 stage. They found that the RC47 complex contains an intact electron-transfer chain from the primary donor P_{680} through reduction of Q_A, but that the final transfer step from Q_A to Q_B is blocked. An earlier study on intact cells from the CP43-deletion strain of *Synechocystis* 6803 also demonstrates poor electron transfer from Q_A to Q_B (Diner et al., 1991). Sharply decreased binding of DCMU to the Q_B site suggests a structural alteration around the Q_B site (Carpenter et al., 1990; Diner et al., 1991), and this could explain the impaired electron transfer. By binding in close proximity to the Q_B site, Psb28 may contribute to the structural perturbations in this area that protects RC47 by blocking electron transfer from Q_A to Q_B. Studying the electron transfer properties of RC47 purified from a strain containing and a strain lacking Psb28 could probe further this possibility.

3) Psb28 may prevent binding of the phycobilisome to the RC47 complex. The phycobilisome is a large light-harvesting antenna protein complex that binds on the cytoplasmic surface of PSII and supplies it with excitation energy (Mullineaux, 2008; Komenda et al., 2012). PSII assembly intermediates, however, do not contain an active water-splitting complex, and accepting excitation

energy in this state leads to photodamage (Shinopoulos and Brudvig, 2012). Phycobilisome binding during assembly, therefore, would be undesirable. Recently, an electron microscopic structure of the phycobilisome associated with PSII (Chang et al., 2015) shows that the phycobilisome attaches to PSII by apparently inserting into a cavity on the cytoplasmic surface formed in part by PsbE and PsbF (Chang et al., 2015). Our results position Psb28 just beside this cavity (Fig. S8). Psb28 binding may prevent phycobilisome attachment, minimizing delivery of harmful excitation energy to the RC47 intermediate. Indeed, fluorescence emission spectra from the CP43-deletion strain of *Synechocystis* 6803 (which accumulates RC47) do not indicate coupled energy transfer from the phycobilisome to assembled PSII subunits (Shimada et al., 2008). Similar experiments in a strain containing and a strain lacking Psb28 could determine if Psb28 binding is responsible for this effect.

Combining evidence from previous research with the current results positions us to refine a model summarizing the steps in PSII assembly (Fig. 5). Psb28 binding occurs upon formation of the RC47 complex, and has been positioned in Fig. 5 in a location consistent with our cross-linking results. Given the significant interface between Psb28 and CP47, PsbX, and PsbY, Psb28 would unlikely bind stably before the RC47 stage. Indeed, Psb28 has not been detected in the preceding assembly intermediate, the RC complex (Komenda et al., 2008; Knoppová et al., 2014). Psb28 dissociation occurs in the next step, prior to binding of Psb27, as indicated by the lack of Psb28 in His27-PSII.

In conclusion, we determined the structural location of Psb28, the only cytoplasmic extrinsic protein in PSII, in close proximity to the N-terminal domain of the Cyt b_{559} protein. With the structural information we have gained, we can propose several mechanisms by which Psb28 might exert its protective effect on the PSII subcomplex to which it binds, helping ensure optimal

conversion of this subcomplex into functional PSII. This study also shows that the combination of isotopically-labeled cross-linking with the "mass tags" selection criteria can identify low-abundance cross-linked peptides originating from a large membrane protein complex. Such cross-links provide useful distance restraints for positioning subunits not present in crystal structures of protein complexes. This approach thus holds promise for identifying other transient protein-protein interactions, a burgeoning area of interest in biology.

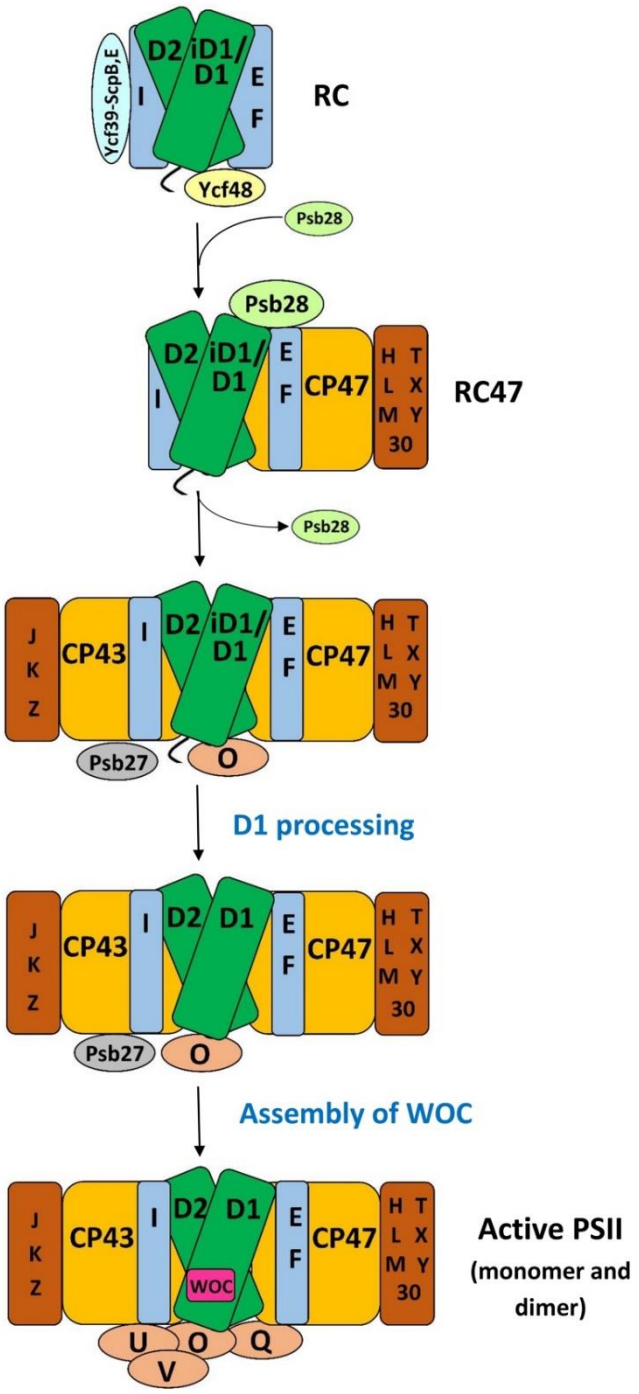


Figure 5. A schematic of the PSII assembly process. This model summarizing previous knowledge (Nickelsen and Rengstl, 2013; Liu et al., 2013; Knoppová et al., 2014; Heinz et al., 2016; Weisz et al., 2016) and our current results. Individual letters (H, L, T, etc.) and "30" represent the PsbH, PsbL, PsbT, Psb30, etc. PSII subunits. iD1 represents the partially-processed "intermediate" form of the D1 protein (Knoppová et al., 2014). RC, "reaction center" complex. WOC, "water-oxidizing complex."

Materials and Methods

Generation of the Psb28-His construct

To generate the Psb28-His construct for expression in *E. coli*, the Psb28 gene was amplified from *Synechocystis* 6803 genomic DNA using the following primers: forward- 5′-GACATATGGCTGAAATTCAATTTTCCAAGG-3′; reverse- 5′-TCGGATCCTTAATGGTGATGGTGATGGTGATGGTGTTCAGATTTGGAAAAACCTAAGCCATTTTCTGCGCCG-3′. The forward primer contains an *Nde*I restriction site and the reverse primer contains a *Bam*HI restriction site. In the reverse primer, sequence encoding an 8x His-tag was inserted before the stop codon. The PCR fragment was cloned into the pET21a vector, and then transformed into the BL21(DE3) strain of *E. coli* for expression of Psb28-His.

Purification of Psb28-His from E. coli

The BL21(DE3) strain of *E. coli* containing the pET21a-Psb28-His vector was grown at 37 °C with constant shaking in 200 mL of LB medium containing 50 μg/mL ampicillin. When the culture reached OD_{600} = 0.6, expression of Psb28-His was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The culture was harvested 3 h after induction by centrifugation at 3000 g for 15 min. Cells were resuspended by vortexing in 10 mL lysis buffer (20 mM Tris, pH 7.9, 100 mM NaCl, 10 mM imidazole), with 1 μg/mL DNase (Sigma, St. Louis, MO), 100 μL protease inhibitor cocktail (Sigma), and 1 mg lysozyme. After allowing cells to rock gently on ice for 15 min, they were lysed by probe sonication on ice for 3 min at 50% duty cycle, then centrifuged at 31,000*g* for 15 min at 4 °C. The supernatant ("lysate") was applied onto a 1 mL pre-packed HisTrap FF column (GE Healthcare, Little Chalfont, Buckinghamshire, UK), pre-equilibrated with 10 mL lysis buffer, by dropwise manual injection. After collecting the flow-through, the column was washed with 10 mL lysis buffer. Elution was

by manual injection of 10 mL elution buffer (lysis buffer containing 200 mM imidazole). Ten 1-mL fractions were collected, and analyzed by SDS-PAGE for Psb28 content. The most concentrated fraction contained 13 mg/mL Psb28 and was used for the subsequent experiments. A portion of this fraction was submitted to Cocalico Biologicals (Reamstown, PA) for generation of anti-Psb28 antisera in rabbit.

Cyanobacterial culture and PSII purification

Generation of the Δ*psb*O-His47 (Liu et al., 2014), Δ*ctp*A-His47 (Liu et al., 2011b), Δ*ctp*A:Δ*psb27*-His47 (Liu et al., 2011a), and His27 (Liu et al., 2011b) strains has been reported previously. The HT3 (His47) strain was a kind gift from Dr. Terry Bricker (Louisiana State University, Baton Rouge, LA) (Bricker et al., 1998). Cyanobacterial strains were grown in BG11 medium at 30 °C under 30 μmol photons $m^{-2} \cdot s^{-1}$. The growth media were supplemented with 10 μg/mL spectinomycin and 5 μg/mL kanamycin (Δ*psb*O-His47); 5 mM glucose, 10 μM 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU), 5 μg/mL kanamycin, and 3 μg/mL erythromycin (Δ*ctp*A-His47); 5 mM glucose, 10 μM DCMU, 10 μg/mL chloramphenicol, and 5 μg/mL kanamycin (Δ*ctp*A:Δ*psb27*-His47); and 5 μg/ml gentamicin (His27). Histidine-tagged PSII complexes were purified as described previously (Kashino et al., 2002a) with minor modifications. PSII samples were stored in 25% glycerol, 10 mM $MgCl_2$, 5 mM $CaCl_2$, 50 mM MES buffer (pH 6.0) (RB buffer). After FPLC purification of Δ*psb*O-His47-PSII and His47-PSII, these complexes were purified further by glycerol gradient ultracentrifugation (180,000 *g*, 18 h, 4 °C) using a 5-30% linear glycerol gradient. The gradient was made by using stock solutions of 5% and 30% glycerol in RB buffer. PSII samples in RB buffer were diluted to 5% glycerol before ultracentrifugation, and 50 or 100 μg Chl *a* containing samples were loaded in each tube. PSII monomer and dimer represented by their bands were recovered after ultracentrifugation and were

concentrated using Vivaspin 500 centrifugal concentrators (100,000 molecular weight cutoff) (Vivaproducts, Littleton, MA). Protein gel electrophoresis and immunoblotting were performed as described previously (Kashino et al., 2001; Kashino et al., 2002b). Immunoblot quantification analysis was performed with the ImageQuant TL software package (GE Healthcare, Pittsburgh, PA).

Cross-linking and proteolytic digestion

PSII samples were cross-linked using unlabeled BS^3 (Thermo Fisher Pierce, Rockford, IL) or a 1:1 mixture of unlabeled BS^3 and BS^3 labeled with 12 deuteriums (Creative Molecules, Victoria, Canada). The cross-linker was dissolved in a stock solution of RB buffer. This solution was added to aliquots of PSII samples containing 1 or 2 μg Chl *a* (in RB buffer after ultracentrifugation and concentration) at a cross-linker:PSII molar ratio of 50-300:1, and incubated in the dark at room temperature for 50 min. The reaction was quenched by addition of a solution containing Tris (pH 7.5) to 50 mM final concentration. Salt and detergent were removed from the cross-linked samples by precipitation using the 2D-Clean-Up kit (GE Healthcare). Cross-linked protein pellets were resuspended in 20 μL 8 M urea, 50 mM ammonium bicarbonate. A two-step digestion with lysyl endopeptidase (LysC) and trypsin was used to increase final yield of tryptic peptides (Leitner et al., 2012). Lysyl endopeptidase (MS-grade, Wako Chemicals USA, Richmond, VA) was added to the protein sample at a 1:50 LysC:protein ratio (weight:weight) and the samples were incubated at 37 °C for 2 h. After 2 h, samples were diluted 1:8 in 50 mM ammonium bicarbonate to a final urea concentration of 1 M. Trypsin (Sigma, St. Louis, MO) was added to the samples at a 1:25 trypsin:protein ratio (weight:weight) and the samples were incubated at 37 °C overnight. After digestion, samples were acidified to a final concentration of 1% formic acid and analyzed by LC-MS/MS.

LC-MS/MS

Aliquots (5-µL) of the peptide samples were loaded onto an Ultimate 3000 Nano LC system (Thermo Scientific Dionex, Sunnyvale, CA) attached in-line to a Q Exactive Plus mass spectrometer (Thermo Fisher, Waltham, MA). Peptide samples initially flowed through a guard column (Acclaim PepMap100, 100 µm × 2 cm, C18, 5 µm, 100 Å; Thermo Scientific Dionex) in Solvent A (water with 0.1% formic acid) and were separated on a C_{18} reversed-phase column (Magic, 0.075 mm ×150 mm, 5 μm, 120 Å, Michrom Bioresources, Inc., Auburn, CA) packed in house, at 4.5 µL/min. Peptides were eluted using a linear 90-min gradient from 5-95% solvent B (80% acetonitrile, 20% water, 0.1% formic acid), followed by a 10-min hold at 95% solvent B. Eluted samples were flowed directly into the mass spectrometer via a PicoView Nanospray Source (PV550, New Objective, Inc., Woburn, MA) with a spray voltage of 1.8 kV. The instrument was operated in positive-ion mode with a scan range from *m/z* 380-1500. Full mass spectra were acquired at 70,000 resolving power for ions of *m/z* 200, with automatic gain control set at 3×10^6 ions and a maximal injection time of 200 ms. Data-dependent product-ion spectra were acquired at 17,500 mass resolving power for ions of *m/z* 200, with automatic gain control set at 1×10^5 ions and a maximal injection time of 100 ms. The top 15 precursor ions were fragmented by HCD with an isolation window of 3.0 *m/z* and normalized collision energy of 30%. For samples run with the "mass tags" feature enabled (Petrotchenko et al., 2014), only precursor ion pairs with a mass-to-charge shift selection criteria of ±6.03762, ±4.02508, ±3.01881, ±2.41505, ±2.01254, or ±1.72503 Da, corresponding to peptide charge states of 2-7, respectively, were selected for fragmentation. Up to the ten top precursor ions matching these criteria were fragmented by HCD with an isolation window of 3.0 *m/z* and normalized collision energy of 25%. Non-cross-linked samples were analyzed excluding charge states other than 2-7, and each cross-linked sample was analyzed in

duplicate or triplicate, in runs that excluded charge states other than either 2-7, 3-7, and 4-7. Data presented in Figs. S2 and the upper spectra in Figs. S4-S6 were acquired on an LTQ-Orbitrap XL (Thermo Fisher, Waltham, MA), as described previously (Liu et al., 2014), with the "mass tags" feature enabled as described above.

Cross-link data analysis

The raw LC-MS/MS files were searched for proteins in the *Synechocystis* proteome using PEAKS (ver. 7.0, Bioinformatics Solutions, Inc., Waterloo, ON, Canada). Peptides were identified with a 0.1% false discovery rate. Mono-linked peptides were identified by PEAKS, as well as by ICC-CLASS described below, with the light and heavy forms of the cross-linker as user-defined modifications (+156.0079 and +168.1540 Da, respectively). The identified proteins served as the database for cross-link detection using the ICC-CLASS software suite (Petrotchenko et al., 2010; Petrotchenko et al., 2014). Raw LC-MS/MS files were converted to MGF format using Proteome Discoverer (Thermo Scientific) and to Xtract text files using the Xcalibur software File Converter Tool (Thermo Scientific). Isotopic doublets were detected in the MS data with the ICC-CLASS programs, and cross-link candidates were identified with DXMSMSMatchESI (Petrotchenko et al., 2010; Petrotchenko et al., 2014). Search settings were as follows: Crosslinker- DSS; M_{ip}- 137.06025; Mclrest1- 0 Da; Mclrest2- 0 Da; Digest sites- KR; Including CL site- No; Missed Digests- up to 2; CL sites- K; Dead-end peptides only- No; Intra-peptide only- No; Systematic error- 0 ppm; Mass tolerance- ± 10 ppm; Precursor tolerance- ± 10 ppm; Retention time tolerance- ± 2 s; Fragments tolerance- ± 10-500 ppm. After identification of cross-links using these search conditions, the search tolerance was expanded to identify cross-links that may have been missed initially by including CL site- Yes; Missed digests- up to 4. Product-ion spectra of candidate cross-links were inspected manually to verify the identification, using the predicted fragment ion masses

in DXMSMSMatchESI or the PeptidesCL program (www.creativemolecules.com). Additional cross-link data analysis was performed using Protein Prospector (Trnka et al., 2014; Chalkley et al., 2014).

Intact protein LC-MS

His47-PSII samples were purified as described above, and a sample containing 1.2 μg Chl *a* was precipitated using the 2D Clean-up Kit (GE Healthcare). The sample was resuspended in 250 μL 90% formic acid and spun down to remove insoluble material. The supernatant was filtered, and a portion (containing ~1.6 μg protein) was injected onto a PLRP-S column (Agilent, 2.1x150mm, 300Å pore size) controlled by an Agilent 1200 HPLC. Buffer A, Buffer B, and the LC gradient were described previously (Thangaraj et al., 2010). Eluted proteins were analyzed online by a Maxis quadrupole-time-of-flight 4G mass spectrometer (Bruker Daltonics, Billerica, MA).

Protein-protein docking

We used the DOT 2.0 docking program (Roberts et al., 2013) to predict the binding interface of PSII (PDB ID: 4PJ0) (Hellmich et al., 2014)) and Psb28 (PDB ID: 2KVO) (Yang et al., 2011) with the 3D structures obtained from Protein Data Bank. Because Psb28 is known to bind to the RC47 subcomplex (Dobáková et al., 2009; Boehm et al., 2012; Sakata et al., 2013), the subunits that are not components of this assembly intermediate (CP43, PsbJ, PsbK, PsbO, PsbU, PsbV, Psb30, and PsbZ) (Boehm et al., 2012) were removed from the structure prior to docking. The DOT program uses convolution methods to perform a systematic rigid-body translational and rotational search. The van der Waals and electrostatic energies of both the molecules were mapped onto grids to predict the interactions between two macromolecules. One of the two molecules was rotated and translated around the other to predict the energetically favorable complexes. The Psb28

molecule was selected as the moving molecule and used a cubic grid of 192 Å on a side with 1 Å spacing between points for the translational search. A set of 54,000 rotational orientations, with a reduced over-sampling at the corners of the cubical grid, provided a resolution of 6° for the rotational search. Combined translational and rotational search resulted in over 382 billion configurations (192^3 * 54,000) of the PSII and Psb28 molecules. The top 4,000 docked conformations based on calculated interaction energy were saved and analyzed. These complexes formed four distinct clusters of which three were located in the membrane-spanning regions or on the lumenal surface of PSII, making these clusters physiologically irrelevant; they were not considered further. The cluster localized at the cytosolic surface of PSII (Cluster 1) was composed of 1,282 Psb28 conformations. The top 100 conformations from Cluster 1 were investigated further as described in the Results section. Analysis of the docked conformations was conducted in VMD (Humphrey et al., 1996).

References

Boehm M., et al. (2012). Subunit composition of CP43-less photosystem II complexes of *Synechocystis* sp. PCC 6803: implications for the assembly and repair of photosystem II. *Phil. Trans. R. Soc. B.* 367(1608):3444-3454.

Bricker T.M., Morvant J., Masri N., Sutton H.M., Frankel L.K. (1998). Isolation of a highly active Photosystem II preparation from *Synechocystis* 6803 using a histidine-tagged mutant of CP 47. *Biochim. Biophys. Acta* 1409(1):50–57.

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

Bricker T.M., Mummadisetti M.P., Frankel L.K. (2015). Recent advances in the use of mass spectrometry to examine structure/function relationships in photosystem II. *J. Photochem. Photobiol. B* 152:227-246.

Buncherd H., Roseboom W., de Koning L.J., de Koster C.G., de Jong L. (2014). A gas phase cleavage reaction of cross-linked peptides for protein complex topology studies by peptide fragment fingerprinting from large sequence database. *J. Proteomics* 108:65-77.

Burnap R.L., Shen J.-R., Jursinic P.A., Inoue Y., Sherman L.A. (1992). Oxygen yield and thermoluminescence characteristics of a cyanobacterium lacking the manganese-stabilizing protein of Photosystem II. *Biochemistry* 31(32):7404-7410.

Cardona T., Sedoud A., Cox N., Rutherford A.W. (2012). Charge separation in Photosystem II: A comparative and evolutionary overview. *Biochim. Biophys. Acta* 1817(1):26-43.

Carpenter S.D., Charité J., Eggers B., Vermaas W. (1990). Characterization of site-directed and hybrid *psbC* mutants of *Synechocystis* 6803. *Current Research in Photosynthesis*, ed. Baltcheffsky M. (Kluwer Academic Publishers, Dordrecht), pp. 359-362.

Chalkley R.J., Trnka M.J., Michael N., Baker, P.R. (2014). Identifying cross-linked peptides using Protein Prospector. 62nd ASMS Conference on Mass Spectrometry and Allied Topics, Baltimore, MD, June 15.

Chang L., et al. (2015). Structural organization of an intact phycobilisome and its association with photosystem II. *Cell Res.* 25(6):726-737.

Chu H.-A., Chiu H.-F. (2016). The roles of cytochrome b_{559} in assembly and photoprotection of Photosystem II revealed by site-directed mutagenesis studies. *Front. Plant. Sci.* 6:1261.

Chu F., Mahrus S., Craik C.S., Burlingame A.L. (2006). Isotope-coded and affinity-tagged cross-linking (ICATXL): An efficient strategy to probe protein interaction surfaces. *J. Am. Chem. Soc.* 128(32):10362-10363.

Cormann K.U., Möller M., Nowaczyk M.M. (2016). Critical assessment of protein cross-linking and molecular docking: An updated model for the interaction between Photosystem II and Psb27. *Front. Plant. Sci.* 7:157.

Diner B.A., Petrouleas V., Wendoloski .J.J. (1991). The iron-quinone electron-acceptor complex of Photosystem-II. *Physiol. Plant* 81(3):423-436.

Dobáková M., Sobotka R., Tichý M., Komenda J. (2009). 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(2):1076-1086.

Fischer L., Chen Z.A., Rappsilber J. (2013). Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. *J. Proteomics* 88:120-128.

Fritzsche R., Ihling C.H., Götze M., Sinz A. (2012). Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun. Mass Sp.* 26(6):653-658.

Grasse N., et al. (2011). Role of novel dimeric Photosystem II (PSII)-Psb27 protein complex in PSII repair. *J. Biol. Chem.* 286(34):29548-29555.

Heinz S., Liauw P., Nickelsen J., Nowaczyk M. (2016). Analysis of photosystem II biogenesis in cyanobacteria. *Biochim. Biophys. Acta,* 1857(3):274-287.

Hellmich J., et al. (2014). Native-like Photosystem II superstructure at 2.44 Å resolution through detergent extraction from the protein crystal. *Structure* 22(11):1607-1615.

Herzog F., et al. (2012). Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* 337(6100):1348-1352.

Humphrey W., Dalke A., Schulten K. (1996). VMD - Visual Molecular Dynamics. *J. Molec. Graphics* 14(1):33-38.

Ikeuchi M., Inoue Y. (1988). 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(7):1233-1239.

Ikeuchi M., Inoue Y., Vermaas W. (1995). Characterization of Photosystem II subunits from the cyanobacterium *Synechocystis* sp. PCC 6803. *Photosynthesis: From light to biosphere: Proceedings of the Xth International Photosynthesis Congress, Montpellier, France, 20-25 August 1995*, ed Mathis P(Kluwer, Dordrecht), pp. 297-300.

Järvi S., Suorsa M., Aro, E.-M. (2015). Photosystem II repair in plant chloroplasts - Regulation, assisting proteins and shared components with photosystem II biogenesis. *Biochim Biophys Acta* 1847(9):900-909.

Johnson G.N., Rutherford A.W., Krieger A. (1995). A change in the midpoint potential of the quinone Q_A in Photosystem II associated with photoactivation of oxygen evolution. *Biochim. Biophys. Acta* 1229(2):202-207.

Kahraman A., et al. (2013). Cross-link guided molecular modeling with ROSETTA. *PLoS ONE* 8(9):e73411.

Kang S., et al. (2009). Synthesis of biotin-tagged chemical cross-linkers and their applications for mass spectrometry. *Rapid Commun. Mass Spectrom.* 23(11):1719-1726.

Kashino Y., Koike H., Satoh K. (2001). An improved sodium dodecyl sulfate-polyacrylamide gel electrophoresis system for the analysis of membrane protein complexes. *Electrophoresis* 22(6):1004–1007.

Kashino Y., et al. (2002). Low-molecular-mass polypeptide components of a Photosystem II preparation from the thermophilic cyanobacterium *Thermosynechococcus vulcanus*. *Plant Cell Physiol.* 43(11):1366–1373.

Kashino Y., et al. (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(25):8004-8012.

Knoppová J., et al. (2014). Discovery of a chlorophyll binding protein complex involved in the early steps of Photosystem II assembly in *Synechocystis*. *Plant Cell* 26(3):1200-1212.

Komenda J., Masojidek J. (1998). The effect of Photosystem II inhibitors DCMU and BNT on the high-light induced D1 turnover in two cyanobacterial strains Synechocystis PCC 6803 and Synechococcus PCC 7942. *Photosynth. Res.* 57(2):193-202.

Komenda J., et al. (2008). 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(33):22390-22399.

Komenda J., et al. (2012). The Psb27 assembly factor binds to the CP43 complex of Photosystem II in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant. Physiol.* 158(1):476-486.

Leitner A., et al. (2012). Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol. Cell. Proteomics* 11(3):M111.014126. DOI: 10.1074/mcp.M111.014126.

Leitner A., Faini M., Stengel F., Aebersold R. (2016). Crosslinking and mass spectrometry: An integrated technology to understand the structure and function of molecular machines. *Trends Biochem. Sci.* 41(1):20-32.

Liu H., Huang R.Y.-C., Chen J., Gross M.L., Pakrasi H.B. (2011a). Psb27, a transiently associated protein, binds to the chlorophyll binding protein CP43 in photosystem II assembly intermediates. *Proc. Natl. Acad. Sci. U.S.A.* 108(45):18536-18541.

Liu H., Roose J.L., Cameron J.C., Pakrasi H.B. (2011b). A genetically tagged Psb27 protein allows purification of two consecutive photosystem II (PSII) assembly intermediates in *Synechocystis* 6803, a cyanobacterium. *J. Biol. Chem.* 286(28):24865-24871.

Liu H., Chen J., Huang R.Y.-C., Weisz D., Gross M.L., Pakrasi H.B. (2013). Mass spectrometry-based footprinting reveals structural dynamics of Loop E of the chlorophyll-binding protein CP43 during Photosystem II assembly in the cyanobacterium *Synechocystis* 6803. *J. Biol. Chem.* 288(20):14212-14220.

Liu H., et al. (2014). MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. *Proc. Natl. Acad. Sci. U.S.A.* 111(12):4638-4643.

Mabbitt P.D., Wilbanks S.M., Eaton-Rye J.J. (2014). Structure and function of the hydrophilic Photosystem II assembly proteins: Psb27, Psb28, and Ycf48. *Plant Physiol. Biochem.* 81:96-107.

Merkley E.D., et al. (2014). Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci.* 23(6):747-759.

Müller D.R., et al. (2001). Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal. Chem.* 73(9):1927–1934.

Mullineaux C.W. (2008). Phycobilisome-reaction centre interaction in cyanobacteria. *Photosynth. Res.* 95(2-3):175-182.

Nanba O., Satoh K. (1987). Isolation of a photosystem II reaction center consisting of D-1 and D-2 polypeptides and cytochrome *b*-559. *Proc. Natl. Acad. Sci. U.S.A.* 84(1):109-112.

Nickelsen J., Rengstl B. (2013). Photosystem II Assembly: From cyanobacteria to plants. *Annu. Rev. Plant Biol.* 64:609-635.

Nowaczyk M.M., et al. (2012). Deletion of *psbJ* leads to accumulation of Psb27-Psb28 photosystem II complexes in *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta* 1817(8):1339-1345.

Paramelle D., Miralles G., Subra G., Martinez J. (2013). Chemical cross-linkers for protein structure studies by mass spectrometry. *Proteomics* 13(3-4):438-456.

Pearson K.M., Pannell L.K., Fales H.M. (2002). Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Commun. Mass Spectrom.* 16(3):149–159.

Petrotchenko E.V., Borchers C.H. (2010). ICC-CLASS: isotopically-coded cleavable crosslinking analysis software suite. *BMC Bioinformatics* 11:64.

Petrotchenko E.V., Serpa J.J., Borchers C.H. (2011) An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol. Cell. Proteomics* 10(2): 10.1074/mcp.M110.001420.

Petrotchenko E.V., Makepeace K.A.T., Serpa J.J., Borchers C.H. (2014). Analysis of protein structure by cross-linking combined with mass spectrometry. *Methods Mol. Biol.* 1156:447-463.

Rappsilber J. (2011). The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* 173(3):530–540.

Roberts V.A., Thompson E.E., Pique M.E., Perez M.S., Ten Eyck L.F. (2013). DOT2: Macromolecular docking with improved biophysical models. *J. Comput. Chem.* 34(20):1743-1758.

Roose J.L., Pakrasi H.B. (2004). Evidence that D1 processing is required for manganese binding and extrinsic protein assembly into Photosystem II. *J. Biol. Chem.* 279(44):45417-45422.

Sakata S., Mizusawa N., Kubota-Kawai H., Sakurai I., Wada H. (2013). Psb28 is involved in recovery of photosystem II at high temperature in *Synechocystis* sp. PCC 6803. *Biochim. Biophys. Acta* 1827(1):50-59.

Shimada Y., et al. (2008). Spectral properties of the CP43-deletion mutant of *Synechocystis* sp. PCC 6803. *Photosynth. Res.* 98(1-3):303-314.

Shinopoulos K.E., Brudvig G.W. (2012). Cytochrome b_{559} and cyclic electron transfer within photosystem II. *Biochim. Biophys. Acta* 1817(1):66-75.

Sinz A. (2014). The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert Rev. Proteomics.* 11(6):733-743.

Shi Y., et al. (2015). A strategy for dissecting the architectures of native macromolecular assemblies. *Nat. Methods* 12(12):1135-1138.

Suga M., et al. (2015). Native structure of photosystem II at 1.95Å resolution viewed by femtosecond X-ray pulses. *Nature* 517(7532):99-103.

Thangaraj B., et al. (2010). Data-directed top-down Fourier-transform mass spectrometry of a large integral membrane protein complex: Photosystem II from *Galdieria sulphuraria*. *Proteomics* 10(20):3644-3656.

Trnka M.J., Baker P.R., Robinson P.J.J., Burlingame A.L., Chalkley, R.J. (2014). Matching cross-linked peptide spectra: Only as good as the worse identification. *Mol. Cell. Proteomics* 13(2):420-434.

Umena Y., Kawakami K., Shen J.R., Kamiya N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* 473(7345):55-61.

Walzthoeni T., et al. (2011). False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat. Methods* 9(9):901-903.

Weisz D.A., Gross M.L., Pakrasi H.B. (2016). The use of advanced mass spectrometry to dissect the life-cycle of Photosystem II. *Front. Plant. Sci.* 7:617. doi: 10.3389/fpls.2016.00617

Yang Y.H., et al. (2011). Solution NMR structure of photosystem II reaction center protein Psb28 from *Synechocystis sp*. strain PCC 6803. *Proteins* 79(1):340-344.

Zelter A., et al. (2015). The molecular architecture of the Dam1 kinetochore complex is defined by cross-linking based structural modelling. *Nat. Commun.* 6:8673.

Zheng C., et al. (2013). XLink-DB: Database and software tools for storing and visualizing protein interaction topology data. *J. Proteome Res.* 12(4):1989-1995.

Supplementary Results and Discussion

Isotope-labeled cross-linker enabled confident cross-link identification

Cross-link data analysis is notoriously difficult owing to the large number of potential cross-linked peptides that might have formed and that must be included in the database search. This number rises roughly with the square of the number of proteins in the database. Distinguishing real cross-linked peptides from false positives thus becomes a key challenge. Our use of a 1:1 mixture of labeled and unlabeled cross-linker (Müller et al., 2001; Pearson et al., 2002; Petrotchenko et al., 2014), based on a strategy described in (Müller et al., 2001), addressed this challenge by providing an easily-observed "fingerprint" in the liquid chromatogram, MS spectra, and MS/MS spectra indicating a real cross-link (see Figure 3). In the chromatogram, the light and heavy forms eluted off the LC column within ~10 s of each other; in the MS spectra, the light and heavy forms occurred as characteristic doublets with a 12 Da peak shift; and in the MS/MS spectra, the light and heavy forms were identical, except that product ions containing the cross-linker molecule occurred with +12 Da shift for the heavy form. This shift in the MS/MS spectra also helped determine the sequence of product ions, by revealing whether or not a particular product ion contained the cross-linker molecule or not. All cross-linked peptides presented in this study displayed all of the characteristic fingerprints, and nearly all major product ions matched ones predicted for that peptide, resulting in highly confident cross-link identification.

The characteristic doublet observed in MS spectra of real cross-links addressed another key challenge in cross-link identification, namely that there tend to be many more non-linked peptides than cross-linked peptides present in a typical protein digest after cross-linking. With the typical "highest-abundance ion" criterion for selecting ions for fragmentation by MS/MS (a prerequisite for peptide sequence determination), many more non-linked peptides are selected than cross-linked peptides. This is a diversion of precious instrument time towards sequencing uninformative

peptides. Enrichment of cross-linked peptides before LC-MS/MS helps, and chromatographic (Leitner et al., 2012; Fritzsche et al., 2012; Buncherd et al., 2014) or affinity (Chu et al., 2006; Kang et al., 2009; Petrotchenko et al., 2011; Paramelle et al., 2013) enrichment has been used successfully for this purpose. In this study, we performed effective enrichment of cross-linked peptides *during* MS using the "mass tags" feature of the instrument to screen each mass spectrum for isotopic doublets on-the-fly, and select only those doublets for fragmentation, as described by the Borchers group (Petrotchenko et al., 2014). This strategy permitted the instrument to ignore most uninformative non-linked peptides. Using this approach we identified more PSII cross-links than we and others have found using the same cross-linker without this feature enabled (Cormann et al., 2016).

Evaluation of cross-link data quality

As mentioned in the "Results," we measured the distance between the C_α's of linked residues in the PSII crystal structure (Hellmich et al., 2014) to assess data quality, using a 30 Å upper threshold for cross-links to be considered consistent with the crystal structure (Fig. S1, Table S2). This value was obtained as follows: with a lysine side chain distance of 6.5 Å and cross-linker arm length of 11.4 Å, the C-α's of linked residues should be within 24.4 Å of each other in the crystal structure. An additional allowance of several angstroms is typically made to account for protein flexibility, and a 30 Å threshold is used commonly (Walzthoeni et al., 2011; Herzog et al., 2012; Fischer et al., 2013; Shi et al., 2015; Zelter et al., 2015). A recent survey study (Merkley et al., 2014) confirmed this value to be appropriate.

Of the thirteen measurable cross-link distances, eleven (85%) were less than 30 Å (Fig. S1, Table S2), which matches well with the 89% value obtained in the recent large-scale study (Merkley et al., 2014) of cross-links found in the *XLdb* (Kahraman et al., 2013). The remaining

two cross-links had distances of 34 and 38 Å. In both of these cases, at least one of the two linked residues lies on a flexible loop of the protein, which could facilitate cross-linking over the slightly larger distance between the sites.

Table S1. Mono-linked peptides identified in this study.

	[M+H]+	Charge	Protein	Peptide sequence	Linked residue
1	3872.7867	+4	PsbA	ETTEVESQNYGY**K**FGQEEETYNIVAAHGYFG R	K238
2	947.5002	+2	PsbB	**G**LPWYR	G2 (N-term)
3	919.5618	+2	PsbB	LY**K**ALR	K227
4	2045.9609	+3	PsbB	YQWD**K**GYFQEEIQR	K277
5	2628.2678	+3	PsbB	TGAMNSGDGIAQEWIGHPIF**K**DK	K347
6	3752.8297	+3	PsbB	SES**K**FSVEQTGVTVSFYGGALDGQTFSNPSDV KK	K389
7	3320.6418	+3	PsbB	FSVEQTGVTVSFYGGALDGQTFSNPSDV**K**K	K418
8	2636.2192	+3	PsbB	**K**AQLGEGFDFDTETFNSDGVFR	K423
9	3137.5638	+3	PsbB	DVFAGVDPGLEEQVEFGVFA**K**VGDLSTR	K497
10	2174.1668	+3	PsbC	LGANIASAQGPTGLG**K**YLMR	K338
11	2302.2587	+3	PsbC	GPWLEPLRGPNGLDLD**K**LR	K378
12	1361.7065	+2	PsbC	AAAAGFE**K**GIDR	K456
13	1979.0105	+3	PsbE	**S**GTTGERPFSDIVTSIR	S2 (N-term)
14	2102.0938	+3	PsbF	**A**TQNPNQPVTYPIFTVR	A2 (N-term)
15	1730.9315	+3	PsbH	LGDILRPLNSEYG**K**	K20
16	3440.7029	+3	Psb27	**K**GDAGGL**K**SFTTMQTALNSLAGYYTSYGAR	K56, K63
17	980.6027	+2	Psb27	PIPE**K**LK	K90
18	1264.7966	+3	Psb27	PIPEKL**K**KR	K92
19	978.5148	+2	Psb28	**A**EIQFSK	A2 (N-term)
20	2115.1369	+3	Psb28	AEIQFS**K**GVAETVVPEVR	K8
21	2272.1636	+3	Psb28	NGQSGMA**K**FYFLEPTILAK	K32
22	2968.5532	+3	Psb28	G**K**FINGRPTAIEATVILNSQPEWDR	K69

Table S2. Cross-linked peptides identified in this study.

	[M+H]+	Charge	Protein 1	Protein 2	Cross-linked peptide sequence	Linked residues	Distance (Å)
1	4645.1652	+4	PsbA	PsbB	ETTEVESQNYGY**K**FGQEEETYNIVAAHGYF GR—**G**LPWYR	K238-G2 (N-term)	17
2	4391.9917	+4	PsbA	PsbI	ETTEVESQNYGY**K**FGQEEETYNIVAAHGYF GR—**K**DFE	K238-K35	34
3	3882.0473	+4	PsbB	PsbB	LY**K**ALR— DVFAGVDPGLEEQVEFGVFA**K**VGDLSTR	K227-K497	20
4	4338.1007	+4	PsbB	PsbB	SES**K**FSVEQTGVTVSFYGGALDGQTFSNPS DVKK—D**K**EGR	K349-K389	14
5	3677.8117	+4	PsbB	-	FSVEQTGVTVSFYGGALDGQTFSNPSDV**KK** FAR	K418-K419	4
6	3138.5333	+4	PsbB	PsbB	**K**FAR—**K**AQLGEGFDFDTETFNSDGVFR	K419-K423	7
7	2475.3975	+5	PsbB	PsbH	LY**K**ALR—LGDILRPLNSEYG**K**	K227-K20	12
8	1321.7312	+3	PsbB	PsbL	LY**K**ALR—**M**DR	K227-M1 (N-term)	38
9	1349.6652	+2	PsbB	PsbL	**G**LPWYR—**M**DR	G2 (N-term)-M1 (N-term)	24
10	1880.9339	+3	PsbC	PsbI	AAAAGFE**K**GIDR—**K**DFE	K456-K35	17
11	3906.0209	+4	PsbE	PsbF	**S**GTTGERPFSDIVTSIR— **A**TQNPNQPVTYPIFTVR	S2 (N-term)-A2 (N-term)	-
12	3919.0502	+4	PsbE	Psb28	**S**GTTGERPFSDIVTSIR— AEIQFS**K**GVAETVVPEVR	S2 (N-term)-K8	-
13	4042.1354	+4	PsbF	Psb28	**A**TQNPNQPVTYPIFTVR— AEIQFS**K**GVAETVVPEVR	A2 (N-term)-K8	-
14	4042.1502	+4	PsbF	Psb28	**A**TQNPNQPVTYPIFTVR— **A**EIQFSKGVAETVVPEVR	A2 (N-term)-A2 (N-term)	-
15	3267.6082	+3	Psb27	-	**K**GDAGGL**K**SFTTMQTALNSLAGYYTSYGA R	K56, K63	14
16	962.59099	+2	Psb27	-	PIPE**K**L**K**	K90, K92	6
17	2058.0421	+3	Psb28	Psb28	LSKS**K**—YGAENGLGFS**K**SE	K24-K110	28
18	4273.2531	+4	Psb32	Psb32	TGVYDLPILSPGS**K**TFLVDQAEAISLANENR —LNSDL**K**K	K60-K83	-

The Psb28-PsbE and Psb28-PsbF cross-links, as well as the cross-linked residues for each entry, are shown in red. All distances are linear (Euclidean) distances between the C-α's of the two linked residues, as measured in PDB 4PJ0 (PSII), 2KND (Psb27) and 2KVO (Psb28) with adjustments for the following entries:

2, 10- PDB 3WU2 was used because PsbI-K35-C-α is not resolved in PDB 4PJ0.

3-PsbB-K497 corresponds to K498 in 4PJ0, so that residue was used for distance measurement.

5,6-PsbB-K419 corresponds to S419 in PDB 4PJ0, and that residue was used for distance measurement.

8,9- PsbL-M1 was not resolved in PDB 4PJ0, but was resolved in PDB 3WU2, so 3WU2 was used instead for these distance measurements.

10-PsbC-K456 corresponds to K457 in PDB 4PJ0, and that residue was used for distance measurement.

Table S3. Energetics obtained from the Psb28-RC47 docking calculations for the top 100 Cluster-1 conformations. All energy values are given in kcal/mol.

Conformation #	Interaction Energy	E_elec[1]	E_vdw[2]	Conformation #	Interaction Energy	E_elec[1]	E_vdw[2]
1	-56.11	-52.41	-3.70	**51**	-45.41	-39.51	-5.9
2	-55.04	-44.14	-10.90	**52**	-45.31	-43.31	-2
3	-53.91	-44.91	-9.00	**53**	-45.25	-39.85	-5.4
4	-52.37	-45.17	-7.20	**54**	-45.23	-33.43	-11.8
5	-52.27	-51.77	-0.50	**55**	-45.16	-40.46	-4.7
6	-52.20	-40.20	-12.00	**56**	-45.09	-36.79	-8.3
7	-51.64	-43.04	-8.60	**57**	-45.04	-37.74	-7.3
8	-50.92	-45.72	-5.20	**58**	-45.01	-40.21	-4.8
9	-50.86	-41.46	-9.40	**59**	-45.01	-37.51	-7.5
10	-50.68	-43.38	-7.30	**60**	-44.99	-37.19	-7.8
11	-50.52	-47.22	-3.30	**61**	-44.94	-37.94	-7
12	-49.57	-38.77	-10.80	**62**	-44.87	-39.17	-5.7
13	-49.50	-42.40	-7.10	**63**	-44.84	-41.74	-3.1
14	-49.25	-42.05	-7.20	**64**	-44.77	-37.77	-7
15	-49.05	-40.55	-8.50	**65**	-44.67	-38.97	-5.7
16	-48.68	-45.58	-3.10	**66**	-44.63	-35.73	-8.9
17	-48.66	-47.36	-1.30	**67**	-44.48	-38.78	-5.7
18	-48.59	-36.99	-11.60	**68**	-44.48	-33.38	-11.1
19	-48.54	-39.84	-8.70	**69**	-44.46	-37.06	-7.4
20	-48.17	-39.67	-8.50	**70**	-44.45	-44.05	-0.4
21	-47.84	-43.24	-4.60	**71**	-44.45	-33.65	-10.8
22	-47.39	-40.49	-6.90	**72**	-44.44	-35.74	-8.7
23	-47.38	-35.18	-12.20	**73**	-44.33	-41.53	-2.8
24	-47.36	-45.16	-2.20	**74**	-44.33	-37.33	-7
25	-47.32	-41.02	-6.30	**75**	-44.30	-39.90	-4.4
26	-47.27	-44.57	-2.70	**76**	-44.25	-34.95	-9.3
27	-47.23	-42.63	-4.60	**77**	-44.22	-41.52	-2.7
28	-47.07	-44.97	-2.10	**78**	-44.20	-41.30	-2.9
29	-46.97	-43.07	-3.90	**79**	-44.16	-34.56	-9.6
30	-46.91	-39.11	-7.80	**80**	-44.14	-42.44	-1.7
31	-46.55	-35.05	-11.50	**81**	-44.12	-42.82	-1.3
32	-46.53	-42.73	-3.80	**82**	-44.10	-43.40	-0.7

33	-46.49	-39.59	-6.90	**83**	-44.08	-35.48	-8.6
34	-46.39	-43.19	-3.20	**84**	-44.05	-37.95	-6.1
35	-46.32	-42.32	-4.00	**85**	-44.04	-37.24	-6.8
36	-46.23	-41.03	-5.20	**86**	-43.93	-40.53	-3.4
37	-46.15	-40.45	-5.70	**87**	-43.91	-36.51	-7.4
38	-46.13	-43.53	-2.60	**88**	-43.91	-39.21	-4.7
39	-46.04	-35.94	-10.10	**89**	-43.89	-41.79	-2.1
40	-45.98	-42.38	-3.60	**90**	-43.85	-37.95	-5.9
41	-45.89	-37.29	-8.60	**91**	-43.82	-35.62	-8.2
42	-45.87	-40.47	-5.40	**92**	-43.82	-34.42	-9.4
43	-45.78	-37.98	-7.80	**93**	-43.79	-38.19	-5.6
44	-45.74	-42.64	-3.10	**94**	-43.77	-36.27	-7.5
45	-45.72	-43.22	-2.50	**95**	-43.77	-41.37	-2.4
46	-45.63	-40.33	-5.30	**96**	-43.76	-40.16	-3.6
47	-45.51	-40.11	-5.40	**97**	-43.76	-40.56	-3.2
48	-45.48	-41.08	-4.40	**98**	-43.74	-34.34	-9.4
49	-45.45	-31.65	-13.80	**99**	-43.73	-38.33	-5.4
50	-45.43	-40.53	-4.90	**100**	-43.69	-38.19	-5.5

[1]Electrostatic energy
[2]van der Waals energy

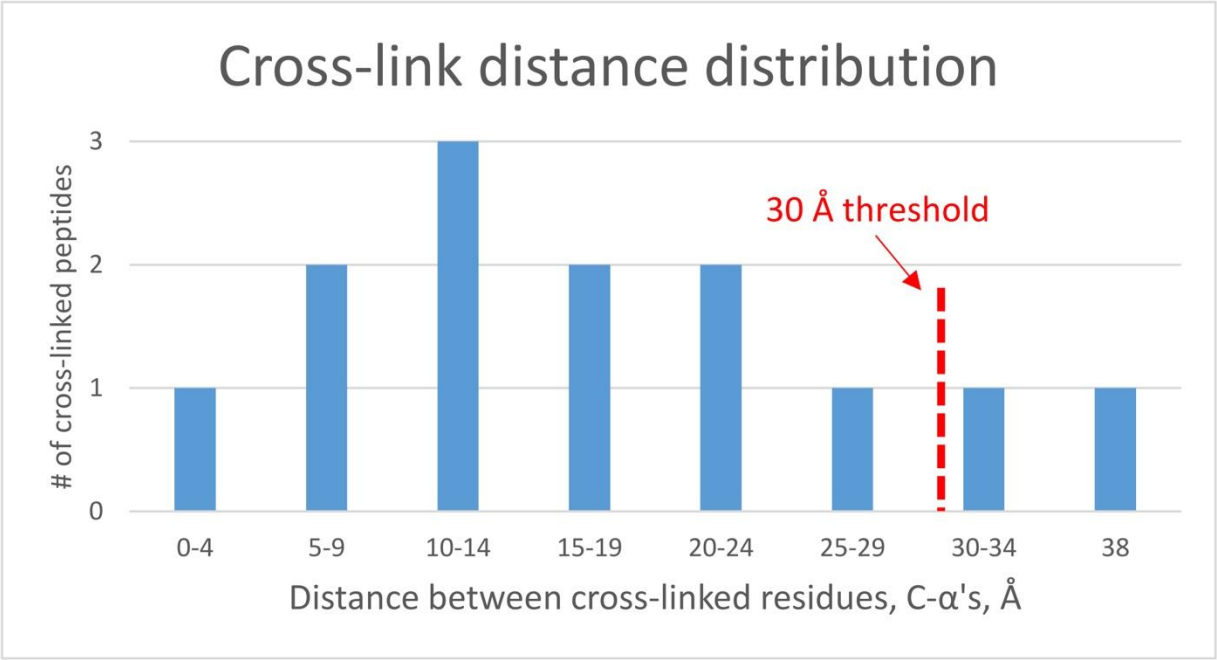


Figure S1. Distribution of linear (Euclidean) distances between C-α's of cross-linked residues in Δ*psbO*-PSII.

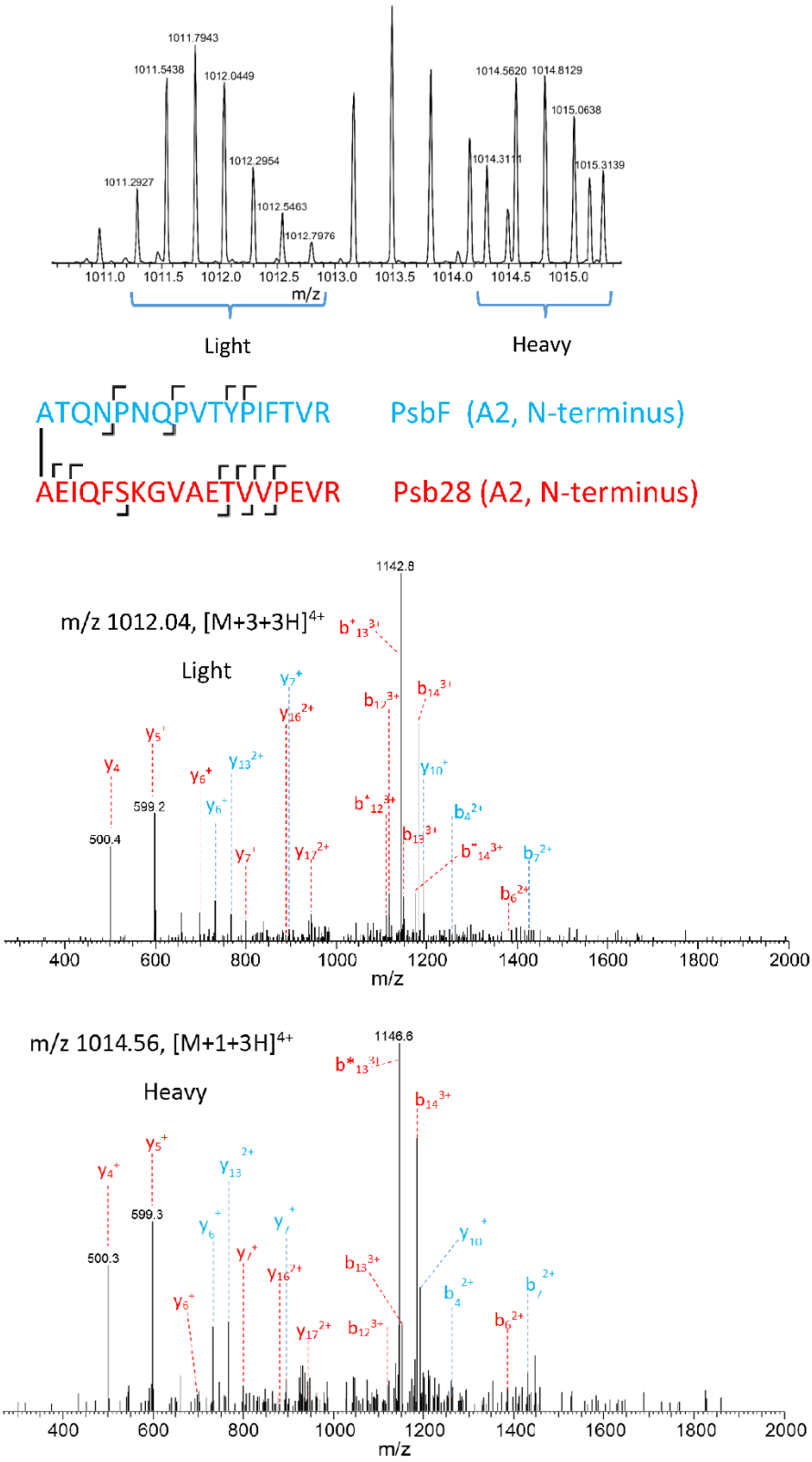


Figure S2. Mass spectrometric data showing a cross-link between Psb28-A2 and PsbF-A2. See Fig. 3 legend for information on interpreting these spectra.

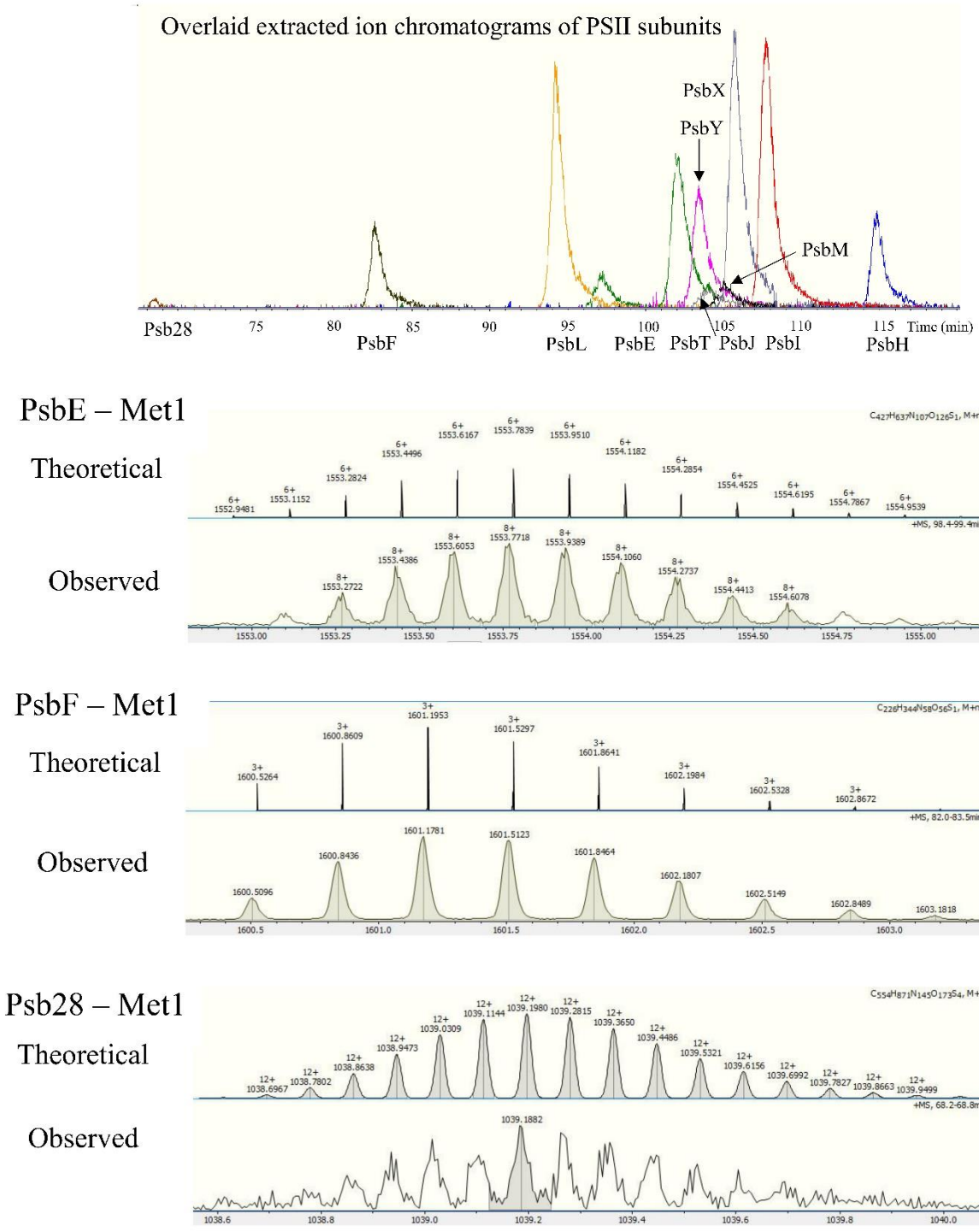


Figure S3. Intact-mass spectra of PsbE, PsbF, and Psb28. These spectra show that the N-terminal methionine has been cleaved for all three proteins, rendering PsbE-S2, PsbF-A2, and Psb28-A2 as the N-terminal residues of the mature proteins. Theoretical mass spectra of each protein are shown for comparison. Mass accuracy of 8, 11, and 10 ppm (within 0.1 Da) was achieved for PsbE, and PsbF, and Psb28, respectively.

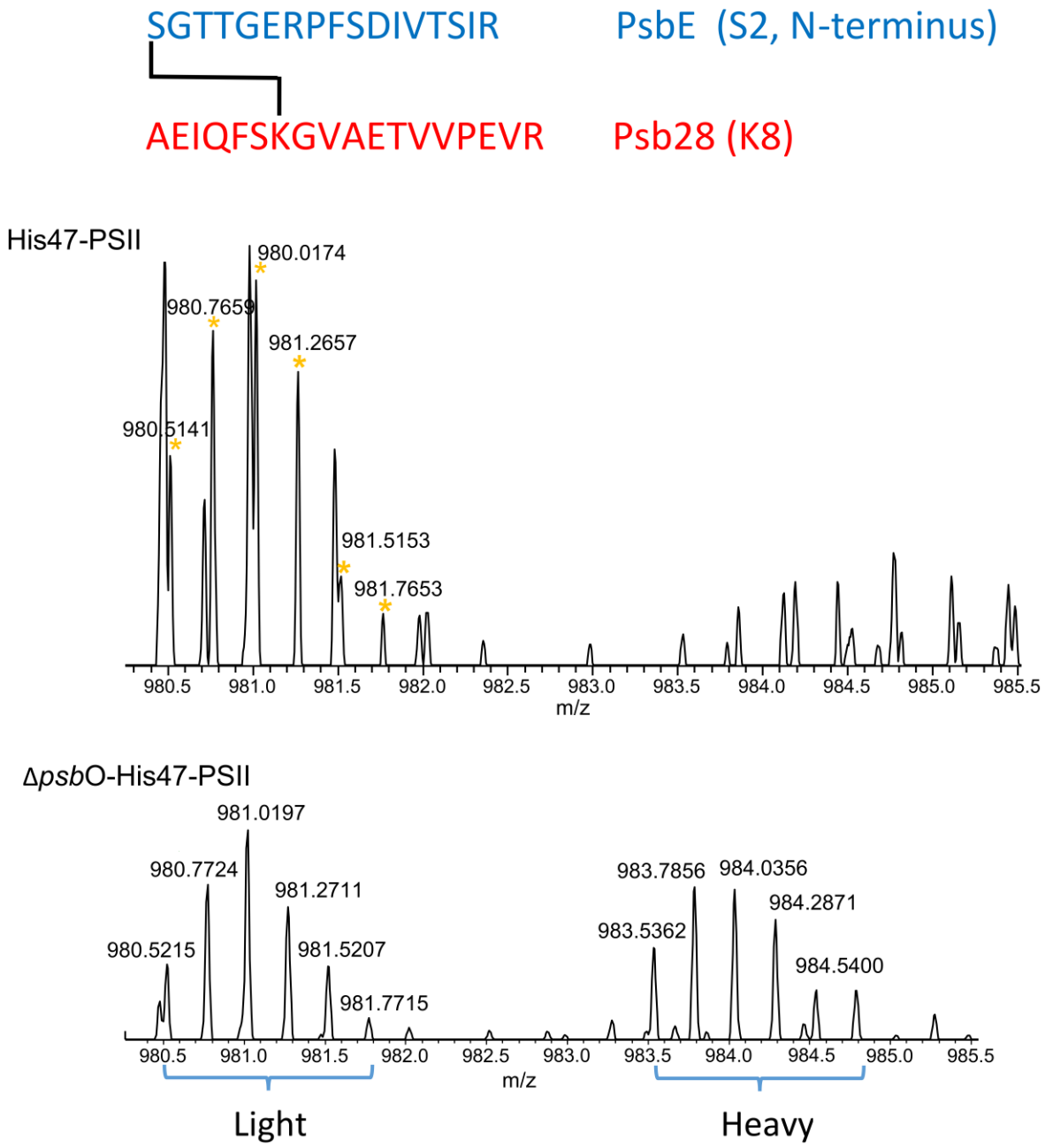


Fig. S4. **PsbE-Psb28 cross-link detected in His47 sample.** The PsbE-Psb28 cross-link identified in the Δ*psb*O-His47 strain of *Synechocystis* 6803 used in this study (lower spectrum), was also detected in the His47 strain (wild-type with a polyhistidine tag on the CP47 protein) (upper spectrum, 3 ppm MS-1 accuracy). This demonstrates that the association between Psb28 and PsbE we describe is not an artifact of the absence of *psb*O. The upper spectrum reflects a technical replicate experiment in which non-isotope-encoded BS^3 was used.

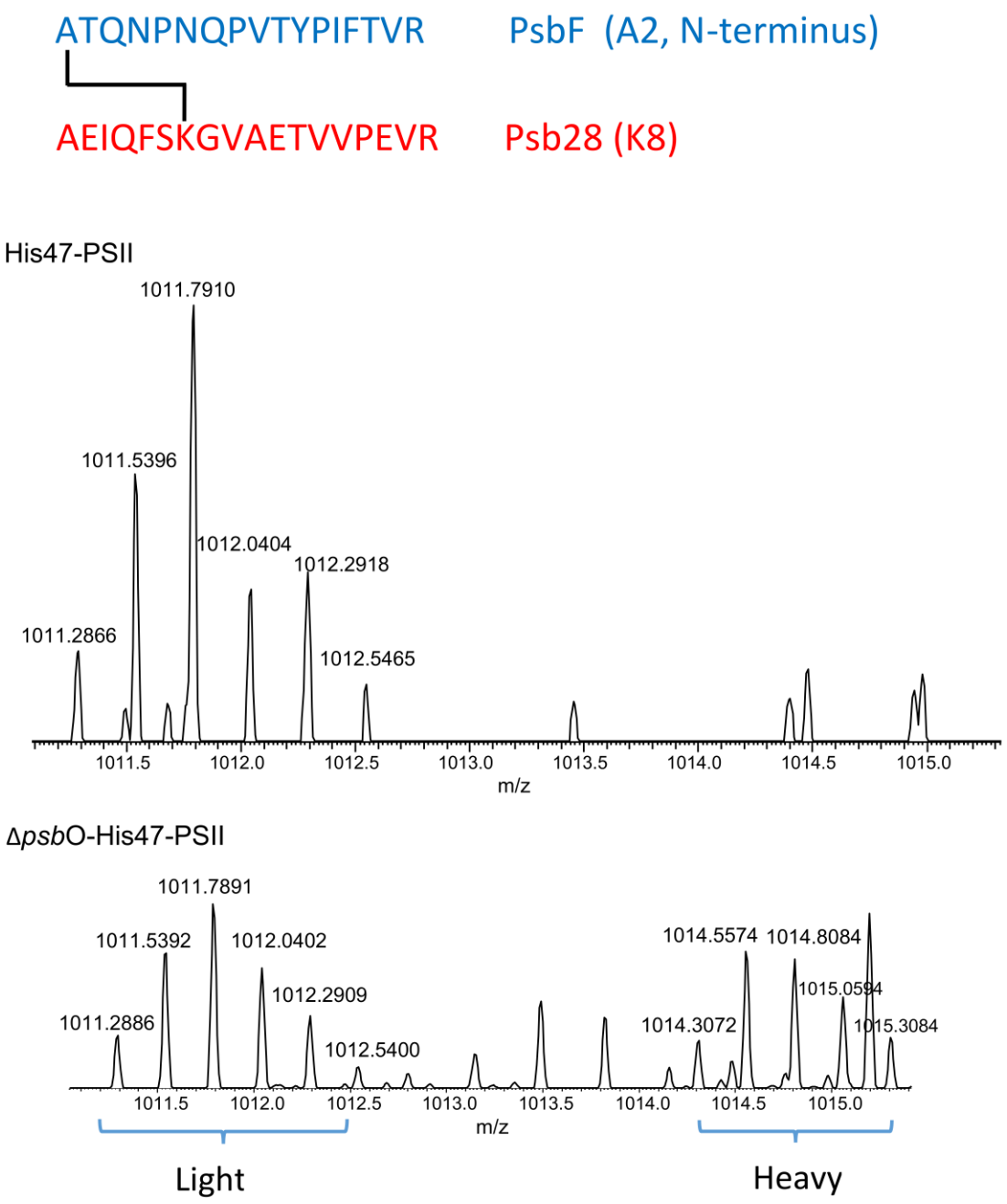


Fig. S5. **PsbF-Psb28 cross-link detected in His47 sample.**The PsbF-Psb28 cross-link identified in the Δ*psb*O-His47 strain of *Synechocystis* 6803 used in this study (lower spectrum), was also detected in the His47 strain (wild-type with a polyhistidine tag on the CP47 protein) (upper spectrum, 1 ppm MS-1 accuracy). This demonstrates that the association between Psb28 and PsbF we describe is not an artifact of the absence of *psb*O. The upper spectrum reflects a technical replicate experiment in which non-isotope-encoded BS^3 was used (the same LC-MS/MS experiment from which the upper spectrum in Fig. S4 was found).

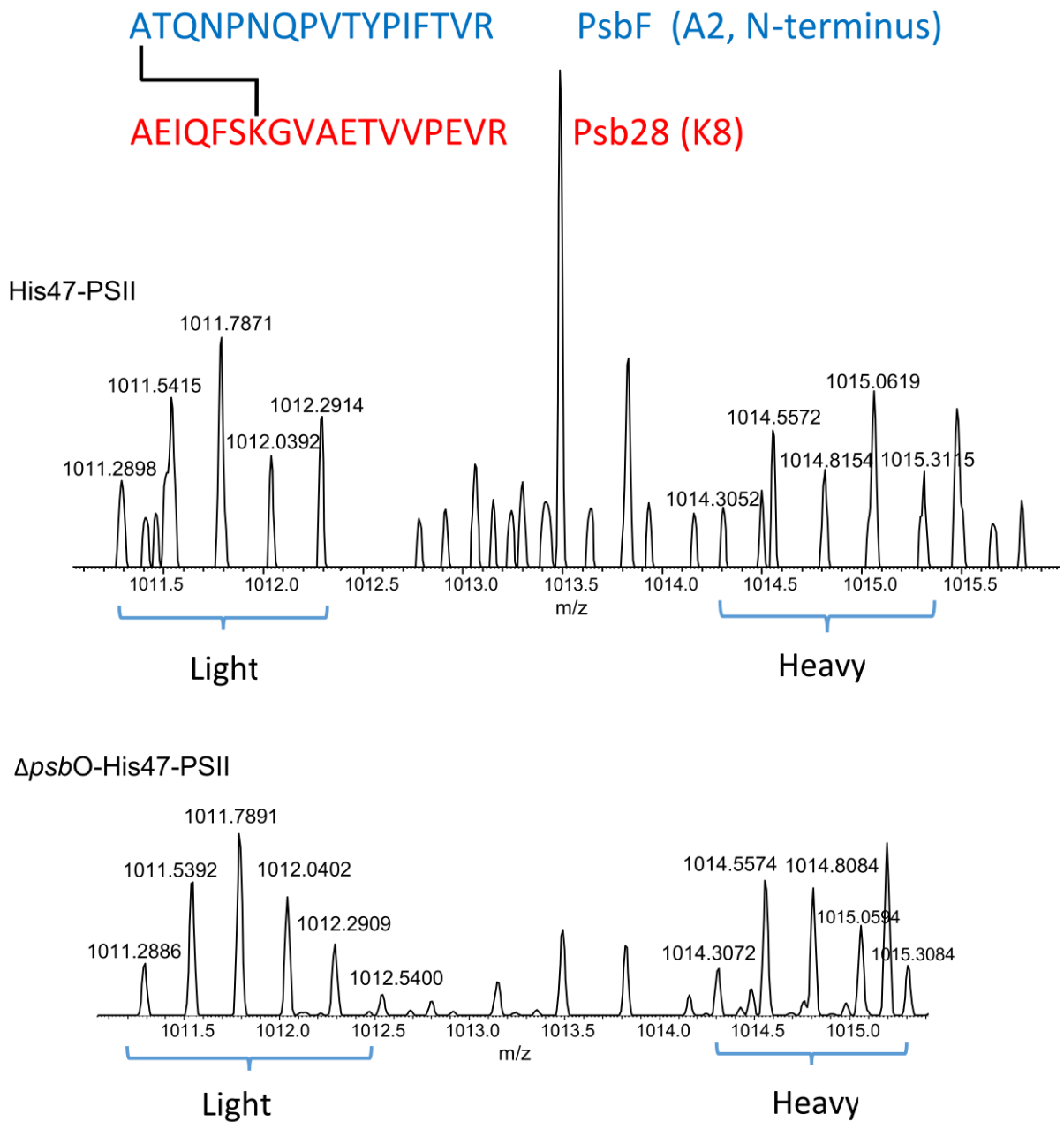


Fig. S6. **Isotope-encoded PsbF-Psb28 cross-link detected in His47 sample.** Same result as Fig. S5, with the upper spectrum demonstrating that the MS-1 spectrum of the PsbF-Psb28 cross-link was detected in the His47 sample (2 ppm mass accuracy), but in this figure the upper spectrum reflects a technical replicate experiment in which isotope-encoded BS^3 was used.

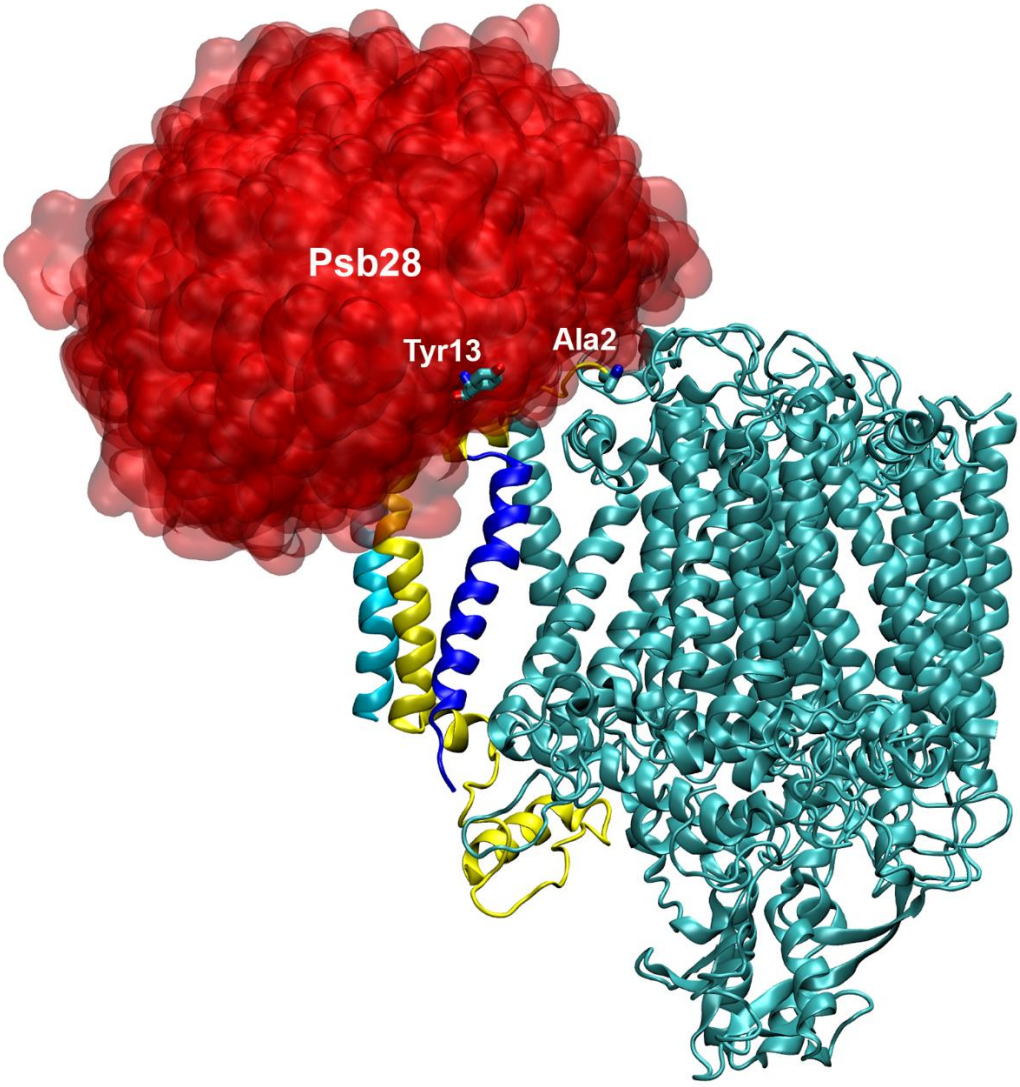


Figure S7. Top docked conformations of Psb28 to RC47. The top 100 Cluster-1 conformations show Psb28 binding to the RC47 complex on the cytosolic surface above PsbE (yellow) and PsbF (blue). PsbE-A2 and PsbF-Y13 are shown. PsbF-Y13 is actually obscured behind part of the Psb28 cloud, but its position has been brought to the fore in this image for clarity.

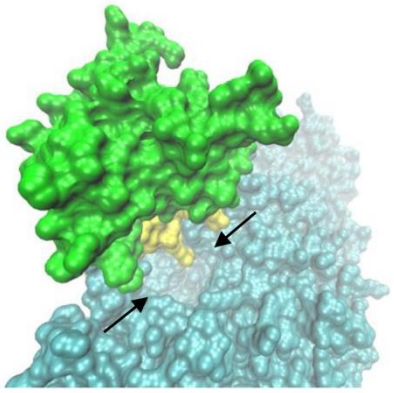


Figure S8. **Psb28 binding above the PSII cytosolic-surface cavity.** In our model Psb28 binds just beside the cavity on the cytosolic surface of PSII (shown by arrows) that is the site of phycobilisome attachment to PSII (Chang et al., 2015). Green, Psb28-Conformation 1; yellow, PsbE; Cyan, RC47 proteins.

Chapter 3

Oxidative modifications of PSII detected by mass spectrometry

Summary

Photosystem II undergoes frequent light-induced damage, prompting an intricate cycle of repair and reassembly. It is known that PSII can produce reactive oxygen species (ROS) by reaction of molecular oxygen with redox-active cofactors, and these ROS may be responsible for the observed PSII damage. This study uses high-resolution tandem MS to identify oxidative modifications found on the PSII proteins D1, D2, and CP43. It was found that 18 residues were modified on the cytosolic side of PSII, and 42 were modified on the lumenal side. Of the 18 cytosolic-side residues, eight cluster around the redox-active metal centers Q_A or Q_B. Only three of the 18 residues are buried in the protein complex, but all three are within 8 Å of Q_A. Our results suggest that Q_A and Q_B are sources of ROS, which is consistent with previous findings, and that the surrounding protein residues can be oxidized in a light-dependent manner. The 42 lumen-side residues cluster mainly into two nearly continuous, roughly linear "arm" formations, both leading from the Mn_4Ca cluster all the way to the surface of PSII. These two formations appear to track pathways of travel for oxygen/ROS to leave PSII; as has been recognized in the past, such pathways would be a valuable mechanism for minimizing overall ROS damage to the complex.

Introduction

Photosystem II (PSII) is a membrane-protein complex found in the thylakoid membrane of organisms that perform oxygenic photosynthesis. Fully assembled PSII consists of approximately twenty protein subunits and multiple redox-active cofactors that allow PSII to function as a catalyst for the light-driven oxidation of water and concomitant reduction of plastoquinone. This conversion of sunlight to chemical energy is critical for sustaining nearly all life on Earth, but it is a taxing reaction for the PSII complex itself. PSII undergoes frequent photo-induced damage,

prompting an intricate cycle of repair and reassembly (Nickelsen and Rengstl, 2013; see chapter 1 for a description of the PSII life-cycle).

The remarkable water-splitting reaction requires formation of the highest-potential (~1.25 V) species in all of biology, P_{680}^{+} (Barber and Andersson, 1992; Grabolle and Dau, 2005). The multiple redox cofactors in the electron transfer chain of PSII react with widely varying kinetics, providing opportunity for reaction with molecular oxygen (a byproduct of water oxidation) and formation of reactive oxygen species (ROS) such as H_2O_2, $O_2^{\cdot -}$, and $OH^{\cdot}$ (Pospíšil, 2009). In addition, charge recombination reactions can produce triplet chlorophyll states that can react with molecular oxygen to produce ROS. For reviews of these processes, see Krieger-Liszkay et al. (2008) and Pospíšil (2009). Oxidative damage of PSII residues by reaction with ROS is one of the major molecular mechanisms proposed as the source of PSII photodamage (Vass and Cser, 2009; Takahashi and Badger, 2011; Vass, 2012).

Modern high-resolution protein mass spectrometry allows a detailed characterization of oxidative modifications on PSII, although relatively few studies have explored this topic in depth. In an early MS study on pea PSII, Barber and co-workers (Sharma et al., 1997) noticed that D1 and D2 peptides near predicted metal cofactor sites were preferentially oxidized compared to other peptides, but the exact oxidized residues could not be determined. Barry and co-workers (Dreaden et al., 2011; Kasson et al., 2012) found that spinach CP43-W365 and D1-W317 are oxidized to N-formylkynurenine in a light-induced manner, and that these modifications are correlated with decreased oxygen evolution. These two residues are located 17 and 14 Å, respectively, from the Mn_4Ca cluster.

Using tandem MS, Bricker and co-workers (Frankel et al., 2012; Frankel et al., 2013b) characterized the oxidative modifications found in spinach PSII. Their results are discussed further

below, but they found numerous oxidized residues near the Mn_4Ca cluster and the redox-active Q_A and Q_B metal centers, as well as other oxidized residues that were mostly surface-exposed. Based on their results, they proposed the existence of an oxygen/ROS exit channel involving CP43 residues near the Mn_4Ca cluster, though they did not have evidence that it leads all the way to the surface of PSII. A channel in that area had not previously been identified in computational studies (Murray and Barber, 2007; Ho and Styring, 2008; Gabdulkhakov et al., 2009).

In this study, we used tandem MS to characterize the oxidative modifications present in PSII from *Synechocystis* 6803. We sought to explore the hypothesis that residues near redox-active metal centers are particularly susceptible to oxidative modification. In addition, we compared the results for light-incubated and dark-incubated samples, to investigate the effect of light exposure on oxidative modifications to PSII. Our results support the idea that light exposure leads to oxidative damage of PSII residues, particularly those in close proximity to redox active centers. Our results also describe two pathways of ROS travel all the way from the Mn_4Ca cluster to the surface of PSII.

Results and Discussion

In analyzing our MS data, we focused on the D1, D2, and CP43 proteins since these proteins are most closely associated with the Mn_4Ca cluster, Q_A, and Q_B, which have been previously identified as likely sites of ROS generation (Frankel et al., 2012; Frankel et al., 2013b). We obtained 35%, 37%, and 46% sequence coverage of the D1, D2, and CP43 proteins, respectively. Our coverage is appreciably higher than has been typically reported for these proteins (Nakamura et al., 2004; Aro et al., 2005; Liu et al., 2013a,b), including in the study of oxidative modifications in higher plant PSII by Bricker and co-workers (Frankel et al., 2013a), in which 24%, 27%, and 26% coverage was reported for D1, D2, and CP43, respectively. Our higher

coverage likely can be attributed to the increased sensitivity and speed of the Q-Exactive Plus (QE) mass spectrometer we employed, compared to the LTQ-FTICR, earlier Orbitrap, or MALDI-TOF instruments used previously. The high mass accuracy (~0.02 Da) of MS/MS spectra recorded by the QE enabled highly confident peptide identification and residue-level localization of PTMs, to an extent not possible on earlier instruments lacking high mass-accuracy MS/MS. All of the oxidative modifications presented below were localized unambiguously to the single residue reported. Several example MS/MS spectra are shown in Fig. 1 to indicate the quality of the data obtained in this study. See Table 1 for a list of the types of oxidative modifications included in the searches.

A total of 472 residues on D1, D2, and CP43 were covered by MS, and oxidative modifications were identified on 60 (13%) (Table 2). Of these 60 modified residues, 22 were detected only in the light-incubated samples (while only five were detected in the dark-incubated samples but not in the light-incubated samples). This finding supports the idea that PSII photochemistry leads to oxidative damage of its protein components. Of all the PSII proteins, D1 is believed to be the prime site of photodamage (Nickelsen and Rengstl, 2013) and as a result is turned over the fastest (Yao et al., 2012). Vermaas and co-workers found that in *Synechocystis*, D1, D2, and CP43 have increasing half-lives of <1, 3.3 ±1, and 6.5 ± 1.5 h, respectively (Yao et al., 2012), with a longer half-life presumably reflecting a lower damage rate. Matching this trend, D1 showed the most dramatic response to light incubation, with a 3.3-fold increase in number of oxidized residues detected after light incubation, followed by D2 (2-fold increase), followed by CP43 (1.2-fold increase) (Table 2).

A.

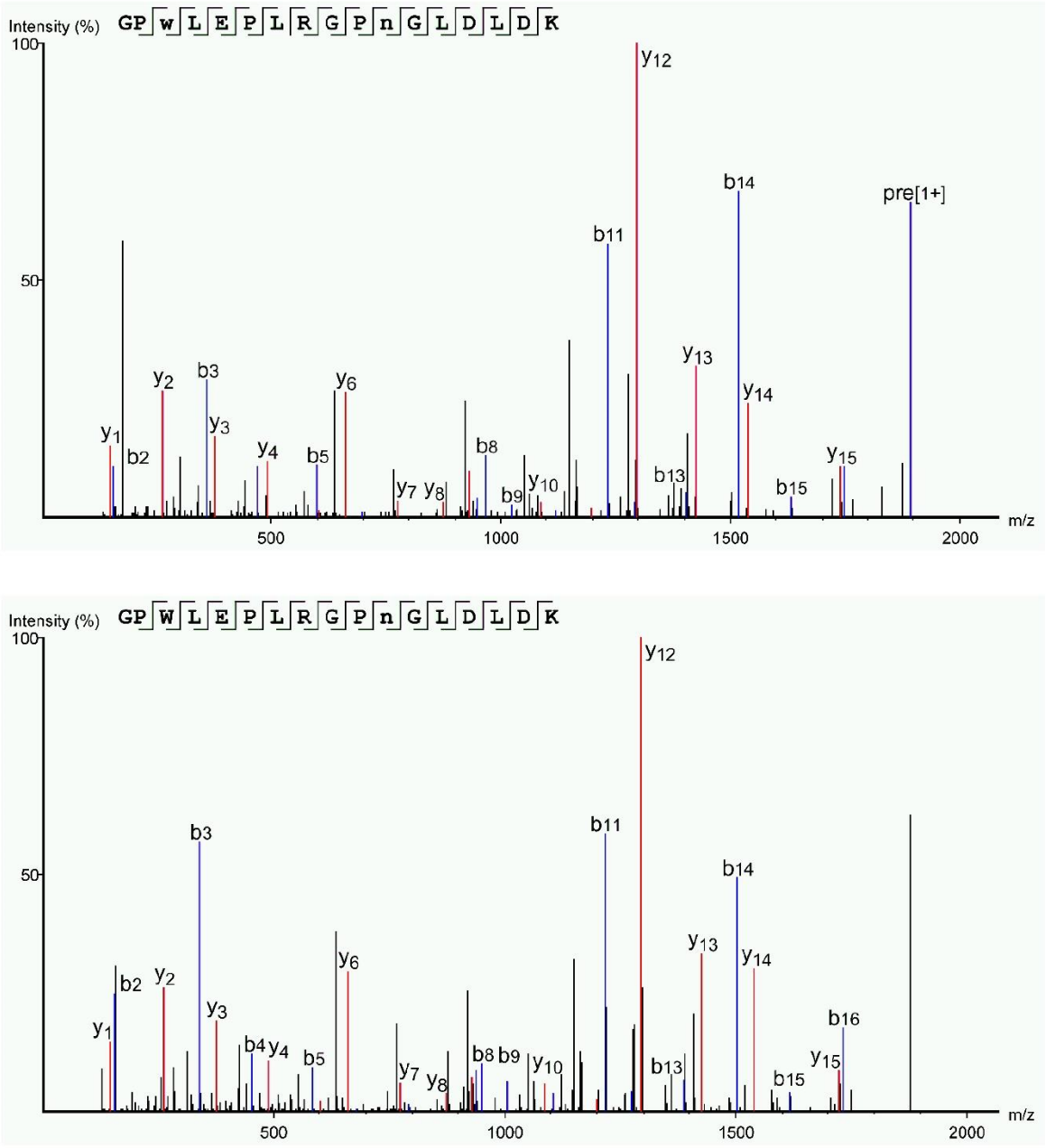


Figure 1A. (legend follows)

B.

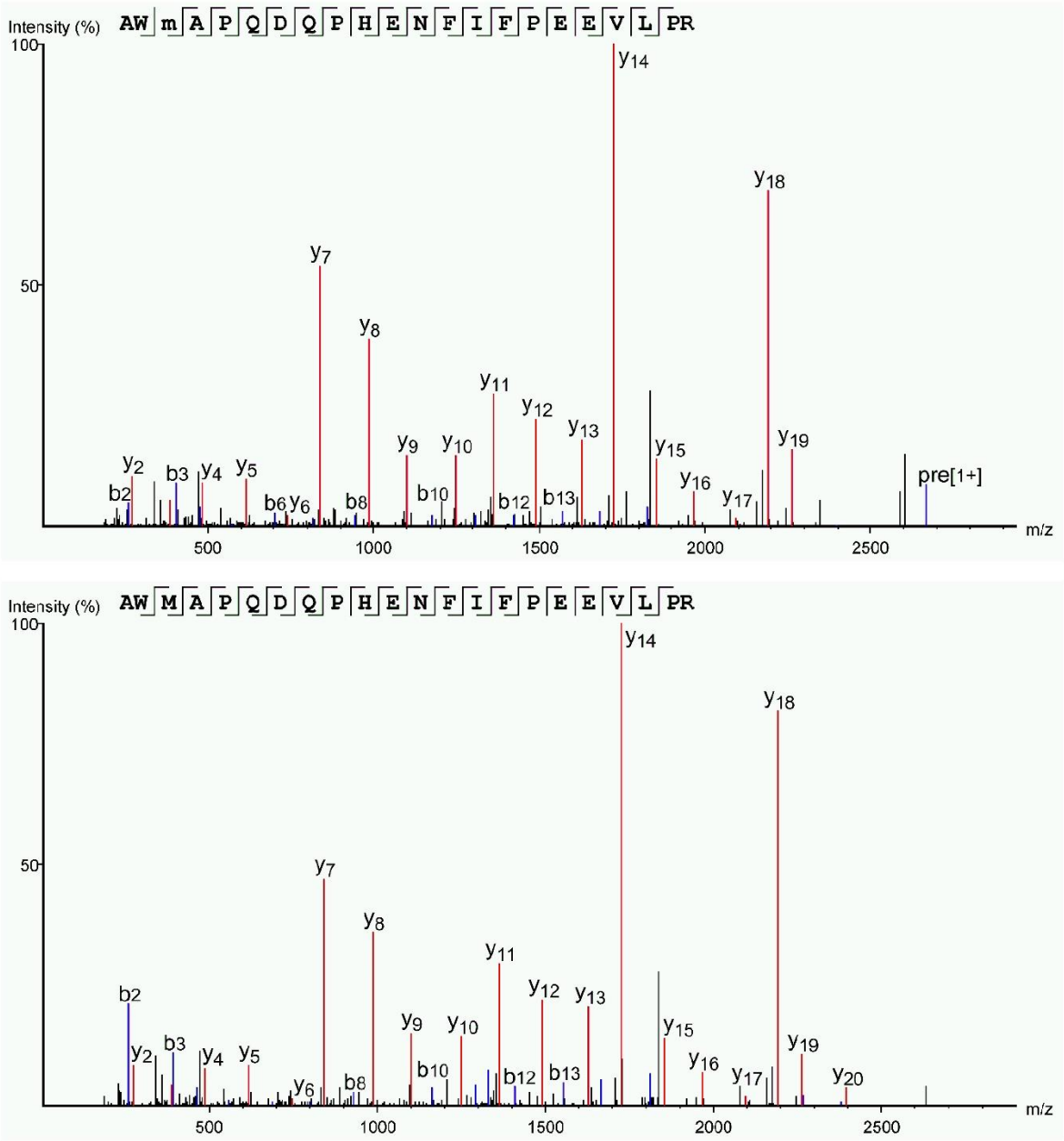


Figure 1B. (legend follows)

C.

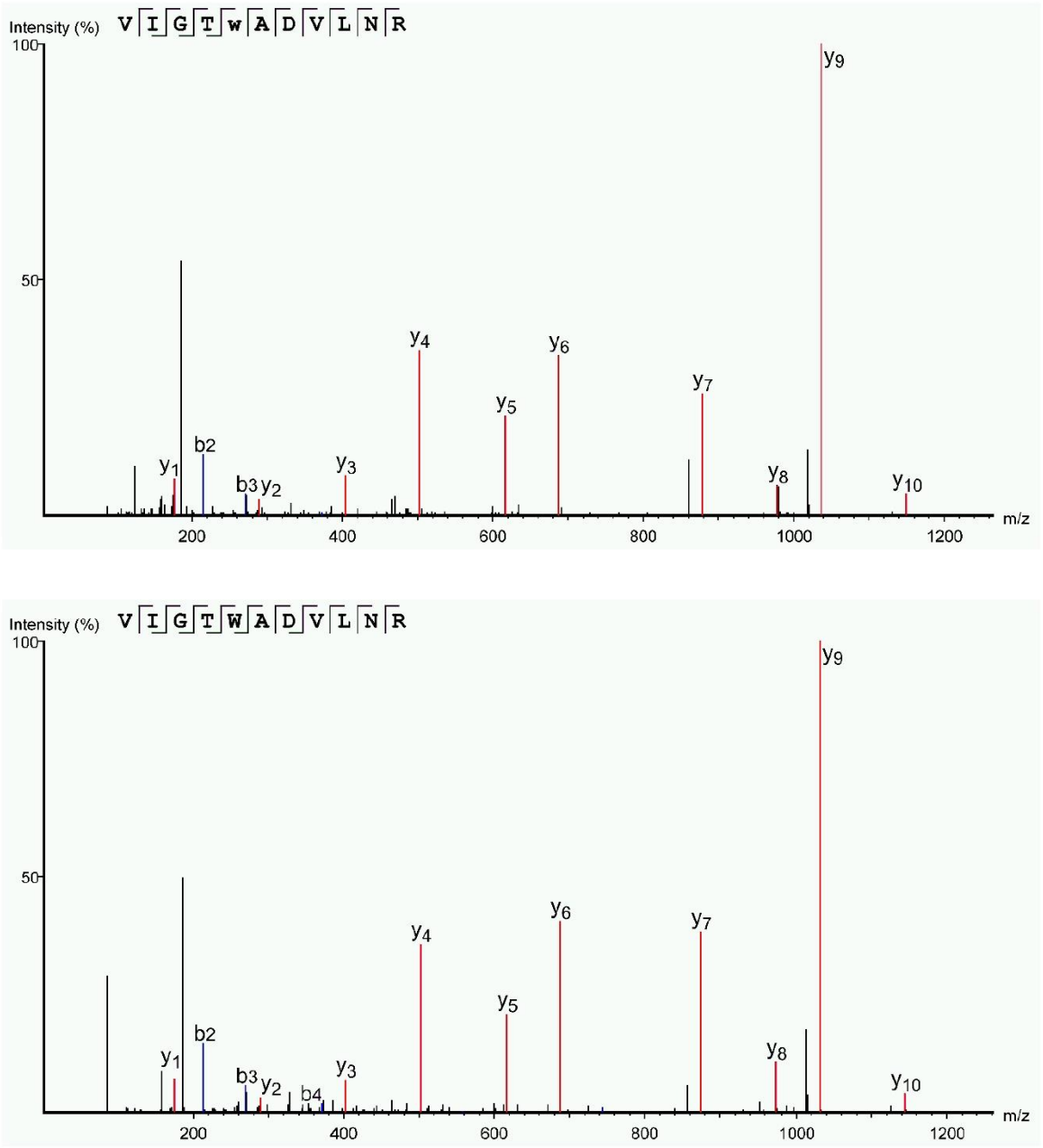


Figure 1C. (legend follows)

D.

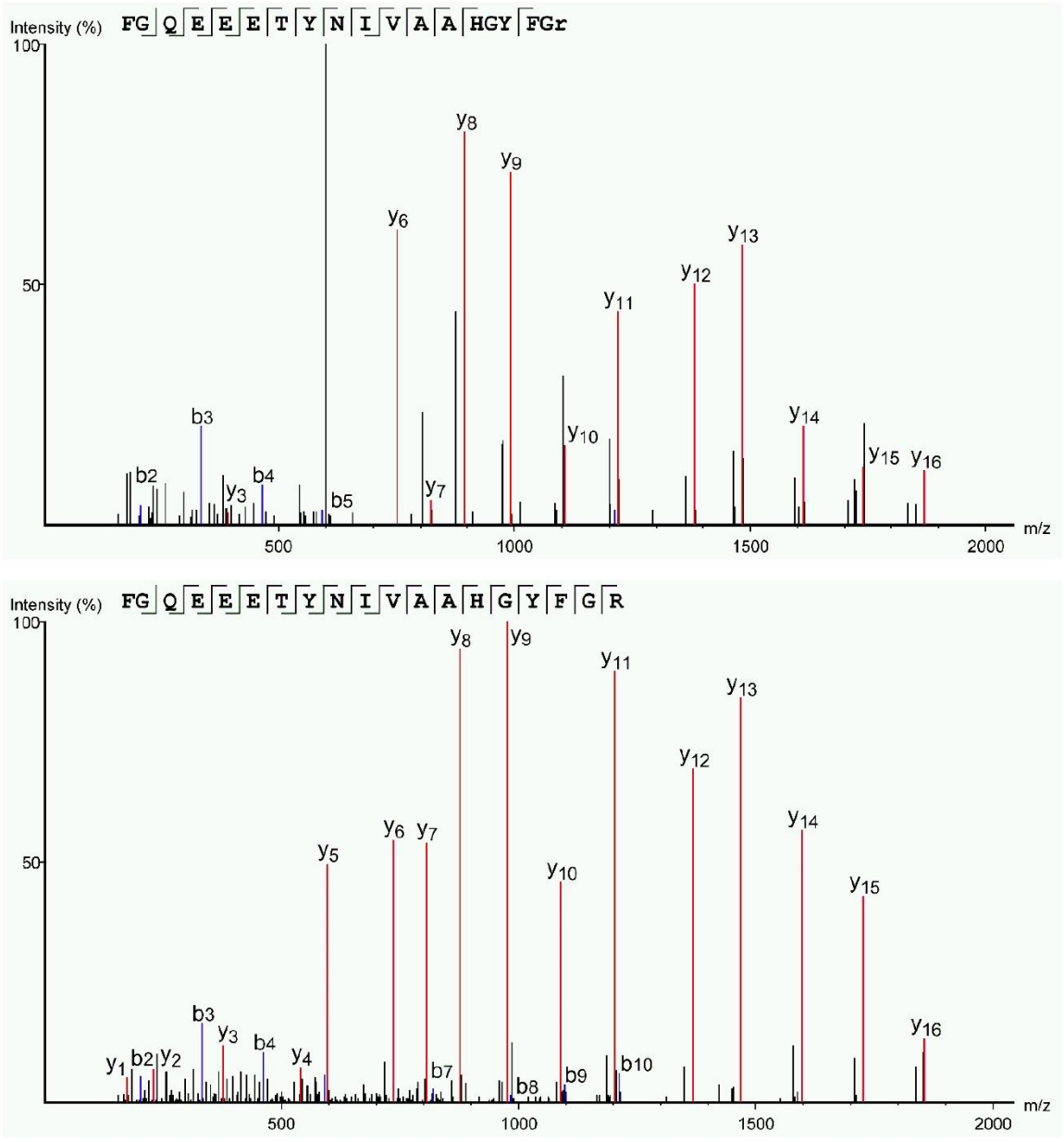


Figure 1D. (legend follows)

Figure 1. Examples of MS/MS spectra detected that identify oxidative modifications of PSII residues.

The fragment maps and corresponding labeled b- and y-ions show the excellent fragmentation series obtained at high mass accuracy, which allowed highly confident peptide identification and unambiguous residue-level localization of the oxidative modification. Lower-case lettering below and in the fragment map of each spectrum indicates the site of oxidation.

A. MS/MS spectra of the oxidized CP43 peptide ^{362}GPwLEPLRGPNGLDLD^{378}K (top) and the unmodified form of the peptide (bottom). The oxidation (+15.9949 Da) was localized to ^{364}W. Note that ^{372}N contains the common deamidation modification (+0.9840 Da) in both spectra.

B. MS/MS spectra of the oxidized D2 peptide ^{327}AWmAPQDQPHENFIFPEEVLP^{348}R (top) and the unmodified form of the peptide (bottom). The oxidation (+15.9949 Da) was localized to ^{329}M.

C. MS/MS spectra of the oxidized D1 peptide ^{313}VIGTwADVLN^{323}R (top) and the unmodified form of the peptide (bottom). The +3.9949 Da modification was identified as an oxidation of ^{317}W to kynurenin.

D. MS/MS spectra of the oxidized D1 peptide ^{239}FGQEEETYNIVAAHGYFG^{257}r (top) and the unmodified form of the peptide (bottom). The +13.9793 Da modification was identified as carbonylation of ^{257}R. Note that although the fragmentation pattern appears to localize the modification only to one of ^{255}F, ^{256}G, or ^{257}R, the latter residue is the only one susceptible to this modification (see Table 1).

Table 1. Oxidative modifications included as variable modifications in the MS database searches. Based on Renzone et al., 2007.

Oxidative modification	Abbreviation	Mass change (Da)	Modifiable residues
Methionine aldehyde	mal	-32.0085	M
Decarboxylation	dcar	-30.0105	D,E
Cysteine hydroxylation	cysh	-15.9772	C
Serine, threonine carbonylation	stcb	-2.0157	S,T
Tryptophan to kynurenine	kyn	3.9949	W
Tryptophan to oxolactone	oxol	13.9793	W
Carbonylation	carb	13.9793	E,I,K,L,P,Q,R,V
General oxidation	go	15.9949	A,D,E,F,H,I,K,L,M,N, P,Q,R,S,T,V,W,Y
Tryptophan to hydroxykynurenine	hkyn	19.9898	W
Dihydroxylation	dihy	31.9898	C,F,K,P,R,W,Y
Methionine sulfone	msul	31.9898	M
Trihydroxylation	trih	47.9847	F,W,Y

Table 2. Oxidative modifications of PSII detected in this study.

Residue	Metal center	Dist. to metal center (Å)	Ox. mod.[a]	B/ SE[b]		Residue	Metal center	Dist. to metal center (Å)	Ox. mod[a]	B/ SE[b]	
D1						**CP43**					
Y235	Q_A	17	go	SE	*	M21[c]	-	-	go	-	
V249	Q_B	7	carb	SE	*	D26	Q_B	28	dcar	SE	
R257	Q_B	10	carb	SE	*	S29	Q_B	30	stcb	SE	*
Y262	Q_B	9	go	SE	*	W34	Q_B	23	kyn	SE	
T316	OEC	25	stcb	B	*	W35	Q_B	22	kyn	SE	
W317	OEC	23	kyn	B		M80	OEC	27	go	SE	
V330	OEC	9	carb	B		W150	Q_A	36	go	SE	*
M331	OEC	9	go	B		T187	OEC	28	stcb	B	
H332	OEC	2	go	B	*	W188	OEC	29	dihy	SE	
E333	OEC	6	carb	B	*	T199	OEC	42	stcb	SE	
						S329	OEC	24	stcb	B	*
D2						L336	OEC	16	go	B	
W14	Q_B	35	go	SE	*	F350	OEC	14	go	B	
W21	Q_B	25	dihy	SE	*	E353	OEC	6	carb	B	*
T221	Q_A	8	stcb	B		T354	OEC	8	go	B	
T238	Q_B	12	stcb	SE		M355	OEC	10	go	B	
S245	Q_A	8	go	B	*	W358	OEC	14	go	SE	
M246	Q_A	7	go	SE		W364	OEC	23	go	B	
W253	Q_A	5	dihy	B	*	E366	OEC	24	dcar	SE	
F314	OEC	12	go	B	*	P371	OEC	28	dihy	SE	
Y315	OEC	15	go	B	*	P385	OEC	28	go	SE	*
M325	OEC	15	go	B		W386	OEC	27	kyn	SE	
R326	OEC	19	carb	B		E393	OEC	20	carb	B	
A327	OEC	19	go	B	*	M395	OEC	14	go	B	
W328	OEC	17	kyn	SE		T396	OEC	16	stcb	B	
M329	OEC	19	go	B		S402	OEC	11	stcb	B	
A330	OEC	22	go	B	*	L403	OEC	18	carb	B	
P331	OEC	24	carb	SE		S405	OEC	19	stcb	B	
Q332	OEC	24	carb	SE	*	T411	OEC	12	stcb	B	
D333	OEC	24	dcar	SE		M468	Q_A	18	go	SE	
P335	OEC	29	go	SE							
E337	OEC	30	go	SE	*						

Green, nearest to Q_A site; purple, nearest to Q_B site; red, nearest to Mn_4Ca cluster (OEC)

* only detected in light-exposed sample

[a] Ox. mod., oxidative modification. For many residues, multiple types of oxidative modifications were observed, though only one type is listed for each residue.

[b] B/SE, buried/solvent-exposed.

[c] CP43-M21 was not resolved in the PSII crystal structure (PDB 3WU2)

The oxidized residues were mapped onto the 1.9 Å crystal structure of PSII from *T. vulcanus* (PDB 3WU2) (Fig. 2). This comparison is enabled by the extremely high degree of conservation between PSII protein sequences of *Synechocystis* and *T. vulcanus* (and in fact across all oxygenic photosynthetic organisms). The D1, D2, and CP43 proteins of these two organisms share 86%, 90%, and 87% sequence identity, and 94%, 99%, and 94% sequence similarity, respectively. Of the 60 oxidized residues, 18 were found on the cytosolic side of PSII, and 42 were found on the lumenal side. These two groups of oxidized residues will be considered separately.

Cytosolic-side oxidized residues

In order to be oxidized, a residue must have been exposed to ROS. As has been pointed out (Frankel et al., 2012; Frankel et al., 2013a), it is reasonable that surface-exposed residues could be oxidized by ROS in the bulk solvent, but buried oxidized residues should reflect the location of ROS production within the complex. Of the 18 cytosolic-side oxidized residues we detected (Fig. 3A), three are buried and 14 are surface-exposed (one, CP43-M21, is not resolved in the crystal structure, but based on its position is very likely to also be surface-exposed) (Table 2). Interestingly, each of the three buried oxidized residues (D2-T221, S245, and W253) are located in the immediate vicinity (5-8 Å) of Q_A, which, like Q_B, is a metal center that is a likely source of ROS production (Frankel et al., 2013b). Five additional residues located in the vicinity (within 15 Å) of Q_A or Q_B were found to be oxidized as well. Overall, eight of the 18 cytosolic-side oxidized residues are located within 15 Å of Q_A or Q_B (Fig. 3B), supporting the idea that these redox-active metal centers are sites of ROS production.

Bricker and co-workers (Frankel et al., 2013b) detected oxidized residues by MS in PSII from spinach, and we compared their results with ours. The results are quite complementary; the Bricker group detected three oxidized residues within 15 Å of Q_A (D1-F239, Q241, and E242),

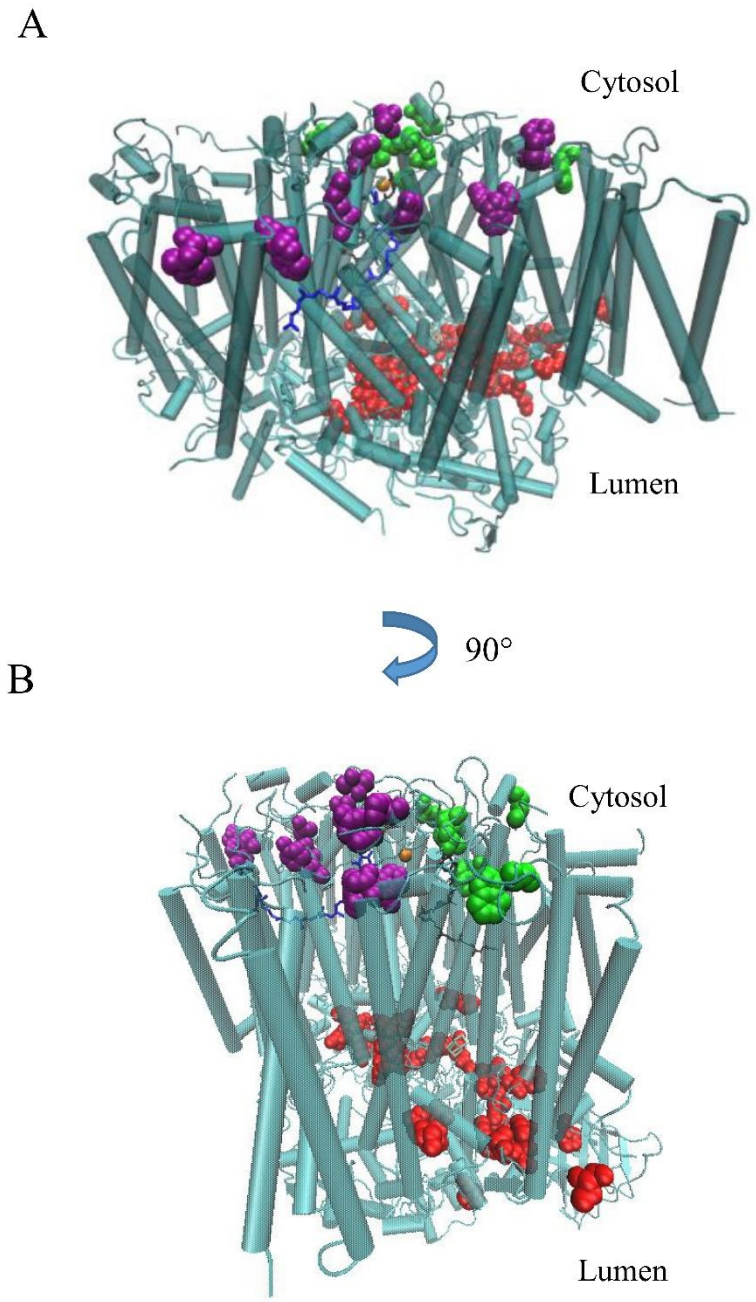


Figure 2. Residues with oxidative modifications detected in this study.

The modified residues on D1, D2, and CP43 were mapped onto the 1.9 Å crystal structure of PSII from *T. vulcanus* (PDB 3WU2). Red- modified residues on the lumenal side of PSII; green- modified residues on the cytosolic side of PSII, in the vicinity of Q_A; purple- modified residues on the cytosolic side of PSII, in the vicinity of Q_B.

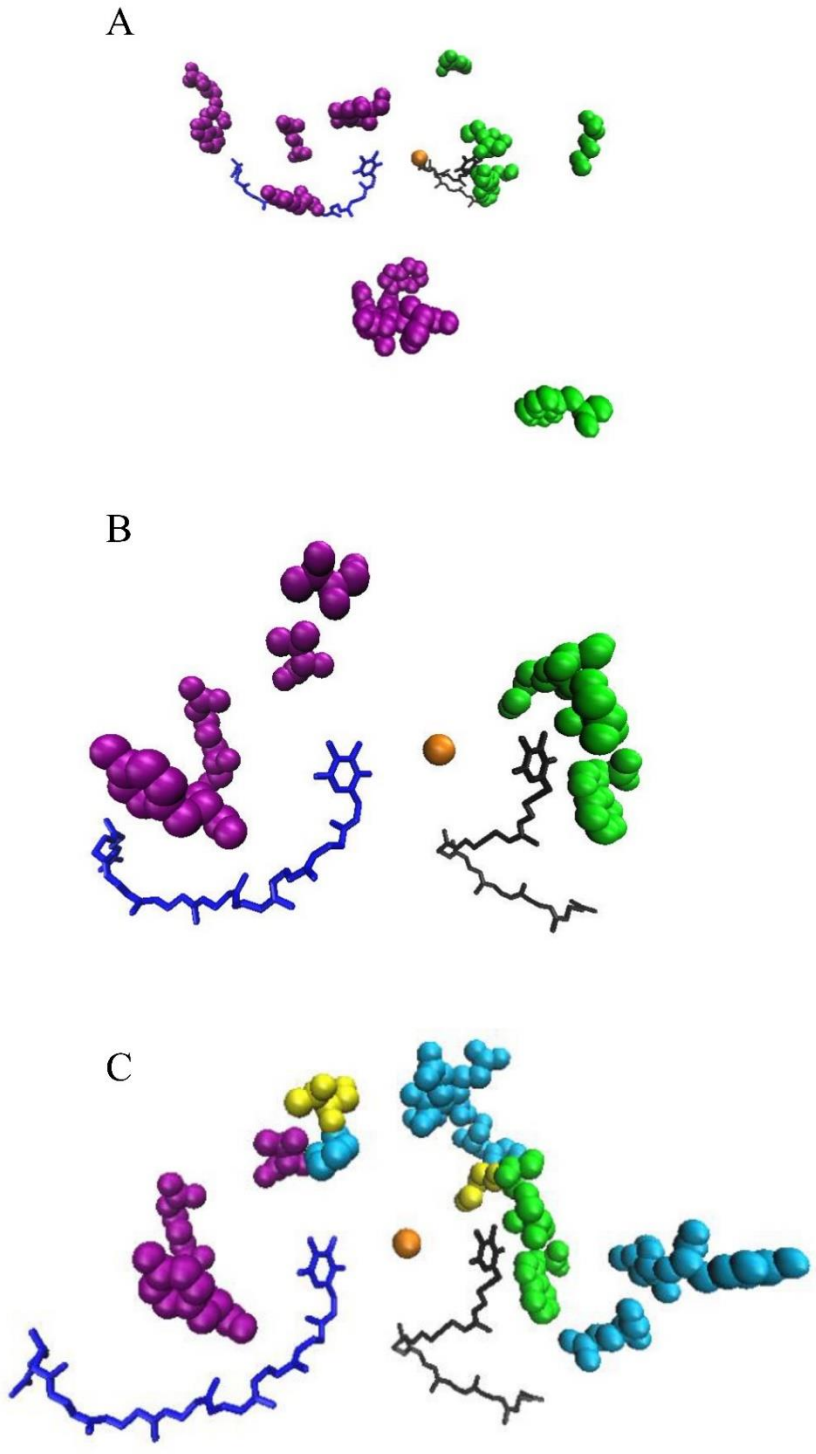


Figure 3. (legend follows)

Figure 3. Cytosolic-side oxidative modifications detected in this study.

Color scheme throughout figure: Black, Q_A; blue, Q_B; orange, non-heme iron; purple, oxidized residues detected nearer to Q_B than Q_A; green, oxidized residues detected nearer to Q_A than Q_B.

A. Overview of the 18 cytosolic-side oxidized residues detected in this study.

B. View of just the oxidized residues within 15 Å of Q_A or Q_B detected in this study.

C. Same residues displayed as in **B**, with the addition of the cytosolic-side oxidized residues detected in Frankel et al. (2012). Cyan, residues only detected in Frankel et al. (2012); yellow, residues detected in both studies.

and one within 15 Å of Q_B (D2-P237), that were not detected in the current study. Conversely, the current study detected three oxidized residues within 15 Å of Q_A (D2-T221, S245, and W253) and three within 15 Å of Q_B (D1-V249, R257, and Y262) that were not detected by the Bricker group. One oxidized residue in this vicinity of Q_A (D2-M246), and one in that of Q_B (D2-T238), were detected in both studies. Overall the two studies show that the residues near Q_A and Q_B are particularly susceptible to oxidative modification (Fig. 3C). Notably, none of the oxidized residues in this vicinity detected by the Bricker group are buried beneath the protein surface, while three detected in our study are. Our study thus strengthens the evidence that Q_A and Q_B are sources of ROS that cause damage to the surrounding PSII residues.

Lumen-side oxidized residues

As mentioned above, our hypothesis was that the oxidized PSII residues would cluster particularly around redox-active metal centers, which are likely sources of ROS. The 42 lumen-side oxidized residues are shown in Fig. 4A. Examination of these results shows that 15 of these residues are indeed in the vicinity (within 15 Å) of the Mn_4Ca cluster, but the remaining 27 are not (the full range of distances is 2-42 Å from the cluster). Interestingly, however, most of the oxidized residues lie in a nearly continuous formation that is centered on the Mn_4Ca cluster and consists of two roughly linear "arms" that radiate outward in opposite directions (Fig 4A). One

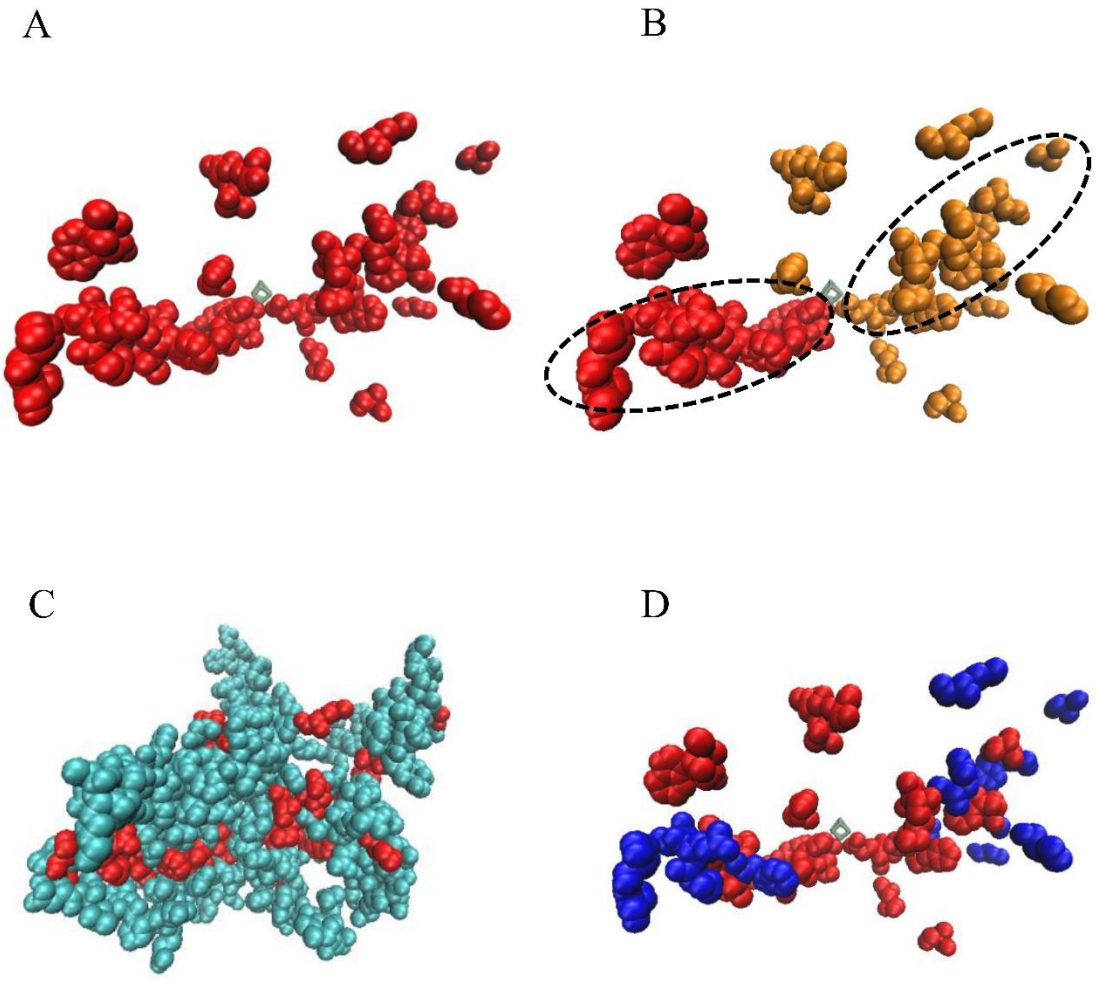


Figure 4. Lumen-side oxidized residues detected in this study.

A. Depiction of the 42 lumen-side oxidized residues detected in this study (red).

B. Same depiction as **A** but the D1/D2 "arm" of oxidized residues is shown in red and the CP43 "arm" is shown in orange. A dashed oval indicates each arm.

C. All of the 255 lumen-side residues that were covered by MS in this study. Some residues are obscured behind others and are not visible in this view. Red- oxidized residues; cyan- non-oxidized residues.

D. Same depiction as **A**, except that in this view, the surface-exposed residues are colored in blue and buried residues are colored in red. VMD (Humphrey et al., 1996) was used to determine if a residue is surface-exposed or buried.

In **A,B,** and **D**, the Mn_4Ca cluster is shown in silver.

arm consists only of D1/D2 residues, and the other arm consists only of CP43 residues (Fig. 4B). To ensure this distinct pattern was not skewed by limited MS coverage of lumen-side residues, all 255 covered lumen-side residues were mapped onto the crystal structure (Fig. 4C). This depiction shows that broad coverage of lumen-side residues was indeed achieved, so it does not appear that the observed geometrical arrangement of the oxidized residues is simply a quirk of limited MS coverage.

An explanation for this observed formation is that the oxidized residues line the walls of a channel through which ROS exit PSII after being generated near the Mn_4Ca cluster. The outermost oxidized residues on both sides of the linear formation are surface-exposed and in contact with bulk solvent (Fig. 4D), so the two arms appear to track ROS travel from the Mn_4Ca cluster all the way out of PSII.

Several computational studies have searched for water, H^+, and/or O_2 channels in PSII, with the assumption that such channels must exist to supply the buried Mn_4Ca cluster active site with substrate water, and for removal of O_2 to minimize oxidative damage to the complex (Anderson, 2001; Murray and Barber, 2007; Ho and Styring, 2008; Gabdulkhakov et al., 2009). These studies have identified around five different channel systems, with nearly the same channel having been identified by multiple groups in several instances (reviewed in Ho 2012). Each of these five channel systems were compared with the oxidized residue formation observed in this study to identify any potential overlap (Fig. 5). Channel G (Gabdulkhakov et al., 2009) contains a stretch of residues (consisting of D2-T316, K317, L320, L321, and E323, and CP47-E364 and S365) that are in continuous contact with the D1/ D2 arm detected in this study, and travel alongside it for a portion of its extension (Fig. 5E-F). Based on its narrow size and amino acid composition, Channel G was proposed to conduct protons, not oxygen species, though water

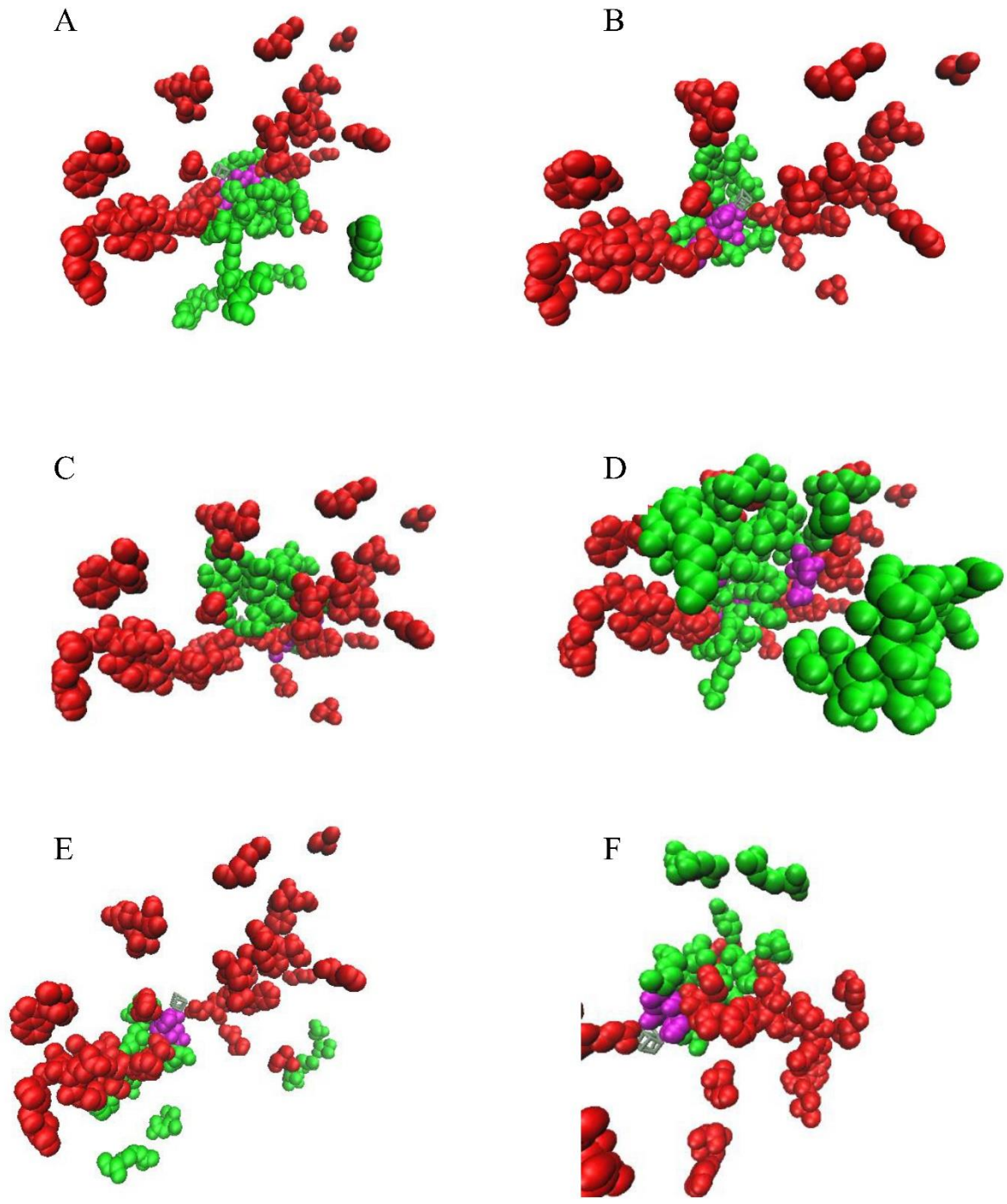


Figure 5. (legend follows)

Figure 5. Comparison of the lumen-side oxidative modified residues with channels previously identified in PSII by computational studies.

The channels depicted in **A, B, C,** and **D** correspond to the "narrow," "broad," "back," and "large" channels, respectively, identified by Ho and Styring (2008). The channel depicted in **E** and **F** corresponds to Channel G identified by Gabdulkhakov et al., 2009. For correspondences of these channels to those identified in the other computational studies cited in the text, see Ho 2012. Green- channel residues not found oxidized in this study; red- oxidized residues not part of the channel; purple- channel residues that were found oxidized in this study; silver- Mn_4Ca cluster.

molecules could also be modeled into it. No other channel overlaps significantly with the D1/D2 arm of oxidized residues, and no channel at all identified in these computational studies overlaps significantly with the CP43 arm (Fig. 5).

As Bricker and co-workers have pointed out (Frankel et al., 2012), these computational studies have important drawbacks to keep in mind. First, they used a 2.9 Å PSII crystal structure, at which resolution amino acid side chain positions cannot be fully resolved. Slightly altered positioning of several side chains could be enough to open or close pathways in such a computational analysis. Second, these studies can only detect pathways present in the static PSII crystal structure, while PSII dynamics could give rise to pathways that are partially blocked in the crystal structure. It is known, e.g., that PSII conformational changes occur throughout the S-state cycle (Bricker et al., 2015).

Bricker and co-workers (Frankel et al., 2012) identified lumen-side oxidative modifications in spinach PSII in a similar MS analysis to ours, and we compared our results with theirs. The Bricker group detected fewer oxidized residues in the vicinity of the Mn_4Ca cluster than in this study (Fig. 6A), but interestingly, when taken together, the results of the two studies support each other. Four oxidized residues on the CP43 arm detected in this study were also found oxidized by Bricker's group (CP43-E353, T354, M355, and E366), and several additional residues were found oxidized on or near the CP43 arm that were not identified in our study (Fig. 6A). Based on their

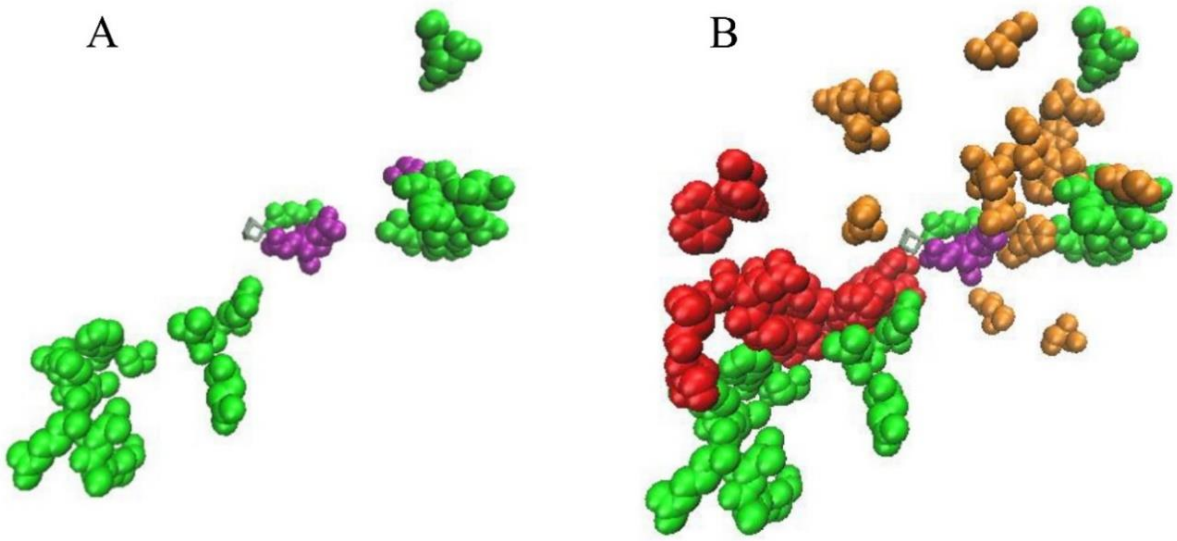


Figure 6. Comparison of lumen-side oxidized residues detected by Frankel et al. (2012) and the current study.

A. Lumen-side oxidized residues detected by Frankel et al. (2012). Green- only detected by Frankel et al. (2012). Purple- also detected in the current study. Silver- Mn_4Ca cluster.

B. Lumen-side oxidized residues detected in both studies. Green- only detected by Frankel et al. (2012). Red- D1/D2 residues only detected in this study. Orange- CP43 residues only detected in this study. Purple- residues detected in both studies. Silver- Mn_4Ca cluster.

Note: Some additional lumen-side residues detected by Frankel et al. (2012) that are not relevant to this discussion are not displayed in this figure.

results, Bricker and co-workers proposed that the four oxidized CP43 residues they detected located within 15 Å of the Mn_4Ca cluster, may constitute part of an ROS exit channel leading from the cluster to the surface of PSII. However, additional oxidized residues that could complete this channel by leading all the way to the surface of the complex, were not detected to support their hypothesis. In contrast, the oxidized CP43 residues we identified do lead all the way from the

Mn_4Ca cluster to the PSII surface, in a way that supports and elaborates on the original Bricker proposal. Superimposing the two studies' results on each other shows a formation of oxidized residues that originates at the Mn_4Ca cluster and widens somewhat as it radiates outward towards the surface of PSII (Fig. 6B).

No common oxidized residues on the D1/D2 arm were detected in both studies, but Bricker and co-workers found three D2 and ten CP47 oxidized residues located in two clusters nearby (D2-E344, L346, and R348, and CP47-R358, M359, P360, T361, F363, E364, T365, R423, A424, and Q425) (Fig. 6A). Of these, the nearest residue to the Mn_4Ca cluster is R348, 13 Å away. The positioning of these residues and their distance from the Mn_4Ca cluster do not clearly indicate a ROS pathway leading away from it, and the authors did not propose such a pathway (Frankel et al., 2012). However, our study detected more total oxidized residues, and as described above, a D1/D2 arm leading from the Mn_4Ca cluster all the way to the PSII surface. Interestingly, superimposing the two studies' results on each other shows that the oxidized D1/D2 residues we detected "fill in the gaps" between the Bricker group's D2 and CP47 oxidized residues and the Mn_4Ca cluster (Fig. 6B). Taking the two studies together, the Bricker residues are in fact part of a near-continuous path of oxidized residues leading from the Mn_4Ca cluster to the PSII surface. Similar to the CP43 arm, the combined set of oxidized D1/D2/CP47 residues shows a formation that originates at the Mn_4Ca cluster and widens somewhat as it radiates outward towards the surface of PSII (Fig. 6B).

Bricker and co-workers (Frankel et al., 2013a) studied their system further by performing synchrotron radiolysis of water and exposing their PSII sample to the resulting hydroxyl radical species for 0, 4, 8, or 16 s. They reasoned that for a residue to be oxidized in this experiment, it must have been in contact with water, from which the hydroxyl radicals were generated. Mapping

buried oxidized residues on the PSII crystal structure could therefore indicate the location of water channels within PSII. As mentioned above, only four of the 42 oxidized residues (10%) that we detected were found to be oxidized by Bricker and co-workers before the radiolysis experiment (0 s irradiation). However, after 4, 8, or 16 s of irradiation, they found that 24 of the 42 oxidized residues (57%) that we detected, were oxidized. Of these, 16 are buried, suggesting that at least 16 of the 27 buried oxidized lumen-side residues that we detected are indeed in contact with a water-filled channel in PSII. This finding is consistent with the possibility that the oxidized residue arms we detected are ROS exit pathways, as an ROS exit channel that is exposed to bulk solvent would be expected to contain water as well.

It should be noted that a quantitative analysis was not performed on the modifications identified in this study; that is, the absolute or relative *extent of modification* of a given peptide has not been determined. An absolute quantification would require knowing the relative ionization efficiency or "flyability" in the mass spectrometer of the modified and unmodified forms of each peptide; see chapter 1 for a further discussion of this issue. Quantification of the fold-change of a modified peptide between light- and dark-incubated samples, relative to the unmodified form, similar to analysis of a footprinting experiment, is feasible with this dataset using the label-free techniques described in chapter 1. Such an analysis is currently ongoing, and preliminary results show that the relative modification extent of many oxidized residues increases significantly following light exposure.

In conclusion, we have found that the total number of oxidized residues in the PSII proteins D1, D2, and CP43 increases by around 50% following light incubation. The D1 protein showed the most pronounced increase (3-fold), consistent with the fact that it undergoes photodamage fastest (Nickelsen and Rengstl, 2013) and has the shortest half-life of all PSII proteins (Yao et al.,

2012). The oxidized residues we identified are arranged in two nearly continuous, roughly linear arms originating at the Mn_4Ca cluster and radiating outward all the way to the surface of PSII. One arm consists of D1/D2 residues, and the other of CP43 residues. Both arms are consistent with the oxidized residue pattern detected previously by Bricker and co-workers (Frankel et al., 2012), but the current results provide a significantly more detailed depiction and extend the finding by showing a nearly continuous path on both arms from the Mn_4Ca cluster all the way to the PSII surface. We propose that these arms represent exit pathways for ROS generated at the Mn_4Ca cluster. Our results provide the most detailed molecular evidence to date for the existence of ROS exit pathways within PSII, a critical mechanism long believed to exist for protecting PSII from photodamage.

Materials and methods

Cell culture and PSII purification

PSII from the His47 strain of *Synechocystis* 6803 was used in this study. For cell culture and PSII purification procedures, see chapter 2.

Sample preparation and proteolytic digestion

Two His47 PSII samples (two biological replicates) consisting of 1 μg Chl *a* were precipitated using the 2D-cleanup kit (GE Healthcare) to remove salt and detergent. For the light-incubated condition, the PSII samples were incubated at 30° C and 55 μmol photons $m^{-2}s^{-1}$ for 6 h prior to precipitation. The dark-incubated samples were treated identically to the light-incubated samples, but were wrapped in silver foil during the incubation. After precipitation, the protein pellets were resuspended in 15 μL 8 M urea. Disulfide reduction was performed by incubation in a final concentration of 2.5 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 min at 37 °C.

Alkylation of cysteines was then performed by addition of iodoacetamide to a final concentration of 5 mM, with incubation at room temperature in the dark for 30 min. Digestion was performed in two stages; first, lysyl endopeptidase was added at around 1:20 μg/μg protein, followed by incubation for 2 h at 37 °C. After 2 h, samples were diluted 1:7 in 100 mM Tris pH 8.5, 10 mM $CaCl_2$ trypsin buffer, with a final ratio of around 1 μg trypsin/20 μg protein, and incubated at at 37 °C overnight. The digestion was stopped by acidifying the sample to a final concentration of 1% formic acid. Samples were then analyzed by LC-MS/MS.

LC-MS/MS

Samples were analyzed as described in chapter 2, with the following adjustments. For LC, a linear 82-min gradient from 2-43% solvent B (80% acetonitrile, 20% water, 0.1% formic acid) was used, followed by a linear 30-min gradient from 43-98% solvent B, followed by a 5-min hold at 95% solvent B. Automatic gain control for MS/MS was set at 2.5×10^5 ions and a maximal injection time of 100 ms. Charge states other than 2-7 were excluded, and each sample was analyzed in triplicate.

Data analysis

The raw LC-MS/MS files were searched for *Synechocystis* PSII proteins using PEAKS (ver. 7.0, Bioinformatics Solutions, Inc., Waterloo, ON, Canada). The oxidative modifications listed in Table 1 were included in the search, as well as carbamidomethylation (C,D,H,K,E, peptide N-terminus), +57.0215 Da; deamidation (N,Q), +0.9840 Da; acetylation of protein N-terminus, +42.0106 Da; and carbamylation (K, peptide N-terminus), +43.0058 Da. Peptides were identified with a 0.1% false discovery rate. For every candidate oxidized residue that met the 0.1% false discovery threshold, MS/MS spectra were inspected manually to confirm data quality, and spectra

were rejected if the oxidation could not be localized to a single residue. All protein visualization was performed using VMD (Humphrey et al., 1996).

References

Anderson, J.M. (2001). Does functional photosystem II complex have an oxygen channel? *FEBS Lett.* 488, 1-4.

Aro, E.-M., Suorsa, M., Rokka, A., Allahverdiyeva, Y., Paakkarinen, V., Saleem, A., et al. (2005). Dynamics of photosystem II: a proteomic approach to thylakoid protein complexes. *J. Exp. Bot.* 56, 347-356.

Barber, J., and Andersson, B. (1992). Too much of a good thing—light can be bad for photosynthesis. *Trends Biochem. Sci.* 17, 61-66.

Bricker, T.M., Mummadisetti, M.P., and Frankel, L.K. (2015). Recent advances in the use of mass spectrometry to examine structure/function relationships in photosystem II. *J. Photochem. Photobiol. B* 152, 227-246.

Dreaden, T.M., Chen, J., Rexroth, S., and Barry, B.A. (2011). N-formylkynurenine as a marker of high light stress in photosynthesis. *J. Biol. Chem.* 286, 22632-22641.

Frankel, L.K., Sallans, L., Bellamy, H., Goettert, J.S., Limbach, P.A., and Bricker, T.M. (2013a). Radiolytic mapping of solvent-contact surfaces in Photosystem II of higher plants: Experimental identification of putative water channels within the photosystem. *J. Biol. Chem.* 288, 23565-23572.

Frankel, L.K., Sallans, L., Limbach, P.A., and Bricker, T.M. (2013b). Oxidized amino acid residues in the vicinity of Q_A and $Pheo_{D1}$ of the Photosystem II reaction center: Putative generation sites of reducing-side reactive oxygen species. *PLoS ONE* 8, e58042.

Frankel, L.K., Sallans, L., Limbach, P.A., and Bricker. T.M. (2012). Identification of oxidized amino acid residues in the vicinity of the Mn_4CaO_5 cluster of Photosystem II: Implications for the identification of oxygen channels within the photosystem. *Biochemistry* 51, 6371-6377.

Gabdulkhakov, A., Guskov, A., Broser, M., Kern, J., Mueh, F., Saenger, W., et al., (2009). Probing the accessibility of the Mn_4Ca cluster in Photosystem II: Channels calculation, noble gas derivatization, and cocrystallization with DMSO. *Structure* 17, 1223-1234.

Grabolle, M., and Dau, H. (2005). Energetics of primary and secondary electron transfer in Photosystem II membrane particles of spinach revisited on basis of recombination-fluorescence measurements. *Biochim. Biophys. Acta* 1708, 209-218.

Ho, F.M. (2012). Structural and mechanistic investigations of Photosystem II through computational methods. *Biochim. Biophys. Acta* 1817, 106-120.

Ho, F.M., and Styring, S. (2008). Access channels and methanol binding site to the $CaMn_4$ cluster in photosystem II based on solvent accessibility simulations, with implications for substrate water access. *Biochim. Biophys. Acta* 1777, 140-153.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD - Visual Molecular Dynamics. *J. Molec. Graphics.* 14, 33-38.

Kasson, T.M.D., Rexroth, S., and Barry, B.A. (2012). Light-induced oxidative stress, N-formylkynurenine, and oxygenic photosynthesis. *PLoS ONE* 7, 42220.

Krieger-Liszkay, A., Fufezan, C., and Trebst, A. (2008). Singlet oxygen production in photosystem II and related protection mechanism. *Photosynth. Res.* 98, 551-564.

Liu, H., Chen, J., Huang, R.-C., Weisz, D., Gross, M.L., and Pakrasi, H.B. (2013a). Mass spectrometry-based footprinting reveals structural dynamics of Loop E of the chlorophyll-binding protein CP43 during Photosystem II assembly in the cyanobacterium *Synechocystis* 6803. *J. Biol. Chem.* 288, 14212-14220.

Liu, H., Zhang, H., Niedzwiedzki, D., Prado, M., He, G., Gross, M.L., et al. (2013b). Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. *Science* 342, 1104-1107.

Murray, J.W., and Barber, J. (2007). Structural characteristics of channels and pathways in photosystem II including the identification of an oxygen channel. *J. Struct. Biol.* 159, 228-237.

Nakamura, T., Dohmae, N., and Takio, K. (2004). Characterization of a digested protein complex with quantitative aspects: An approach based on accurate mass chromatographic analysis with Fourier transform-ion cyclotron resonance mass spectrometry. *Proteomics* 4, 2558-2566.

Nickelsen, J., and Rengstl, B. (2013). Photosystem II assembly: From cyanobacteria to plants. *Annu. Rev. Plant Biol.* 64, 609-635.

Pospíšil, P. (2009). Production of reactive oxygen species by photosystem II. *Biochim. Biophys. Acta* 1787, 1151-1160.

Renzone, G., Salzano, A.M., Arena, S., D'Ambrosio, C., and Scaloni, A. (2007). Mass spectrometry-based approaches for structural studies on protein complexes at low-resolution. *Curr. Proteom.* 4, 1-16.

Sharma, J., Panico, M., Shipton, C.A., Nilsson, F., Morris, H.R., and Barber, J. (1997). Primary structure characterization of the photosystem II D1 and D2 subunits. *J. Biol. Chem.* 272, 33158-33166.

Takahashi, S., and Badger, M.R. (2011). Photoprotection in plants: A new light on photosystem II damage. *Trends Plant Sci.* 16, 53-60.

Vass, I. (2012). Molecular mechanisms of photodamage in the Photosystem II complex. *Biochim. Biophys. Acta* 1817, 209-217.

Vass, I., and Cser, K. (2009). Janus-faced charge recombinations in photosystem II photoinhibition. *Trends Plant Sci.* 14, 200-205.

Yao, D.C.I., Brune, D.C., and Vermaas, W.F.J. (2012). Lifetimes of photosystem I and II proteins in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett*. 586, 169-173.

Chapter 4

Structural localization of PsbQ in Photosystem II using chemical cross-linking and mass spectrometry

This chapter was adapted from:

Liu, H., Zhang, H., Weisz, D.A., Vidavsky, I., Gross, M.L., Pakrasi, H.B. (2014). MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. *Proc. Natl. Acad. Sci. USA* 111, 4638-4643.

Liu, H., Weisz, D.A., and Pakrasi, H.B. (2015). Multiple copies of the PsbQ protein in a cyanobacterial photosystem II assembly intermediate complex. *Photosynth. Res.* 126, 375-383.

Summary

PsbQ is a lumenal extrinsic protein component that regulates the water splitting activity of Photosystem II in plants, algae and cyanobacteria. However, PsbQ is not observed in the currently available crystal structures of PSII from thermophilic cyanobacteria. The structural location of PsbQ within the PSII complex has, therefore, remained unknown. In this study, we purified PSII complexes from *Synechocystis* 6803 by means of a polyhistidine tag on PsbQ. To determine the binding site of PsbQ within PSII, we subjected the purified Q-His-PSII to chemical cross-linking followed by immunodetection and LC-MS/MS analysis. Our results demonstrate that PsbQ is closely associated with the PsbO and CP47 proteins, as revealed by cross-links identified between PsbQ-^{120}K and PsbO-^{180}K, between PsbQ-^{120}K and PsbO-^{59}K, and between PsbQ-^{102}K and CP47-^{440}D. We further show that genetic deletion of the *psbO* gene results in the complete absence of PsbQ in PSII complexes as well as the loss of the dimeric form of PSII. Overall, our data provide the first molecular-level description of the enigmatic binding site of PsbQ in PSII in a cyanobacterium. These results also help us understand the sequential incorporation of the PsbQ protein during the PSII assembly process as well as its stabilizing effect on the oxygen evolution activity of PSII. Interestingly, during purification of Q-His-PSII, we isolated a novel PSII assembly intermediate in addition to the mature PSII complex used in the cross-linking experiment. This new complex, which we refer to as PSII-Q4, contained four copies of the PsbQ protein per PSII monomer, instead of the expected one copy. PSII-Q4 lacked two other lumenal extrinsic proteins, PsbU and PsbV, which are present in the mature PSII complex. We suggest that PSII-Q4 is a late PSII assembly intermediate that is formed just before the binding of PsbU and PsbV, and incorporate these results into an updated model of PSII assembly.

Introduction

Photosystem II (PSII) functions as a light-driven, water-plastoquinone oxidoreductase in oxygenic photosynthesis. PSII is a membrane protein complex containing more than 20 protein subunits. Early biochemical investigations established that at least seven major intrinsic proteins are required for oxygen evolution: CP47, CP43, D1, D2, the α and β subunits of cytochrome b_{559} and PsbI (Bricker and Ghanotakis, 1996). Additionally, a number of low molecular mass intrinsic polypeptides are associated with these seven major polypeptides.

Despite the relatively conserved overall functioning of PSII, the lumenal extrinsic PSII protein comlement varies significantly across different phyla (Kashino et al., 2002; Thornton et al., 2004; Roose et al., 2007b; Bricker et al., 2012). In higher plants and green algae, four extrinsic proteins, PsbO, PsbP, PsbQ, and PsbR, are required to support maximal rates of oxygen evolution under physiological conditions. In contrast, PsbO, PsbU, and PsbV (cytochrome c_{550}) play analogous roles in cyanobacteria and red algae (Roose et al., 2007b). More recently, it was discovered that homologs of PsbQ and PsbP exist in cyanobacteria (Shen and Inoue, 1993; Kashino et al., 2002; Thornton et al., 2004). These differences are especially noteworthy given that the oxygen-evolving complex (OEC) itself is practically unaltered from cyanobacteria to green algae and higher plants (Umena et al., 2011).

The presence of PsbQ in cyanobacterial PSII was discovered by Kashino et al. (2002) when analyzing the complete protein complement of isolated PSII complexes from the HT3 strain of *Synechocystis* sp. PCC 6803, which contains a C-terminal His_6-tag on the CP47 protein (Bricker et al., 1998). The physiological role of PsbQ was subsequently investigated in several labs through phenotypic analysis of either single or double mutants lacking PsbQ as well as other extrinsic PSII proteins (Thornton et al., 2004; Summerfield et al., 2005; Kashino et al., 2006). Direct evidence

for the binding of PsbQ to PSII was obtained using isolated PSII complexes from the Q-His strain that contains a C-terminal His_8-tag on PsbQ (Roose et al., 2007a). It was shown that PsbQ defines cyanobacterial PSII complexes with high activity and stability. Although the x-ray crystal structure of isolated cyanobacterial PsbQ has been solved to near-atomic resolution, PsbQ has not been identified in any of the crystal structures of cyanobacterial PSII from the thermophilic cyanobacteria *Thermosynechococcus elongatus* BP-1 (Guskov et al., 2009) and *Thermosynechococcus vulcanus* (Umena et al., 2011). Owing to the absence of PsbQ in these structures, the binding site of this protein within the PSII complex has remained unknown.

In the current study, we have used chemical cross-linking followed by immunodetection and mass spectrometry to investigate the structural location of PsbQ in PSII from *Synechocystis* 6803. Based on our analysis, we propose a model in which PsbQ binds to PSII through its close association with PsbO and CP47, thereby stabilizing the PSII dimer and permitting high rates of oxygen evolution in this cyanobacterium. During purification of Q-His-PSII, we also identified a novel PSII complex that contains four copies of PsbQ per PSII monomer. Based on its subunit composition, we propose that it is a late PSII assembly intermediate formed after the binding of PsbO, but before the binding of PsbU and PsbV.

Results

Using a BLASTp search, we examined the prevalence of PsbQ across the cyanobacterial phylum. Homologs to *Synechocystis* 6803 PsbQ were detected in 97 other cyanobacterial species (Fig. 1 and Fig. S1). PsbQ is present in an evolutionarily diverse variety of diazotrophic and non-diazotrophic strains, but was not found in *Gloeobacter* species, an ancient lineage of cyanobacteria that do not have thylakoid membranes. Therefore, we propose that the PsbQ gene evolved in cyanobacteria concurrently with, or soon after, the development of the thylakoid system. Notably,

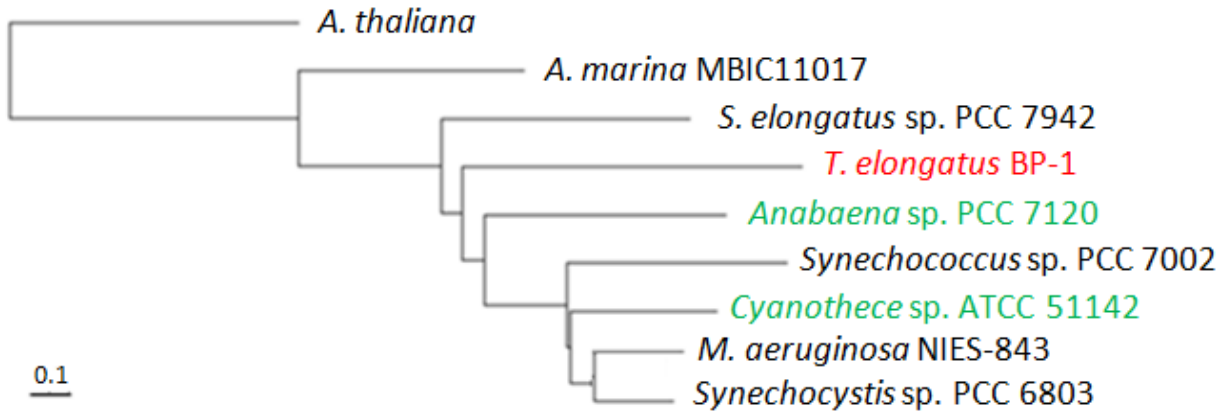
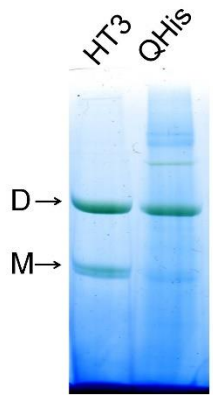


Figure 1. Phylogenetic tree showing evolutionary relationship between PsbQ protein sequences from selected cyanobacterial species. *A. thaliana* serves as an outgroup. Green, nitrogen-fixing; red, thermophilic. *A. thaliana*, *Arabidopsis thaliana*; *A. marina* MBIC11017, *Acaryochloris marina* MBIC11017*; T. elongatus* BP-1, *Thermosynechococcus elongatus* BP-1; *M. aeruginosa* NIES-843, *Microcystis aeruginosa* NIES-843. Scale bar indicates substitutions per nucleotide.

PsbQ homologs are present in the thermophiles *T. elongatus* and *T. vulcanus*, although crystallized PSII from them lack PsbQ (see above, Guskov et al., 2009; Umena et al., 2011).

We have previously described the isolation and characterization of highly active and stable PSII from the Q-His strain of *Synechocystis* 6803 (Roose et al., 2007a). In the present study, we took advantage of its availability to pursue the structural location of PsbQ within functional PSII. Oxygen-evolving PSII is usually found as both a dimer and monomer, with the isolated dimer being more active than the monomer (Nowaczyk et al., 2006). Blue-native gel (BN-gel) and SDS-PAGE protein profile analysis of HT3-PSII and Q-His-PSII (Fig. 2) showed that the HT3-PSII preparation could be resolved into two major green bands (Fig. 2A), corresponding to PSII dimer

A.



B.

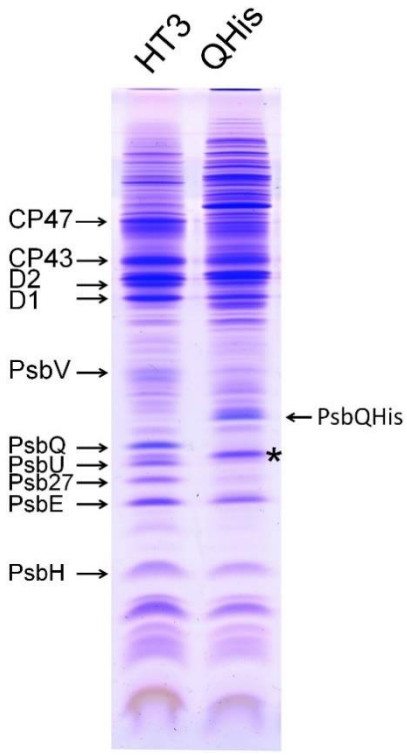


Figure 2. Native-gel and SDS-PAGE analysis of HT3-PSII and Q-His-PSII.

A. Blue native gel analysis of HT3-PSII and Q-His-PSII. Arrows indicate the dimeric (D) and monomeric (M) forms of PSII.

B. Polypeptide profiles of isolated HT3PSII and QHisPSII. The positions of major PSII protein components are indicated on the left. The C-terminally His-tagged PsbQ and the Psb28 protein (asterisk) are indicated. Each lane was loaded with sample containing 5 µg of Chl *a*.

and PSII monomer, respectively. Interestingly, only the dimeric form was observed for the Q-His-PSII preparation. This observation may partially explain the higher oxygen evolving activity of Q-His-PSII compared to HT3-PSII (Roose et al., 2007a), presumably because the His_6-tag on CP47 in the HT3 strain enables the isolation of a mixture of both a more active PSII dimer and a less active PSII monomer.

Cross-linking results

As discussed in Chapters 1 and 2, chemical cross-linkers can covalently link amino acid pairs found in close proximity to each other in a protein or a protein complex (Sinz, 2006; Leitner et al., 2010; Rappsilber, 2011; Zheng et al., 2011; Tabb, 2012). In this study, we used 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide (EDC), a zero-length cross-linker that can link carboxylate groups (from aspartate (D), glutamate (E) side chains, and protein C-termini) to primary amine groups (from lysine (K) or protein N-termini). We also used the thiol-cleavable 12-Å cross-linker 3,3´-dithiobis(sulfosuccinimidylpropionate) (DTSSP) and the non-cleavable bis(sulfosuccinimidyl)suberate (BS^3), both of which can link two amino acid residues containing primary amine groups. Similar cross-linking approaches have been used to reveal a close association between PsbO and CP47 in PSII (Bricker et al., 1988; Seidler, 1996), a result that was subsequently validated by crystallographic studies (Zouni et al., 2001). To identify the binding site of PsbQ in the PSII dimer, we chose to cross-link Q-His-PSII instead of HT3-PSII, because the former contains a more biochemically homogeneous population than HT3-PSII (see above).

We first probed our cross-linked samples with anti-PsbQ antibodies and found that treatment of Q-His-PSII with EDC and DTSSP generated similar, but not identical, cross-linked products, with apparent molecular masses of 23 kDa (a), 27 kDa (b), 36 kDa (c), 52 kDa (d), 57 kDa (e), 75 kDa (g), 80 kDa (f), and 96 kDa (h) (Fig. 3A, 3B). In the absence of EDC and DTSSP,

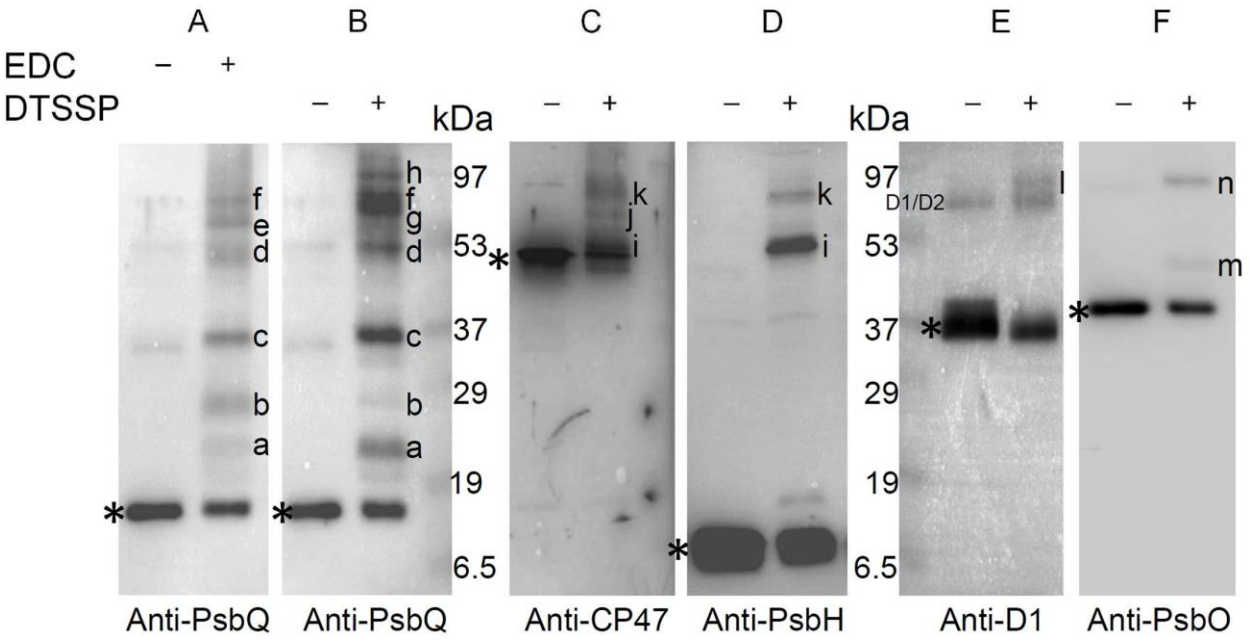


Figure 3. Chemical cross-linking of PsbQ with other PSII subunits.

A. Detection of cross-linked species of PsbQ. Lane 1: QHisPSII; Lane 2: QHisPSII+EDC. Immunoblots were probed with anti-PsbQ antibodies.

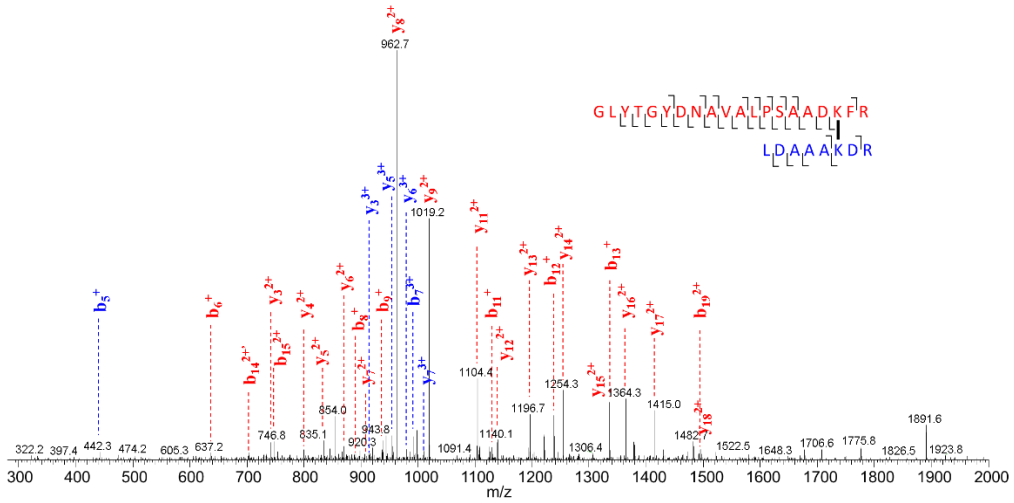
B, C, D, E, F. DTSSP cross-linking in Q-His-PSII. Immunoblots were probed with anti-PsbQ, anti-CP47, anti-PsbH, anti-D1, and anti-PsbO antibodies, respectively. Lane 1: Q-His-PSII; Lane 2: Q-His-PSII+DTSSP. Each lane was loaded with sample containing 0.2 µg of Chl*a*. Uncross-linked proteins are labeled with asterisks.

C-terminally His_8-tagged PsbQ (PsbQ-His) migrated at 17 kDa (Fig. 3A, 3B). PsbQ is N-terminally lipid-modified and, thus, anchored to the thylakoid membrane or to the transmembrane domain of PSII (Thornton et al., 2004; Kashino et al., 2006; Ujihara et al., 2008). However, the observed PsbQ-containing cross-linked products (Fig. 3A, 3B) indicated that PsbQ is in close contact with other proteins and possibly bind to PSII via inter-protein interactions. We further examined the DTSSP cross-linked products by using anti-CP47 (Fig. 3C), anti-PsbH (Fig. 3D), anti-D1 (Fig. 3E), and anti-PsbO (Fig. 3F) antibodies. Products d and i may be identical and

represent a cross-link between PsbQ and CP47, or products d and m may be identical and represent a cross-link between PsbQ and PsbO. Similarly, products h and n may also represent a cross-linked complex containing PsbQ and PsbO, and alternatively any of the PsbQ-containing cross-linked products observed in Fig. 3B may be linked to any of the numerous other PSII proteins. Based on these data, we suggest structural interactions between PsbQ and CP47 and/or PsbQ and PsbO.

To determine the cross-link partner(s) of PsbQ more definitively, we used HPLC coupled with mass spectrometry (LC/MS/MS). As discussed in Chapters 1 and 2, MS-based techniques can be used to identify cross-linked species with high resolution and mass accuracy. Careful interpretation of the mass and the product-ion spectra of candidates identified in database searches can reveal the exact amino acid pairs that are cross-linked (Petrotchenko and Borchers, 2010). To minimize sample losses, we subjected our cross-linked products to direct in-solution trypsin digestion instead of in-gel digestion, followed by LC/MS/MS (see details in Materials and Methods). We detected two cross-links between PsbQ and PsbO, as seen in the product-ion (MS/MS) spectra for a representative PsbO-PsbQ cross-linked species (Fig. 4A) identified by a database search (Fig. S2). Examining the product-ion spectrum, we could verify that the identified link between residues ^{120}K of PsbQ and ^{180}K of PsbO is a confident search result (Fig. 4A). (Note the peptide-sequence numbering used in this study hereafter is based on the PSII structure (Umena et al., 2011) (PDB ID: 3ARC) and the PsbQ structure (Jackson et al., 2010) (PDB ID: 3LS1) for literature consistency, unless otherwise stated). Residue ^{120}K of PsbQ was also linked to residue ^{59}K of PsbO (Fig. S2). Considering the cross-linker arm span of DTSSP (12 Å), we were able to localize ^{120}K of PsbQ approximately between ^{59}K and ^{180}K of PsbO.

A.



B.

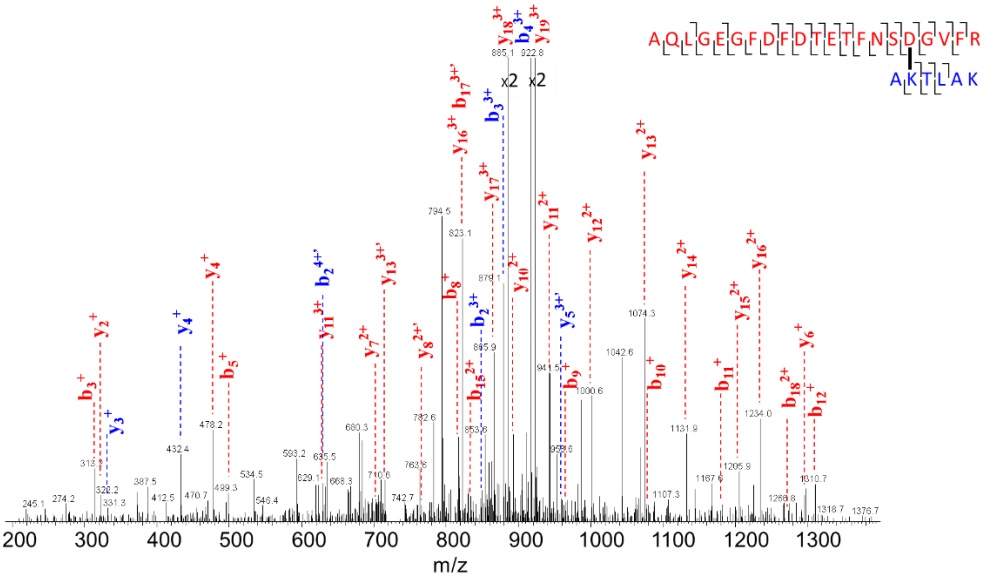


Figure 4. Product-ion (MS/MS) spectra obtained for cross-linked peptides.

A. Product-ion (MS/MS) spectra of the intermolecular cross-link between peptide GLYTGYDNAVALPSAAD^{180}KFR (PsbO) and LDAAA^{120}KDR (PsbQ). Cross-links were formed specifically between ^{180}K of PsbO (in peptide 163-182) and ^{120}K of PsbQ (in peptide 115-122) upon treatment with DTSSP.

B. Product-ion (MS/MS) spectra of the intermolecular cross-link between peptides AQLGEGFDFDTETFNS^{440}DGVFR (CP47) and A^{102}KTLAK (PsbQ). ^{440}D of CP47 was cross-linked to ^{102}K of PsbQ upon treatment with EDC.

In addition to the inter-protein cross-links between PsbQ and PsbO, we detected two inter-protein cross-links between CP47 and PsbO, as well as many intra-protein cross-links within CP47 and PsbO (See Table S1). Considering the difference in protein sequence between PsbO in the crystal structure (PDB ID: 3ARC) from *T. vulcanus* and PsbO from *Synechocystis* 6803, we generated a homology model of *Synechocystis* 6803 PsbO ($PsbO_{sc}$) by using I-TASSER (Zhang, 2008) and compared the model to the crystal structure version by using the PyMOL and APBS programs (Baker et al., 2001; Delano, 2002) (Fig. S4). Our cross-linking results are consistent with the structural relationship between CP47 and PsbO observed in the PSII crystal structures from the two thermophilic cyanobacteria (Zouni et al., 2001; Umena et al., 2011). We consider these results as positive controls. Using immunological methods, Bricker and coworkers have earlier determined that PsbO forms cross links to CP47 in PSII (Bricker et al., 1988). With current mass-spectrometry-based cross-link analysis and the availability of PSII crystal structures, we have confirmed this early finding and added finer molecular detail.

Analysis of the EDC-treated sample revealed a cross-link between ^{102}K of PsbQ and ^{440}D on Loop E of CP47 (Fig. 4B, Fig. S5). Given that EDC is a carboxylate-to-amine zero-length cross-linker, this cross-link implies that these two residues interact by complementary charges and are within van der Waals contact of each other. This result particularly helps to define a close association between PsbQ and CP47.

psbO deletion mutant

To gain further insight into the *in vivo* structural and functional relationships between PsbO and PsbQ, we generated a *psbO* deletion mutation in the genetic background of HT3 (henceforth HT3-ΔO). Using the HT3 (Bricker et al., 1998), HT3-ΔQ (Kashino et al., 2006) and the newly generated HT3ΔO strains, we then isolated three types of PSII complexes, with the normal

complement of extrinsic proteins (HT3), with PsbQ deleted (HT3-ΔQ), and with PsbO deleted (HT3-ΔO), respectively. BN gel analysis (Fig. 5A) of the HT3-PSII, HT3Δ-Q-PSII, and HT3-ΔO-PSII preparations showed that HT3-PSII was resolved to two major green bands (Fig. 5A, lane 1, and Liu et al., 2011b). The PSII dimer/monomer ratio in HT3-ΔQ-PSII was not notably different than that of HT3-PSII (Fig. 5A, lane 2), except that RC47 (the CP43-less PSII precursor) was more pronounced than in HT3PSII, indicative of the possible role of PsbQ in protection of functional PSII or during PSII biogenesis. Both SDS-PAGE polypeptide profile (Fig. 5B) and immunodetection (Fig. 5C) analysis support that the PsbO protein level did not significantly change in the absence of PsbQ. Deletion of *psbO* (HT3ΔO), however, resulted in complete elimination of dimeric PSII (Fig. 5A) and the absence of PsbQ (Fig. 5B, Fig. 5C). It appears that binding of PsbO to PSII is independent of PsbQ, but PsbQ is unable to bind to PSII in the absence of PsbO, a protein which presumably contributes to the inter-monomer interactions with CP47 and stabilization of the PSII dimer (De Las Rivas and Barber, 2004; Bentley and Eaton-Rye, 2008).

Purification and biochemical characterization of the PSII-Q4 complex ("Complex 2")

The PSII-Q4 complex was purified separately from Q-His-PSII using nickel-affinity chromatography and a two-step elution gradient of histidine-containing buffer (Fig. 6a, red line). The two elution peaks labeled in the chromatogram correspond to two distinct protein omplexes, at this stage referred to as Complex 1 and Complex 2 (later identified as the PSII dimer (PSII-D) and PSII-Q4, respectively). They were collected separately and subjected to biochemical characterization. Complexes 1 and 2 represent approximately 53% and 47% of the total PSII yield, respectively. In one of the previous studies from our lab in which PSII from the Q-His strain was purified, elution was performed with a single linear gradient (Roose et al., 2007a), and the different elution conditions used in this study likely account for the isolation of a second PSII complex.

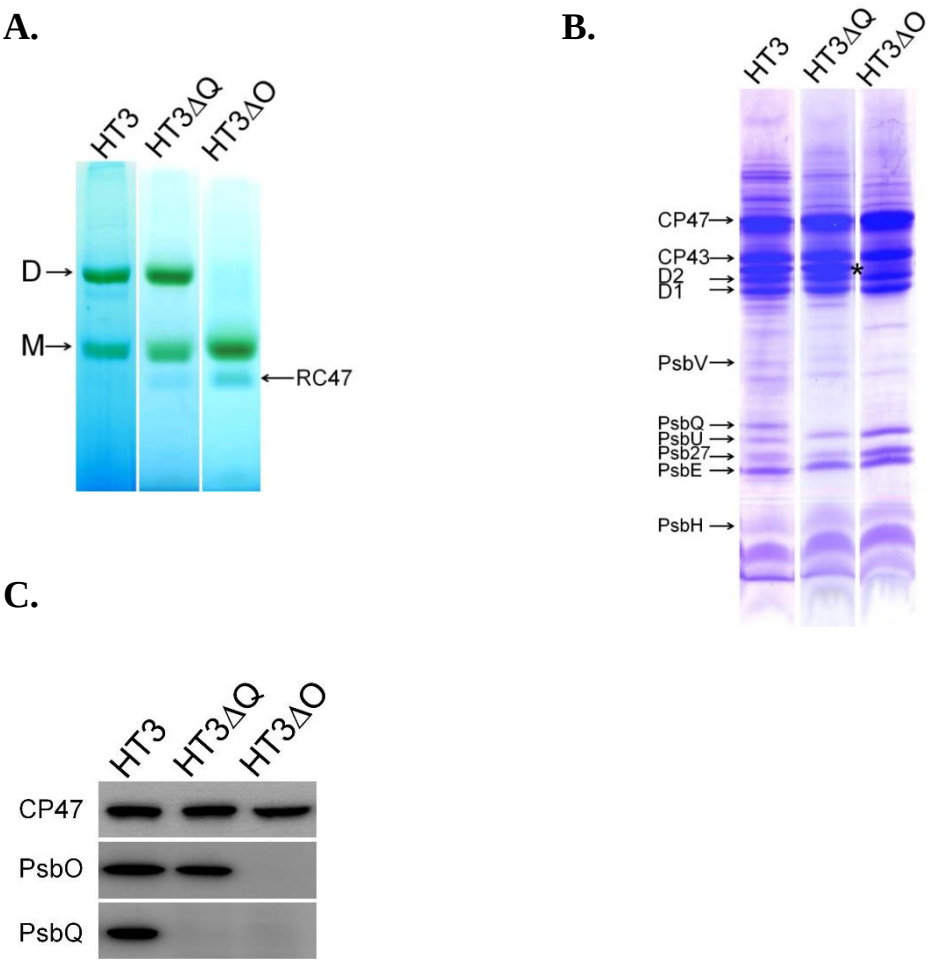


Figure 5. Gel electrophoresis and immunoblot analysis of HT3-PSII, HT3-ΔQ-PSII, and HT3-ΔO-PSII.

A. Blue native gel analysis of HT3-PSII, HT3-ΔQ-PSII, and HT3-ΔO-PSII. Dimeric (D) and monomeric (M) PSII are indicated.

B. Polypeptide profiles of isolated HT3-PSII, HT3-ΔQ-PSII, and HT3-ΔO-PSII. The positions of the major PSII protein components are indicated on the left. PsbO is marked with an asterisk. Each lane was loaded with sample containing 5 µg of Chl *a*.

C. Immunodetection of PSII polypeptides in the isolated complexes. Samples containing 0.2 µg of Chl*a* were fractionated by SDS-PAGE, transferred to a PVDF membrane, and probed by using anti-CP47, anti-PsbO, and anti-PsbQ antibodies.

To determine if both elution peaks correspond to PSII complexes, 77K fluorescence emission spectra were obtained. Both samples displayed characteristic "F685" (at 683 nm) and "F695" (at 692 nm) fluorescence signatures indicative of an assembled PSII reaction center (Fig. 6b) (Satoh, 1980; Liu et al., 2011). The relative intensities of the two peaks were, however, slightly different, reflecting a slight structural difference between the two PSII complexes. A larger difference between the two complexes was observed by measuring their oxygen evolution activity; Complex 2 evolved oxygen at only 66% of the saturated rate of Complex 1 (841 and 1270 μmol $O_2 \cdot$mg Chl $a^{-1} \cdot h^{-1}$, respectively).

SDS-PAGE analysis was performed to analyze the protein compositions of the two complexes. The resulting gel (Fig. 7A) shows that both complexes contained a nearly identical profile of the core PSII proteins compared to the control HT3-PSII sample (which contains a His-tag on the CP47 protein). However, Western blot analysis showed that Complex 2 lacked the PsbU and PsbV proteins (Fig. 7B). In contrast, PsbO, the other lumenal extrinsic protein observed in the PSII crystal structure, was present in equal levels across all three complexes.

Surprisingly, the gel indicated that on a per-chlorophyll basis, Complex 2 contained significantly more PsbQ than Complex 1. Immunoblot analysis of these three complexes using a PsbQ-specific antibody confirmed this result (Fig. 7B). The Western blot indicates that Complex 1 and HT3-PSII contained roughly equal levels of PsbQ, implying that Complex 2 contains more copies of PsbQ per PSII monomer than both Complex 1 and HT3-PSII.

To determine if the additional copies of PsbQ in Complex 2 were a contamination from free copies of the protein that are in fact unassociated with PSII, blue-native gel electrophoresis followed by denaturing gel electrophoresis and immunodetection was performed. The mild blue-native gel conditions (Schägger and von Jagow, 1991) keep PSII complexes intact during

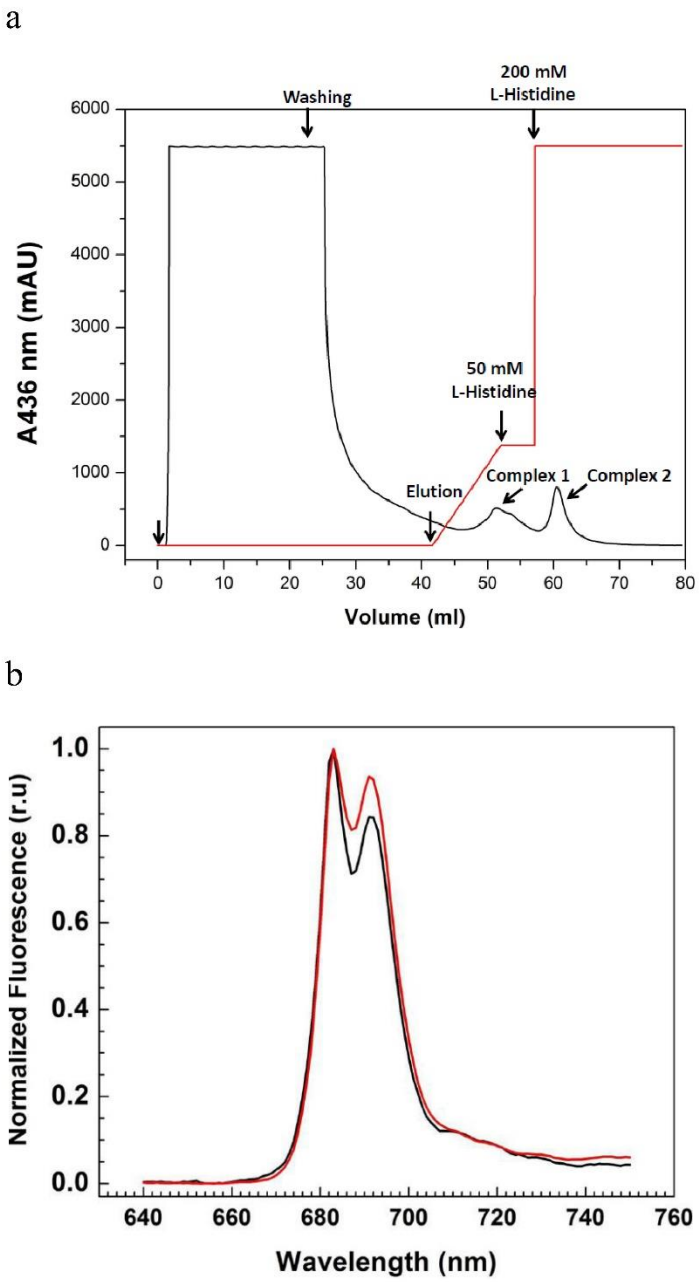


Figure 6. Isolation of two kinds of PSII complexes from the Q-His strain of *Synechocystis* 6803.

A. FPLC chromatogram for isolation of Complex 1 and Complex 2. The first arrow indicates the sample loading. Sample washing, first-step gradient elution and second-step gradient elution are also indicated.

B. Fluorescence emission spectrum at 77K for Complex 1 (red) and Complex 2 (black), with excitation at 435 nm. PSII peaks are located at 683 and 692 nm, respectively. Spectra were normalized at 683 nm.

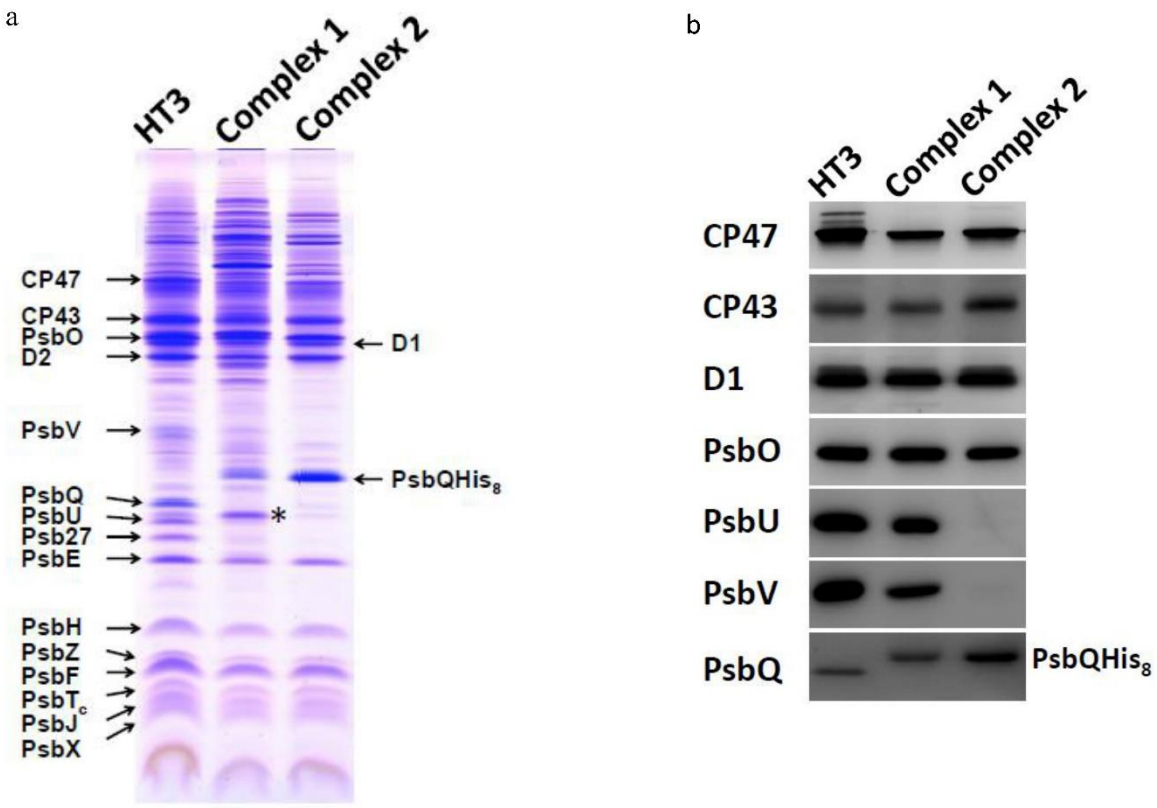


Figure 7. Polypeptide compositions of Complex 1 and Complex 2.

A. SDS-PAGE protein profile with Coomassie Brilliant Blue staining. Each lane contained sample with 2.4 μg of Chl *a*. HT3, a PSII complex with a C-terminally polyhistidine-tagged version of CP47, is used as a reference. Major PSII subunits are labeled. PsbQHis8 is indicated on the right. * indicates a protein band containing Psb28 and Sll1130. Assignment of the PSII subunits is based on Kashino et al. (2002);

B. Immunodetection of PSII polypeptides in the isolated complexes after SDS-PAGE. Each lane contained sample with 0.2 μg of Chl *a*. Specific antibodies against CP47, CP43, D1, PsbO, PsbU, PsbV, and PsbQ were used for immunodetection of the corresponding proteins.

electrophoresis, while unassociated proteins migrate separately due to their smaller size. Subsequent denaturing gel electrophoresis on individual excised native-gel bands allows characterization of the proteins that are present in a particular PSII complex. The native gel (Fig. 8A) showed that both Complex 1 and Complex 2 were present almost exclusively as dimers, consistent with our results presented above. The dimer band was excised and analyzed by denaturing gel electrophoresis followed by silver staining (Fig. 8C and 8D), as well as by immunodetection (Fig. 8B and 8D). Both techniques confirmed the initial observation that PsbQ was present at an elevated level in Complex 2 compared to Complex 1. These results indicate that the additional copies of PsbQ found in Complex 2 are indeed bound to the PSII complex, and are not unassociated-protein contaminants in the sample.

Quantification of PsbQ levels in PSII-Q4 ("Complex 2")

To quantify the increased level of PsbQ in Complex 2, PsbQ content in a dilution series of Complex 2 was detected by Western blot (Fig. 9A). Pixel densitometry analysis was performed to obtain a calibration curve of band intensity vs. PsbQ content. By fitting PsbQ band intensity from Complex 1 (from the same blot) to the calibration curve (Fig. 9B), we concluded that Complex 1 contains 23% of the PsbQ content of Complex 2, on a per chlorophyll basis; *i.e.*, Complex 2 contains four times as many copies of PsbQ as Complex 1.

It is conceivable that the Q-His strain could produce a significantly higher quantity of PsbQ than the WT strain, potentially leading to an artifactual Q-His-PSII complex with higher PsbQ content. This possibility was ruled out by comparing PsbQ levels in the total membrane fraction of the Q-His and WT strains (Fig. 9C) by immunoblotting. After normalizing to D1 levels in the two strains, we determined that the PsbQ content of the Q-His strain was 1.1 (+/- 0.3)-fold of that of the WT strain. This result indicates that the Q-His strain does not synthesize dramatically

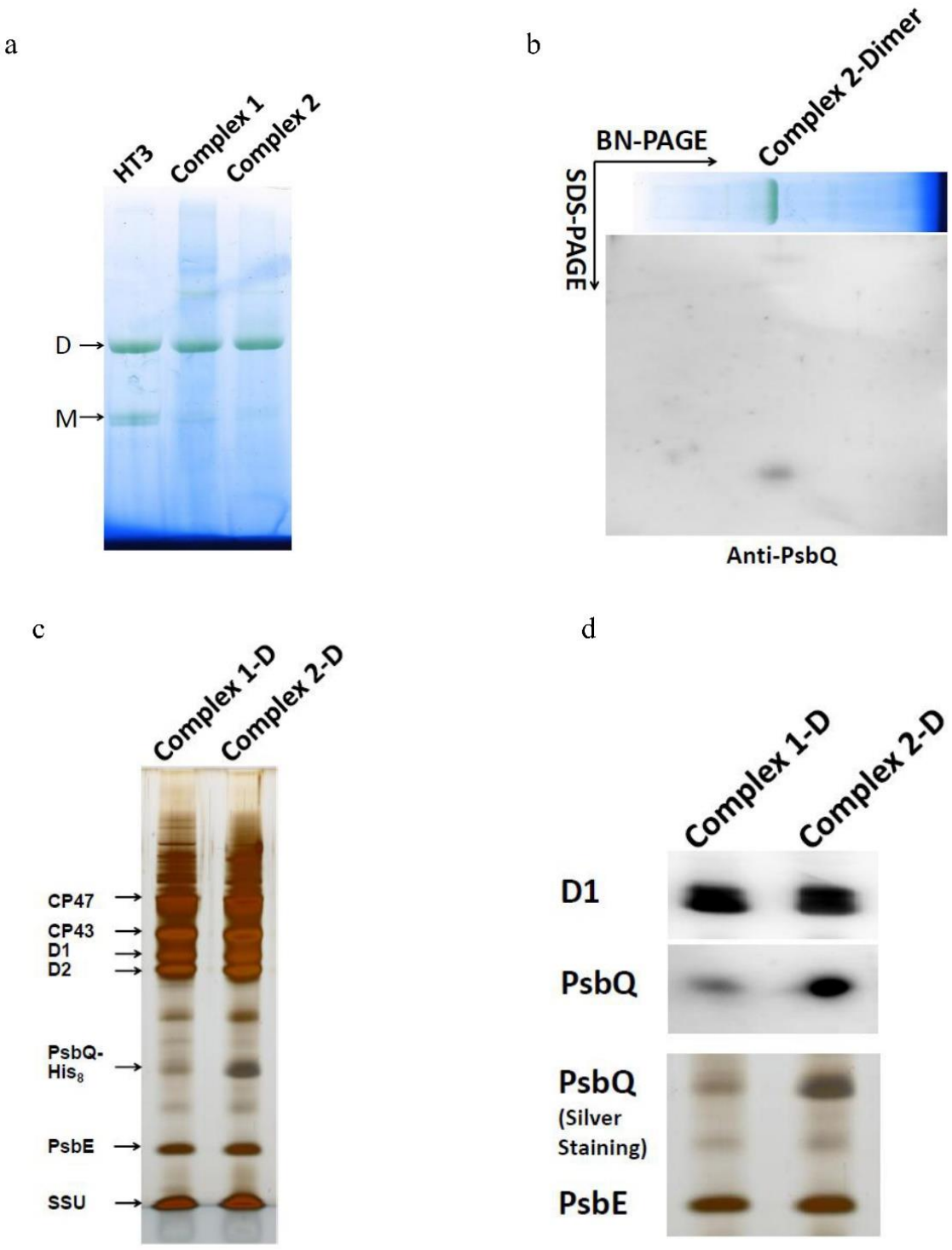


Figure 8. **Two-dimensional Blue-Native-PAGE and immunoblot analysis of Complex 1 and Complex 2**.

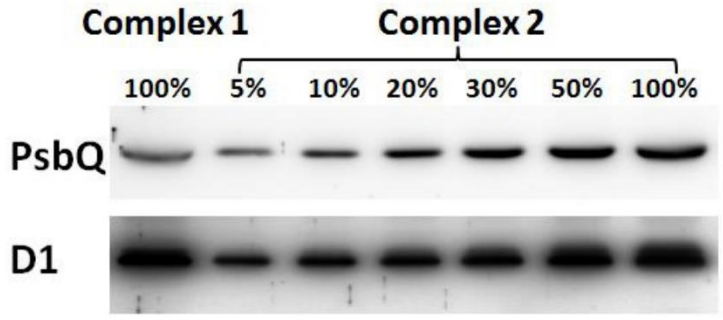
A. Blue-Native gel electrophoresis of HT3-PSII, Complex 1, and Complex 2. M, PSII monomer; D, PSII dimer.

B. Blue-Native-PAGE was followed by fractionation of Complex 2 in the second dimension by SDS-PAGE, immunoblotting, and probing with anti-PsbQ antibody.

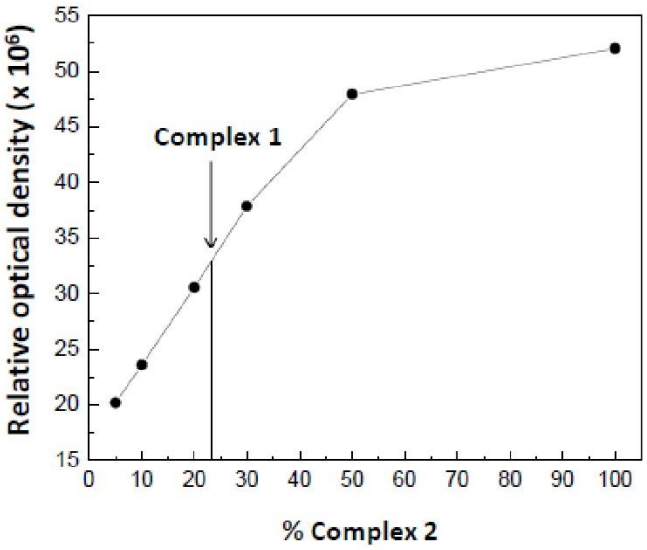
C. Blue-Native-PAGE was followed by fractionation of Complex 1 and Complex 2 dimers in the second dimension by SDS-PAGE and silver staining. SSU = small protein subunits of PSII. Assignment of the distinct subunits of the Q-His protein complexes is based on Kashino et al. (2002), where we observed that the PsbF band also contains PsbI and PsbL, and the PsbX band also contains PsbM.

D. immunodetection of D1 and PsbQ in the BN-PAGE-isolated Complex 1 dimer and Complex 2 dimer. The silver-stained bands corresponding to PsbQ and PsbE are also shown.elevated levels of PsbQ compared to WT. Thus, the additional PsbQ content observed in Complex 2 cannot be attributed to a change in the expression level of PsbQ in the Q-His strain.

a



b



c

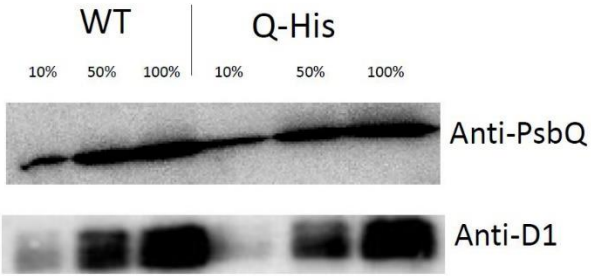


Figure 9. Relative quantification of the PsbQ protein in Complex 1 and Complex 2.

A. SDS-PAGE followed by immunodetection, using anti-PsbQ and anti-D1 (as internal standard) antibodies, of a dilution series of Complex 2 compared to a fixed quantity of Complex 1. Loading ranged from sample containing 0.01 µg Chl *a* (5%) to sample containing 0.2 µg Chl *a* (100%). After immunodetection, pixel densitometry analysis of the PsbQ bands in Complex 2 was used to construct the standard curve shown in **B**; the amount of PsbQ protein present in the Complex 1 sample containing 0.2 µg Chl *a* was then determined from the standard curve. The arrow indicates the amount of PsbQ protein present in Complex 1.

C. SDS-PAGE followed by immunodetection, using anti-PsbQ and anti-D1 (as internal standard) antibodies, of a dilution series of the total membrane fraction from WT and Q-His cells. Pixel densitometry analysis was performed as in **B** to quantify PsbQ content relative to D1 levels in the two strains (not shown).

elevated levels of PsbQ compared to WT. Thus, the additional PsbQ content observed in Complex 2 cannot be attributed to a change in the expression level of PsbQ in the Q-His strain.

Discussion

The mechanism of water oxidation in PSII, first developed in ancient cyanobacteria, has remained virtually unaltered during the evolution of algae and plants. Various extrinsic proteins, such as the well-defined PsbO, PsbU, and PsbV proteins in cyanobacteria, are believed to fine-tune the specific needs of PSII activity, helping modulate its inorganic cofactor (manganese, calcium, and chloride) requirements and optimize water oxidation activity. In cyanobacterial PSII, PsbQ is an enigmatic protein. Homologs of PsbQ are present in all thylakoid-containing cyanobacterial strains (Fig. 1 and Fig. S1) and it is a stoichiometric component of highly active PSII preparations from *Synechocystis* 6803 (Roose et al., 2007a). Thus, the absence of PsbQ in the PSII crystal structures from *T. vulcanus* and *T. elongatus* might have resulted from the loss of this protein during the purification of PSII from these thermophilic cyanobacteria.

Cross-linking results

Our current study demonstrates that PsbQ is closely associated with PSII via interactions with PsbO and CP47. Based on our cross-linking data (Fig. 4A, Fig. 4B, and discussion below), we propose a molecular model of the location of PsbQ within a PSII (PDB ID 3ARC) dimer (Fig. 10A). In this model, PsbQ is positioned in a valley formed by the long lumenal Loop E of CP47 and PsbO from each monomer. In this valley, the positively charged surface formed by helices 2

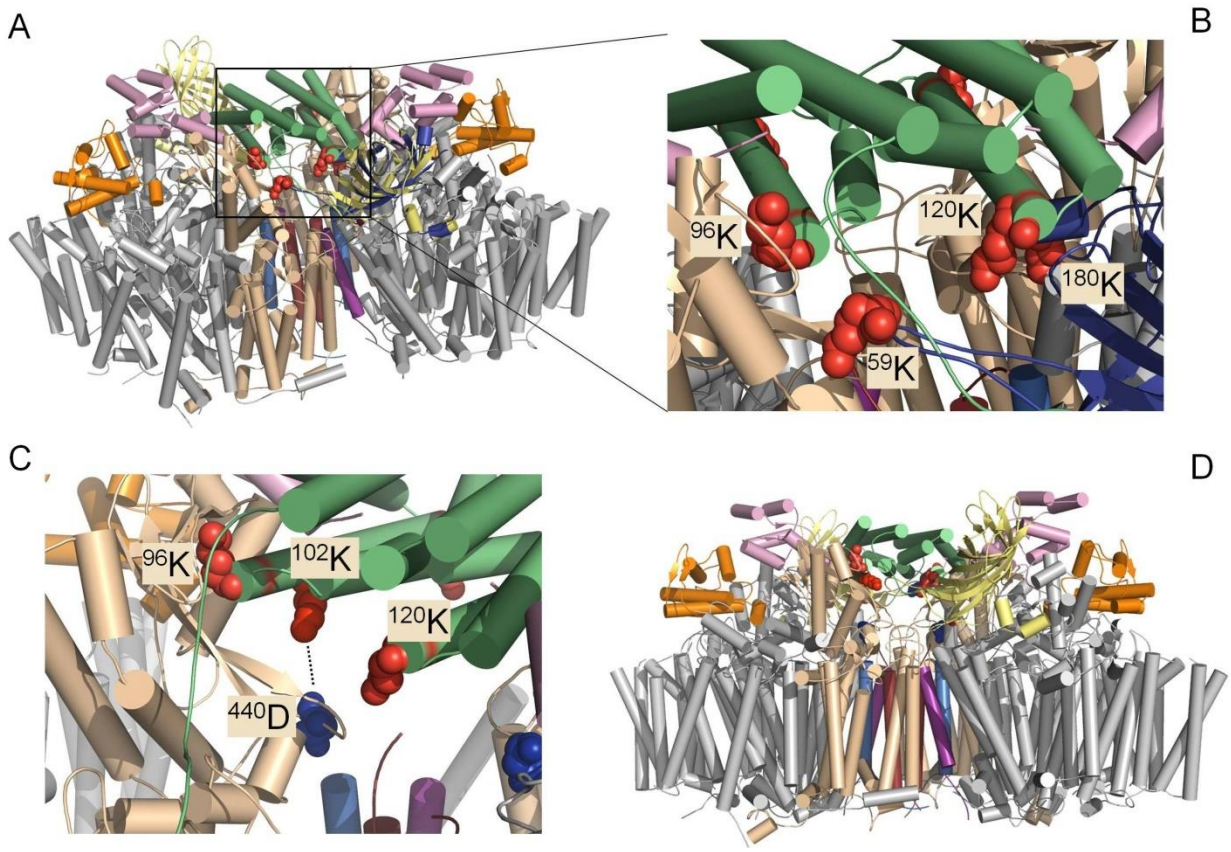


Figure 10. Schematic model for the binding of PsbQ to the interface of dimeric PSII. PsbQ: PDB ID 3LS1; PSII: PDB ID 3ARC. PsbQ, lime-green; CP47, wheat; PsbO, yellow and blue; PsbU, pink; PsbV, orange; PsbM, magenta; other PSII components, gray.

A. Lumenal side view (tilted) showing PsbQ binding to the cove formed by CP47 and PsbO.

B. A magnified view of the interactions between PsbQ and PsbO shown in panel A. ^{120}K of PsbQ as well as ^{59}K and ^{180}K of PsbO are represented as spheres. ^{96}K from the second PsbQ protein in the PSII dimer is also shown.

C. Same as **B** with clockwise rotation (60°), showing interactions between PsbQ and CP47, as deduced from Fig. 4B. ^{102}K of PsbQ and ^{440}D of CP47 are represented as spheres.

D. Side view of the interactions between the two PsbQ proteins and the PSII dimer. All images were prepared by using PyMOL software (30).

and 3 (H2 and H3) of PsbQ (Jackson et al., 2010) interacts with the negatively charged surface of the PsbO protein (Fig. S4C). By adopting this orientation, two conserved helices (H2a and H2b) from the two PsbQ proteins in the dimer are arranged in an anti-parallel configuration, with their N-termini pointing toward the interface of Loop C of CP47 and PsbO, which is close to the peripheral interface of the two PSII monomers. This structural location of PsbQ does not cause any apparent structural conflicts between PsbQ and the other extrinsic proteins, *i.e.*, PsbO, PsbU and PsbV. An implication of the spatial relationship between PsbQ, PsbO and Loop E of CP47, however, is that only after PsbO is recruited to the lumenal side of PSII, is PsbQ able to bind. This hypothesis is supported by our observation that deletion of PsbO leads to complete absence of PsbQ in PSII (Fig. 5B, 5C). In contrast, loss of PsbV reduces the PsbQ level to nearly 40% of that in the wild type cells (Wegener et al., 2008). Our model is consistent with the experimental observation that PsbO is the first lumenal extrinsic protein recruited to PSII (Liu et al., 2011b; Nowaczyk et al., 2012). It is also supported by the earlier observation that deletion of the *psbO* gene abolishes the dimerization process of PSII *in vivo* (Bentley and Eaton-Rye, 2008), as observed in our PSII preparations as well (Fig. 5A).

The present study addresses whether the lumenal domains of CP47, CP43 and three known extrinsic proteins are able to accommodate additional extrinsic proteins. The data using the zero-length cross-linker EDC (Fig. 3A, Fig. 4B) suggest that regions of complementary charge exist between PsbQ and other PSII proteins. It was previously suggested that PsbQ is associated directly with a small (4 kDa), intrinsic PSII core subunit in a green-algal PSII complex (Nagao et al., 2010).

The identity of this protein, however, remains unknown. It seems likely that the small cross-linked products observed in Fig. 3A (a, b) and Fig. 3B (a, b) could result from cross-links between PsbQ and small subunits located in the interface between two PSII monomers (*e.g.*, PsbM, PsbL

and PsbT). This assumes, however, a closer contact of PsbQ to the transmembrane domain of PSII, or the floor of the valley formed by CP47 and PsbO. Such a close contact seems reasonable (Fig. 4B) for a compact association of PsbO and its binding partners including CP47, CP43, D1, D2, and PsbU. To address such a hypothesis, however, more elegant protein foot-printing experiments *e.g.*, hydrogen-deuterium exchange (HDX), or fast photochemical oxidation of proteins (FPOP) (Chen et al., 2010), need to be designed.

PsbO interacts with a number of other PSII subunits (Bricker et al., 2012). Our MS-based cross-linking analysis also detected two CP47-PsbO cross-links that are consistent with the available crystal structures (Guskov et al., 2009; Umena et al., 2011) (Table S1). Moreover, PsbO also participates in an inter-monomer interaction with CP47, stabilizing the PSII dimer. *In silico* analysis indicates that ^{59}K of PsbO could form a charge-pair interaction with ^{307}E of CP47 (Darnell et al., 2007; Bricker et al., 2012). Taken together with the structural proximity of PsbQ to PsbO and the close association between PsbQ and CP47 (Fig. 4A, 4B), we speculate that PsbQ increases the stability of the PSII dimer by interacting with PsbO and CP47, thus decreasing the solvent exposure of those interaction interfaces. This structural model is consistent with the results of previous studies that deletion of the *psbQ* gene in *Synechocystis* 6803 results in photosynthetic defects under Ca^{2+} and Cl^- limiting conditions (Thornton et al., 2004; Summerfield et al., 2005). The phenotype of this deletion mutant, however, was less severe relative to that of other cyanobacterial extrinsic protein mutants, indicative of the auxiliary role of PsbQ in PSII photochemistry under nutrient-replete conditions.

Genetic and physiological data indicate that PsbQ stabilizes the binding of PsbV in PSII (Kashino et al., 2006; Summerfield and Eaton-Rye, 2006), and *in silico* protein docking analysis seem to support this idea as well (Fagerlund and Eaton-Rye, 2011). Our LC-MS/MS analysis,

however, failed to detect any confident PsbQ-PsbV cross-links (data not shown). Interestingly, PsbQ-PsbQ cross-linked species (DMLGLASSLLP^{96}KDQDK, LDAAA^{120}KDRNGSQAK) (cross-link between ^{96}K and ^{120}K; Fig. S3) were consistently detected in our cross-linked samples. It is unlikely that this cross-linked species originated from one copy of PsbQ (*i.e.*, an intra-protein link), because the arm span of the cross-linker used in this study is 12 Å and the distance between ^{96}K and ^{120}K is 41 Å (Fig. S6). More importantly, the two Ks are located on opposite ends of Helix 3 in the crystal structure (Jackson et al., 2010) rather than in a loop area that might have more structural flexibility. Our immunodetection of cross-linked species seems also to support a dimeric PsbQ from both EDC and DTSSP cross-linker results (Fig. 3A, label **c** and Fig. 3B, label **c**). In our model, Helix 3 is buried in the interface formed by Loop E of CP47 and PsbO. Our data indicate that the two copies of PsbQ are located in structural proximity to each other and probably share the same symmetrical axis as the two PSII monomers. In line with earlier work (De Las Rivas and Barber, 2004), our model (Fig. 10A) assumes there is only one copy of PsbQ per PSII monomer, but we cannot completely exclude a more complicated scenario of more than one copy per PSII monomer. In this context, it is noteworthy that the Chl *c*- containing diatom *Chaetoceros gracilis* has two different PsbQ homologs, PsbQ' and Psb31, in each PSII monomer (Nagao et al., 2013). Our observation of a PsbQ-PsbQ interaction in Q-His-PSII also does not necessarily exclude the possibility that PsbQ could associate with monomeric PSII, in which, after the recruitment of PsbO, PsbO and PsbQ synergistically facilitate a rapid PSII dimerization process during the dynamic assembly of PSII.

PSII-Q4 in the context of the PSII life-cycle

We isolated two types of PSII complex were isolated from the Q-His strain of *Synechocystis* by nickel-affinity chromatography using a stepped linear gradient of histidine during

elution. Both complexes displayed the 77K fluorescence signatures of an assembled PSII reaction center complex, and both were found almost exclusively as dimers (Fig. 8A). Both contained a nearly identical profile of PSII proteins, compared to each other and to the control HT3-PSII complex. However, Complex 2 contained four times as many copies of PsbQ as Complex 1, and was lacking the extrinsic proteins PsbU and PsbV. The additional histidine tags in Complex 2 caused increased affinity for the His-tag resin and explain its later elution time. Complex 2 exhibited 66% of the oxygen evolution activity of Complex 1, which is understandable since PsbU and PsbV are known to optimize oxygen evolution capability (Bricker et al., 2012). Based on these data, we conclude that Complex 1 is the fully-assembled active PSII dimer (PSII-D), and propose that Complex 2 (renamed PSII-Q4) is a newly-identified PSII assembly intermediate.

Our cross-linking results indicated that fully-assembled PSII contains one copy of PsbQ per PSII monomer. Hence, PSII-Q4 contains four copies. PsbU and PsbV are absent in PSII-Q4, and are roughly the same size as PsbQ (14.2, 17.9, and 15.7 kDa, respectively). Considering that all three proteins bind near each other on the lumenal surface of PSII (Umena et al., 2011), the extra copies of PsbQ in the PSII-Q4 assembly intermediate complex may temporarily occupy the binding regions of PsbU and PsbV. PsbQ contains a lipid modification that anchors it to the thylakoid membrane (Thornton et al., 2004; Kashino et al., 2006). This lipid anchor may position PsbQ near the lumenal surface of PSII, enabling facile, although relatively weak, binding of several copies of PsbQ to PSII during the assembly process, which are then replaced by PsbU and PsbV once they are able to diffuse to their binding sites.

The PsbQ deletion mutant in *Synechocystis* 6803 shows only a slight reduction in growth rate compared to the wild-type strain (Thornton et al., 2004; Summerfield et al., 2005) while the PsbV deletion mutant requires additional Ca^{2+} and Cl^{-} in the growth medium to enable

photoautotrophic growth (Shen et al., 1998). However, the Δ*psbV*: Δ*psbQ* double-deletion mutant is unable to grow photoautotrophically (Summerfield et al., 2005). It is possible that one or more copies of PsbQ in PSII-Q4 occupy the PsbV binding site, serving as an imperfect and temporary substitute for PsbV.

Enami and co-workers found that the PsbQ-type protein Psb31 is able to substitute partially for the role of PsbO in the centric diatom *Chaetoceros gracilis* (Nagao et al., 2010). In addition to Psb31, *C. gracilis* contains a second PsbQ-type protein, PsbQ´, whose lumenal binding site and role in optimizing oxygen evolution is distinct from those of Psb31 (Nagao et al., 2010; Nagao et al., 2013). While PsbO is present in the PSII-Q4 complex, both of these results reinforce the idea of the binding promiscuity of the PsbQ protein in PSII and its ability to substitute for other lumenal extrinsic proteins. The ability of the PSII-Q4 complex to evolve oxygen, but at a lower rate (66%) than fully-assembled PSII, can be understood in this context.

Previous work has shown that PsbO is the first of the lumenal extrinsic PSII proteins to bind to PSII (Liu et al., 2011; Nowaczyk et al., 2012), and this is consistent with its presence in the PSII-Q4 complex despite the absence of PsbU and PsbV. The observation that both Q-His PSII complexes were found almost exclusively in the dimeric state suggests that PsbQ binds after PsbO, and that PsbQ binding may facilitate dimerization. This is consistent with the cross-linking results presented above that show that PsbQ helps stabilize the PSII dimer interface. PSII-Q4 thus appears to be a late PSII assembly intermediate that is formed just before binding of PsbU and PsbV. Although we cannot exclude the possibility that PSII-Q4 may also form during the disassembly of PSII (after the dissociation of PsbU and PsbV), the relatively high oxygen evolution rate of PSII-Q4 (66% of fully-assembled PSII) suggest that the majority of PSII-Q4 complexes contain an undamaged D1 protein.

We have incorporated our results into an updated model of PSII assembly (Fig. 11). We suggest that PSII dimerization occurs after dissociation of Psb27 and formation of the WOC. Four copies of PsbQ bind during or immediately after this step, stabilizing this active dimer. Though we did not observe a monomeric PSII-Q4 intermediate, it is possible that such a complex forms transiently in between Psb27 dissociation and dimerization. The PSII-Q4 dimer is the first intermediate during PSII assembly that is capable of oxygen evolution, albeit at around two-thirds the rate of mature PSII. An active WOC increases PSII vulnerability to oxidative damage; as discussed above, due to the presence of a lipid anchor of PsbQ, multiple copies of PsbQ in PSII-Q4 serve as readily-available substitutes for PsbU and PsbV, helping to stabilize the WOC and protect PSII as soon as it gains oxygen-evolving capability. The additional copies of PsbQ must bind relatively weakly, since PsbU and PsbV replace them on the lumenal surface of PSII once they are able to diffuse to their binding site, forming the fully-assembled, fully-protected, PSII dimer (PSII-D). Though we hypothesize that formation of PSII-Q4 stabilizes the active dimer until PsbU and PsbV bind, we found in this study that PSII-M and PSII-D are still able to form in the absence of PsbQ. It is therefore possible that a portion of the complexes bypass PSII-Q4 formation during assembly, as indicated by a dashed line in Fig. 11. Light-induced D1 damage triggers partial disassembly of PSII-M and PSII-D, possibly to the RC47 stage (Nickelsen and Rengstl, 2013) (see Fig. 11). It is possible that the PSII-Q4 complex also forms after the dissociation of PsbU and PsbV during PSII disassembly. A new copy of D1 replaces the damaged copy and reassembly of active PSII occurs.

Changing environmental conditions can alter the rate of PSII damage, making it difficult for the cell to maintain its required rate of real-time energy production. The PSII-Q4 dimer may assist

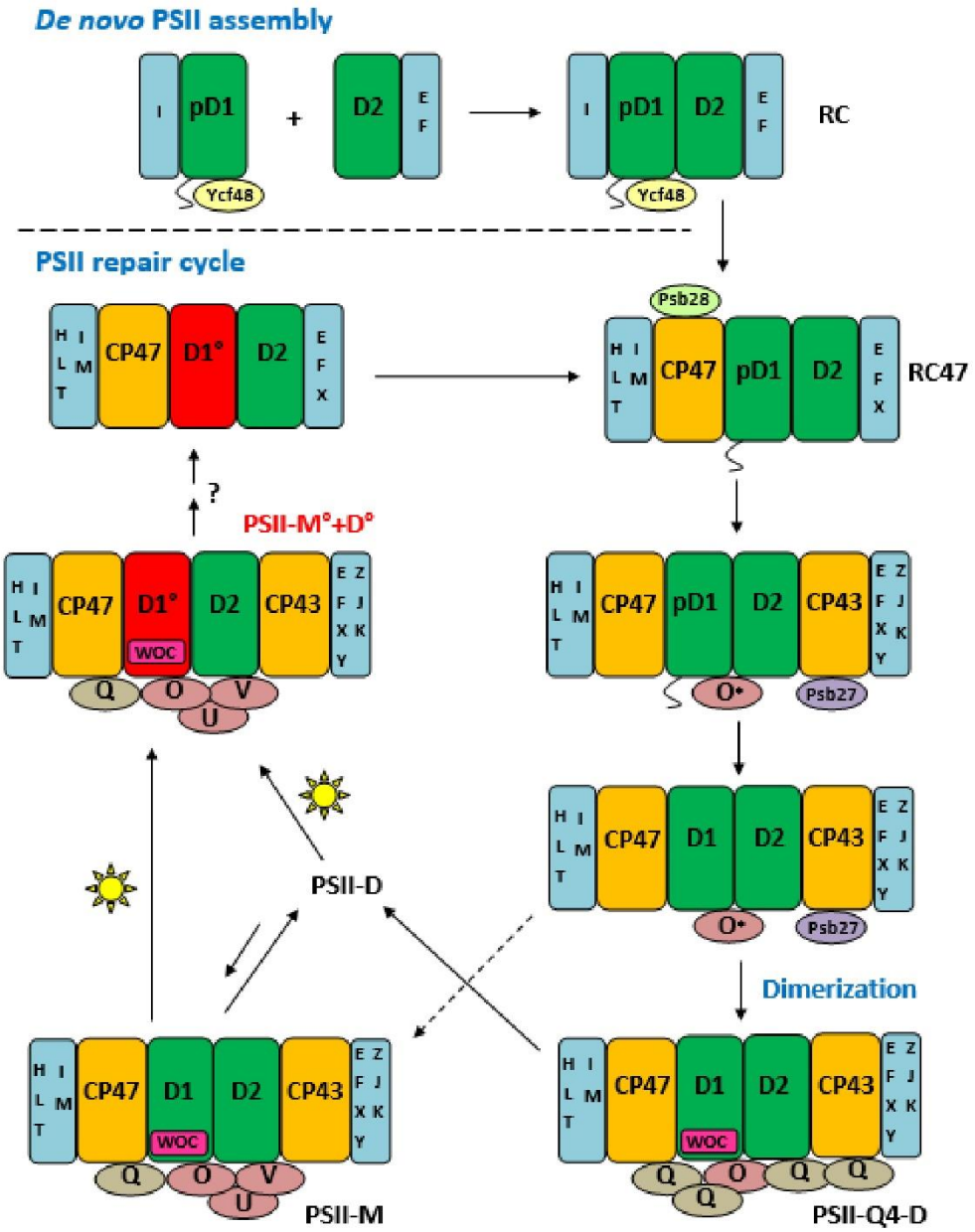


Figure 11. A schematic model for PSII assembly and repair.

During *de novo* assembly, the pD1-PsbI-Ycf48 subcomplex joins with the D2-PsbE-PsbF subcomplex to form the reaction center (RC) subcomplex. The RC47 complex is formed after the attachment of CP47 and several low-molecular-weight proteins to RC. Next, CP43 and several more low-molecular-weight proteins bind to the complex. Psb27 associates at this stage, and PsbO is also bound sub-stoichiometrically (indicated by a *). After pD1 processing, Psb27 dissociates, PsbO binds stoichiometrically, the WOC is assembled, and dimerization occurs. PSII-Q4 forms at this stage, followed by PSII-D. PSII-D can interconvert with PSII-M. In a parallel pathway, PSII-M may also form directly after Psb27 dissociation, bypassing formation of PSII-Q4. After light-induced D1 damage of PSII-M and PSII-D (represented in red by "D1°"), PSII is partially disassembled. The damaged D1 protein is removed and degraded, a newly-

synthesized copy is inserted, and re-assembly of active PSII occurs. M, monomer; D, dimer. The low-molecular-weight proteins and extrinsic protein names have been shortened from , e.g., "PsbH" to "H," "PsbO" to "O," etc. * indicates sub-stoichiometric binding. The relative positions of different proteins in this diagram are based roughly on the PSII crystal structures and cross-linking results (Liu et al., 2011; Umena et al., 2011; Liu et al., 2013; Suga et al., 2015).

by serving as a pool of nearly-assembled PSII that can be rapidly converted to PSII-D in response to shifts in the equilibrium concentration of PSII-D and PSII-M.

To summarize, PsbQ is present in a highly active and stable form of PSII in the mesophilic cyanobacterium *Synechocystis* 6803, but its location within the complex has remained unknown owing to its absence in the PSII crystal structures obtained from thermophilic cyanobacteria. This discrepancy may reflect the different environmental conditions to which the organisms adapt. Our study presents the first molecular-level model for the binding site of PsbQ to dimeric PSII. In addition, we detected a cross-link between two copies of PsbQ, which suggests strongly the presence of PsbQ in the PSII dimer interface. Our results help elucidate the stage of incorporation of PsbQ during the PSII assembly process and its stabilizing effect on the PSII dimer, and, thus, provide a basis for further investigation of its role in optimizing PSII function. In addition, we have isolated a PSII subcomplex (PSII-Q4) with four copies of the PsbQ protein. Based on our characterization of the complex, we conclude that it is a late PSII assembly intermediate, formed after the binding of PsbO and before the binding of PsbU and PsbV. This complex helps to stabilize PSII immediately after it becomes capable of water oxidation. Our results provide further evidence for the binding promiscuity of PsbQ and its ability to substitute for other lumenal extrinsic PSII proteins.

Materials and methods

BLAST analysis, sequence alignment and phylogenetic analysis

NCBI-BLASTp software (Altschul et al., 1990) was used to detect homologs to the *Synechocystis* 6803 PsbQ protein (Sll1638) in other sequenced cyanobacterial strains. A protein was considered a homolog if its E-value was $<10^{-6}$. Search parameters were as follows: database- non-redundant protein sequences; organism- cyanobacteria; word size- 3; max matches in a query range- 0; matrix- BLOSUM62; gap costs- Existence:11 Extension:1; compositional adjustments- conditional compositional score matrix adjustment. BLAST results were aligned using Clustal Omega (Goujon et al., 2010; Sievers et al., 2011) (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalo) using the default parameters. Phylogenetic trees were viewed using TreeView v.1.6.6 (Page, 1996).

Cyanobacterial culture and PSII purification.

Cyanobacteria strains were grown in BG11 medium. Generation of the Q-His strain was previously reported (Roose et al., 2007). The HT3 strain was a generous gift from Dr. Terry M. Bricker (Bricker et al., 1998). The HT3-ΔQ strain was previously reported (Kashino et al., 2006). The HT3-ΔO strain was generated by transforming the HT3 strain by using the ΔO construct previously reported (Chandler et al., 2003). Purification of histidine-tagged PSII complexes was performed as described in Kashino et al. (2002), with minor modifications. For purification of PSII-Q4, a two-step elution gradient was used as follows: the first step was a linear increase from 0-50 mM histidine for ten minutes at 0.5 mL/min, followed by a hold at 50 mM histidine for 5 minutes. The histidine concentration in the elution buffer was then switched to 200 mM histidine, and held at this concentration for an additional 23.5 min.

77K fluorescence spectroscopy

Fluorescence emission spectra at 77K were recorded on a Fluoromax-2 fluorometer (Jobin Yvon, Longjumeau, France) with excitation at 435 nm. PSII samples were diluted to 0.1 mg Chl

a/mL in a buffer previously reported (Bricker et al. 1998). Fluorescence emission curves were normalized at 683 nm.

Protein analysis

Protein electrophoresis was performed as in Kashino et al., (2001) and Kashino et al. (2002), unless otherwise indicated. After transfer of the fractionated proteins onto PVDF membranes (Millipore, Billerica, MA), PSII subunits were identified by using specific antisera. Bands were visualized by using enhanced chemiluminescence reagents (Westpico; Pierce, Rockford, IL) on an ImageQuant LAS-4000 imager (GE healthcare, Pittsburgh, PA), and image quantification was performed using ImageQuant TL software. Levels of PSII monomers and dimers were determined by Blue-Native (BN-gel) electrophoresis (Schägger and von Jagow, 1991). Silver staining of protein gels following SDS-PAGE was performed using metallic silver (Ag) protein stain according to the manufacturer's protocol (Thermo Scientific, Rockford, IL, USA).

Oxygen evolution measurements

The steady-state rate of oxygen evolution by PSII was measured on a Clark-type electrode at 5 μg Chl *a*/mL in 50 mM MES-NaOH (pH 6.5), 20 mM $CaCl_2$, 0.5 M sucrose, at 30°C. Buffer contained 1 mM potassium ferricyanide and 0.5 mM 2,6-dichloro-*p*-benzoquinone as electron acceptors. Samples were incubated in the dark at 30°C for 1 min before the measurement. Irradiance of 8250 μmol photons $m^{-2} \cdot s^{-1}$ was used during the measurement.

Chemical cross-linking

PSII preparations were resuspended at 0.2 mg/mL chlorophyll *a* in 25% glycerol, 10 mM $MgCl_2$, 5 mM $CaCl_2$, and 50 mM MES buffer (pH 6.5). Cross-linking of PSII samples with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, Sigma, St Louis, MO) was performed

essentially as described previously (Liu et al., 2011). DTSSP and BS^3 (Thermo Scientific) cross-linking was performed according to the manufacturer's protocol (with minor modification) followed by desalting using Zeba spin columns (Thermo Scientific).

Proteolytic Digestion and Peptide Clean-up

Modified samples were precipitated using acetone, and the resuspended samples were directly subjected to trypsin digestion. One reason in-solution digestion of the cross-linked samples was preferred after chemical cross-linking and quenching was that cross-linked species usually represent only very small fractions of total proteins, especially for protein complexes like PSII that contain at least 40 subunits per PSII dimer (Zouni et al., 2001). Thus, the goal was to avoid over crosslinking the sample, which could lead to protein aggregation and denaturation. In-solution digestions also avoided the loss of cross-linked products during the post-gel handling, and reduced the artifacts of gel adduction.

LC-MS/MS

The peptide mixture from trypsin digestion was dissolved in water with 0.1% formic acid. The peptide samples were analyzed by using our LC-MS proteomics workflow (Zhang et al., 2010). The peptide sample was loaded onto an Ultimate 3000 Nano LC system (Thermo Scientific Dionex, Sunnyvale, CA) coupled with an LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA). The peptides were trapped by a guard column (Acclaim PepMap100, 100 µm x 2 cm, C18, 5 µm, 100 Å, Thermo Scientific Dionex) through which solvent A (water with 0.1 % formic acid, Sigma-Aldrich) was pumped at 6 µL/min. The peptide mixture was fractionated on a custom-packed Michrom Magic C 18 RP column, as previous reported. The peptides were eluted at a flow rate of 260 nL/min, ramping a gradient from 5% to 60% solvent B (80% acetonitrile, 20% water and 0.1% formic acid) in 110 min. The eluted samples were directly

introduced into mass spectrometer via a PicoView nano electrospray source (New objective, Woburn, MA). Ion source and other parameters of the mass spectrometer were optimized by tuning with peptide standards. The mass spectrometer was operated in data-dependent mode by using previously reported parameters (Zhang et al., 2010).

LC-MS/MS data analysis

LC-MS/MS data in Thermo Xcalibur .raw files were converted into mzXML and mgf format by MM File Conversion from the Mass Matrix package**.** Product-ion mass spectra were searched against the UniProt database to generate the protein list for cross-linked peptide identification. The cross-linked peptides were identified by using the search algorithm, Mass Matrix (Xu et al., 2010). For Mass Matrix, each protein sequence pair was established and searched against all LC/MS data. The Mass Matrix search parameters were as follows: Variable modification: Oxidation of Met, Max # variable PTM/peptide: 1, Peptide tolerance: 15 ppm, MS/MS tolerance: 0.8 Da, Mass type: Monoisotopic, C13 isotope ions: Yes, Enzyme: Trypsin, Missed cleavages: 3, Fixed modification: none, Peptide length: from 3 to 50, Cross-link search mode: Exploratory, Cross-link sites cleavability: Non-cleavable by enzyme, Max # cross-links/ peptide: 2. The search results were viewed using XMAP (v 0.5.1, Mass Matrix).

References

Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215(3):403-410.

Baker N.A., Sept D., Joseph S., Holst M.J., and McCammon J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* 98(18):10037-10041.

Bentley F.K. and Eaton-Rye J.J. (2008). The effect of removing Photosystem II extrinsic proteins on dimer formation and recovery from photodamage in *Synechocystis* sp. PCC 6803. *Photosysthesis Energy from the Sun: 14th International Congress on Photosynthesis.*, eds Allen J.F., Gantt E., Golbeck J.H., & Osmond B. Dordrecht), pp. 715-717.

Bricker, T.M. and Ghanotakis D.F. (1996). Introduction to oxygen evolution and the oxygen-evolving complex, *in: D.R. Ort, C.F. Yocum (Eds), Oxygenic photosynthesis: The light reactions,* (Dordrecht: Kluwer Academic Publishers).

Bricker T.M., Morvant J., Masri N., Sutton H.M., and Frankel L.K. (1998). Isolation of a highly active photosystem II preparation from *Synechocystis* 6803 using a histidine-tagged mutant of CP 47. *Biochim. Biophys. Acta* 1409(1):50-57.

Bricker T.M., Odom W.R., and Queirolo C.B. (1988). Close association of the 33-Kda extrinsic protein with the apoprotein of CPa1 in Photosystem-II. *FEBS Lett.* 231(1):111-117.

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(1):121-142.

Chandler L.E., Bartsevich V.V., and Pakrasi H.B. (2003). Regulation of manganese uptake in *Synechocystis* 6803 by RfrA, a member of a novel family of proteins containing a repeated five-residues domain. *Biochemistry* 42(18):5508-5514.

Chen J., Rempel D.L., and Gross M.L. (2010). Temperature jump and fast photochemical oxidation probe submillisecond protein folding. *J. Am. Chem. Soc.* 132(44):15502-15504.

Darnell S.J., Page D., and Mitchell J.C. (2007). An automated decision-tree approach to predicting protein interaction hot spots. *Proteins* 68(4):813-823.

Delano W.L. (2002). The PyMol molecular graphics system, Software. *The PyMol Molecular Graphics System (Delano Scientific, Palo Alto, CA).*

De Las Rivas J. and Barber J. (2004). Analysis of the Structure of the PsbO Protein and its Implications. *Photosynth. Res.* 81(3):329-343.

Fagerlund R.D. and Eaton-Rye J.J. (2011). The lipoproteins of cyanobacterial photosystem II. *J. Photochem. Photobiol. B* 104(1-2):191-203.

Goujon M., *et al.* (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* 38:W695-699.

Iglesias A.H., Santos L.F., and Gozzo F.C. (2010). Identification of cross-linked peptides by high-resolution precursor ion scan. *Anal. Chem.* 82(3):909-916.

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(13):2765-2767.

Kashino Y., *et al.* (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(25):8004-8012.

Kashino Y., Inoue-Kashino N., Roose J.L., and Pakrasi H.B. (2006). Absence of the PsbQ protein results in destabilization of the PsbV protein and decreased oxygen evolution activity in cyanobacterial photosystem II. *J. Biol. Chem.* 281(30):20834-20841.

Kashino Y., Koike H., and Satoh K. (2001). An improved sodium dodecyl sulfate-polyacrylamide gel electrophoresis system for the analysis of membrane protein complexes. *Electrophoresis* 22(6):1004-1007.

Leitner A., *et al.* (2010). Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* 9(8):1634-1649.

Liu H., Huang R.Y., Chen J., Gross M.L., and Pakrasi H.B. (2011a). Psb27, a transiently associated protein, binds to the chlorophyll binding protein CP43 in photosystem II assembly intermediates. *Proc. Natl. Acad. Sci. U.S.A.* 108(45):18536-18541.

Liu H., Roose J.L., Cameron J.C., and Pakrasi H.B. (2011b). A genetically tagged Psb27 protein allows purification of two consecutive photosystem II (PSII) assembly intermediates in *Synechocystis* 6803, a cyanobacterium. *J. Biol. Chem.* 286(28):24865-24871.

Nagao R., *et al.* (2010). Topological analysis of the extrinsic PsbO, PsbP and PsbQ proteins in a green algal PSII complex by cross-linking with a water-soluble carbodiimide. *Plant Cell. Physiol.* 51(5):718-727.

Nagao R., *et al.* (2013). Crystal structure of Psb31, a novel extrinsic protein of photosystem II from a marine centric diatom and implications for its binding and function. *Biochemistry* 52(38):6646-6652.

Nowaczyk M.M., *et al.* (2006). Psb27, a cyanobacterial lipoprotein, is involved in the repair cycle of photosystem II. *Plant Cell* 18(11):3121-3131.

Nowaczyk M.M., *et al.* (2012). Deletion of *psbJ* leads to accumulation of Psb27-Psb28 photosystem II complexes in *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta* 1817(8):1339-1345.

Page R.D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12(4):357-358.

Petrotchenko E.V. and Borchers C.H. (2010). Crosslinking combined with mass spectrometry for structural proteomics. *Mass Spectrom. Rev.* 29(6):862-876.

Rappsilber J. (2011). The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* 173(3):530-540.

Rinner O., *et al.* (2008). Identification of cross-linked peptides from large sequence databases. *Nat. Methods* 5(4):315-318.

Roose J.L., Kashino Y., and Pakrasi H.B. (2007a). The PsbQ protein defines cyanobacterial Photosystem II complexes with highest activity and stability. *Proc. Natl. Acad. Sci. U.S.A.* 104(7):2548-2553.

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

Schägger H. and von Jagow G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* 199(2):223-231.

Seidler A. (1996). Intermolecular and intramolecular interactions of the 33-kDa protein in photosystem II. *Eur. J. Biochem.* 242(3):485-490.

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(7):1825-1832.

Sievers F., *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539.

Sinz A. (2006). Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* 25(4):663-682.

Summerfield T.C. and Eaton-Rye J.J. (2006). *Pseudocyphellaria crocata*, *P. neglecta* and *P. perpetua* from the Northern and Southern Hemispheres are a phylogenetic species and share cyanobionts. *New. Phytol.* 170(3):597-607.

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(2):805-815.

Tabb D.L. (2012). Evaluating protein interactions through cross-linking mass spectrometry. *Nat. Methods* 9(9):879-881.

Thornton L.E., *et al.* (2004). Homologs of plant PsbP and PsbQ proteins are necessary for regulation of photosystem ii activity in the cyanobacterium *Synechocystis* 6803. *Plant Cell* 16(8):2164-2175.

Ujihara T., Sakurai I., Mizusawa N., and Wada H. (2008). A method for analyzing lipid-modified proteins with mass spectrometry. *Anal. Biochem.* 374(2):429-431.

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(7345):55-60.

Wegener K.M., *et al.* (2008). High sensitivity proteomics assisted discovery of a novel operon involved in the assembly of photosystem II, a membrane protein complex. *J. Biol. Chem.* 283(41):27829-27837.

Xu H., Hsu P.H., Zhang L., Tsai M.D., and Freitas M.A. (2010). Database search algorithm for identification of intact cross-links in proteins and peptides using tandem mass spectrometry. *J. Proteome Res.* 9(7):3384-3393.

Zhang Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40.

Zheng C., *et al.* (2011). Cross-linking measurements of *in vivo* protein complex topologies. *Mol. Cell. Proteomics* 10(10):M110 006841.

Zouni A., *et al.* (2001). Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 Å resolution. *Nature* 409(6821):739-743.

Zhang H., *et al.* (2010). Improved mass spectrometric characterization of protein glycosylation reveals unusual glycosylation of maize-derived bovine trypsin. *Anal. Chem.* 82(24):10095-10101.

Table S1. Cross-links relevant to this study identified using either DTSSP (12 Å spacer arm) or its non-cleavable homolog BS^3 (11.4 Å spacer arm).

Cross-links	Link type	PSII subunit(s)	Distance (Å)
419K-423K	Intra-link	CP47-CP47	14.6
227K-505K	Intra-link	CP47-CP47	28
497K-505K	Intra-link	CP47-CP47	17.6
180K-419K	Inter-link	PsbO-CP47	14.6
186K-418K	Inter-link	PsbO-CP47	15.8
185K-180K	Intra-link	PsbO-PsbO	15.8
185K-186K	Intra-link	PsbO-PsbO	13.2
186K-180K	Intra-link	PsbO-PsbO	12.8
100K-96K	Intra-link	PsbQ-PsbQ	11.1
106K-100K	Intra-link	PsbQ-PsbQ	16.6
102K-106K	Intra-link	PsbQ-PsbQ	8.3
120K-180K	Inter-link	PsbQ-PsbO	*
120K-59K	Inter-link	PsbQ-PsbO	*
120K-96K	Inter-link	PsbQ-PsbQ	*

* = Crystallographic data not available

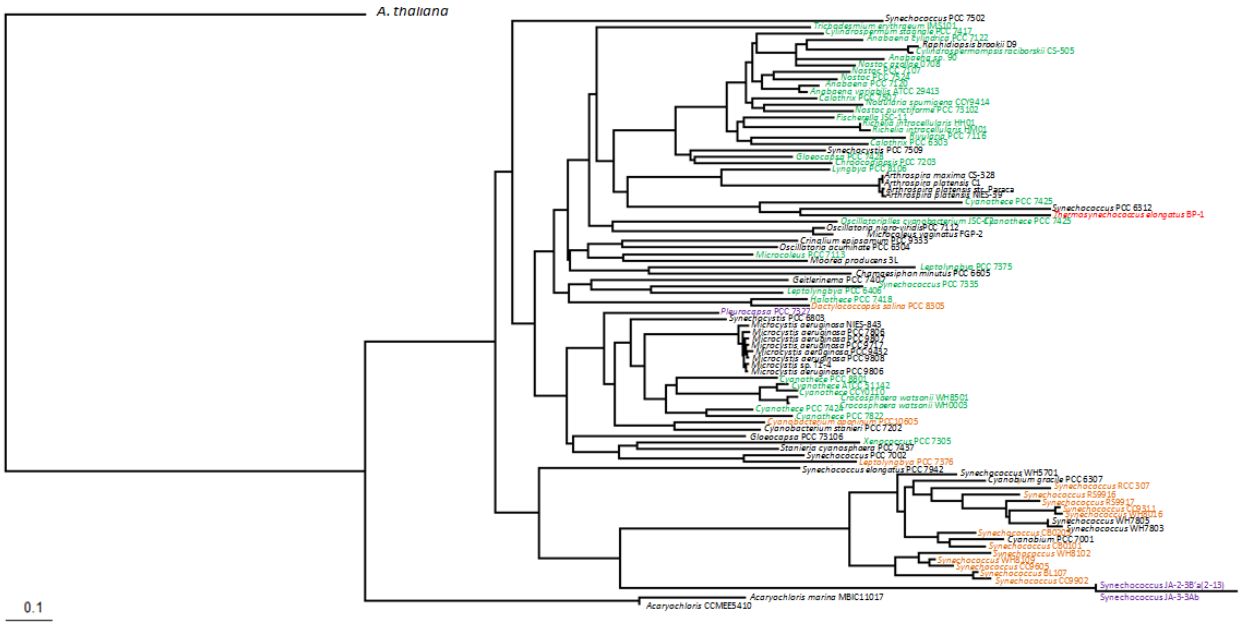


Figure S1. Phylogenetic tree showing evolutionary relationship between PsbQ sequences from different cyanobacterial species.

Black = mesophilic and non-nitrogen-fixing. Green = mesophilic and nitrogen-fixing; red = thermophilic and non-nitrogen-fixing; purple = thermophilic and nitrogen-fixing. Orange = mesophilic, and nitrogen-fixing ability not provided in Joint Genome Institute Integrated Microbial Genomes database (http://img.jgi.doe.gov/cgi-bin/m/main.cgi) or the respective culture collection's database. Scale bar units: substitutions per nucleotide site. The PsbQ sequence of *M. aeruginosa* sp. PCC 9809 was identical to that of *M. aeruginosa* NIES-843; the PsbQ sequences of *M. aeruginosa* sp. PCC 9443, *M. aeruginosa* sp. PCC 7941, *M. aeruginosa* sp. PCC 9701, *M. aeruginosa* TAIHU98, and *M. aeruginosa* SPC 777 were identical to that of *M. aeruginosa* sp. PCC 9717; the PsbQ sequence of *M. aeruginosa* DIANCHI905 was identical to that of *M. aeruginosa* sp. PCC 7806; the PsbQ sequence of *S. elongatus* sp. PCC 7942 was identical to that of *S. elongatus* sp. PCC 6301; the PsbQ sequence of *Cyanothece* sp. PCC 8802 was identical to that of *Cyanothece* sp. PCC 8801; the PsbQ sequence of *Cyanothece* sp. ATCC 51472 was identical to that of *Cyanothece* sp. ATCC 51142; the PsbQ sequence of *Calothrix parietina* was identical to that of *Calothrix* sp. PCC 6303; for simplicity, only the latter member of each of these groups of species was listed in the phylogenetic tree.

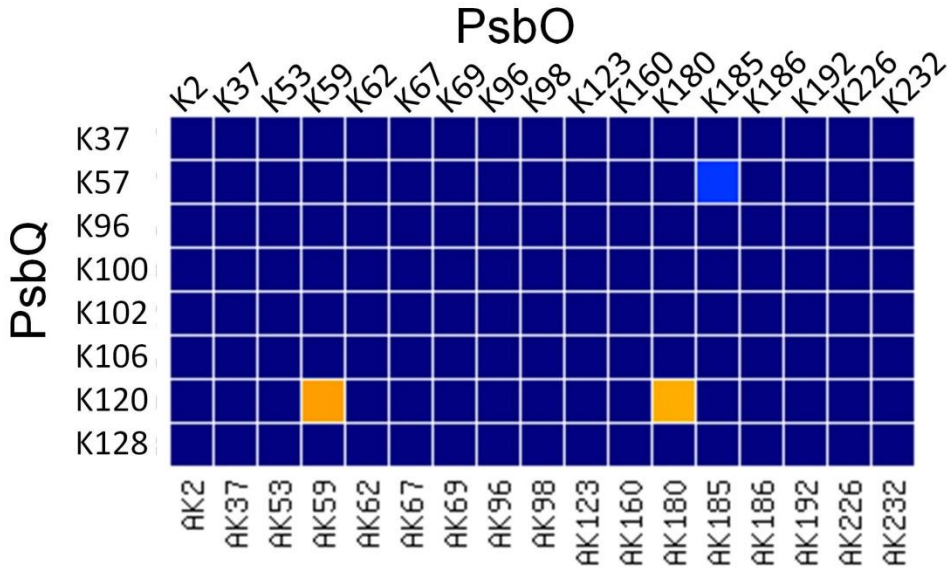


Figure S2. BS^3-induced cross-links between PsbQ and PsbO determined by tandem LC-MS and a subsequent database search using MassMatrix. Each cell in the heat map represents a cross-link between two lysine residues. The cross-linking reagent to PSII mole ratio was 20:1 and the final concentration of PSII was 0.2 mg/mL chlorophyll a. The cross-link between PsbQ:K57-PsbO:K185 was considered to be a false positive because of its high charge state (+8) and m/z value of 380, a known instrument-specific "phantom" peak.

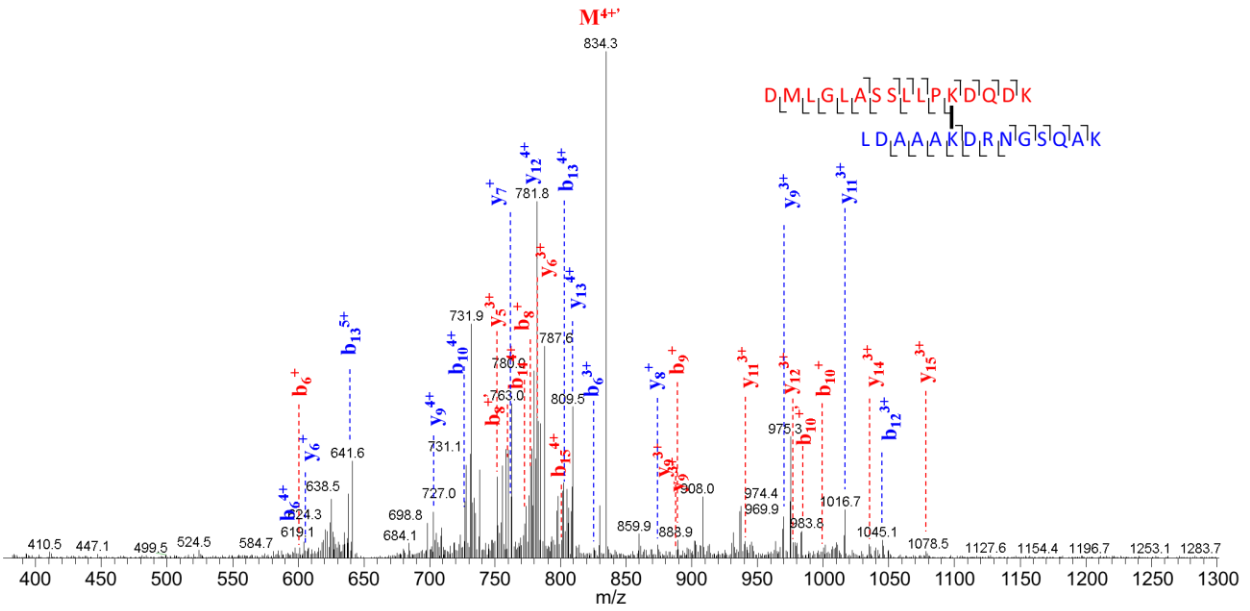


Figure S3. Mass spectrometric data showing an intermolecular PsbQ-PsbQ cross-linked peptide. Product ion (MS^2) spectrum identified as an intermolecular cross-link between two PsbQ peptides, DMLGLASSLLPK*DQDK and LDAAAK*DRNGSQAK, using DTSSP as cross-linker.

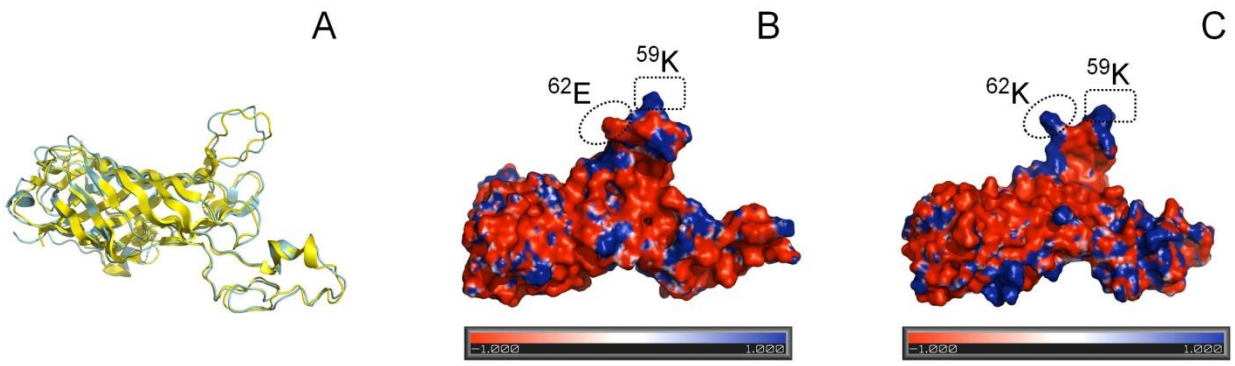


Figure S4. Structural and electrostatic-potential comparison of PsbO from *Synechocystis* 6803 and *T. vulcanus*.

A. The alignment of the predicted 3-D structure of PsbO (*Synechocystis* 6803, blue) and PsbO from 3ARC (yellow). Electrostatic potential surface of PsbO from *T. vulcanus* (**B**) and *Synechocystis* 6803 (**C**). K59 and K62 are labeled to show the homology and electrostatic potential differences.

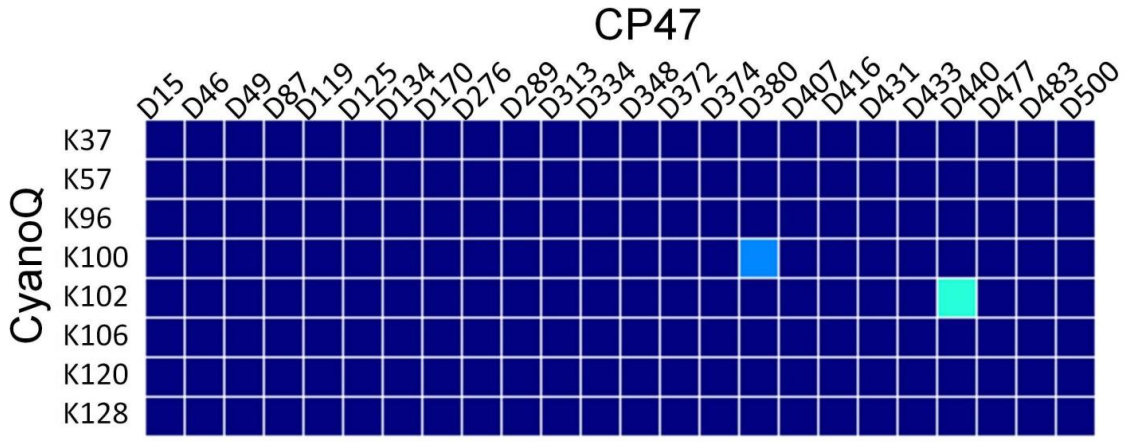


Figure S5. EDC-induced cross-links between PsbQ and CP47 determined by tandem LC-MS/MS and a subsequent database search in MassMatrix.

The turquoise-shaded cell in the heat map represents a cross-link between K102 of PsbQ and D440 of CP47. The light-blue-shaded cell indicating a cross-link between K100 of PsbQ and K380 of CP47 was regarded as a false positive because of its high charge state (+8) and m/z value of 380, a known instrument-specific "phantom" peak.

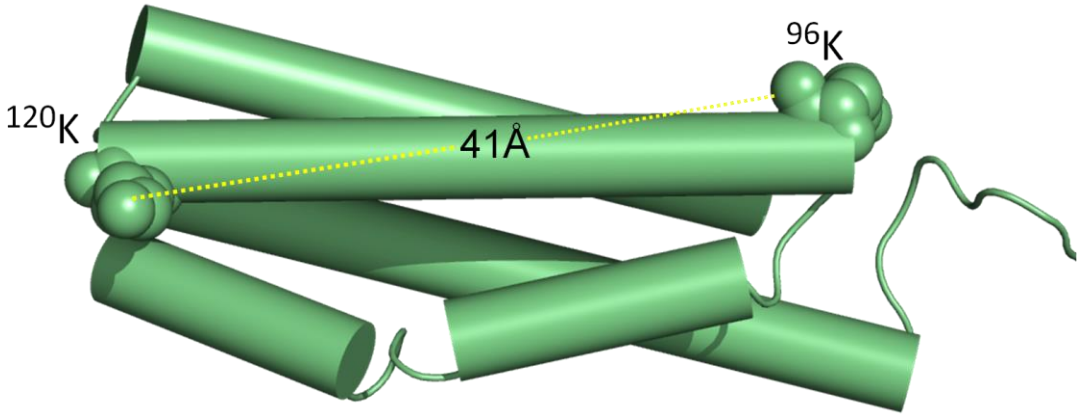


Figure S6. Distance between K96 and K120 in the *Synechocystis* 6803 PsbQ crystal structure. The distance is 41 Å. PDB ID: 3LS1. K96 and K120, spheres.

Chapter 5

Conclusions, future directions, and additional work

Summary and conclusions of this work

The work in Chapter 2 used cross-linking and mass spectrometry to demonstrate that Psb28, a PSII protein that binds to the RC47 assembly intermediate, is located on the cytosolic surface of PSII in close association with PsbE and PsbF, the α- and β-subunits of cytochrome b_{559}. Though Psb28 binds only transiently to PSII during assembly (and possibly again during disassembly), it exerts a protective effect, assisting efficient conversion of RC47 into fully assembled PSII (Sakata et al., 2013). Based on its structural location within the complex that we have determined, we have proposed three possible mechanisms for its protective effect, which can be probed in future experiments. The common thread in these mechanisms is that they minimize unproductive light-induced electron transfer reactions within RC47.

PSII assembly intermediates past the RC stage have a dilemma: *enough* of the PSII electron transfer chain has been assembled to enable light-induced primary charge separation, but *not enough* of the chain has yet been assembled to enable a complete, productive cycle that oxidizes water, generates molecular oxygen, and reduces plastoquinone. The in-between PSII assembly intermediates can thus generate highly oxidizing species similar to fully assembled PSII, but cannot quench them as effectively as fully assembled PSII can. As seen in the work in Chapter 3, even fully assembled PSII is highly susceptible to oxidative damage; how much more so can we expect the assembly intermediates to be. Previous work in this laboratory has shown that the Psb27 protein protects assembly intermediates from harmful premature electron transfer chemistry, and binds on the lumenal surface of the complex. Psb28 appears to serve a similar purpose, binding on the cytosolic surface of the complex. Several dozen accessory proteins are now known to associate with PSII assembly intermediates, but the binding site and function of nearly all of them

is unknown. It is possible that some of these proteins are also involved in protecting their respective subcomplexes from their own powerful electron transfer capabilities.

The work in Chapter 3 characterized oxidative modifications that can form on D1, D2, and CP43. It was found that around 50% more residues are modified following light exposure than in control, dark-incubated samples, strengthening the long-held belief that PSII proteins undergo oxdiative damage by ROS produced as an inevitable byproduct of PSII photochemistry. The D1 protein showed the most dramatic response to light (3.3-fold more oxidized residues detected in the light-incubated sample), consistent with the observation that D1 is turned over the fastest of all PSII proteins.

By mapping the lumen-side oxidative modifications detected onto the PSII crystal structure, it was found that they form two distinct, nearly continuous, roughly linear "arms" radiating outward from the buried Mn_4Ca cluster all the way to the surface of PSII. Since in order to be oxidized, these residues must have been exposed to reactive oxygen species (ROS), this geometrical formation appears to track ROS travel from the buried Mn_4Ca cluster (a site of ROS generation), away from the complex and into the bulk solvent. Existence of a channel that allows diffusion of oxygen/ROS away from PSII was proposed earlier as a critical means of minizming ROS-induced damage to the complex (Anderson, 2001). Several of the residues involved in one of the two putative channels identified here (on the CP43 "arm") were proposed earlier to be part of an ROS exit channel (Frankel et al., 2012). In this earlier study, additional oxidized residues nearby were not detected, and therefore such a channel could not be verified and its position could not be described. The current work is consistent with that earlier finding but extends it by describing a full path for this putative channel. The other putative ROS channel identified here has not previously been proposed; earlier results (Frankel et al., 2012; Frankel et al., 2013),

however, are consistent with its location. This study highlights another way in which modern high-resolution tandem MS can reveal structural information about a protein complex.

Like the work in Chapter 2, the study in Chapter 4 used MS-based cross-linking to identify the location of PsbQ within the PSII complex. The results revealed that PsbQ binds on the lumenal surface of PSII in close association with CP47 and PsbO. An inter-protein PsbQ-PsbQ cross-link shows that PsbQ binds at the dimer interface of PSII, consistent with the other two cross-links and with biochemical evidence that its binding stabilizes dimer formation. While it was known that PsbQ is a member of active, fully-assembled cyanobacterial PSII, its binding site was previously unknown since it is not found in any cyanobacterial PSII crystal structure. The work in Chapters 2 and 4 thus highlight how MS-based cross-linking can provide valuable information complementary to, and not obtainable by, higher-resolution structural methods.

Future directions for MS contributions to PSII life-cycle research

MS technology and associated sample preparation techniques are evolving rapidly. Increasing sensitivity and speed of instruments for bottom-up proteomics allows better coverage of transmembrane PSII proteins; for example, coverage of the core D1, D2, CP47, and CP43 proteins is appreciably higher on a Thermo Q-Exactive Plus instrument than has been reported on LTQ-Orbitrap, LTQ-FTICR, and MALDI-TOF instruments (with increased coverage ranging from ~10-40% of the total protein sequence) (Aro et al., 2005; Frankel et al., 2012; Liu et al., 2013a,b). This increased coverage will mean that more PTMs and cross-linked peptides can be identified, and a larger portion of the PSII complex can be mapped by footprinting. PTM analysis, especially using quantitative techniques to compare complexes exposed to different conditions, may help elucidate signals (largely unknown in cyanobacteria) that trigger D1 degradation. The increasing availability of high-sensitivity instruments that can achieve high sequence coverage is

enabling detailed quantitative and non-quantitative global proteomic studies. The new challenge is to reduce the large amounts of information becoming available into specific testable hypotheses for targeted follow-up studies.

Improvements at all stages of the cross-linking workflow are occurring, from linker design to linked-peptide enrichment and software analysis. Specifically, isotope-labeled and MS-cleavable linkers are powerful tools that are just beginning to be applied to PSII research. *In-vivo* cross-linking is a promising approach to detect transient or unstable interactions that are difficult to capture after cell lysis. Cross-linking may enable binding site identification for at least some of the approximately 30 accessory proteins now known or believed to associate with PSII during its life-cycle (Nickelsen and Rengstl, 2013; Järvi et al., 2015). Detecting interactions between PSII subcomplexes and, e.g., D1 degradation proteases or proteins involved in chlorophyll loading would also be of prime interest.

Intact-mass measurements of the large core PSII proteins D1, D2, CP47, and CP43 were reported in several studies (Sharma et al., 1997b; Whitelegge et al., 1998; Huber et al., 2004; Thangaraj et al., 2010), with detection of the phosphorylated form of D1 as well in some cases (Whitelegge et al., 1998; Huber et al., 2004). However, their top-down analysis has not yet been achieved. Top-down technology is continuously developing, especially methods for increased product-ion sequence coverage (Frese et al., 2012; Shaw et al., 2013; Brunner et al., 2015) and analysis of larger integral membrane proteins and their PTMs (Ryan et al., 2010; Howery et al., 2012). Such analysis will make it easier to identify nearly-stoichiometric (and potentially important) PTM events from the many trace ones found under different conditions, not an easy task using bottom-up MS.

Native MS is capable of analyzing certain intact membrane protein complexes, although the technology is still developing, and no one approach works for all protein complexes (reviewed in Mehmood et al., 2015). Native MS analysis of PSII has not yet been demonstrated, but the technique could in theory serve as a complementary method to native gels to characterize the distribution of PSII subcomplexes under various conditions, and their components. This might be particularly useful to address the stoichiometry of accessory proteins and cofactors, and could add a new tool to address the long-standing question of chlorophyll loading in PSII.

General conclusion

The use of MS in PSII life-cycle research has been fueled by improvements in sample preparation methods for analysis of membrane proteins, increasing availability of MS instrumentation, and significant advances in instrument sensitivity, speed, and mass accuracy. Techniques from each of the four pillars of proteomics will continue to be employed to study the PSII life cycle. These techniques have addressed a wide range of questions regarding the composition of PSII complexes, the time-dependent dynamic changes of individual subunits and complexes under different environmental conditions, and the tertiary and quaternary structure of PSII complexes.

Modern MS techniques provide a higher level of detail and confidence than previous methods; examples are identification of a protein's phosphorylation site instead of mere detection of a phosphorylated protein, and identification of specific cross-linked residues instead of only suggestive evidence that two particular proteins are cross-linked to each other. For other applications, the use of MS permits entirely new questions to be asked (e.g., what proteins are present on a proteome-wide scale for a purified PSII complex).

The new information has opened up new questions about function. For example, what are the physiological roles of the many new PTMs that have been identified? What purpose does an accessory protein serve by binding at this particular location on a PSII complex? Put in the context of this study, are some of the many detectable oxidative modifications more important than others, e.g. for signaling D1 degradation? What are the functional implications of the newly identified Psb28 binding site? (See above for discussion of this latter question). In some cases, the sensitivity of MS is a potential pitfall: identification of a protein in a PSII sample does not necessarily mean it is a stoichiometric component, or that it associates specifically with the complex at all. Thus, information from MS should be a starting point for more targeted genetic and biochemical studies, and MS is one component of an expanding toolbox for PSII life-cycle research. Rapidly developing MS technology promises continued contributions to this field, which has a wide range of fascinating questions about membrane protein complex composition, dynamics, and structure yet to be answered.

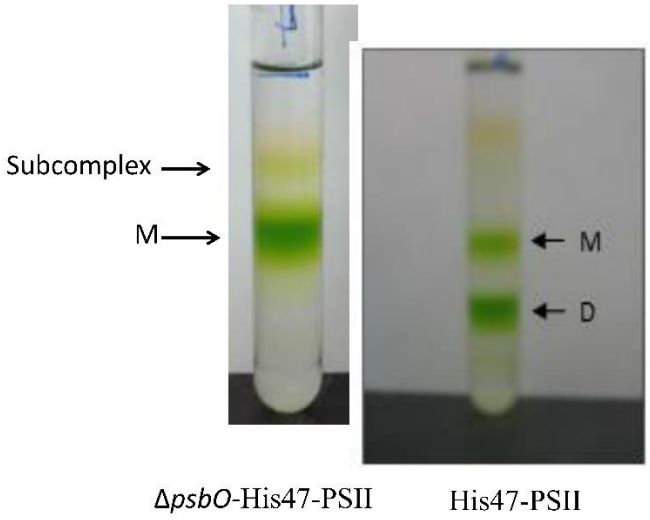
Additional work: Preliminary characterization of a PSII complex lacking the RC subunits

Results and Discussion

After glycerol gradient ultracentrifugation of Δ*psbO*-His47-PSII as described in Chapter 2, an unidentified light green band was observed above the PSII monomer (Fig. 1A). This band was collected and analyzed by gel electrophoresis and immunoblotting (Fig. 2). Surprisingly, no D1 or D2 gel band was observed for this complex. Confirming this observation, MS analysis of this subcomplex detected no D1 or D2 peptides, and also did not detect PsbE, PsbF, and PsbI; these five proteins comprise the core "reaction center" (RC) subcomplex of PSII, the minimal unit necessary for primary charge separation (Nanba and Satoh, 1987; Ikeuchi and Inoue, 1988). We routinely identify all five of these proteins by MS in our PSII preparations, and indeed they were all identified in the PSII monomer band. MS did identify the PSII proteins CP47, CP43, PsbH, PsbL, PsbQ, PsbT, and PsbY, however, in the subcomplex, which are also all proteins routinely identified by bottom-up MS in our PSII preparations. It appears from these results that the newly identified subcomplex is specifically lacking the five RC components, D1, D2, PsbE, PsbF, and PsbI. We refer to it as the "no-RC" or NRC complex.

To check if the NRC complex formed as an artifact of the solubilization/ultracentrifugation procedure (perhaps due to destabilization caused by the Δ*psbO* mutation, since this band was not observed in the His47-PSII sample), we collected just the Δ*psbO*-His47-PSII monomer band after ultracentrifugation, resolubilized it in 2.4% dodecyl maltoside (DM) (3x the concentration used originally), and performed glycerol gradient ultracentrifugation again (Fig. 1B). The NRC band was not observed in this complex, indicating it is not an artifact of the solubilization/ultracentrifugation procedure. This control experiment will be repeated, however,

A.



B.

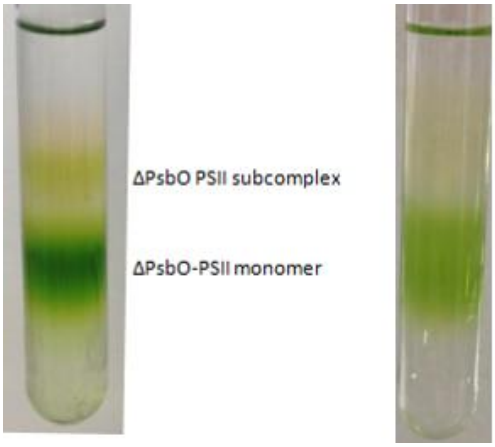


Figure 1. Glycerol gradient ultracentrifugation of purified PSII complexes reveals an unidentified subcomplex in the Δ*psbO*-His47-PSII sample.

A. M, PSII monomer; D, PSII dimer. 100 µg Chl *a* loaded per tube.

B. Left, after ultracentrifugation of Δ*psbO*-His47-PSII as in **A.** Right, identical ultracentrifugation after solubilization of just the isolated lower band (PSII monomer) in the tube on the left.

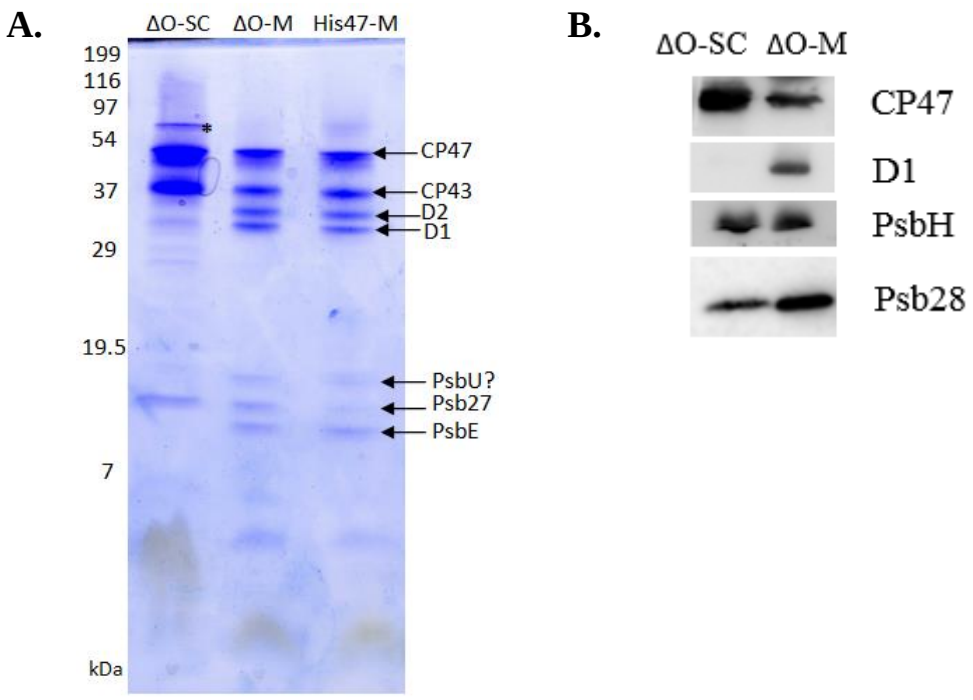


Figure 2. Analysis of the components of the newly identified subcomplex by gel electrophoresis (A) and immunoblotting (B). 1.5 µg Chl *a* loaded on the gel in **A.** MS analysis of the band marked with an asterisk showed it contained CP47 and CP43.

using 0.8% DM instead of 2.4%, since it is conceivable that the NRC complex might be further dissociated in presence of this increased amount of DM.

Despite loading equal levels of chlorophyll on the gel (Fig. 1A), the CP47 and CP43 bands were significantly more pronounced than for the control PSII samples loaded on the same gel. Since CP47 and CP43 contain the large majority of Chl *a* molecules found in PSII, our preliminary assessment is that the absence of D1 and D2 is not enough to account for the more pronounced CP47 and CP43 bands, given equal Chl *a* loading on the gel. The explanation for this observation

is that CP47 and CP43 are under-chlorophyll-loaded in this subcomplex. Experiments are ongoing to investigate this possibility and to characterize this NRC complex further.

Materials and methods

Cell culture, PSII purification, and protein analysis

The Δ*psbO*-His47 strain of *Synechocystis* 6803 was cultured, and PSII was purified, as described in Chapter 2. Besides the major PSII band obtained following glycerol gradient ultracentrifugation, an additional lighter green band was observed above it. This band was collected and concentrated using a Vivaspin 500 centrifugal concentrator (50 kDa molecular weight cutoff) (Vivaproducts, Littleton, MA). Protein gel electrophoresis and immunoblotting were performed as described previously (Kashino et al., 2001; Kashino et al., 2002).

Proteolytic digestion and LC-MS/MS

PSII precipitation and digestion were carried out as described in Chapter 2. LC-MS/MS was carried out as described in Chapter 4, and the Mascot search engine was used to identify proteins in the sample based on the entire *Synechocystis* 6803 genome.

References

Anderson, J.M. (2001). Does functional photosystem II complex have an oxygen channel? *FEBS Lett.* 488, 1-4.

Aro, E.-M., Suorsa, M., Rokka, A., Allahverdiyeva, Y., Paakkarinen, V., Saleem, A., et al. (2005). Dynamics of photosystem II: a proteomic approach to thylakoid protein complexes. *J. Exp. Bot.* 56, 347-356.

Brunner, A.M., Lossl, P., Liu, F., Huguet, R., Mullen, C., Yamashita, M., et al. (2015). Benchmarking multiple fragmentation methods on an Orbitrap Fusion for top-down phospho-proteoform characterization. *Anal. Chem.* 87, 4152-4158.

Frankel, L.K., Sallans, L., Bellamy, H., Goettert, J.S., Limbach, P.A., and Bricker, T.M. (2013). Radiolytic mapping of solvent-contact surfaces in Photosystem II of higher plants: Experimental identification of putative water channels within the photosystem. *J. Biol. Chem.* 288, 23565-23572.

Frankel, L.K., Sallans, L., Limbach, P., and Bricker, T.M. (2012). Identification of oxidized amino acid residues in the vicinity of the Mn_4CaO_5 cluster of Photosystem II: Implications for the identification of oxygen channels within the photosystem. *Biochemistry* 51, 6371-6377.

Frese, C.K., Altelaar, A.F.M., van den Toorn, H., Nolting, D., Griep-Raming, J., Heck, A.J.R., et al. (2012). Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal. Chem.* 84, 9668-9673.

Howery, A.E., Elvington, S., Abraham, S.J., Choi, K.H., Dworschak-Simpson, S., Phillips, S., et al. (2012). A designed inhibitor of a CLC antiporter blocks function through a unique binding mode. *Chem. Biol.* 19, 1460-1470.

Huber, C.G., Walcher, W., Timperio, A.M., Troiani, S., Porceddu, A., and Zolla, L. (2004). Multidimensional proteomic analysis of photosynthetic membrane proteins by liquid extraction-ultracentrifugation-liquid chromatography-mass spectrometry. *Proteomics* 4, 3909-3920.

Ikeuchi, M., and Inoue, Y. (1988). A new 4.8-kDa polypeptide instrinsic 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.

Järvi, S., Suorsa, M., and Aro, E.-M. (2015). Photosystem II repair in plant chloroplasts - Regulation, assisting proteins and shared components with photosystem II biogenesis. *Biochim. Biophys. Acta* 1847, 900-909.

Kashino Y, Koike H, Satoh K (2001) An improved sodium dodecyl sulfate-polyacrylamide gel electrophoresis system for the analysis of membrane protein complexes. *Electrophoresis* 22(6):1004–1007.

Kashino Y, et al. (2002) Low-molecular-mass polypeptide components of a Photosystem

II preparation from the thermophilic cyanobacterium *Thermosynechococcus vulcanus*. *Plant Cell Physiol* 43(11):1366–1373.

Liu, H., Chen, J., Huang, R.-C., Weisz, D., Gross, M.L., and Pakrasi, H.B. (2013a). Mass spectrometry-based footprinting reveals structural dynamics of Loop E of the chlorophyll-binding protein CP43 during Photosystem II assembly in the cyanobacterium *Synechocystis* 6803. *J. Biol. Chem.* 288, 14212-14220.

Liu, H., Zhang, H., Niedzwiedzki, D., Prado, M., He, G., Gross, M.L., et al. (2013b). Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. *Science* 342, 1104-1107.

Mehmood, S., Allison, T.M., and Robinson, C.V. (2015). Mass spectrometry of protein complexes: From origins to applications. *Annu. Rev. Phys. Chem.* 66, 453-474.

Nanba, O., and Satoh, K. (1987). Isolation of a photosystem II reaction center consisting of D-1 and D-2 polypeptides and cytochrome *b*-559. *Proc. Natl. Acad. Sci. USA* 84, 109-112.

Nickelsen, J., and Rengstl, B. (2013). Photosystem II assembly: From cyanobacteria to plants. *Annu. Rev. Plant Biol.* 64, 609-635.

Ryan, C.M., Souda, P., Bassilian, S., Ujwal, R., Zhang, J., Abramson, J., et al. (2010). Post-translational modifications of integral membrane proteins resolved by top-down Fourier transform mass spectrometry with collisionally activated dissociation. *Mol. Cell. Proteomics* 9, 791-803.

Sakata S, Mizusawa N, Kubota-Kawai H, Sakurai I, Wada H (2013) Psb28 is involved in recovery of photosystem II at high temperature in *Synechocystis* sp. PCC 6803. *BBA-Bioenergetics* 1827(1):50-59.

Sharma, J., Panico, M., Barber, J., and Morris, H.R. (1997b). Purification and determination of intact molecular mass by electrospray ionization mass spectrometry of the photosystem II reaction center subunits. *J. Biol. Chem.* 272, 33153-33157.

Shaw, J.B., Li, W., Holden, D.D., Zhang, Y., Griep-Raming, J., Fellers, R.T., et al. (2013). Complete protein characterization using top-down mass spectrometry and ultraviolet photodissociation. *J. Am. Chem. Soc.* 135, 12646-12651.

Thangaraj, B., Ryan, C.M., Souda, P., Krause, K., Faull, K.F., Weber, A.P.M., et al. (2010). Data-directed top-down Fourier-transform mass spectrometry of a large integral membrane protein complex: Photosystem II from *Galdieria sulphuraria*. *Proteomics* 10, 3644-3656.

Whitelegge, J.P., Gundersen, C.B., and Faull, K.F. (1998). Electrospray-ionization mass spectrometry of intact intrinsic membrane proteins. *Protein Sci.* 7, 1423-1430.