Architecture of the Photosynthetic Oxygen-Evolving Center

Kristina N. Ferreira,[1]* **Tina M. Iverson,**[2]* **Karim Maghlaoui,**[1] **James Barber,**[1]† **So Iwata**[1,2,3]†

Photosynthesis uses light energy to drive the oxidation of water at an oxygen-evolving catalytic site within photosystem II (PSII). We report the structure of PSII of the cyanobacterium *Thermosynechococcus elongatus* at 3.5 angstrom resolution. We have assigned most of the amino acid residues of this 650-kilodalton dimeric multisubunit complex and refined the structure to reveal its molecular architecture. Consequently, we are able to describe details of the binding sites for cofactors and propose a structure of the oxygen-evolving center (OEC). The data strongly suggest that the OEC contains a cubane-like Mn_3CaO_4 cluster linked to a fourth Mn by a mono-μ-oxo bridge. The details of the surrounding coordination sphere of the metal cluster and the implications for a possible oxygen-evolving mechanism are discussed.

Photosynthesis uses light energy to couple the formation of molecular dioxygen to the fixation of CO_2. This process simultaneously generates an aerobic atmosphere and produces a readily usable carbon source, both of which act to sustain almost all life on this planet. Central to this process is photosystem II (PSII), which catalyzes one of the most thermodynamically demanding reactions in biology: the photoinduced oxidation of water. In this way it provides a primary source of reducing equivalents (water-derived electrons and protons), which, with an additional input of energy from photosystem I (PSI), are used to convert carbon dioxide to biomass, food, and fuel. Molecular dioxygen is released as a by-product. Given the essential role of PSII in maintaining the biosphere, it is of considerable importance to elucidate its catalytic mechanisms, particularly those involved in the water oxidation process.

Much knowledge of the catalytic mechanisms of PSII has been obtained from a wide range of approaches (*1*–*3*), but the precise molecular details of water oxidation catalyzed by the oxygen-evolving center (OEC) remain elusive. Two x-ray studies at 3.8 Å (*4*) and 3.7 Å (*5*) have given the first direct hints as to the structure of PSII, but these

models revealed neither the complete details of the OEC nor the surrounding protein residues, information vital for formulating a reaction mechanism for water oxidation.

Here we report the crystal structure of a cyanobacterial PSII complex at 3.5 Å resolution with more than 90% of its amino acid residues traced. We have assigned all the subunits of the PSII complex to specific gene products and provide a description of the protein environment of the various redox-active cofactors and pigments within this complex. We conclude that the OEC is a cubane-like Mn_3CaO_4 cluster with a mono-μ-oxo bridge to a fourth Mn ion. On the basis of our proposed structure of the metal center and its protein environment, we discuss the possible mechanism of the oxygen-evolving reaction.

Structure determination and overall architecture. Dimeric PSII complexes were isolated from *Thermosynechococcus elongatus*. Details of sample preparation, crystallization, and structure determination are provided in (*6*). The molecular replacement trials using the coordinates of earlier works [1FE1, 3.8 Å resolution, C_α atoms only, *R* factor 59% (*4*); and 1IZL, 3.7 Å resolution, mainly polyalanine, *R* factor 53% (*5*)] produced density maps that were highly biased to their models and therefore could not be used. For this reason, the experimental phases for the structure presented here were calculated by the method of multiple isomorphus replacement with six heavy atom derivatives (Table 1). After fitting to the experimental map, the model was refined at 3.5 Å resolution (Table 1 and fig. S1).

The crystallographic asymmetric unit contains a dimer of PSII, where the two mono-

mers in the dimer are almost identical. In the model, the monomer contains 19 protein subunits that have been clearly assigned to electron density, with the exception of the tentative assignment of PsbN, where the side chains are not included in the model as a result of disorder. Each monomer contains 36 chlorophyll a (Chl) and 7 all-trans carotenoids assumed to be β-carotene molecules. However, two of the Chl molecules (Chl^{35} and Chl^{36}) are loosely attached to the complex and disordered. Additionally, there is unassigned electron density remaining that could be carotenoid molecules or the fatty acid tails of lipid molecules. Each monomer model also includes one OEC, one heme *b*, one heme *c*, two plastoquinones, two pheophytins, one nonheme Fe, and two bicarbonates (one is tentatively assigned as an unknown nonprotein ligand at the OEC).

The PS II dimer (Fig. 1) has dimensions of 105 Å depth (45 Å in membrane), 205 Å length, and 110 Å width. The overall structure (Fig. 1) is similar to those reported previously (*4*, *5*); the root mean square (RMS) deviations of the C_α backbone between the dimer structure and 1FE1 (*4*) and 1IZL (*5*) are 2.1 Å (for 1359 C_α atoms) and 1.9 Å (for 3519 C_α atoms), respectively. However, compared with previous models (*4*, *5*), the sequence assignment has been substantially improved. We have newly assigned or reassigned 3916 residues and built the side chains. The improvement is evident in the models for the extrinsic subunits, extrinsic domains of D1, D2, CP43, and CP47, and the small transmembrane subunits.

Within the dimer, the monomers are related by a noncrystallographic twofold axis perpendicular to the membrane plane. Our model of PSII consists of 16 integral membrane subunits composed of 35 transmembrane helices and three lumenal subunits. The monomer is characterized by pseudo-twofold symmetry, which relates the D1, CP47, and PsbI subunits to the D2, CP43, and PsbX subunits (Fig. 1B).

Protein subunits. The D1 and D2 subunits form the center of the PSII monomer complex. Each of these subunits comprises five transmembrane helices (A to E) organized in a manner almost identical to that of the L and M subunits of the reaction center of photosynthetic purple bacteria (bRC) (*7*, *8*), with RMS deviation between PSII and bRC from *Rhodopseudomonas viridis* of 1.9 Å for 395 C_α atoms. However, the C-terminal domains and the loops joining the transmembrane helices are more extended in the case of the D1 and D2 subunits compared with bRC, especially on the lumenal side close to the OEC. Flanking the opposite sides of the D1/D2 heterodimer are the CP43 (PsbC) and CP47 (PsbB) subunits, each having six trans-

membrane helices (I to VI) arranged in a circular manner similar to the N-terminal domains of the PsaA and PsaB proteins of PSI (*9*). The transmembrane helices and chlorophyll molecules of each of these subunits exhibit an internal pseudo-threefold symmetry. Violating this internal symmetry in each subunit is a large lumenal domain between transmembrane helices V and VI. The large domain of CP43 contains two long and three short helices (fig. S2). In the case of CP47, the lumenal domain contains two long and four short helices and three β sheets (fig. S3).

Outside the pseudo-symmetric CP43/D1–D2/CP47 core, there are 13 transmembrane helices in the map, which are assigned to specific low-molecular-weight subunits based on tracing their amino acid sequences. (Fig. 1B). The assignment of three of these helices, PsbE, PsbF, and PsbK, where PsbE and PsbF are the α and β subunits, respectively, of cytochrome b559 (Cyt *b*559), agrees with earlier studies (*4*, *5*, *10*). The remaining 10 transmembrane heli-

ces are newly assigned to 9 low-molecular-weight subunits, including PsbZ, which has two transmembrane helices (I and II). One transmembrane helix reported previously (*4*, *5*) adjacent to Cyt *b*559, and designated PsbX (*4*), was not present in our electron density map. With the exception of Cyt *b*559, the functions of most of these small subunits are unclear. Our assignments, however, indicate that PsbL, PsbM, and PsbT are involved in dimer formation. The symmetrically related PsbI and PsbX proteins stabilize the peripheral chlorophylls of the D1 and D2 proteins (Chl_{ZD1} and Chl_{ZD2}). Four small subunits (PsbJ, PsbK, putative PsbN, and PsbZ) adjacent to CP43 may facilitate carotenoid binding, because four of the seven assigned β-carotene molecules are found in their vicinity (Fig. 1B).

Single copies of each of the three extrinsic proteins, PsbO, PsbU, and PsbV, are located on the lumenal surface (Fig. 1A and fig. S4). Taken together, these extrinsic proteins and the C terminus of the D2 protein form a "cap" over the OEC. None are involved directly in the

ligation of the OEC. However, PsbO, which forms an eight-stranded β barrel (fig. S4A), stabilizes the backbone conformation of the AB loop and C terminus of the D1 protein, which provide the majority of the ligands to the OEC. Additionally, our sequence assignment of the PsbO subunit revealed that a large loop between strands 5 and 6 forms a part of a hydrophilic pathway connecting the OEC with the lumenal surface. Although the PsbO β barrel resembles the β barrel of porins (*5*), it is not an open tube but contains a number of bulky hydrophobic side chains.

Electron transfer. The reactions of PSII are powered by light-driven primary and secondary electron transfer processes across the reaction center (RC), composed of the D1 and D2 subunits. Figure 2A shows the cofactors involved in these electron transfer reactions.

Upon illumination, an electron is ejected from the excited primary electron donor P680, a Chl located toward the lumenal surface (P_{D1} in Fig. 2A is likely to be P680, as discussed

Table 1. Data collection, refinement, and phasing statistics for PSII structure determination. SLS, Swiss Light Source; ESRF, European Synchrotron Radiation Facility; EMP, ethylmercury phosphate; PCMBS, p-chloromercuribenzoylsulfonate.

Data set				Data collection and phasing	
	Native	Native (Mn-edge)	Native (for Ca detection)	CdCl$_2$	K$_3$[Au(Cn)$_6$]
Beam line	SLS X06SA	SLS X06SA	SLS X06SA	ESRF ID29	ESRF ID29
Wavelength (Å)	1.0076	1.89340	2.25430	1.0052	1.0052
Resolution (Å)	40.0–3.5	40.0–3.8	40.0–3.8	40.0–3.8	40.0–3.8
Total observation	298,731	168,066	212,461	245,580	221,188
Unique reflections	103,604	71,063	79,041	78,778	73,595
Completeness (%)*	87.3 (80.9)	74.6 (63.8)	83.6 (62.3)	82.0 (55.1†)	77.8 (61.7†)
Redundancy	2.88	2.37	2.68	3.11	3.00
R$_{sym}$ *‡	0.08 (0.43)	0.11 (0.50)	9.0 (0.55)	0.08 (0.43)	0.09 (0.43)
I/σ(I)*	10.4 (2.0)	7.3 (2.2)	7.4 (2.1)	9.9 (1.6)	9.0 (2.0)
Phasing power§	N/A	N/A	N/A	0.65	0.53

Data set	EMP1	EMP2	PCMBS	TaBr$_6$
Beam line	SLS X06SA	SLS X06SA	SLS X06SA	SLS X06SA
Wavelength (Å)	0.96863	0.96863	0.96863	0.96863
Resolution (Å)	40.0–4.1	40.0–5.1	40.0–3.8	40.0–3.8
Total observation	215,331	68,313	179,833	172,328
Unique reflections	80,342	26,529	75,502	71,395
Completeness (%)*	73.0 (58.5†)	64.4 (48.9†)	73.5 (31.9†)	74.6 (53.8†)
Redundancy	2.68	2.57	2.38	2.41
R$_{sym}$ *‡	0.10 (0.40)	0.08 (0.44)	0.12 (0.50)	0.08 (0.33)
I/σ(I)*	6.1 (1.3)	11.6 (1.8)	5.9 (2.1)	11.2 (2.6)
Phasing power§	0.90	0.41	0.57	0.62

	Refinement
Resolution (Å)*	20.0–3.5 (3.56–3.50)
Number of reflections	103,485 (85.7%)
R factor*‖	0.303 (0.340)
R$_{free}$ *¶	0.346 (0.384)
Average B values (Å²)	74.2
RMS deviations from ideal values	
Bond length (Å)	0.014
Bond angles (°)	2.11
Dihedral angles (°)	19.4
Improper torsion angles (°)	2.04

*Values in parentheses are for the highest resolution shell. †The completeness of the highest shell is decreased from the overlap of the spots caused by the increased mosaicity pronounced in heavy atom soaking experiments. ‡$R_{sym} = \Sigma_h \Sigma_i |I_i(h) - <I(h)>|/\Sigma_h \Sigma_i I_i(h)$, where $I_i(h)$ is the *i*th measurement. §Phasing power is the RMS value of F_h divided by the RMS lack-of-closure error. ‖R factor $= \Sigma_h ||F(h)_{obs}| - |F(h)_{calc}||/\Sigma_h |F(h)|$. ¶$R_{free}$ was calculated for 1% of reflections randomly excluded from the refinement.

below). The electron is quickly transferred toward the stromal surface to the final electron acceptor, plastoquinone Q_B, through Chl_{D1}, pheophytin ($Pheo_{D1}$), and plastoquinone Q_A. After accepting two electrons and undergoing protonation, Q_B is released from PSII into the membrane matrix. The cationic radical P680$^{\bullet+}$

is reduced by a redox-active tyrosine, known as Tyr_Z (Tyr[161] of D1 subunit), to generate a neutral tyrosine radical $Tyr_Z^{\bullet}$, which acts as an oxidant for the water oxidation process at the OEC. The oxidation of two water molecules to produce a dioxygen proceeds in a stepwise manner as described in the S-state cycle (11).

With the exception of Q_A, all the redox-active cofactors involved in the electron transfer processes are located on the D1 side of the RC.

The radical cation P680$^{\bullet+}$ has a very high oxidizing potential, recently estimated to be 1.3 to 1.4 V (12), which is required for the water-splitting reaction. This contrasts with 0.4 V produced by the bacterial equivalent consisting of a "special pair" of bacterio-chlorophylls (BChl), which is reduced by a cytochrome after photooxidation. The two chlorophyll molecules in PSII equivalent to the special pair of the bRC are denoted P_{D1} and P_{D2} (4). Although the overall organization of P_{D1} and P_{D2} is very similar to the special pair, the relative orientation angle between their tetrapyrrole heads differs slightly from their bacterial equivalents. The head groups of P_{D1} and P_{D2} are in direct van der Waals contact, with a Mg-Mg distance of 8.2 Å (Fig. 2B). This is slightly longer than the 7.5 Å observed for the special pair of bRC, which could, together with the difference in the head-group orientation, explain why P680 shows more monomeric character and weaker electronic coupling than its bacterial counterpart (13).

P_{D1} and P_{D2} are located close to Chl_{D1} and Chl_{D2} (Fig. 2B), which are equivalent to the accessory BChls of the bRC. The P680 excited state is delocalized over the four chlorophylls (14, 15), and Chl_{D1}, which is the chlorophyll closest to the active $Pheo_{D1}$, is involved in the initial primary charge separation (16, 17). Unlike accessory BChls in the bRC, which are ligated by histidines, there is no protein ligand for Chl_{D1} and Chl_{D2}; the closest residue to Chl_{D1} is D1 Thr[179] and to Chl_{D2} is D2 Ile[178].

The electrons are transferred from $Pheo_{D1}$ to Q_A, which is a firmly bound plastoquinone. Even though PSII and bRCs use different primary quinone acceptors, the Q_A binding pockets are structurally similar. In PSII, the Q_A is hydrogen bonded to the main-chain amide group of D2 Phe[261] and D2 His[214], where the latter also serves as a ligand to the nonheme Fe (Fig. 3, A and B). The quinone ring is accommodated within a hydrophobic cavity composed of residues, including D2 Ile[213], D2 Thr[217], D2 Met[246], D2 Ala[249], D2 Trp[253], D2 Ala[260], and D2 Leu[267].

The nonheme Fe, which mediates electron transfer from Q_A to Q_B, is positioned on a pseudo-twofold axis of the D1/D2 heterodimer. Like the bRC, the nonheme Fe has four histidine ligands: D1 His[215], D1 His[272], D2 His[214], and D2 His[268] (Fig. 3A). In bRC, a glutamate serves as a fifth ligand, whereas in PSII this ligand is not from the protein. It has been suggested that bicarbonate may function as the fifth ligand to the nonheme Fe in PSII (18) and that bicarbonate has a regulatory function involving electron flow from Q_A to Q_B as well as facilitating the protonation of Q_B (2, 19). In our PSII structure, nonheme Fe is associated with an electron density that is sufficient to accom-
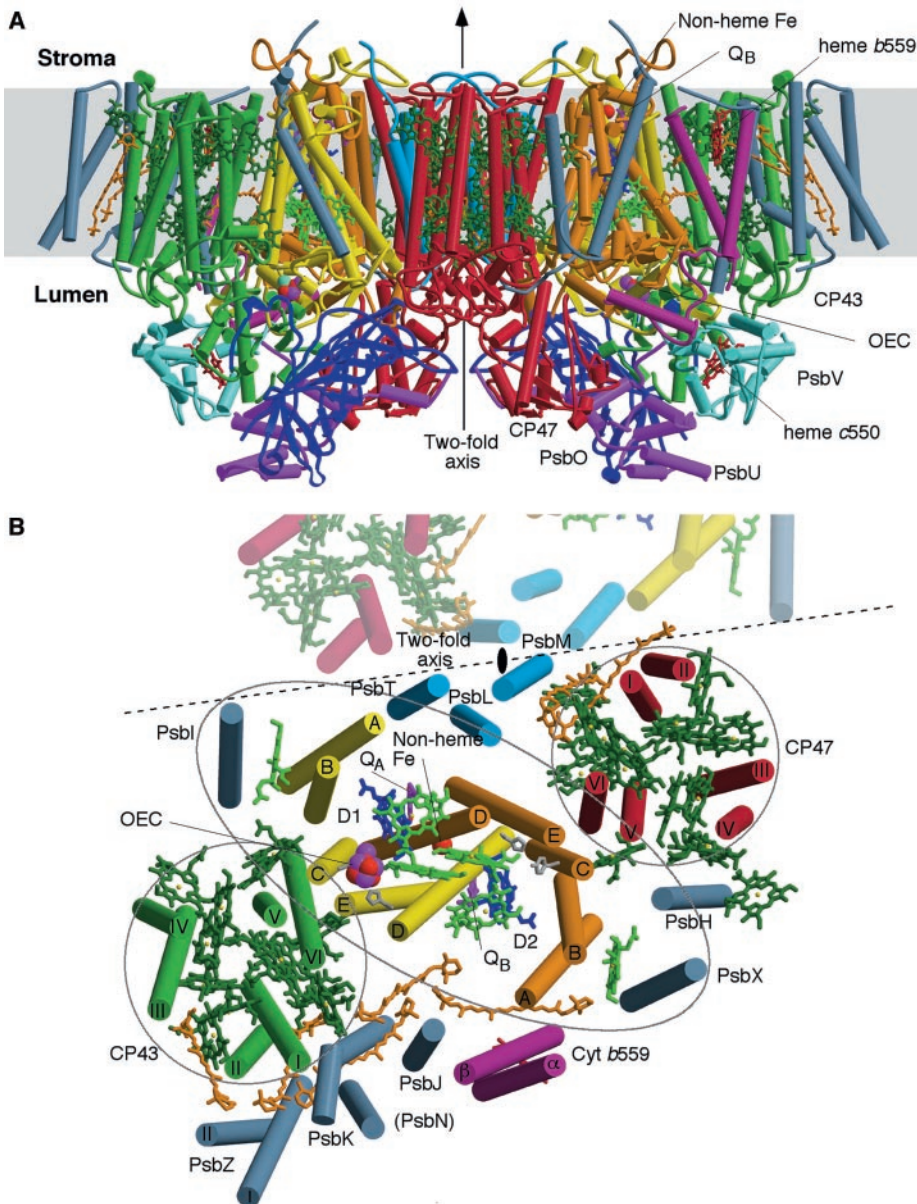

Fig. 1. Overall structure of PSII. (**A**) View of the PSII dimer perpendicular to the membrane normal. Helices are represented as cylinders with D1 in yellow; D2 in orange; CP47 in red; CP43 in green; cyt b559 in wine red; PsbL, PsbM, and PsbT in medium blue; and PsbH, PsbI, PsbJ, PsbK, PsbX, PsbZ, and the putative PsbN in gray. The extrinsic proteins are PsbO in blue, PsbU in magenta, and PsbV in cyan. Chlorophylls of the D1/D2 reaction center are light green, pheophytins are blue, chlorophylls of the antenna complexes are dark green, β-carotenes are in orange, hemes are in red, nonheme Fe is red, Q_A and Q_B are purple. The oxygen-evolving center (OEC) is shown as the red (oxygen atoms), magenta (Mn ions), and cyan (Ca^{2+}) balls. (**B**) View of the PSII monomer along the membrane normal from the lumenal side. A part of the other monomer in the dimmer is shown to emphasize the region of monomer/monomer interaction along the dotted line. The pseudo-twofold axis perpendicular to the membrane plane passing through the nonheme Fe relates the transmembrane helices of the D1/D2 heterodimer, the low molecular subunits, PsbI and PsbX, and CP43 and CP47 as emphasized by the black lines encircling these subunits. Coloring is the same as in (A).
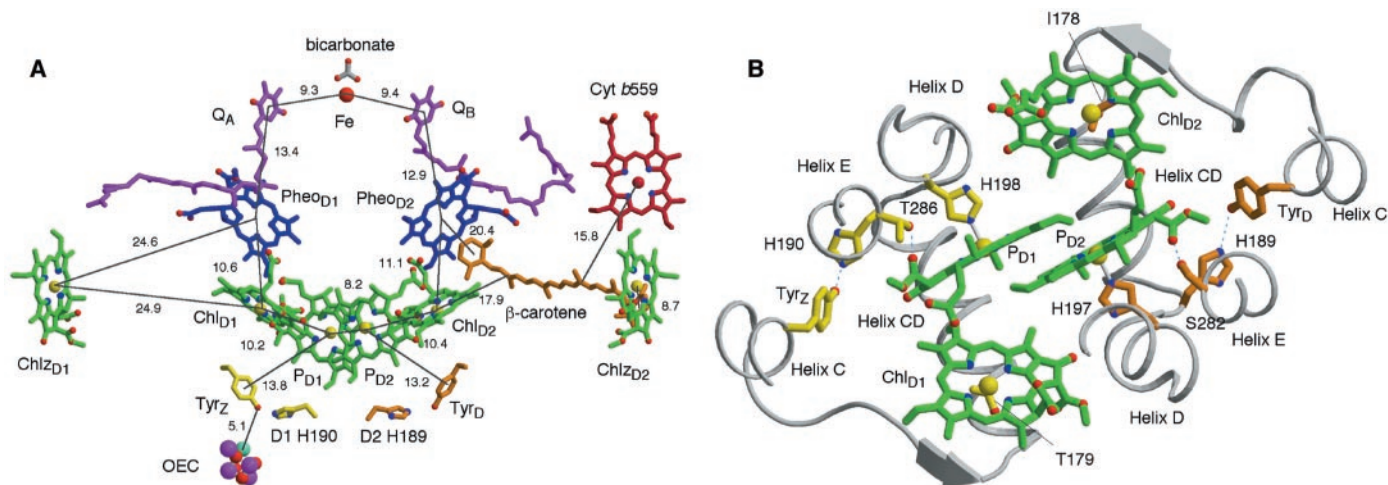
Fig. 2. Cofactors involved in electron transfer. (**A**) Electron transfer cofactors shown perpendicular to the internal pseudo-twofold. Coloring scheme is the same as in Fig. 1. The phytol tails of the chlorophylls and pheophytins have been removed for clarity. The side chains of Tyr$_Z$ (D1 Tyr[161]) and D1 His[190] are shown in yellow, and Tyr$_D$ (D2 Tyr[160]) and D2 His[189] are in orange. The four chlorophylls comprising P680 are in direct van der Waals contact, and other electron transfer distances are given in

Å. (**B**) The P680 dimer of chlorophylls (P$_{D1}$ and P$_{D2}$) and accessory Chls (Chl$_{D1}$ and Chl$_{D2}$). Coloring scheme is the same as in Fig. 1, except that the protein main chain is depicted in light gray, whereas the side-chain bonds and carbon atoms follow the coloring of the protein subunits (D1, yellow; D2, orange). The histidine ligands D1 His[198] and D2 His[197] are shown, as well as the redox-active Tyr$_Z$–D1 His[190] and Tyr$_D$–D2 His[189] pairs. The view is down the pseudo-twofold axis from the stromal side.

modate this anion. Close to this nonprotein density are D1 Tyr[246] and D2 Lys[264], positioned and oriented such that they could stabilize the bicarbonate by hydrogen bonding (Fig. 3A).

There is less conservation between the Q$_B$ sites of PSII and bRC than for the Q$_A$ sites. There are similarities in that the PSII Q$_B$ is hydrogen bonded to D1 Ser[264] and D1 His[215], which is also a ligand for the nonheme Fe, and possibly to the main-chain amide group of D1 Phe[265]. However, the Q$_B$ sites in PSII and the bRC show a 3.5 Å displacement when superimposed using C$_\alpha$ atoms of D1/D2 and the L/M subunits. This is because there is a conformational difference in the loops containing D1 Ser[264] and the equivalent in bRC, which is caused by the insertion of one residue. As a consequence, the volume of the Q$_B$ binding pocket, composed of the residues including D1 Met[214], D1 Leu[218], D1 Ala[251], D1 Phe[255], and D1 Leu[271], is slightly larger in PSII than in the bRC. This could explain the difference in herbicide specificity between PSII and bRC (20). In addition, the proton pathway connecting the Q$_B$ site to the stromal space is considerably different in PSII compared with bRC. The PSII Q$_B$ site is much closer to the surface than in bRC because of the absence of an equivalent to the bRC H subunit. Because D1 His[252] is within hydrogen-bonding distance of D1 Ser[264], this residue could aid Q$_B$ protonation (Fig. 3).

The antenna system and carotenoids. We have identified 14 and 16 Chls bound to the transmembrane regions of CP43 and CP47, respectively (Fig. 4). Although most Chls are arranged in two layers on opposite sides of the membrane, two (Chl[9] of CP43 and Chl[20] of CP47) are positioned at almost an equal distance from both surfaces. As a result of this, they form a stack with adjacent Chls positioned

on the stromal and lumenal sides (Fig. 4A). The function of these stacked Chls is not clear, but they may facilitate the fast energy transfer known to occur within CP43 and CP47 and could be the origin of the long-wavelength–absorbing Chls of these proteins (21).

Many Chls in CP43 and CP47 are approximately related by an internal pseudo-twofold axis of the PSII monomeric complex and, within each protein, Chls are arranged around the internal pseudo-threefold axis. Ten Chls of CP43 and 13 Chls of CP47 are ligated by histidine, and one in CP43 has an asparagine ligand. Both proteins have two Chls in equivalent positions associated with either methionine or serine side chains. These side chains are not close enough to form direct ligands but could facilitate such an interaction by means of bound water molecules. The Mg^{2+} of two loosely attached Chls (Chl[35] and Chl[36]) are close to the main-chain carbonyl oxygen and cysteine side chain, respectively, but it is not clear whether these groups are ligands because of disorder.

The overall organization of Chls in PSII differs considerably from that in PSI. In PSI, the N-terminal domains of PsaA and PsaB are equivalent to CP43 and CP47, and the C-terminal domains of PsaA and PsaB are equivalent to D1 and D2 subunits (22). The distribution of the peripheral antenna Chls of PSI bound to the N-terminal domains of PsaA and PsaB is similar to that for CP43 and CP47. However, the C-terminal domains of PsaA and PsaB, together with other subunits, coordinate 43 Chls, which form a central antenna surrounding the electron transfer system of the PSI RC. This arrangement is notably different from PSII, which only coordinates two peripheral RC Chls (Chl$_{ZD1}$ and Chl$_{ZD2}$). Figure 4B clearly shows that the locations of these Chls are not
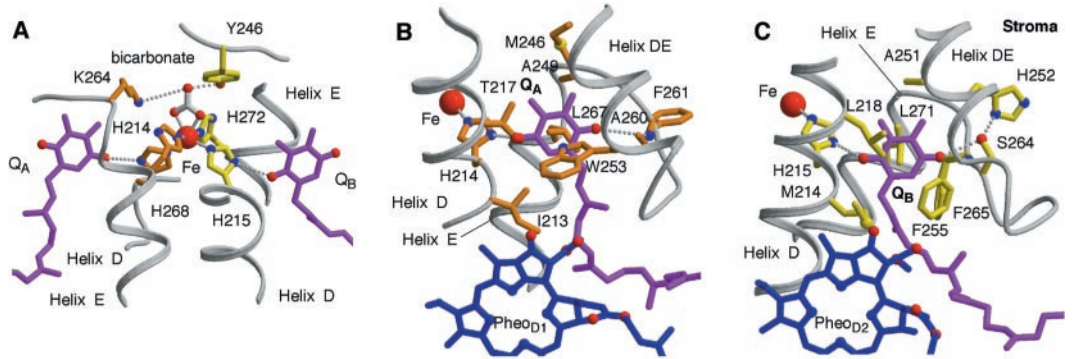
optimized to mediate energy transfer from CP43 and CP47 to the RC as is the case for the central antenna system of PSI. This difference probably explains the well-known slow trapping of excitation energy in PSII compared with PSI (21). However, Chl$_{ZD1}$ and Chl$_{ZD2}$ may have other functions, including protection of PSII against photoinduced damage (23).

Seven β-carotenes are assigned to the density, although it is possible that there are more as indicated by biochemical analyses (24); potential β-carotene densities are observed at the D1/CP43 and the dimer interfaces. β-carotene can be photooxidized when water splitting is inhibited with a msec time constant (25, 26), which suggests that at least one carotenoid must be positioned about 20 Å from P680. Moreover, there is spectroscopic evidence indicating that β-carotene facilitates long-distance electron flow from Cyt b559 to P680$^{•+}$ (27) and from Chl$_{ZD2}$ or Chl$_{ZD1}$ (28, 29). According to our assignment, the head group of one all trans–β-carotene is in direct contact with Chl$_{ZD2}$ and is located between Cyt b559 and the RC Chls (Figs. 2A and 4B). It is therefore likely that this β-carotene is involved in electron transfer from Cyt b559 and Chl$_{ZD2}$ to P680, a conclusion also made by Kamiya and Shen, who placed one all trans– and one cis–β-carotene in a similar position in their structure (5).

Of the remaining assigned carotenoids, four are positioned between Cyt b559 and CP43, possibly stabilized by the cluster of small transmembrane subunits in that region, and two are located on the CP47 side. The electron density map also indicates that CP47 may contain 2 to 3 β-carotenes close to the monomer/monomer interface in the dimer. Six of the seven assigned β-carotene mole-

Fig. 3. Electron acceptor side of PSII. (**A**) Overall view of the nonheme iron, Q_A and Q_B. Coloring scheme is as in Fig. 1, with protein main chains depicted in gray and with side-chain bonds and atoms following the coloring of the protein subunit as used in Fig. 1. The bicarbonate completing the coordination sphere of the nonheme Fe is shown with magenta bonds and is probably hydrogen bonded to D2 Lys[264] and D1 Tyr[246]. (**B**) The Q_A binding pocket. The hydrophobic residues forming this pocket are shown. The O_1 of the plastoquinone head group is likely to be hydrogen bonded to the nonheme Fe ligating D2 His[214] by its δ-nitrogen, whereas the O_4 atom may hydrogen bond to the backbone amide nitrogen of D2 Phe[261]. (**C**) The Q_B binding pocket. Q_B binds deep into a cavity lined with the hydrophobic residues. O_1 is likely to be hydrogen bonded to the δ-nitrogen of D1 His[215], which also forms a ligand to the nonheme Fe, whereas O_4 may form hydrogen bonds with the amide nitrogen of D1 Phe[265] and the side chain γ-oxygen of D1 Ser[264]. D1 Ser[264] appears to make further hydrogen-bonding contact with D1 His[252]. Probable hydrogen bonds are shown as dotted lines; solid lines represent ligands.

cules are in close contact with Chl head groups, as required for facilitating energy transfer from β-carotene to Chl and for quenching Chl triplets. Some of these carotenoids, and perhaps others not yet assigned, may serve to quench singlet oxygen known to be produced by P680 triplets formed by the recombination of P680$^{\bullet+}$ Pheo$^{\bullet-}$ (*30*).

The oxygen-evolving center. The density assigned to the OEC, consisting of a large globular domain connected to an extended region (Fig. 5A), is located close to the surface of the CD lumenal helix of the D1 subunit. The location and shape of the density are similar to that reported in previous x-ray studies and interpreted to have three Mn ions in the larger domain and one in the extended region (*4, 5*). However, in our electron density map, it is clear that four metal ions the size of Mn can be accommodated at the corners of a tetrahedron in the large globular density, whereas one metal ion is located in the center of the extended region. This arrangement is strongly supported by the difference density map (simulated-annealing omit map), which clearly shows the peak for the respective metal when each metal atom of the cluster is omitted from the model (fig. S5A). One of the five metal ions is likely to be Ca^{2+}, which is associated with the OEC (*2*). We have identified the metals using x-ray anomalous difference maps at the Mn absorption edge (1.89 Å) and at 2.25 Å wavelength, where Ca^{2+} has an anomalous difference (f″) 3.9 times as strong as Mn (Fig. 5A and fig. S5B). The Mn anomalous map only correlates with one metal in the small domain and three in the large globular domain, whereas the 2.25 Å wavelength map covers the remaining one metal ion in the large domain. On the basis of the approximate positions between the metal ions and their coordination properties, we suggest that the OEC is a cubane-like $Mn_3CaO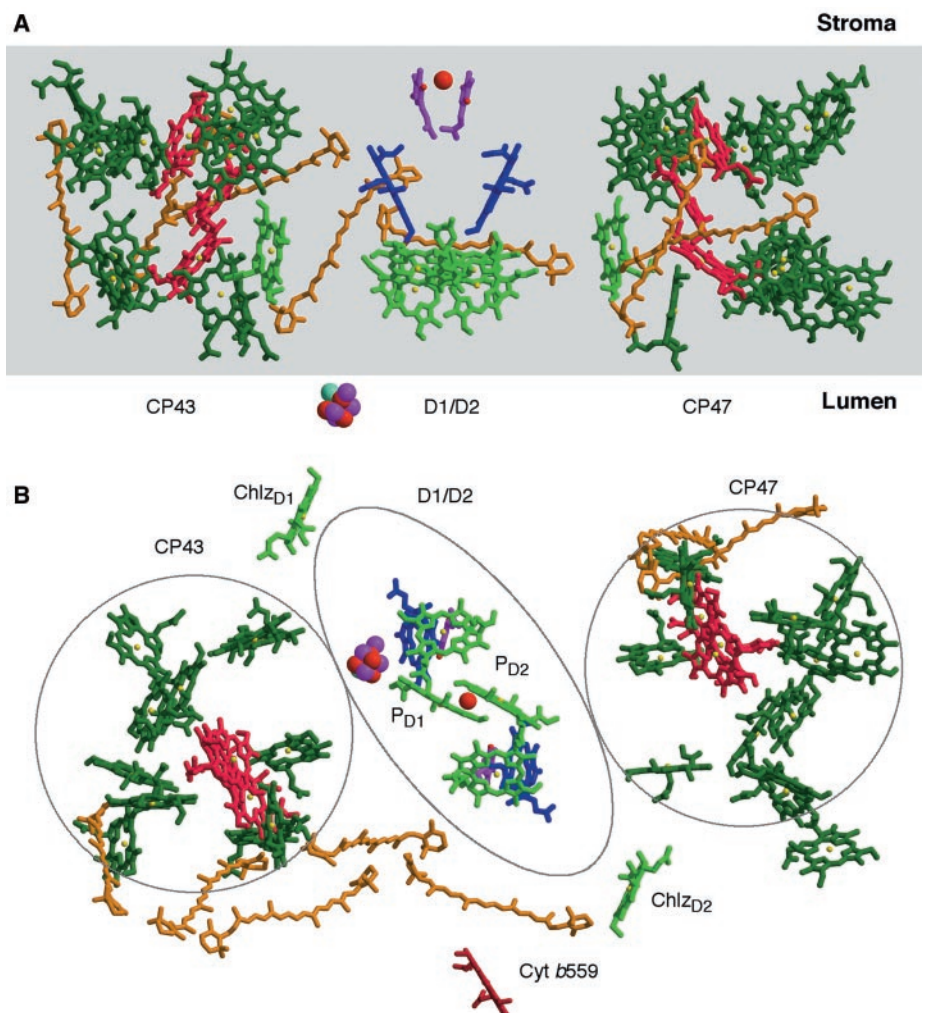_4$ cluster with each metal ion in this cluster having three-μ-oxo bridges (the large domain) connected to another Mn ion by a mono-μ-oxo bridge in the extended region (Fig. 5). In our model of the OEC, the metal-to-metal distances within the cubane-like cluster are ~2.7 Å for Mn-Mn

Fig. 4. Pigment organization in the PSII monomer complex. (**A**) Energy transfer in the PSII monomer complex. View is perpendicular to the membrane normal. Chlorophylls (green from RC and dark green from antenna) have the phytol tail omitted for clarity. Chlorophylls that are stacked with the planes of the porphyrin head group parallel and within van der Waals contact are highlighted in red. β-carotenes (orange) are shown as ball-and-stick models and tend to have one head group in direct van der Waals contact with a chlorophyll. (**B**) View of the pigments along the pseudo-twofold axis, perpendicular to the view in Fig. 2A.

and ~ 3.4 Å for Mn-Ca^{2+}, which are typical for the oxo bridges proposed and compatible with distances derived from extended x-ray absorption fine structure (EXAFS) studies (*31*). However, higher resolution data will be required to determine the precise distances and to investigate whether some of the bridging oxygen atoms are protonated. The distance between the Mn ion (Mn^4) in the small domain and the two Mn ions (Mn^2 and Mn^3) in the cubane-like cluster is ~ 3.3 Å, which is a typical distance for a mono-μ-oxo bridge between Mn ions and also is in line with EXAFS measurements (*31*).

Although the arrangement of Mn ions in our model fits one of the possible models derived from EXAFS (*31*), a cubane-like Mn$_3$CaO$_4$ cluster linked to a fourth Mn by a mono-μ-oxo bridge as proposed here has not been specifically suggested. However, a regular cubane structure for the Mn$_4$ cluster was proposed some time ago (*32*), and recently Dismukes and colleagues have given evidence for a distorted cubane core for higher S states (*33*), in line with an earlier paper (*34*). The 3 + 1 arrangement for the Mn ions has been proposed by Peloquin and Britt (*35*).

The electron density shows that the Mn$_3$CaO$_4$ cluster has four side chains as ligands: D1 Asp^{342} for Mn$_1$, D1 Glu^{189} and D1 His^{332} for Mn$_2$, and CP43 Glu^{354} for Mn$_3$ (fig. S5, C and D, and Fig. 5, A and B). Additionally, the carboxyl group of C-terminal D1 Ala^{344} is located close to Ca^{2+}, although in the current density map these are not connected (fig. S1C). This could, however, be a possible ligand of Ca^{2+} at some stage during the S-state cycle. D1 His^{337} seems to interact with the cluster by a hydrogen bond to the O2 oxygen of the cubane-like structure (Fig. 5B). Mn$_4$, located in the small domain, has two direct protein ligands, D1 Asp^{170} and D1 Glu^{333}, with a third residue D1 Asp^{61} possibly interacting through a water molecule. Furthermore, between this metal and Ca^{2+}, a nonprotein density indicative of an anionic ligand is observed (Fig. 5A and fig. S5D). According to the coordination geometry of Ca^{2+} and Mn and difference Fourier studies, we tentatively fit a bicarbonate molecule to this density, where it seems to be acting as a tridentate bridge between Mn^4 and Ca^{2+}; bicarbonate has been suggested to play a role in the assembly of the OEC (*36*). It is very likely that this is the water oxidation site and, in the active state, this nonprotein ligand is replaced by water molecules, X$_1$ for Mn^4, and X$_{21}$ and X$_{22}$ for Ca^{2+} (Fig. 5B), although one of the Ca^{2+} ligands could be a Cl$^-$, which is a cofactor for the OEC (*1*–*3*). The coordination number is usually six or seven for Ca^{2+} and six for Mn. It is possible to fulfill these coordination patterns assuming water molecules or hydroxide are associating with the Mn and Ca^{2+} in addition to the ligands mentioned above. Although water or hydroxide cannot be observed at the current resolution, the proposed OEC model can accommodate these molecules without conflicting with other atoms in the cluster. The hydroxide ions could play a role in maintaining the electroneutrality of the cluster.

The assignment of the metal ligands is consistent with a wide range of mutational studies (*37, 38*). The involvement of the C-terminal domain of the D1 protein in the assembly and stabilization of the OEC was first emphasized by Diner and colleagues (*39*), followed by several studies suggest-
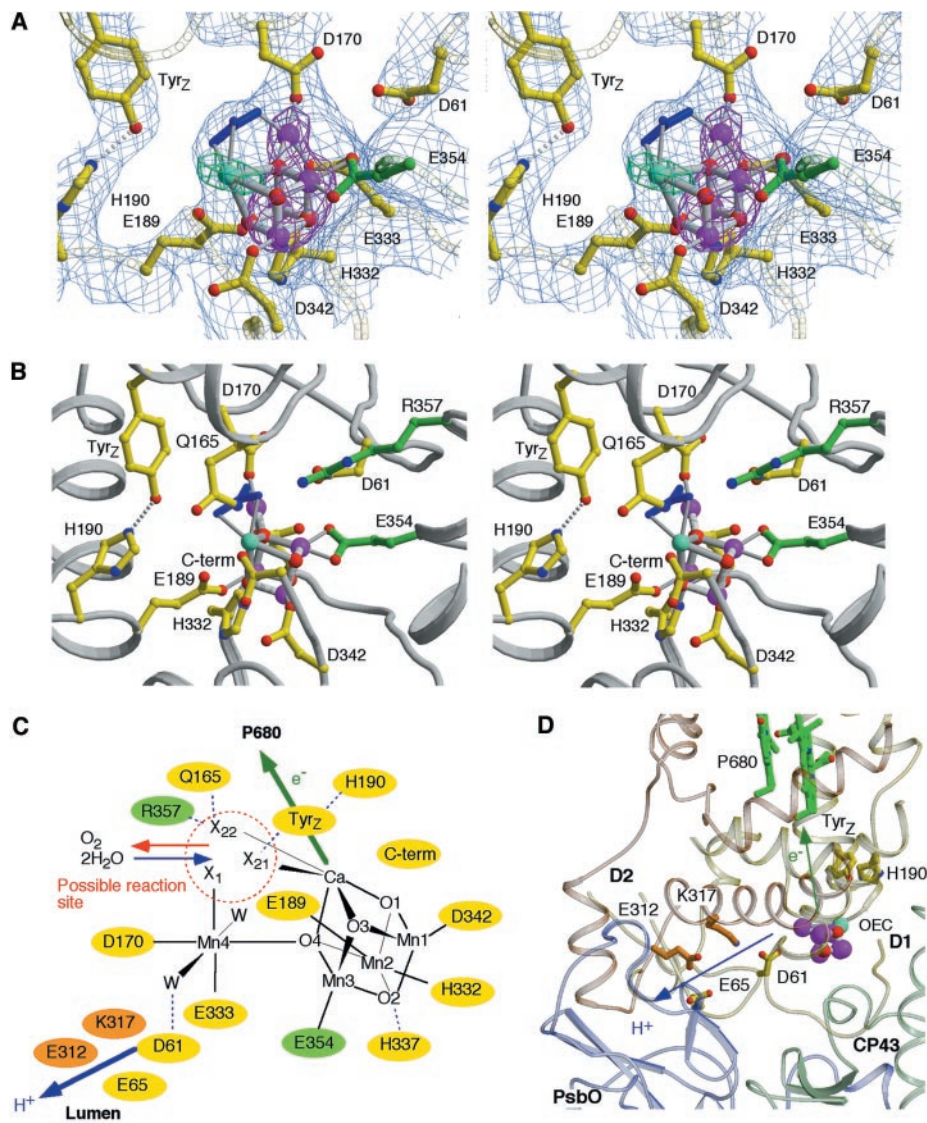

Fig. 5. The oxygen-evolving center (OEC). (**A**) Stereo view of the OEC with side-chain ligands and possible catalytically important side-chain residues. Mn ions, Ca^{2+}, and oxygen atoms are shown in magenta, cyan, and red, respectively. One unidentified nonprotein ligand to the OEC is colored in green. The protein main chain is depicted in light gray; the side-chain bonds and carbon atoms follow the coloring of the protein subunits (D1, yellow; CP43, green). σ_A weighted $2|F_o| - |F_c|$ density is shown as a light-blue wire mesh contoured at 1.5σ. Anomalous difference Fourier maps at 1.89340 Å (Mn edge, contoured at 10σ) and 2.25430 Å (highlights Ca^{2+}, contoured at 7σ) wavelengths are shown in magenta and blue-green, respectively. (**B**) The same as (A) but with a rotation around the y axis of 40° and without electron density and anomalous difference maps. (**C**) Schematic view of the OEC. Residues in D1, D2, and CP43 subunits are shown in yellow, orange, and green, respectively. X$_{11}$, X$_{21}$, and X$_{22}$ are possible substrate water binding positions to Mn^4 (X$_{11}$) and to Ca^{2+} (X$_{21}$ and X$_{22}$), identified from the position of the nonprotein ligand and coordination pattern of Mn and Ca^{2+} ions. Possible water molecules, which are not visible at the current resolution, are indicated as W. Possible hydrogen bonds are shown as light-blue dotted lines. (**D**) Hydrophilic pathway between the active site and lumen (blue arrow). The residue coloring is the same as in (B).

ing that D1 Asp[342] and the D1 C-terminal carboxy group of D1 Ala[344] could be possible metal ligands (*37*). All mutations of D1 His[332] prevent photoautotrophic growth and, again, this residue has been suggested to be a Mn ligand (*40*). Site-directed mutagenesis of D1 Glu[189], D1 His[332], D1 Glu[333], and D1 His[337] have all indicated that these residues are involved directly or indirectly in stabilizing the OEC (*40*). There have been many mutations of D1 Asp[170], indicating its involvement in the assembly of the Mn$_4$ cluster by providing a ligand to the high-affinity site that binds the first Mn ion during the assembly (*39*). To our knowledge, there has been no suggestion that CP43 Glu[354] could be a potential ligand for the Mn$_4$ cluster, but site-directed mutagenesis of the corresponding residue in the cyanobacterium *Synechocystis* significantly reduces PSII activity (*41*). This glutamate is a part of a 3_{10} helix contained within the motif Gly-Gly-Glu-Thr-Met-Arg-Phe-Trp-Asp, which is conserved in all known CP43 sequences. This motif is located in the large lumenal domain joining helices V and VI of CP43 (fig. S2) and seems to form a "lid" over the OEC. This finding emphasizes the important role played by the large extrinsic domain of this chlorophyll-binding protein in water oxidation as predicted in (*41*).

Tyr$_Z$ (D1 Tyr[161]) and Glu[189] (both on the X$_{21}$ side), and CP43 Arg[357] and D1 Gln[165] (both on the X$_{22}$ side) are closely associated with the nonprotein ligand binding site and thus possibly form a hydrogen bond to substrate water molecules during the reaction cycle. Tyr$_Z$ is oriented to form a hydrogen bond to D1 His[190] (Fig. 5A), in contrast to the conclusions of Fromme *et al.* (*42*). This hydrogen bond has been suggested to be important in stabilizing the tyrosine radical Tyr$_Z^{\bullet}$ generated by the reduction of P680$^{\bullet+}$ (*43*). The site is deep in the cavity holding the OEC and does not have an obvious connection to the lumenal space.

In the D2 protein, there is an equivalent to Tyr$_Z$, which is denoted Tyr$_D$ (D2 Tyr[160]). This tyrosine is also oxidized by P680$^{\bullet+}$ to generate a long-lived tyrosine radical (Tyr$_D^{\bullet}$), which is not involved directly in water oxidation but may help bias the electron transfer reactions to the D1 side of the RC (*44*). It is oriented in a similar way as Tyr$_Z$ and positioned such that it is likely to form a hydrogen bond with its neighboring histidine residue (D2 His[189]), which is assumed to stabilize Tyr$_D^{\bullet}$. The environment of Tyr$_D$ is very hydrophobic compared with that of Tyr$_Z$. In the D2 protein, the site equivalent to the OEC in the D1 protein is occupied by the bulky side chains of D2 Phe[169], D2 Phe[184], D2 Phe[185], D2 Phe[188], CP47 Phe[362], and CP47 Phe[363] (fig. S3). The two

CP47 residues are contained in the highly conserved motif Phe-Phe-Glu-Thr-Phe-Ser-Val-Leu-Val of the extrinsic domain linking helices V and VI.

Mechanism of water oxidation. A tightly bound Ca^{2+} located close to a Mn ion is a prerequisite for the mechanism of water oxidation suggested by Siegbahn and colleagues (*45*, *46*) and Brudvig and colleagues (*47*). They propose that during the S-state cycle only one of the Mn ions binds a water substrate molecule and, before dioxygen formation, produces a highly reactive electrophilic intermediate, either a Mn(IV) oxyl radical (*46*) or a Mn(V)oxo (*47*). The involvement of Mn(V) in water oxidation has also been advocated by Pecoraro *et al.* (*48*). Our structure for the OEC strongly suggests that Mn$_4$ is this reactive Mn ion. Indeed, the X$_1$ site of Mn$_4$ detected in our electron density map could be occupied by a water molecule or an oxygen intermediate during the water oxidation reaction.

According to the mechanisms cited above, the O=O bond formation is proposed to occur by a nucleophilic attack from a second-substrate water molecule ligated to Ca^{2+}, where this metal ion may also act in part as a weak Lewis acid, possibly aided by the Mn$_3$ cluster and by a Cl$^-$ ligand (*47*). In our structure, the X$_{21}$ or X$_{22}$ ligands of Ca^{2+}, which are close to the X$_1$ site, seem an ideal binding niche for the second-substrate water molecule. The nearby residues, D1 Gln[165] and CP43 Arg[357], may provide hydrogen bonds for stabilizing intermediates prior to O=O bond formation. Indeed, mutation of CP43 Arg[357] to a serine totally inhibited the evolution of oxygen and prevented photoautotrophic growth (*49*). Thus, the reaction schemes proposed by Siegbahn, Brudvig, Pecoraro, and colleagues are broadly consistent with our proposed structure of the OEC. However, their reaction mechanisms also incorporate the proposal of Hoganson and Babcock (*50*) that the oxidation of water involves proton-coupled electron transfer to the Tyr$_Z$ radical, with the hydrogen bond between Tyr$_Z$ and D1 His[190] being important not only for stabilizing Tyr$_Z^{\bullet}$ but also in providing the exit pathway for protons derived from water oxidation. Although Tyr$_Z$ is close to the water oxidation site and could also provide a hydrogen bond to reaction intermediates, our structure suggests that protons are more likely to exit by another route, as suggested by Junge and colleague (*51*). Water molecules and some residues, including Glu[189], could link the proposed water oxidation site with a proton channel formed by polar residues connecting D1 Asp[61] (which is closely associated with Mn$_4$) to the lumenal surface (Fig. 5, B and C) without the involvement of Tyr$_Z$ and

D1 His190, which are positioned on the other side of the OEC.

References and Notes
1. J. Barber, *Q. Rev. Biophys.* **36**, 71 (2003).
2. B. A. Diner, G. T. Babcock, in *Advances in Photosynthesis: The Light Reactions*. D. R. Ort, C. F. Yocum, Eds. (Kluwer Academic Publ., Dordrecht, The Netherlands, 1996), vol. 4, pp. 213–247.
3. B. A. Diner, F. Rappaport, *Annu. Rev. Plant Biol.* **53**, 551 (2002).
4. A. Zouni *et al.*, *Nature* **409**, 739 (2001).
5. N. Kamiya, J. R. Shen, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 98 (2003).
6. Materials and methods are available as supporting material on *Science* Online.
7. J. Deisenhofer *et al.*, *Nature* **318**, 618 (1985).
8. J. P. Allen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8487 (1988).
9. P. Jordan *et al.*, *Nature* **411**, 909 (2001).
10. B. Hankamer *et al.*, *J. Struct. Biol.* **135**, 262 (2001).
11. B. Kok, B. Forbush, M. McGloin, *Photochem. Photobiol.* **11**, 457 (1970).
12. F. Rappaport *et al.*, *Biochemistry* **41**, 8518 (2002).
13. L. M. C. Barter, J. R. Durrant, D. R. Klug, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 946 (2003).
14. J. R. Durrant *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4798 (1995).
15. J. P. Dekker, R. van Grondelle, *Photosyn. Res.* **63**, 195 (2000).
16. J. Barber, M. D. Archer, *Photochem. Photobiol. A-Chem.* 142, 97 (2001).
17. B. A. Diner *et al.*, *Biochemistry* **40**, 9265 (2001).
18. R. Hienerwadel, C. Berthomieu, *Biochemistry* **34**, 16288 (1995).
19. Govindjee, J. J. S. van Rensen, in *The Photosynthetic Reaction Center*, vol.1 (Academic Press, San Diego, CA, 1993), pp. 357–389.
20. W. Oettmeier, in *The Photosystems: Structure, Function, and Molecular Biology*, J. Barber, Ed. (Elsevier, Amsterdam, The Netherlands, 1992), vol. 11, pp. 349–408.
21. F. L. de Weerd *et al.*, *Biophys. J.* **82**, 1586 (2002).
22. W. D. Schubert *et al.*, *J. Mol. Biol.* **280**, 297 (1998).
23. D. H. Stewart, G. W. Brudvig, *Biochim. Biophys. Acta* **1367**, 63 (1998).
24. C. A. Tracewell *et al.*, *Arch. Biochem. Biophys.* **385**, 61 (2001).
25. A. Telfer, J. De Las Rivas, J. Barber, *Biochim. Biophys. Acta* **1060**, 106 (1991).
26. J. Hanley *et al.*, *Biochemistry* **38**, 8189 (1999).
27. P. Faller, A. Pascal, A. W. Rutherford, *Biochemistry* **40**, 6431 (2001).
28. C. A. Tracewell *et al.*, *Biochemistry* **40**, 193 (2001).
29. C. A. Tracewell, G. W. Brudvig, *Biochemistry* **42**, 9127 (2003).
30. A. Telfer *et al.*, *Biochemistry* **33**, 14469 (1994).
31. J. H. Robblee, R. M. Cince, V. K. Yachandra, *Biochim. Biophys. Acta* **1503**, 7 (2001).
32. J. C. de Paula *et al.*, *J. Am. Chem. Soc.* **108**, 4002 (1986).
33. T. G. Carrell, A. Tyryshkin, G. C. Dismukes, *J. Biol. Inorg. Chem.* **7**, 2 (2002).
34. J. B. Vincent, G. Christou, *Inorg. Chim. Acta* **136**, L41 (1987).
35. J. M. Peloquin, R. D. Britt, *Biochim. Biophys. Acta* **1503**, 96 (2001).
36. S. V. Baranov *et al.*, *Biochemistry* **39**, 6060 (2000).
37. R. J. Debus, *Biochim. Biophys. Acta* **1503**, 164 (2001).
38. B. A. Diner, *Biochim. Biophys. Acta* **1503**, 147 (2001).
39. B. A. Diner, P. J. Nixon, J. W. Farchaus, *Curr. Opin. Struct. Biol.* **1**, 546 (1991).
40. H.-A. Chu, P. Nguyen, R. J. Debus, *Biochemistry* **34**, 5859 (1995).
41. C. Rosenberg *et al.*, *Biochemistry* **38**, 15994 (1999).
42. P. Fromme *et al.*, *Philos. Trans. R. Soc. London Ser. B* **357**, 1337 (2002).
43. C. Tommos, G. T. Babcock, *Biochim. Biophys. Acta* **1458**, 199 (2000).
44. P. Faller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14368 (2001).
45. P. E. Siegbahn, R. H. Crabtree, *J. Am. Chem. Soc.* **121**, 117 (1999).
46. P. E. Siegbahn, *Curr. Opin. Chem. Biol.* **6**, 227 (2002).

47. J. S. Vrettos, J. Limburg, G. W. Brudvig, *Biochim. Biophys. Acta* **1503**, 229 (2001).
48. V. L. Pecoraro et al., *Pure Appl. Chem.* **70**, 925 (1998).
49. N. Knoepfle, T. M. Bricker, C. Putnam-Evans, *Biochemistry* **38**, 1582 (1999).
50. C. W. Hoganson, G. T. Babcock, *Science* **277**, 1953 (1997).
51. M. Haumann, W. Junge, *Biochim. Biophys. Acta* **1411**, 86 (1999).
52. The authors dedicate this paper to G. T. Babcock and M. P. Klein. J.B. and S.I. acknowledge support from the Biotechnology and Biological Research Science Council. We thank the Centre for Structural Biology and Bioinformatics Facility at Imperial College London for

The Structure and Receptor Binding Properties of the 1918 Influenza Hemagglutinin

S. J. Gamblin,[1]* L. F. Haire,[1]* R. J. Russell,[1]* D. J. Stevens,[1]
B. Xiao,[1] Y. Ha,[2]† N. Vasisht,[1] D. A. Steinhauer,[1]‡ R. S. Daniels,[1]
A. Elliot,[1] D. C. Wiley,[2] J. J. Skehel[1]§

The 1918 influenza pandemic resulted in about 20 million deaths. This enormous impact, coupled with renewed interest in emerging infections, makes characterization of the virus involved a priority. Receptor binding, the initial event in virus infection, is a major determinant of virus transmissibility that, for influenza viruses, is mediated by the hemagglutinin (HA) membrane glycoprotein. We have determined the crystal structures of the HA from the 1918 virus and two closely related HAs in complex with receptor analogs. They explain how the 1918 HA, while retaining receptor binding site amino acids characteristic of an avian precursor HA, is able to bind human receptors and how, as a consequence, the virus was able to spread in the human population.

The HAs of influenza viruses mediate receptor binding and membrane fusion, the first stages of virus infection (*1*). The receptors that they recognize are sialic acids of cell-surface glycoproteins and glycolipids, and the nature of the interactions involved in determining binding specificity has been described in biochemical, genetic, and structural studies (*2–8*). Sialic acids are usually found in either α2,3 or α2,6 linkages to galactose, the predominant penultimate sugar of N-linked carbohydrate side chains. The binding preference of a given HA for one or other of these linkage types correlates with the species specificity for infection. Thus, the HAs of all 15 antigenic subtypes found in avian influenza viruses bind preferentially to sialic acid in α2,3 linkage (*9*), and it is this form

of the sialosaccharide that predominates in avian enteric tracts where these viruses replicate (*10*). Swine influenza viruses are reported to bind sialic acid in α2,6, and sometimes also α2,3, linkages (*11*), and sialic acid in both linkages is detected in porcine tracheae (*10*). Human viruses of the H1, H2, and H3 subtypes that are known to have caused pandemics in 1918, 1957, and 1968, respectively, recognize α2,6-linked sialic acid (*5*), the major form found on cells of the human respiratory tract (*12, 13*).

Because an avian origin is proposed for the HAs of swine and human viruses (*14*), changes in binding specificity are required for cross-species transfer. The mechanism that human viruses have used to achieve these changes appears to be different for different subtypes. For the HAs of the H2 and H3 human viruses, a minimum of two changes in receptor binding site amino acids, Gln[226] to Leu[226] and Gly[228] to Ser[228], correlates with the shift from avian to human receptor binding (*15, 16*). By contrast, HAs of human H1 viruses acquire the ability to bind to human receptors while retaining Gln[226] and Gly[228] (*11*). To understand how they do this, we determined the structures of HAs from the 1918 pandemic virus (1918-human) with the use of HA expressed from DNA of the sequence recovered from tissues infected with virus in 1918 (*17*) and from the prototype human (1934-human) and swine (1930-

swine) H1 influenza viruses, A/Puerto Rico/8/34 and A/swine/Iowa/30, respectively (Fig. 1). A/Puerto Rico/8/34 was one of the first human influenza viruses isolated in the Americas (*18*) and has been widely used in laboratory investigations of influenza. A/swine/Iowa/30 was the first influenza virus isolated from mammals in 1930 (*19*), 3 years before the first isolate was recovered from humans (*20*).

Overall structure. The structures were solved by molecular replacement, and crystallographic statistics are given in Table 1 and table S1. The overall trimeric structures of the three H1 HAs are similar (Fig. 2), but they show notable differences to HAs of other subtypes with respect to the arrangements of the receptor binding, vestigial esterase, and membrane fusion subdomains, both within the HA trimer and also within individual monomers (*21*) (table S2). As predicted from their placement in the same phylogenetic and structure-based clade, the H1 HAs are most similar to those of the H5 subtype (*21*). We have examined the structures in detail in relation to their receptor binding and membrane fusion activities and to the antigenic variation that occurred in both periods of human H1 virus prevalence, 1918 to 1957 and 1977 to date (Fig. 2), and we will present a detailed description of this analysis elsewhere (*22*). We have, however, concluded that the receptor binding properties are the most distinctive features of the 1918 virus HA, and we focus on these here.

The receptor binding subdomain. The receptor binding sites are located at the membrane-distal tip of each subunit of the HA trimer (Fig. 2). Three secondary structure elements—the 190 helix (residues 190 to 198), the 130 loop (residues 135 to 138), and the 220 loop (residues 221 to 228)—form the sides of each site, with the base made up of the conserved residues Tyr[98], Trp[153], His[183], and Tyr[195] (*1*) (Fig. 2B). The conformations adopted by the 130 and 220 loops of the three H1 HAs are similar, but they are significantly different from those of the equivalent loops in the HAs of other influenza subtypes (*2, 3*) (Figs. 2B, 3, and 4 and table S2). To understand the structural basis of the receptor specificity of H1 HAs, we determined the structures of the 1934-human and the 1930-swine HAs in complex, with α2,3- and α2,6-linked sialopentasaccharides as analogs of avian and human receptors,

[1]Medical Research Council (MRC) National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK. [2]Department of Molecular and Cellular Biology, Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA.

*These authors contributed equally to this work.
†Present address: Department of Pharmacology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA.
‡Present address: Department of Microbiology and Immunology, Emory University School of Medicine, 1510 Clifton Road, Atlanta, GA 30322, USA.
§To whom correspondence should be addressed. E-mail: mbrenna@nimr.mrc.ac.uk