Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution

Zhenfeng Liu[1], Hanchi Yan[1], Kebin Wang[2], Tingyun Kuang[2], Jiping Zhang[1], Lulu Gui[1], Xiaomin An[1] & Wenrui Chang[1]

[1]*National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing 100101, People's Republic of China*
[2]*Laboratory of Photosynthesis and Environmental Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, 20 Nanxincun, Xiangshan, Beijing 100093, People's Republic of China*

....

The major light-harvesting complex of photosystem II (LHC-II) serves as the principal solar energy collector in the photosynthesis of green plants and presumably also functions in photoprotection under high-light conditions. Here we report the first X-ray structure of LHC-II in icosahedral proteoliposome assembly at atomic detail. One asymmetric unit of a large *R*32 unit cell contains ten LHC-II monomers. The 14 chlorophylls (Chl) in each monomer can be unambiguously distinguished as eight Chl*a* and six Chl*b* molecules. Assignment of the orientation of the transition dipole moment of each chlorophyll has been achieved. All Chl*b* are located around the interface between adjacent monomers, and together with Chl*a* they are the basis for efficient light harvesting. Four carotenoid-binding sites per monomer have been observed. The xanthophyll-cycle carotenoid at the monomer–monomer interface may be involved in the non-radiative dissipation of excessive energy, one of the photoprotective strategies that have evolved in plants.

Light harvesting is the primary process in photosynthesis. In green plants, the function of harvesting solar energy is fulfilled by a series of light-harvesting complexes in the thylakoid membrane of chloroplasts. LHC-II, the most abundant integral membrane protein in chloroplasts, exists as a trimer and binds half of the thylakoid chlorophyll molecules. Every monomeric LHC-II comprises a polypeptide of about 232 amino-acid residues, 13–15 Chl*a* and Chl*b* molecules[1], 3–4 carotenoids[2] and one tightly bound phospholipid[3]. Besides the light-harvesting function, LHC-II has also been shown to function in the non-radiative dissipation of excess excitation energy formed under high-light conditions[4,5]. It has a crucial role in minimizing the damaging effects of excess light by operating this photoprotective mechanism as light intensity becomes increasingly saturating. Moreover, LHC-II also takes part in regulating the distribution of excitation energy to photosystems II and I (ref. 6).

The structure of LHC-II from pea has been determined by electron crystallography at 3.4 Å resolution parallel to the membrane plane, and at about 4.9 Å resolution perpendicular to this plane[7]. This model revealed some basic structural features of LHC-II, including three transmembrane α-helices (helices A, B and C) and a short amphipathic helix (helix D), 12 chlorophyll tetrapyrroles with roughly determined locations and orientations, and two carotenoids. A more detailed structural picture of LHC-II, with an unambiguous determination of the identity of the chlorophylls (Chl*a* or Chl*b*) and the orientation of their transition dipole moments, would be beneficial for a better understanding of the basic functional mechanism of LHC-II. We have obtained this information by solving the structure of LHC-II at higher resolution using X-ray crystallography.

In our model of the X-ray structure of LHC-II at 2.72 Å resolution, we provide the basis for investigating quantitatively the underlying mechanism of the light-harvesting process and its adjustment in LHC-II. We also reveal for the first time an elegant arrangement of membrane proteins in the icosahedral proteoliposome assembly, and show that membrane proteins can be crystallized in a way that differs from those described in ref. 8.

Structure determination

Crystallization of trimeric LHC-II isolated from *Spinacia oleracea* is described briefly in the Methods. Structure determination of

LHC-II organized in an icosahedral particle includes initial phasing by the single isomorphous replacement (SIR) method plus phase refining and extending by the real-space averaging method[9]. Data collection, phasing and refinement statistics are listed in Table 1.

The high-quality electron density map enabled us to trace 94% of

Table 1 **Data collection, phasing and refinement statistics**

Data set	Native (1)	Native (2)	K_2HgI_4 (derivative)
Resolution (Å)	50–3.5	25–2.7	50–3.5
R_{merge}*	0.087 (0.112)	0.082 (0.368)	0.100 (0.131)
Completeness (%)	97.9 (95.5)	90.8 (79.6)	80.5 (83.2)
$<I/\sigma>$	6.1 (2.1)	14.4 (2.5)	6.4 (3.3)
SIR phasing statistics			
No. of heavy atom sites	10		
R_{cullis} (centric/acentric)†	0.66/0.69		
Phasing power (centric/acentric)†	1.13/1.67		
Resolution (Å)	15–5.0		
FOM†	0.36		
Phase refinement and extension statistics			
Resolution (Å)	10–2.72		
FOM†	0.922 (0.801)		
Correlation coefficient‡	0.967 (0.765)		
R factor‡	0.128 (0.468)		
Structure refinement statistics			
Resolution	10–2.72		
Reflections (working set)	179,170		
Reflections (test set)	9,326		
R_{work}/R_{free} (%)	19.4/22.1		
r.m.s.d. bond length (Å)	0.0124		
r.m.s.d. bond angle (°)	2.257		
Coordinate error (Å)§			
Luzzati§	0.28		
SigmaA§	0.28		
Number of non-hydrogen atoms			
Protein	16,619		
Cofactors	11,720		
Water	699		

Numbers in parentheses correspond to values in the highest resolution shell.
*$R_{merge} = \Sigma_j \Sigma_h |I_{j,h} - <I_h>|/\Sigma_j \Sigma_h <I_h>$, where *h* are unique reflection indices, $I_{j,h}$ are intensities of symmetry-related reflections and $<I_h>$ is the mean intensity. Reflections with $I > 2.0 \times \sigma_I$ in native (1) and derivative data sets were used in the R_{merge} calculation, whereas the $-3.0 \times \sigma_I$ cutoff was applied in the native (2) data set.
†FOM (figure of merit), R_{cullis} (centric) and phasing power determined by programs from the CCP4 suite (see Methods).
‡Correlation coefficient= $\Sigma_h (<F_o> - |F_o|_h)(<F_c> - |F_c|_h)/[\Sigma_h (<F_o> - |F_o|_h)^2 \Sigma_h (<F_c> - |F_c|_h)^2]^{1/2}$, R factor = $\Sigma_h |F_o - F_c|/\Sigma_h F_o$, where *h* are the unique reflection indices, F_o are the observed structure factors and F_c are the structure factors calculated from inversion of the non-crystallographic symmetry-averaged map.
§From Luzzati plot and SigmaA analysis, as determined with CNS (see Methods).

   287

the 232 amino acids and accurately locate the 14 chlorophylls and 4 carotenoids within one monomeric LHC-II. For the 14 chlorophylls, assignment of the orientation of the Q_x and Q_y transition dipolar moments was accomplished by proper positioning of the chlorophyll head groups. Ten chlorophylls were modelled with complete phytyl chains, but phytyl chains for the remaining four chlorophylls could be only partially modelled. With the help of $2F_o–F_c$ and $F_o–F_c$ electron-density maps (Fig. 1a, b), all 14 chlorophylls were unambiguously characterized as eight Chl*a* and six Chl*b*. The resulting Chl*a*/*b* molar ratio of 1.33 is consistent with the value determined by earlier biochemical analyses[1,2]. Three carotenoids were identified as two luteins and one neoxanthin, and the fourth member was interpreted as a mixed density involving the xanthophyll-cycle carotenoids. In addition, two lipids, one detergent molecule and about 70 water molecules per monomer have been positioned. Figure 1c, d shows two regions of the electron-density map calculated with the phases at 2.72 Å resolution.

Icosahedral proteoliposome and crystal packing

In the $T = 1$ icosahedral particle (Fig. 2a), 20 LHC-II trimers are organized in a closed '532' point group symmetry, with their central C_3 axis serving as the icosahedral C_3 axis and oriented radially towards the sphere centre. One C_3 axis and two C_2 axes of the icosahedron superpose with the crystallographic axes. These trimers form a spherical shell with an outer diameter of about 261 Å and an inner diameter of about 160 Å. They are oriented in the shell with their flat lumenal surface facing the interior of the sphere and the less flat stromal surface facing outwards, taking part in the contacts with other particles in the crystal. The interactions between two adjacent trimers are mediated mainly by two digalactosyl diacyl-glycerol (DGDG) molecules and two pairs of chlorophylls through van der Waals contacts. They are all located near the icosahedral C_2 axis. The digalactosyl head group of each DGDG is simultaneously hydrogen bonded to the lumenal-surface amino acids from two
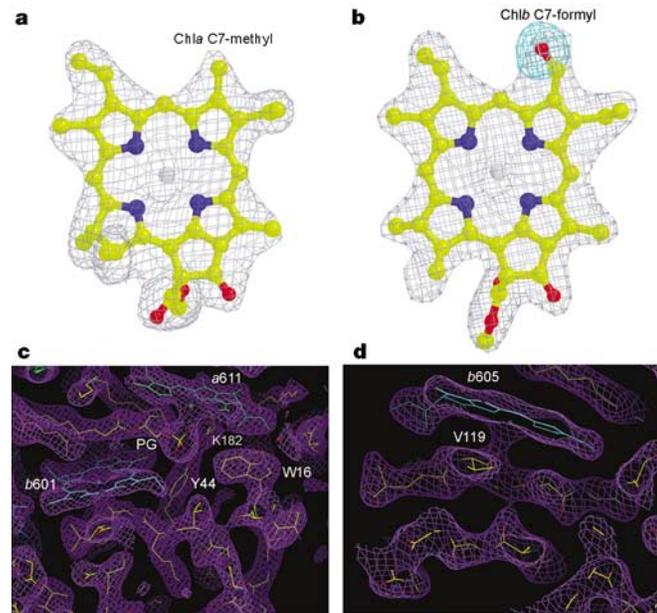
adjacent trimers, functioning as a bridge. The hydrophobic fatty-acid chains of DGDG extend into the membrane interior, interacting with hydrophobic residues and pigments of LHC-II. Other lipids that are expected to fill the gaps between LHC-II trimers and to form a spherical lipid-bilayer vesicle are mostly disordered.

In the crystal lattice (Fig. 2b), LHC-IIs are assembled and packed in a manner different from those in 'Type I' and 'Type II' three-dimensional (3D) crystals of membrane proteins as originally proposed in ref. 8. The LHC–DGDG proteoliposomes assume the shape of closed spheres, presumably originating from curved, small patches of two-dimensional (2D) membrane–protein crystals. Both the outer and inner surfaces of each proteoliposome are hydrophilic. The contacts between two proteoliposomes in the crystal lattice are polar interactions provided by the hydrophilic stromal surfaces of LHC-IIs. The hydrophobic intramembranous surfaces of LHC-II trimers are sheltered from crystal packing by the hydrophobic chains of lipids. We categorize this novel kind of 3D crystal as a 'Type III' membrane-protein crystal.

The apoprotein and LHC-II trimer

The polypeptide main chain of each monomeric LHC-II was continuously traced from Ser 14 to Gly 231. The secondary structure model of spinach LHC-II reported here (Fig. 3a) is similar to the electron crystallographic model of LHC-II from pea[7]. Between the primary structures of spinach and pea LHC-II, 89% of the 232 amino acids are conserved. However, deviations in the residue range, length, turns and orientation between helices in the two species were observed[7] (Fig. 3a). We also found a typical amphipathic short 3_{10}-helix located in the BC loop region and named it helix E. Helix E has a length close to that of helix D and is related to helix D by the internal pseudo-C_2 axis. It is inclined with respect to the membrane plane by an angle of about 30°. In the following EC loop, the polypeptide folds into two short antiparallel strands that are stabilized by an inter-strand ionic pair (Asp 111–His 120) and some hydrogen bonds.

The basic structural and functional unit of LHC-II is the trimer. The whole trimerization region covers the amino-terminal domain, the carboxy terminus, the stromal end of helix B, several hydrophobic residues from helix C and also the pigments and lipid bound to these parts of the polypeptide chain (Fig. 3b). Chl*a* 614, Chl*a* 613, xanthophyll-cycle carotenoid, phosphatidylglycerol (PG), Chl*b* 601 and Chl*a* 602 from one monomer together with Chl*b* 607, Chl*b* 609 and Chl*a* 603 from the neighbouring monomer line up from the periphery of the trimer to the core region near the central C_3 axis at

Figure 1 Electron-density map at 2.72 Å resolution. **a**, Chl*a* and **b**, Chl*b*. Grey cage, $2F_o–F_c$ density (1.5 × σ level); cyan cage, $F_o–F_c$ density (4.0 × σ level). No residual $2F_o–F_c$ or $F_o–F_c$ density appears beside Chl*a* C7-methyl, while strong $2F_o–F_c$ and $F_o–F_c$ densities show up at the position of Chl*b* C7-formyl if it is omitted. **c**, N-terminal region including binding sites for a Chl*b* (cyan) and a phospholipid coordinated to a Chl*a* (green). **d**, Two antiparallel polypeptide strands in the EC loop region with one Chl*b* bound. In **c** and **d**, $2F_o–F_c$ densities (1.5 × σ level) are shown as a purple cage.
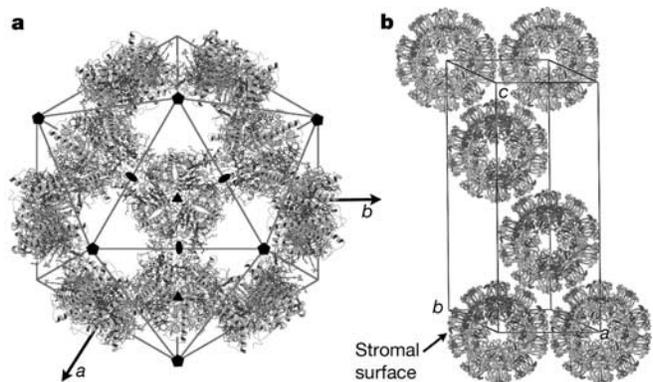

Figure 2 Organization and packing of the icosahedral particles. **a**, Schematic drawing of one-half of the LHC-II–DGDG proteoliposome viewed along the *c* axis of the hexagonal cell. **b**, Packing diagram of 'Type III' membrane-protein crystal, showing the contacts between icosahedral spherical particles in the hexagonal cell. Prosthetic groups are omitted for clarity. The N-terminal domain and AC loop region located at the stromal surface are involved in the crystal packing.

the interface between monomers, forming extensive hydrophobic interactions. Six Chl*a* (Chl*a* 602 and Chl*a* 603 from each monomer) constitute the core of the trimer. Our observations directly reveal the structural role of PG in stabilizing the LHC-II trimer and clearly indicate that hydrophobic interactions dominate the associations between monomers within a trimer. It was shown that removal of the first 49 or 51 amino-acid residues of the polypeptide by proteolytic cleavage led to loss of PG and complete dissociation of the trimer into monomers, and that hydrolysis of the PG by phospholipase A_2 has a similar effect in breaking down the LHC-II trimer[3].

Chlorophyll-binding sites

In a crystallographic asymmetric unit, the individual chlorophyll-binding sites in each LHC-II monomer are occupied by one type of chlorophyll (either Chl*a* or Chl*b*). No mixed binding sites were observed. All central ligands of the 14 chlorophylls have been identified as side chains of seven amino-acid residues, two backbone carbonyls, four water molecules and the phosphodiester group of a PG (Supplementary Table 1; a comparison with a previous model[7] is also included). This coordination mode of Chl*a* 611 to PG is the second case of its kind since its first discovery in photosystem I (ref. 10). On the other side, the phosphodiester group of PG forms a hydrogen-bonding and ionic interaction with the side chains of Tyr 44 and Lys 182 respectively (Fig. 1c).

The polypeptide backbone NH and side chains also form hydrogen bonds with the C7-formyl groups (Chl*b*) and the C13^1-keto groups of several chlorophylls (Supplementary Table 1). These interactions will not only strengthen the linkage between pigments and protein, but also influence the absorption characteristics of chlorophylls as shown previously[11]. Except for Chl*b* 601, nearly all Chl*b* in the complex are selectively hydrogen-bonded to the polypeptide or to the coordinated water of Chl*b* 607 through their C7-formyls. The amide side chain of Gln 131 interacts with three Chl*b* molecules. One hydrogen bond is formed through the interaction of its C = O with the coordinated water of Chl*b* 606, and two additional hydrogen bonds are formed by its NH_2 interacting with the C7-formyls of Chl*b* 607 and Chl*b* 609. Moreover, the C7-formyl of Chl*b* 606 is hydrogen-bonded to the coordinated water of Chl*b* 607. All these interactions bring three Chl*b* into close proximity,

resulting in the clustering of Chl*b* molecules in this region, which may facilitate the efficient energy transfer between these chlorophylls. It was suggested by functional investigations that Gln 131 is involved in the selective binding of Chl*b* molecules to LHC-II[12,13]. As for the selective binding of Chl*a*, we notice that the environment surrounding the C7-methyl groups of Chl*a* molecules is mostly nonpolar. Hydrophobic repulsion or steric hindrance may be the factors affecting the binding affinity of Chl*b* to these Chl*a*-binding sites.

Chlorophyll arrangement for efficient light harvesting

The chlorophylls in LHC-II are vertically distributed into two layers within the membrane, each layer lying close to the stromal or lumenal surface (Fig. 4a). Inside a monomer, the layer close to the stromal surface contains eight chlorophylls (five Chl*a* and three Chl*b*), which surround the central helices A and B more or less evenly to form an elliptical ring (Fig. 4b). The average centre-to-centre distance between two neighbouring chlorophylls is about 11.26 Å, with a maximum of 12.79 Å and a minimum of 9.74 Å. Each chlorophyll inside this layer can find its symmetric mate related by the internal pseudo-C_2 axis. The remaining six chlorophylls (three Chl*a* and three Chl*b*) are arranged in the layer close to the lumenal surface. They form two separate clusters comprising four chlorophylls (three Chl*b* and one Chl*a*) and a Chl*a*–Chl*a* dimer (Fig. 4c). Among them, Chl*b* 606 and Chl*a* 604 are associated with the smallest centre-to-centre distance (8.05 Å) in LHC-II. The shortest distance between two chlorophyll layers is about 13.89 Å (Chl*b* 609 to Chl*b* 606).

Another interesting feature of this chlorophyll arrangement is the enrichment of Chl*b* molecules around helix C and at the interface between monomers (Fig. 4a). All six Chl*b* molecules are located in this region, with five of them belonging to one monomer and the remaining one (Chl*b* 601) from the neighbouring monomer. Chl*b* 601 (II) and Chl*b* 609 (I) (distance, 11.79 Å) are the closest associated couple of chlorophylls between adjacent monomers within a trimeric LHC-II, indicating that this Chl*b*-rich region is of critical importance in energy equilibrating inside a functional trimer.

In the trimeric LHC-II, all 24 chlorophylls from the stromal layer are organized into two irregular circular rings (Fig. 4d). The inner ring located in the core region of a trimer is composed of six Chl*a* molecules that are thought to have an important role in inter-monomeric energy transfer[14]. The remaining nine Chl*a* and nine Chl*b* (those covered by the yellow circular ring in Fig. 4d) form the outer ring and are arranged in a mosaic pattern, with three Chl*b* alternating with three Chl*a*. This new pigment arrangement would favour the efficient absorption of incident light energy from all directions in a broad spectral region and the transfer of the excitation energy to the nearest exit, the putative terminal fluorescence emitter Chl*a* 612 (Supplementary Table 2), in a few steps and at high rates. Energy transfer between two lumenal clusters are much less efficient than those within a stromal layer, as they are separated by larger distances (Fig. 4e). We infer that these lumenal chlorophyll clusters might serve as upstream energy collectors, absorbing energy and transmitting it to the stromal chlorophylls in a relatively independent way. The energy absorbed by the stromal chlorophylls is quickly focused on Chl*a* 612/Chl*a* 611 and is further transmitted to the neighbouring LHCs or reaction centres.

Carotenoids as light-harvesting antennae

The two central carotenoids with all-*trans* configurations are bound in the grooves on both sides of the supercoil (helices A and B) to form a cross-brace. They are assigned as lutein molecules (Fig. 4). Best fit with the electron density is achieved when the β-rings of both lutein molecules are oriented towards the lumenal surface and the ε-rings point to the stromal surface. The polyene chains of lutein 620 and lutein 621 are inclined with respect to the membrane normal
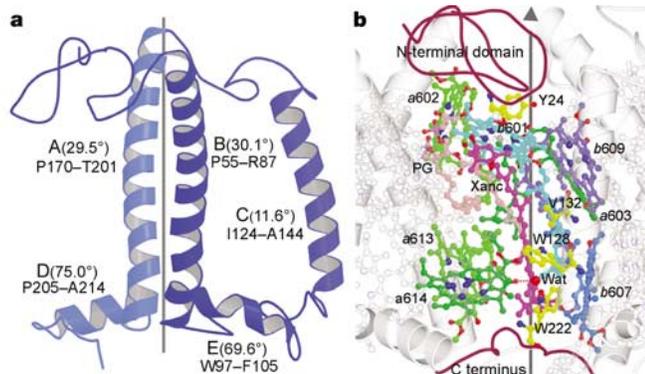

Figure 3 Secondary structure of monomeric LHC-II apoprotein and trimerization. View in parallel with the membrane plane. **a**, The vertical line indicates the approximate direction of the membrane normal and the position of the pseudo-C_2 axis. Helices are labelled A–E. Helix E is newly defined, whereas others are labelled as before[7]. The angle between the central axis of each helix and the membrane normal is shown in parentheses, with the residue range marked below each value. **b**, The interface between two adjacent monomers is shown. Colour code: yellow, amino-acid residues; green, Chl*a*; cyan and blue, Chl*b*; magenta, xanthophyll-cycle carotenoids; pink, PG; red, water; maroon, C_α traces of N-terminal (Ser 14–Asp 54) and C-terminal (Asp 215–Gly 231) polypeptide chain. The vertical line represents the local C_3 axis of an LHC-II trimer.

289

by angles of about 59° and 62°, respectively. Both ring-shaped end groups of these two lutein molecules interact with four internal homologous segments of the polypeptide[15] located on both ends of helices A and B through van der Waals contacts and hydrogen bonds. Their polyene chains are firmly fixed in two elongated narrow hydrophobic cavities on both sides of the supercoil, providing strong and rigid linkage between helices A and B. They are indispensable for proper *in vitro* folding of LHC-II into stable complexes[16–19].

The third carotenoid, shaped like a bent-over hook, is located in the Chl*b*-rich region around helix C and is assigned as 9′-*cis* neoxanthin (Fig. 4). Its polyene chain forms an angle of about 58° with the membrane normal. A value of about 57 ± 1.5° derived from linear dichroism spectra[20] confirms our assignment. The epoxycyclohexane ring of neoxanthin hangs over the chlorin ring of Chl*a* 604 and is hydrogen-bonded to the hydroxyl of Tyr 112 via its C3′-hydroxyl. Side chains of Leu 134, Met 135, Val 138 from helix C and Trp 71 from helix B as well as chlorin rings and phytyl chains of Chl*b* 606 and Chl*b* 608 form a hydrophobic cleft that accommodates the hook-shaped polyene chain of neoxanthin. This binding site has been shown to be highly selective for neoxanthin[17,19]. The cyclohexane ring of neoxanthin on the other end stretches into the exterior solvent region and exhibits weak electron density.

The rate of singlet excitation energy transfer between carotenoids and chlorophylls is correlated with the mutual orientation between them, the centre-to-centre intermolecular distance and the closest distance between two conjugated parts (Supplementary Table 3). Six Chl*a* are found to be in favourable orientations and distances with respect to two luteins for efficient singlet energy transfer from lutein to Chl*a*. The data also show that efficient energy transfer from

neoxanthin to Chl*b* 606 and Chl*b* 608 is highly possible. There is experimental evidence to suggest that singlet excitation energy of luteins is transferred exclusively to Chl*a* molecules and not to Chl*b*[21]. Neoxanthin was found to transfer its energy mostly towards Chl*b*[21,22]. It can be concluded that lutein and neoxanthin found in LHC-II may function as effective accessory light-harvesting antennae, absorbing light in the blue–green spectral region as a complement to Chl*a/b* absorbing in the red region. This is in addition to their obvious structural role as well as their photoprotective role of quenching triplet chlorophylls and singlet oxygen[7].

Structure-based non-photochemical quenching model

The fourth carotenoid we discovered in LHC-II is located at the monomer–monomer interface. The polyene chain of this carotenoid has an all-*trans* configuration and forms a small angle (34°) with the membrane normal. As shown in Fig. 3b, a hydrophobic pocket is formed at the interface by several chlorophylls, hydrophobic residues from the polypeptide and the PG. Part of the polyene chain of this carotenoid together with one of its end groups is accommodated inside this pocket. The opposite end group sticks outside the binding pocket and faces the chlorin plane of Chl*b* 601 at the stromal side. The two ring-shaped end groups of this carotenoid exhibit distinct electron densities (one flat and the other bulgy between C-5 and C-6). This observation led to our original assignment of this carotenoid as an antheraxanthin, an intermediate in the xanthophyll cycle. However, later carotenoid composition analysis revealed that the major component of xanthophyll-cycle carotenoids in the LHC-II preparation used for crystallization is violaxanthin. We propose that the electron density may be interpreted by a mixed binding of different xanthophyll-cycle carotenoids at this site. The end group at the lumenal side points to the cavity formed
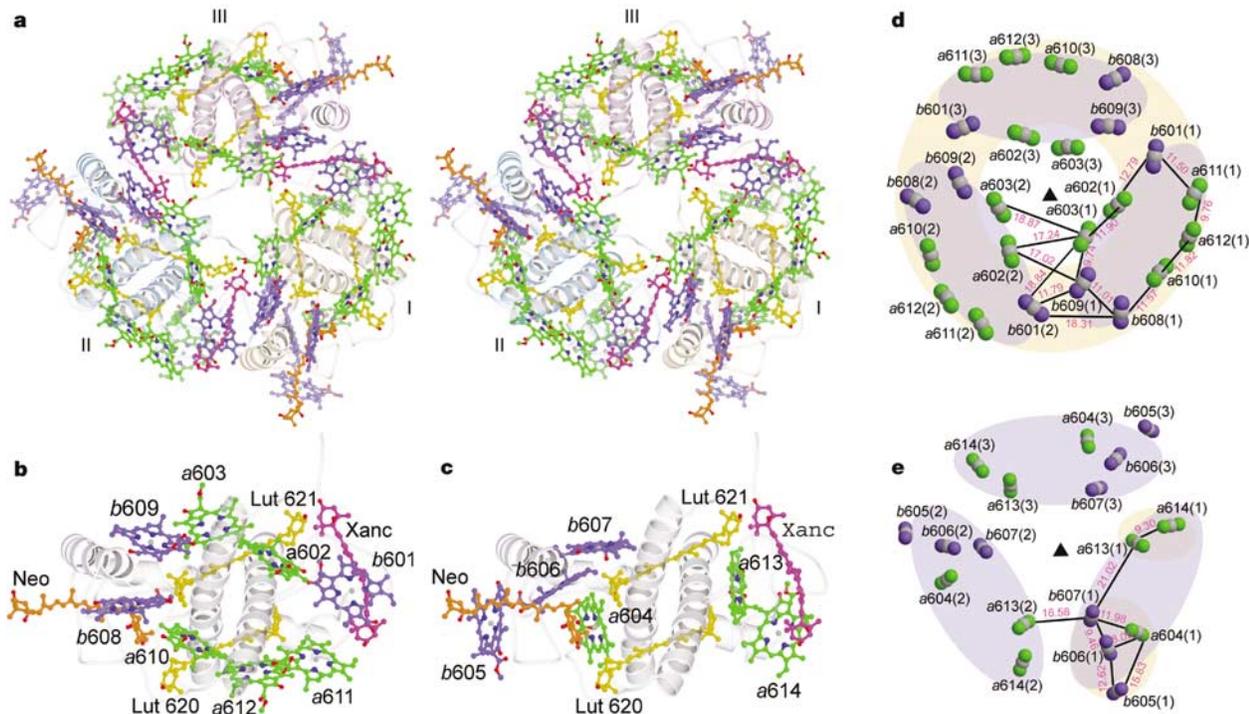

Figure 4 Pigments in the LHC-II trimer and monomer. **a**, Stereo view showing the pigment arrangement pattern in the LHC-II trimer. View along the membrane normal from the stromal side. Monomers are labelled I–III. For clarity, the chlorophyll phytyl chains and lipids are omitted. Green, Chl*a*; blue, Chl*b*; yellow, lutein; orange, neoxanthin; magenta, xanthophyll-cycle carotenoids. **b**, **c**, Pigment pattern in a monomer at the stromal and lumenal sides, respectively. Colour designation the same as in **a**. **d**, **e**, Arrangement of chlorophylls within a LHC-II trimer at the stromal and lumenal sides, respectively.

Chlorophylls are represented by three atoms: the central magnesium atom and two nitrogen atoms. The connecting line between the two nitrogens defines the directions of the Q$_y$ transition dipole. Green, Chl*a* nitrogen; blue, Chl*b* nitrogen; grey, magnesium; purple and blue ellipse, approximate monomer area. The magenta numerical note near the dark line connecting two chlorophylls indicates the centre-to-centre distance (Å) between them.

around the local C_3 axis, suggesting that this cavity might be the docking site for violaxanthin de-epoxidase[23].

The xanthophyll cycle was proposed to have a major role in adjusting the efficiency of light harvesting[4,24,25]. It involves the conversion of violaxanthin into zeaxanthin through antheraxanthin. It was suggested that zeaxanthin molecules can act as direct quenchers of excess excitation by accepting singlet energy transferred from chlorophyll[26,27]. We observed that at least three chlorophylls are close to the xanthophyll-cycle carotenoids and adopt favourable orientations for efficient singlet excitation transfer from chlorophylls to the xanthophyll-cycle carotenoids (Supplementary Table 3). In addition, we noticed that the distance between Chl*a* 613 and Chl*a* 614 is smaller than 10 Å, and their mutual orientation is close to an irregular distribution (Fig. 4e; Supplementary Table 2). These features agree well with the characteristics of statistical pair energy trap[28]. We speculate that this pair of Chl*a* molecules might also function as an excitation energy quencher, which would enhance the quenching effect of the xanthophyll-cycle carotenoids.

Here we propose a structure-based non-photochemical quenching (NPQ) model concerning LHC-II (Fig. 5). Efficient non-photochemical energy-transfer pathways are established upon aggregation of LHC-II trimers mediated by DGDG, so that the excitation energy is able to escape from one trimer to the adjacent trimer via these pathways. The following step is the migration of the excitation to the trapping site (the xanthophyll-cycle carotenoids and/or Chl*a* 613–Chl*a* 614 pair), where the non-radiative dissipation of excitation energy happens. This explains the NPQ observed upon incorporation of LHC-II into the liposomes containing DGDG[29]. Conformational change induced by the protonation of photosystem II proteins including LHC-II was found to be necessary for NPQ[4,25]. Among the seven lumen-exposed acidic residues in our structure, four of them (Glu 94, Asp 111, Glu 207 and Asp 211) form ion pairs with basic residues. The protonation of these acidic residues under low lumenal pH conditions may trigger the conformational changes of helix D and the BC loop. The linker chlorophylls (Chl*a* 614 and Chl*b* 605) at the trimer–trimer interface (Fig. 5) are coordinated to these regions of polypeptide chain. They may be moved and reoriented to promote the non-photochemical energy transfer and/or the quenching effect of the putative energy-trapping sites. As a consequence, the potential damaging effect of excess energy would be avoided. □

Methods

The LHC-II was isolated according to the protocol described previously[30]. A single step of gel filtration chromatography with Hiload 16/60 Superdex 200 pg column (Pharmacia Biotech) was added to improve sample purification, ensuring crystallization reproducibility. The purified LHC-II was solubilized in a solution containing 0.8% *n*-nonyl-β-D-glucoside (BNG) (Anatrace) and 2 mg ml⁻¹ DGDG (Lipid Products) to a final concentration of 4 mg ml⁻¹ chlorophyll (about 8 mg ml⁻¹ protein) and mixed with the crystallization solution containing 66.5 mM HEPES-NaOH pH 7.5, 0.9–1.1 M citrate trisodium and 0.2% *N*,*N*-bis-(3-D-gluconamidopropyl)deoxycholamide (DBC) (Anatrace) in a ratio of 3.0:1.8 (v/v). The resulting drop was equilibrated against a well of 1 ml crystallization solution at 291–293 K using the sitting-drop vapour-diffusion method. Green tabular crystals appeared a week later and grew to a maximum size of about 0.5 × 0.5 × 0.05 mm after one month. Heavy-atom derivative was prepared by soaking the crystal for about 24 h in artificial mother liquor (50 mM HEPES-NaOH pH 7.5, 0.6% BNG, 0.1% DBC, 1.5 mg ml⁻¹ DGDG, 1.09 M citrate trisodium) containing 0.5 mM K₂HgI₄, followed by a backsoaking procedure for about 3–5 h in heavy-atom-free mother liquor. A cryoprotectant (50 mM HEPES-NaOH pH 7.5, 0.4% BNG, 0.15% DBC, 1.0 mg ml⁻¹ DGDG, 1.15 M citrate trisodium, 11% saturated sucrose) was introduced to the crystal by soaking for a few minutes and then the crystal was flash-frozen for the X-ray diffraction experiment. The first native data set and the derivative data set were collected at PF (Tsukuba, Japan) beamline BL6B and the second native data set was collected at BSRF (Beijing, China) beamline 3W1A. A large crystal-to-detector distance (250–350 mm) and a small oscillation (0.5°) are necessary for reducing the overlap of reflections in the large diffraction angle region. Data were processed with Denzo and Scalepack[31]. A typical icosahedral '532' point group symmetry was found by calculating self-rotation function with GLRF[32]. We inferred that there is only one $T = 1$ icosahedral particle residing in a primitive rhombohedral unit cell by analysing the crystal packing. The icosahedron is oriented with one of its '32' subgroups superposing with the crystallographic '32' point group. Each crystallographic asymmetric unit contains one-sixth of the icosahedron.

Ten heavy-atom sites were located using SnB[33] in direct method mode at 5 Å. The arrangement of heavy atoms in the unit cell also abides with the icosahedral '532' symmetry, confirming our judgement based on self-rotation function and crystal packing analysis. The initial SIR phase was calculated with MLPHARE of the CCP4 suite[34] at 5 Å. Phase refinement and extension was performed according to the standard molecular replacement real-space averaging protocol[9]. The NCS matrix was derived from the output of GLRF and improved by IMP[35]. A molecular envelope enclosed by two icosahedral C_3 axes and one icosahedral C_5 axis covering an icosahedral asymmetric unit with a thickness of about 50 Å was generated by MAMA[35]. Electron density was calculated using FFT (CCP4) and density averaging was performed with AVE[35]. The SFALL program (CCP4) was used to calculate F_c and $α_c$ from the averaged density map. The correlation coefficient and R factor were calculated using Rstats (CCP4). The $α_c$ were combined with the initial SIR phases $α_b$ using SigmaA (CCP4) and a new electron-density map was calculated with the F_o and the combined phases. Phases were extended from 5.0 Å to 3.5 Å using a step size of 0.05 Å, and then to 2.72 Å with 0.02 Å step size. Phase refinement was performed for ten iterative cycles in each extension step. The final averaged electron-density map was of high quality, and the backbone was traced without much difficulty. Most of the pigments are clearly defined in the map. The electron-density map was interpreted using the program O[36]. The polypeptide model was built with the help of a published primary sequence of Lhcb1 (ref. 37). The crystals mainly contain two highly homologous polypeptides, Lhcb1 and Lhcb2. Differences between the two polypeptides are confined to the N-terminal region, which may account for the weakness of the electron density in the region before Ser 14.

Structure refinement was performed with CNS[38]. At initial stages, two rounds of NCS-constrained simulated annealing (torsion angle dynamics protocol) were performed at 3.5 Å, followed by positional refinement. After the resolution was extended to 2.72 Å, the NCS constraint was switched to restraint mode and the restraint was gradually released at final stages. The individual B factors were refined using 10–2.72 Å data. Peaks above 3.0 × σ in the F_o–F_c electron density map were picked as candidates for water molecules, after the R factor was reduced to below 23%. Stereochemical restraints were introduced between chlorophylls and their central ligands during refinement. The final model was evaluated with PROCHECK[39] and all ten monomers have good geometry with only one Ramachandran outlier (Val 119) per monomer, the backbone carbonyl of which coordinates a chlorophyll. Secondary structures of polypeptides were analysed with STRIDE[40]. Figures 1c, d and 2b was prepared with program O[36]. All other figures were prepared with Molscript[41] and Raster3d[42].

Figure 5 Structure-based non-photochemical chlorophyll fluorescence quenching model in oligomeric LHC-II. Top view along the icosahedral C_2 axis from the stromal side. DGDG is shown as a yellow transparent ball-and-stick model. Chlorophylls and xanthophyll-cycle carotenoids are represented as in previous figures. Black arrows represent the excitation energy-transfer pathways from one trimer to the neighbouring trimer and the orange arrow shows the possible transfer pathways from chlorophyll Q_y to xanthophyll-cycle carotenoids S_1. The red stars indicate the putative quenching sites. For clarity, characters in one trimer are in black and those of the other are in grey.

1. Peter, G. F. & Thornber, J. P. Biochemical composition and organization of higher plant photosystem II light-harvesting pigment-proteins. *J. Biol. Chem.* **266**, 16745–16754 (1991).
2. Ruban, A. V., Lee, P. J., Wentworth, M., Young, A. J. & Horton, P. Determination of the stoichiometry and strength of binding of xanthophylls to the photosystem II light harvesting complexes. *J. Biol. Chem.* **274**, 10458–10465 (1999).
3. Nußberger, S., Dörr, K., Wang, D. N. & Kühlbrandt, W. Lipid–protein interactions in crystals of plant light-harvesting complex. *J. Mol. Biol.* **234**, 347–356 (1993).
4. Horton, P., Ruban, A. V. & Walters, R. G. Regulation of light harvesting in green plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 655–684 (1996).
5. Elrad, D., Niyogi, K. K. & Grossman, A. R. A major light-harvesting polypeptide of photosystem II functions in thermal dissipation. *Plant Cell* **14**, 1801–1816 (2002).
6. Nilsson, A. *et al.* Phosphorylation controls the three-dimensional structure of plant light harvesting
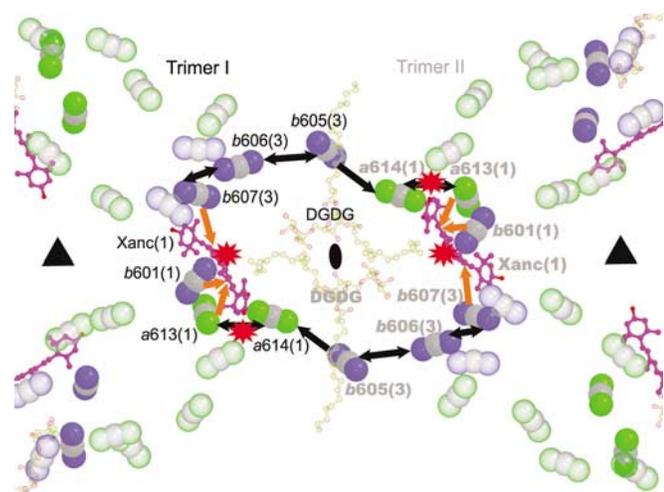
©2004 **Nature Publishing Group**

complex II. *J. Biol. Chem.* **272,** 18350–18357 (1997).

7. Kühlbrandt, W., Wang, D. N. & Fujiyoshi, Y. Atomic model of plant light-harvesting complex by electron crystallography. *Nature* **367,** 614–621 (1994).

8. Michel, H. Crystallization of membrane proteins. *Trends Biochem. Sci.* **8,** 56–59 (1983).

9. Rossmann, M. G. in *Methods in Macromolecular Crystallography* (eds Turk, D. & Johnson, L.) 95–103 (IOS Press, Amsterdam, 2001).

10. Jordan, P. *et al.* Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* **411,** 909–917 (2001).

11. McLuskey, K., Prince, S. M., Cogdell, R. J. & Isaacs, N. W. The crystallographic structure of the B800–820 LH3 light-harvesting complex from the purple bacteria *Rhodopseudomonas acidophila* strain 7050. *Biochemistry* **40,** 8783–8789 (2001).

12. Bassi, R., Croce, R., Cugini, D. & Sandonà, D. Mutational analysis of a higher plant antenna protein provides identification of chromophores bound into multiple sites. *Proc. Natl Acad. Sci. USA* **96,** 10056–10061 (1999).

13. Remelli, R., Varotto, C., Sandonà, D., Croce, R. & Bassi, R. Chlorophyll binding to monomeric light-harvesting complex. A mutation analysis of chromophore-binding residues. *J. Biol. Chem.* **274,** 33510–33521 (1999).

14. Gradinaru, C. C. *et al.* The flow of excitation energy in LHCII monomers: implications for the structural model of the major plant antenna. *Biophys. J.* **75,** 3064–3077 (1998).

15. Bassi, R., Sandonà, D. & Croce, R. Novel aspects of chlorophyll *a/b*-binding proteins. *Physiol. Planta.* **100,** 769–779 (1997).

16. Plumley, F. G. & Schmidt, G. W. Reconstitution of chlorophyll *a/b* light-harvesting complexes: Xanthophyll-dependent assembly and energy transfer. *Proc. Natl Acad. Sci. USA* **84,** 146–150 (1987).

17. Croce, R., Weiss, S. & Bassi, R. Carotenoid-binding sites of the major light-harvesting complex II of higher plants. *J. Biol. Chem.* **274,** 29613–29623 (1999).

18. Paulsen, H., Finkenzeller, B. & Kuhlein, N. Pigments induce folding of light-harvesting chlorophyll *a/b*-binding protein. *Eur. J. Biochem.* **215,** 809–816 (1993).

19. Hobe, S., Niemeier, H., Bender, A. & Paulsen, H. Carotenoid binding sites in LHCIIb. Relative affinities towards major xanthophylls of higher plants. *Eur. J. Biochem.* **267,** 616–624 (2000).

20. Croce, R., Remelli, R., Varotto, C., Breton, J. & Bassi, R. The neoxanthin binding site of the major light harvesting complex (LHCII) from higher plants. *FEBS Lett.* **456,** 1–6 (1999).

21. Gradinaru, C. C., Stokkum, I. H. M. V., Pascal, A. A., Grondelle, R. V. & Amerongen, H. V. Identifying the pathways of energy transfer between carotenoids and chlorophylls in LHCII and CP29. A multicolor, femtosecond pump-probe study. *J. Phys. Chem. B* **104,** 9330–9342 (2000).

22. Croce, R., Müller, M. G., Bassi, R. & Holzwarth, A. R. Carotenoid-to-chlorophyll energy transfer in recombinant major light-harvesting complex (LHC-II) of higher plants. I. femtosecond transient absorption measurements. *Biophys. J.* **80,** 901–915 (2001).

23. Hieber, A. D., Bugos, R. C. & Yamamoto, H. Y. Plant lipocalins: violaxanthin de-epoxidase and zeaxanthin epoxidase. *Biochim. Biophys. Acta* **1482,** 84–91 (2000).

24. Demmig-Adams, B. & Adams, W. W. Photoprotection and other responses of plants to high light stress. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43,** 599–626 (1992).

25. Gilmore, A. M. Mechanistic aspects of xanthophyll cycle-dependent photoprotection in higher plant chloroplasts and leaves. *Physiol. Planta.* **99,** 197–209 (1997).

26. Frank, H. A. *et al.* Photophysics of the carotenoids associated with the xanthophyll cycle in photosynthesis. *Photosynth. Res.* **41,** 389–395 (1994).

27. Ma, Y. Z., Holt, N. E., Li, X. P., Niyogi, K. K. & Fleming, G. R. Evidence for direct carotenoid involvement in the regulation of photosynthetic light harvesting. *Proc. Natl Acad. Sci.* **100,** 4377–4382 (2003).

28. Beddard, G. S., Carlin, S. E. & Porter, G. Concentration quenching of chlorophyll fluorescence in bilayer lipid vesicles and liposomes. *Chem. Phys. Lett.* **43,** 27–32 (1976).

29. Moya, I., Silvestri, M., Vallon, O., Cinque, G. & Bassi, R. Time-resolved fluorescence analysis of the photosystem II antenna proteins in detergent micelles and liposomes. *Biochemistry* **40,** 12552–12561 (2001).

30. Lou, S., Wang, K., Zhao, F., Xu, C. & Kuang, T. A comparative study on PS II light harvesting chlorophyll *a/b* protein complexes between spinach and cucumber. *Acta Bot. Sin.* **37,** 192–197 (1995).

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

32. Tong, L. & Rossmann, M. G. The locked rotation function. *Acta Crystallogr. A* **46,** 783–792 (1990).

33. Weeks, C. M. & Miller, R. The design and implementation of SnB v2.0. *J. Appl. Crystallogr.* **32,** 120–124 (1999).

34. CCP4 Collaborative Computational Project, The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).

35. Jones, T. A. in *Molecular Replacement* (eds Dodson, E. J., Gover, S. & Wolf, W.) 92–105 (SERC Daresbury Laboratory, Warrington, 1992).

36. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).

37. Mason, J. G. Nucleotide sequence of a cDNA encoding the light-harvesting chlorophyll *a/b* binding protein from spinach. *Nucleic Acids Res.* **17,** 5387 (1989).

38. Brünger, A. T. *et al.* Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).

39. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291 (1993).

40. Frishman, D. & Argos, P. Knowledge-based secondary structure assignment. *Proteins Struct. Funct. Genet.* **23,** 566–579 (1995).

41. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24,** 946–950 (1991).

42. Merritt, E. A. & Murphy, M. E. P. Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50,** 869–873 (1994).