Architecture of the Photosynthetic Oxygen-Evolving Center
Kristina N. Ferreira, et al.
Science 303, 1831 (2004);
DOI: 10.1126/science.1093087

The following resources related to this article are available online at www.sciencemag.org (this information is current as of October 18, 2007):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/cgi/content/full/303/5665/1831

Supporting Online Material can be found at:
http://www.sciencemag.org/cgi/content/full/1093087/DC1

A list of selected additional articles on the Science Web sites related to this article can be found at:
http://www.sciencemag.org/cgi/content/full/303/5665/1831#related-content

This article cites 48 articles, 6 of which can be accessed for free:
http://www.sciencemag.org/cgi/content/full/303/5665/1831#otherarticles

This article has been cited by 582 article(s) on the ISI Web of Science.

This article has been cited by 89 articles hosted by HighWire Press; see:
http://www.sciencemag.org/cgi/content/full/303/5665/1831#otherarticles

This article appears in the following subject collections:
Biochemistry
http://www.sciencemag.org/cgi/collection/biochem

Information about obtaining reprints of this article or about obtaining permission to reproduce this article in whole or in part can be found at:
http://www.sciencemag.org/about/permissions.dtl
Architecture of the Photosynthetic Oxygen-Evolving Center

Kristina N. Ferreira,1* Tina M. Iverson,2* Karim Maghlaoui,1 James Barber,1† So Iwata1,2,3†

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 Å 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 Mn3CaO4 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 oxygen to the fixation of CO2. 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 oxygen 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 in 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 Mn3CaO4 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, Cα 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 isomorphous 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 monomers 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 (Chl13 and Chl16) 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 Cα backbone between the dimer structure and 1FE1 (4) and 1IZL (5) are 2.1 Å (for 1359 Cα atoms) and 1.9 Å (for 3519 Cα 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 luminal 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 Cα 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 luminal 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 PsA A and PsB 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 luminal 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 luminal 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 b559), agrees with earlier studies (4, 5, 10). The remaining 10 transmembrane helices 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 b559, and designated PsbX (4), was not present in our electron density map. With the exception of Cyt b559, 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 Psbl and PsbX proteins stabilize the peripheral chlorophylls of the D1 and D2 proteins (ChlZD1 and ChlZD2). 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 luminal 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 luminal 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 luminal surface (P681 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-chloromercuribenzenesulfonate.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Native</th>
<th>Native (Mn-edge)</th>
<th>Native (for Ca detection)</th>
<th>CdCl2</th>
<th>K[Au(Cn)3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam line</td>
<td>SLS X065A</td>
<td>SLS X065A</td>
<td>SLS X065A</td>
<td>ESRF ID29</td>
<td>ESRF ID29</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.0076</td>
<td>1.89340</td>
<td>2.25430</td>
<td>40.0–3.8</td>
<td>40.0–3.8</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>40.0–3.5</td>
<td>40.0–3.8</td>
<td>40.0–3.8</td>
<td>40.0–3.8</td>
<td>40.0–3.8</td>
</tr>
<tr>
<td>Total observation</td>
<td>298,731</td>
<td>212,461</td>
<td>245,580</td>
<td>221,188</td>
<td></td>
</tr>
<tr>
<td>Unique reflections</td>
<td>103,604</td>
<td>79,041</td>
<td>78,778</td>
<td>73,595</td>
<td></td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>87.3 (80.9)</td>
<td>83.6 (62.3)</td>
<td>82.0 (55.1)</td>
<td>77.8 (61.7)</td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>2.88</td>
<td>2.68</td>
<td>3.11</td>
<td>3.00</td>
<td></td>
</tr>
<tr>
<td>Rsym*</td>
<td>0.08 (0.43)</td>
<td>0.11 (0.50)</td>
<td>0.08 (0.43)</td>
<td>0.09 (0.43)</td>
<td></td>
</tr>
<tr>
<td>l/σ(h)*</td>
<td>10.4 (2.0)</td>
<td>7.4 (2.1)</td>
<td>9.1 (2.6)</td>
<td>9.0 (2.0)</td>
<td></td>
</tr>
<tr>
<td>Phasing power§</td>
<td>N/A</td>
<td>N/A</td>
<td>0.65</td>
<td>0.53</td>
<td></td>
</tr>
</tbody>
</table>

**Data collection and phasing**

<table>
<thead>
<tr>
<th>Data set</th>
<th>EMP1</th>
<th>EMP2</th>
<th>PCMB</th>
<th>TabR6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam line</td>
<td>SLS X065A</td>
<td>SLS X065A</td>
<td>SLS X065A</td>
<td>SLS X065A</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.96863</td>
<td>0.96863</td>
<td>0.96863</td>
<td>0.96863</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>40.0–4.1</td>
<td>40.0–3.8</td>
<td>40.0–3.8</td>
<td>40.0–3.8</td>
</tr>
<tr>
<td>Total observation</td>
<td>215,331</td>
<td>179,833</td>
<td>172,328</td>
<td></td>
</tr>
<tr>
<td>Unique reflections</td>
<td>80,342</td>
<td>75,502</td>
<td>71,395</td>
<td></td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>73.0 (58.5)</td>
<td>73.5 (31.9)</td>
<td>74.6 (53.8)</td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>2.68</td>
<td>2.38</td>
<td>2.41</td>
<td></td>
</tr>
<tr>
<td>Rsym*</td>
<td>0.10 (0.40)</td>
<td>0.08 (0.44)</td>
<td>0.08 (0.33)</td>
<td></td>
</tr>
<tr>
<td>l/σ(h)*</td>
<td>6.1 (1.3)</td>
<td>5.9 (2.1)</td>
<td>11.2 (2.6)</td>
<td></td>
</tr>
<tr>
<td>Phasing power§</td>
<td>0.90</td>
<td>0.57</td>
<td>0.62</td>
<td></td>
</tr>
</tbody>
</table>

**Refinement**

| Resolution (Å)* | 20.0–3.5 (3.56–3.50) |
| Number of reflections | 103,485 (85.7%) |
| R factor|| 0.303 (0.340) |
| Rfree* | 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. ‡Rsym = \[\frac{\sum |I(h)| - \langle |I(h)| \rangle}{\sum |I(h)|\], where \(\langle |I(h)| \rangle\) is the i th measurement. §Phasing power is the RMS value of \(F_i\) divided by the RMS lack-of-closure error. †‡R factor = \[\frac{\sum |I(h)| - \langle |I(h)| \rangle}{\sum |I(h)|\]. †Rfree was calculated for 1% of reflections randomly excluded from the refinement. 

---

**Research Articles**

19 MARCH 2004 VOL 303 SCIENCE www.sciencemag.org
protonation, QB is released from PSII into the membrane matrix. The cationic radical P680⁺ is reduced by a redox-active tyrosine, known as TyrZ (Tyr^161 of D1 subunit), to generate a neutral tyrosine radical TyrZ⁻, 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 QA, all the redox-active cofactors involved in the electron transfer processes are located on the D1 side of the RC. The radical cation P680⁺ 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 bacteriochlorophylls (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 PD1 and PD2 (4). Although the overall organization of PD1 and PD2 is very similar to the special pair, the relative orientation angle between their tetrapyrole heads differs slightly from their bacterial equivalents. The head groups of PD1 and PD2 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). Pd1 and Pd2 are located close to ChlD1 and ChlD2 (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 ChlB, which is the chlorophyll closest to the active PheoD1, 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 ChlD1 and ChlD2; the closest residue to ChlD1 is D1 Thr^179 and to ChlD2 is D2 Ile^178.

The electrons are transferred from PheoD1 to QA, which is a firmly bound plastoquinone. Even though PSII and bRCs use different primary quinone acceptors, the QA binding pockets are structurally similar. In PSII, the QA is hydrogen bonded to the main-chain amide group of D2 Phe^261 and D2 His^214, whereas in 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^177, D2 Met^146, D2 Ala^259, D2 Trp^253, D2 Ala^260, and D2 Leu^267.

The nonheme Fe, which mediates electron transfer from QA to QB, 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 QA to QB as well as facilitating the protonation of Qb (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 orange, hemes are in red, nonheme Fe is red, QA and Qb are purple. The oxygen-evolving center (OEC) is shown as the red (oxygen atoms), magenta (Mn ions), and cyan (Ca²⁺) balls. (B) View of the PSII monomer along the membrane normal from the lumenal side. A part of the other monomer in the dimer 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, PsbL and PsbX, and CP43 and CP47 as emphasized by the black lines encircling these subunits. Coloring is the same as in (A).
modulate this anion. Close to this nonprotein density are D1 Tyr256 and D2 Lys254, positioned and oriented such that they could stabilize the bicarbonate by hydrogen bonding (Fig. 3A).

There is less conservation between the Qb sites of PSI and bRC than for the QA sites. There are similarities in that the PSI QA is hydrogen bonded to D1 Ser262 and D1 His275, which is also a ligand for the nonheme Fe, and possibly to the main-chain amide group of D1 Phe265. However, the Qb sites in PSI and the bRC show a 3.5 Å displacement when superimposed using Ca atoms of D1/D2 and the L/M subunits. This is because there is a conformational difference in the loops containing D1 Ser264 and the equivalent in bRC, which is caused by the insertion of one residue. As a consequence, the volume of the Qb binding pocket, composed of the residues including D1 Met214, D1 Leu218, D1 Ala251, D1 Phe255, and D1 Leu272, is slightly larger in PSI than in the bRC. This could explain the difference in herbicide specificity between PSI and bRC (20).

In addition, the proton pathway connecting the Qb site to the stromal space is considerably different in PSI compared with bRC. The PSI Qb site is much closer to the surface than in bRC because of the absence of an equivalent to the bRC H subunit. Because D1 His252 is within hydrogen-bonding distance of D1 Ser264, this residue could aid Qb 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 (Chp of CP43 and Chlβ 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 luminal 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 PSI 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 Mg2+ of two loosely attached Chls (Chlβ15 and Chlδ19) are close to the main-chain carboxyl 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 PSI differs considerably from that in PSI. In PSI, the N-terminal domains of PsA and PsB are equivalent to CP43 and CP47, and the C-terminal domains of PsA and PsB are equivalent to D1 and D2 subunits (22). The distribution of the peripheral antenna Chls of PSI bound to the N-terminal domains of PsA and PsB is similar to that for CP43 and CP47. However, the C-terminal domains of PsA and PsB, 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 PSI, which only coordinates two peripheral RC Chls (Chlδ21 and Chlβ22). 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).

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 mesc 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 Chlβ22 or Chlδ19 (28, 29). According to our assignment, the head group of one all trans-β-carotene is in direct contact with Chlβ22 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 Chlβ22 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, Qₐ, and Qₐ. Coloring scheme is as in Fig. 1, with protein main chains depicted in gray and with side-chain bonds and carbon 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 Lys²⁶⁴ and D1 Tyr²⁴⁶. (B) The QA binding pocket. The hydrophobic residues forming this pocket are shown. The Qₐ of the plastoquinone head group is likely to be hydrogen bonded to the nonheme Fe ligating D2 His²¹⁴ by its δ-nitrogen, whereas the Qₐ atom may hydrogen bond to the backbone amide nitrogen of D2 Phe²⁶¹. (C) The Qₐ binding pocket. Qₐ binds deep into a cavity lined with the hydrophobic residues. O₁ is likely to be hydrogen bonded to the δ-nitrogen of D1 His²¹⁵, which also forms a ligand to the nonheme Fe, whereas O₄ may form hydrogen bonds with the amide nitrogen of D1 Phe²⁶⁵ and the side chain γ-oxygen of D1 Ser²⁶⁴. D1 Ser²⁶⁴ appears to make further hydrogen-bonding contact with D1 His²⁵². Probable hydrogen bonds are shown as dotted lines; solid lines represent ligands.

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 Ca²⁺, which is associated with the OEC (2). We have identified the metal ions using x-ray anomalous difference maps at the Mn absorption edge (1.89 Å) and at 2.25 Å wavelength, where Ca²⁺ 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 Mn₃CaO₄ 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...
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-\(\mu\)-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\(_4\)CaO\(_4\) cluster linked to a fourth Mn by a mono-\(\mu\)-oxo bridge as proposed here has not been specifically suggested. However, a regular cubane structure for the Mn\(_3\) 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\(_4\)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\(_3\), located in the small domain, has two direct protein ligands, D1 Asp\(^{170}\) and D1 Glu\(^{333}\), with a third residue D1 Asp\(^{641}\) 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. SSD). 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\(^n\) 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\(^n\), 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-
ing that D1 Asp142 and the D1 C-terminal carboxy group of D1 Ala344 could be possible metal ligands (37). All mutations of D1 His337 prevent photoautotrophic growth and, again, this residue has been suggested to be a Mn ligand (40). Site-directed mutagenesis of D1 Glu189, D1 His332, D1 Glu333, and D1 His337 have all indicated that these residues are involved directly or indirectly in stabilizing the OEC (40). There have been many mutations of D1 Asp70, indicating its involvement in the assembly of the Mn4 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 Glu514 could be a potential ligand for the Mn4 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 310 helix contained within the motif Gly-Gly-Gly-Thr-Met-Arg-Phg-Trp-Asp, which is conserved in all known CP43 sequences. This motif is located in the large luminal 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).

TyRZ(D1 TyR61) and Glu189 (both on the X21 side), and CP43 Arg357 and D1 Gln165 (both on the X22 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. TyRZ is oriented to form a hydrogen bond to D1 His190 (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 TyRZ generated by the reduction of P680+ (43). The site is deep in the cavity holding the OEC and does not have an obvious connection to the luminal space.

In the D2 protein, there is an equivalent to TyRZ, which is denoted TyRD (D2 TyR160). This tyrosine is also oxidized by P680+ to generate a long-lived tyrosine radical (TyRD), 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 TyRZ and positioned such that it is likely to form a hydrogen bond with its neighboring histidine residue (D2 His189), which is assumed to stabilize TyRD. The environment of TyRD is very hydrophobic compared with that of TyRZ. In the D2 protein, the site equivalent to the OEC in the D1 protein is occupied by the bulky side chains of D2 Phe169, D2 Phe184, D2 Phe185, D2 Phe188, CP47 Phe362, and CP47 Phe463 (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 Ca2+ 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 Mn4 is this reactive Mn ion. Indeed, the X1 site of Mn4 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 Ca2+, where this metal ion may also act in part as a weak Lewis acid, possibly aided by the Mn4 cluster and by a C1 engages in a hydrogen bond to the second-substrate water molecule. The nearby residues, D1 Gln165 and CP43 Arg357, may provide hydrogen bonds for stabilizing intermediates prior to O=O bond formation. Indeed, mutation of CP43 Arg357 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 TyRZ radical, with the hydrogen bond between TyRZ and D1 His319 being important not only for stabilizing TyRZ but also in providing the exit pathway for protons derived from water oxidation. Although TyRZ 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 colleagues (51). Water molecules and some residues, including Glu189, could link the proposed water oxidation site with a proton channel formed by polar residues connecting D1 Asp361 (which is closely associated with Mn4) to the luminal surface (Fig. 5, B and C) without the involvement of TyRZ and D1 His190, which are positioned on the other side of the OEC.
The Structure and Receptor Binding Properties of the 1918 Influenza Hemagglutinin


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, Gln226 to Leu226 and Gly222 to Ser222, 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 Gln226 and Gly222 (17). 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 (22). 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 Tyr298, Trp353, His183, and Tyr195 (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, 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 sialopentaaccharides as analogs of avian and human receptors,

References and Notes

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. 3 These authors contributed equally to this work. 4 Present address: Department of Pharmacology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA. 5 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.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1093087/DC1

Materials and Methods

Figs. S1 to S5

References and Notes

29 October 2003; accepted 22 January 2004

Published online 5 February 2004: 10.1126/science.1093087

Include this information when citing this paper.

The experimental multi-wavelength anomalous dispersion (MAD) phases have been deposited in the Protein Data Bank (entry 155L) and will be available upon publication.