Structural Modeling of the Lhca4 Subunit of LHCI-730 Peripheral Antenna in Photosystem I Based on Similarity with LHCII*

Alexander N. Melkozernov‡ and Robert E. Blankenship

From the Department of Chemistry and Biochemistry and Center for the Study of Early Events in Photosynthesis, Arizona State University, Tempe, Arizona 85287-1604

Peripheral chlorophyll a/b binding antenna of photosystem I (LHCI) from green algae and higher plants binds specific low energy absorbing chlorophylls (red pigments) that give rise to a unique red-shifted emission. A three-dimensional structural model of the Lhca4 polypeptide from the LHCI from higher plants was constructed on the basis of comparative sequence analysis, secondary structure prediction, and homology modeling using LHCII as a template. The obtained model of Lhca4 helps to visualize protein ligands to nine chlorophylls (Chls) and three potential His residues to extra Chls. Central domain of the Lhca4 comprising the first (A) and the third (C) transmembrane (TM) helices that binds 6 Chl molecules and two carotenoids is conserved structurally, whereas the interface between the first and the second TM helices and the outer surface of the second TM helix differ significantly among the LHCl and LHClII polypeptides. The model of Lhca4 predicts a histidine residue in the second TM helix, a potential binding site for extra Chl in close proximity to Chls a5 and b5 (labeling by Kühbrandt). The inter pigment interactions in the formed pigment cluster are suggested to cause a red spectral shift in absorption and emission. Modeling of the LHCl-730 heterodimer based on the model structures of Lhca1 and Lhca4 allowed us to suggest potential sites of pigment-pigment interactions that might be formed upon heterodimerization or docking of the LHCl dimers to the surface of PSI.

In photosynthetic apparatus of green algae and higher plants the peripheral chlorophyll a/b binding antenna (LHCl) increases the light-harvesting capacities of Chl a binding photosystem (PS) core reaction centers I and II (1). In adaptive response to changing environmental conditions, the peripheral antennas undergo dynamic structural changes resulting in regulation of excitation flows between the two photosystems and a photoprotection against excess illumination (2, 3).

In contrast to the PSI-LHCII supercomplex that serves as a source for the low energy shifts in absorption of PSI, biochemical and spectroscopic studies of LHCl suggest that in the PSI-LHCl supercomplexes the peripheral antenna and the PSI core antenna have structurally and spectrally distinct pools of red pigments (9). As in the PSI core antenna, excitonically coupled dimers or trimers of Chl a or Chl b in the LHCl were suggested to form a pool of red pigments in the LHClI (9–11). Recent experiments with antisense inhibited Arabidopsis plants in vitro (12), and reconstitution of the polypeptides in vitro (13, 14) suggests that the presence of low energy pigments is a feature of all four Lhca polypeptides (Lhca1, Lhca2, Lhca3, and Lhca4). Lhca1 seems to possess the less red-shifted spectral forms (684 nm) (14, 15), which is in agreement with its close relatedness to a minor light-harvesting polypeptide of PSI, CP29. On the contrary, the Lhca4 binds the "reddest" pigments (9, 16).

The presence of the strong fluorescence at 730 nm in Lhca4 even at room temperature indicates a specific molecular organization of the pigments that efficiently localizes the excitation and could dissipate the excess excitation energy as a nonphotochemical sink when the LHClI antenna is energetically uncoupled from the PSI (9). Despite a consensus that all cab proteins have common evolutionary origin, similar structure, and protein sequences (17), there are local structural differences that determine the assembly of Lhca polypeptides around PSI as dimers (Lhca2/Lhca2, Lhca3/Lhca3, and Lhca1/Lhca4) (18) in contrast to the trimmer-forming Lhcb polypeptides in LHClII.

The basic structural features of the chlorophyll a/b binding proteins (cab) were revealed by electron crystallography study of the most abundant trimer-forming 25-kDa Lhcb1 protein (19). Each monomer of the LHClII trimer consists of three membrane-spanning α-helices. Two of the helices are highly conserved in protein sequences of different Chl a/b proteins and provide a close contact of pigments. The complex binds about 12 molecules of Chl and 2 molecules of lutein. Although the positions of the pigments are known, the assignments of the Chl a and b molecules are entirely based on spectroscopic data. Two
crossed lutein molecules are positioned in the center of the complex and were suggested to closely interact with chlorophyll a molecules. Each LHCl protein binds about 8 Chl a and 2–3 Chl b molecules (13, 14). Chl b to Chl a energy transfer dynamics in LHCl is largely similar to that in the LHClII (9, 20), although kinetics of Chl a excitation equilibration in LHClII and in Lhca1/Lhca4 heterodimer are significantly different due to the presence of long-wavelength absorbing pigments in Lhca4.

The aim of this work was to gain insight into the structure of the specific local environments in LHCl polypeptides, especially in Lhca4, which could explain the red spectral shifts in LHCl and PSI-LHCl. A three-dimensional structural model of Lhca4 was built based on comparative sequence analysis, motif search, secondary structure prediction, and homology modeling using LHClII as a template (19). The model of Lhca4 predicts a histidine residue in the second transmembrane helix, which might bind an extra Chl in close proximity to Chls a5 and b5 (labeling by Kühlbrandt) resulting in strong interpigment interactions and red spectral shifts. Modeling of the LHCl-730 heterodimer based on the model structures of Lhca1 and Lhca4 allows us to suggest potential sites of strong pigment-pigment interactions that might be formed upon heterodimerization or docking of the LHCl dimers to the surface of PSI.

MATERIALS AND METHODS

Multiple Alignments—The Swiss Protein Database was searched for sequence similarities with Lhca1 (sequence accession number Q01667), Lhca2 (Q43485), Lhca3 (Q52904), Lhca4 (P1R S14305), Lhcb1 (Q39141), and Lhcb4 (Q70473). Alignments of the retrieved 250 sequences were performed using PSI-Blast at the fold recognition server three-dimensional PSSM (www.sbg.bio.ic.ac.uk/ servers/3dpssm/) (21). Multiple alignments were analyzed by ClustalW (22). The search for specific motifs in the retrieved primary sequences was done by using the web-hit-initiated Blast (PHI-Blast) server at bioverbasteur.fr/seqanal/interfaces/phiblast.html (23).

Secondary Structure Prediction—Multiple alignments generated by ClustalW were used for the secondary structure prediction at Network Protein Sequence Analysis server (NPSA) (URL, npsa-phil.ibcp.fr). This server allows building a consensus using a number of the secondary structure prediction methods. All the techniques utilize analysis of physicochemical properties of the residues, hydropathy profiles of the sequences, and accessibility of the residues upon folding from the random coil to a tertiary structure (24).

Molecular Modeling and Docking—A homology model of the Lhca4 has been constructed based on the available structural model of the LHClII (19). Coordinates of the pigments and a-carbon atoms of the transmembrane (TM) helices were kindly provided by W. Kühlbrandt. In the absence of a template Protein Data Bank file in the structural data bases, we have chosen a strategy of semi-automatic and manual adjustments of the orientation of transmembrane helices relative to the available backbone conformation (a-carbon atoms) of TM helices in the LHClII model. A secondary structure of the Lhca4 was generated based on the primary sequence of the polypeptide (P1R AC S14305) using DeepView software (25). Results of the secondary structure prediction consensus were used for the adjustments of conformation of three TM helical regions, helix 4 and the regions of the interhelical loops adjacent to helices. Backbone carbon atoms of the polypeptide chains were superimposed with respective regions of the backbone carbon atoms in the LHClII. At the current stage of the modeling the orientation of the pigments (Chls and carotenoids) in the Lhca4 model has been left unchanged as compared with that in the LHClII model. Extra pigments were added into the structure based on the conservation of potential ligands (Fig. 1) and their fit into the local environment. We note that at this stage the three-dimensional model of Lhca4 neither helps to reassign Chl a and Chl b nor distinguish between orientations of Q, and Q, transitions in Chls. Pigments are labeled according to the LHClII model (19). Labels of the amino acid residues in our model identify a name of the residue, a protein region (N or C termini, helices A–D, and loops ab, bc, and cd) and a number of the residue in the primary sequence.

Adjustments of the conformation of the TM helices were done along and around the helix axes with respect to the conserved pigment-binding sites. Local energy minimization of the TM helices and search

for the side chain rotamers of the amino acid residues were done using the GROMOS96 implementation of the DeepView (25). Conformation of the loop regions was adjusted manually using the Ramachandran plots of the amino acid residues.

Docking of the Lhca1 and Lhca4 in the LHCl-730 heterodimer was done manually based on the three-dimensional structural models of Lhca4, Lhca1, multiple alignments of the primary sequences, and results of mutational analysis in the literature.

Molecular graphics were performed using WebLab ViewerLite 3.10 (Molecular Simulations) or DeepView.

RESULTS AND DISCUSSION

Multiple Alignments, Specific Sequence Motifs

Fig. 1 shows sequence alignments of four dimer-forming LHCl polypeptides (Lhca1, Lhca2, Lhca3, and Lhca4) and two representatives of the cab proteins from the peripheral antenna of PSI, Lhcb1 and Lhcb4. These sequences are part of the multiple alignments of 250 sequences retrieved from Swiss Protein Database using PSI-Blast (data not shown).

In all the sequences of LHCl used in the study, similar patterns conserved in the first and the third parts of the sequences are observed (see shaded sites in Fig. 1). The ProDom Database (URL, prodes.toulouse.inra.fr), which automatically generates families of protein sequence domains with similar patterns, indicates that these most numerous sequences in the Swiss Protein Database reveal significant similarities with a variety of evolutionary distant light-harvesting pigment-binding proteins (reviewed in Ref. 17).

Fig. 1 illustrates that significant differences among the primary sequences of LHCl are observed in the region of the N termini and the middle part of the alignments assigned to the second transmembrane helix. We found that the results of the multiple alignments in the middle most diversified region depend on the number of sequences used in the alignment and parameters of the ClustalW search (not shown). Therefore, the alignments in this region were done manually. Table I summarizes results of the PHI-Blast specific motif search for the higher plant cab proteins. A general motif includes conserved potential pigment-binding sites that are separated by seven predominantly hydrophobic residues. The first pigment-binding site is either Glu or Gln, and the second residue is conserved as Glu. Although the seven residues between the potential pigment-binding sites are different, the PHI-Blast search confirms that they form a specific sequence motif in each group of LHCl (Table I). For Lhca4, exhaustive searches of the sequence data bases detected occurrence of a histidine residue between the two potential ligands (see motif EFILHXXE in Table I), which is unique for Lhca4. The pattern of seven hydrophobic residues between two Chl ligands (Gln and Glu) in Lhcb1 sequences perfectly aligns with the sequences of other LHCl (Fig. 1).

Another important motif in the second TM helix includes a conserved Arg residue, which is located two residues after the second pigment-binding site (conserved Glu). Conservation of the Arg suggests importance of this residue either for pigment binding or stability of the LHCl. The residue before the conserved Arg is either charged, polar, or hydrophobic. Searches for the patterns in non-redundant Swiss Protein Database revealed at this position a positively charged residue (Arg or Lys) in Lhca2, Lhca3, Lhca4, and Lhcb6 (CP24), a polar residue (Gln) in Lhca1 and Lhcb4 (CP29), and a hydrophobic residue (Tyr or Phe) in Lhcb1, Lhcb2, Lhcb3, and Lhcb5 (CP26). Overall, the charged residues at this position are characteristic for LHCl that are bound to the core reaction center complexes (both PSI and PSII).
Secondary Structure Prediction

To gain insight into the secondary structure of the studied sequence motifs, we performed secondary structure prediction for the primary sequences used in the study. Different sequences in each group of the LHCs gave similar prediction results (data not shown). Fig. 2 illustrates a typical consensus on the secondary structure for the Lhca4 from tomato (PIR AC S14305) using eight different prediction methods at the NPSA server (URL, npsa-pbil.ibcp.fr). The data clearly identified regions of four helices. Based on homology with LHCII, we assigned these regions to three transmembrane helices A–C and one small helix D at the C terminus. Topologically, the N terminus is located on the stromal side of the thylakoid membrane, and the axis of the helix D is parallel to the membrane on the luminal side. Conserved regions in helices A, C, and D (Fig. 1) and the predictions for the helical regions (Fig. 2) are in

Fig. 1. Comparison of the representative primary sequences of the LHCI and LHCII polypeptides from higher plants. The sequences were selected based on the multiple alignment of 250 sequences retrieved from the Swiss Protein Databank (accession numbers CB11_LYC5, CB12_LYC5, CB21_MAIZE, and CB13_LYC5), TrEMBL (accession numbers Q01667, Q07473, O65217, Q39141, Q9SYW8, Q07489, and Q32904), and PIR database (S14305). Multiple alignments were performed using ClustalW (v1.8). Pigment-binding sites are shaded. Specific motifs in the region of the second TM helix are boxed. See text for details.

Secondary Structure Prediction

To gain insight into the secondary structure of the studied sequence motifs, we performed secondary structure prediction for the primary sequences used in the study. Different sequences in each group of the LHCs gave similar prediction results (data not shown). Fig. 2 illustrates a typical consensus on the secondary structure for the Lhca4 from tomato (PIR AC S14305) using eight different prediction methods at the NPSA server (URL, npsa-pbil.ibcp.fr). The data clearly identified regions of four helices. Based on homology with LHCII, we assigned these regions to three transmembrane helices A–C and one small helix D at the C terminus. Topologically, the N terminus is located on the stromal side of the thylakoid membrane, and the axis of the helix D is parallel to the membrane on the luminal side. Conserved regions in helices A, C, and D (Fig. 1) and the predictions for the helical regions (Fig. 2) are in
Results of the PHI-Blast search for specific motifs in the second TM helix in LHCI and LHCII polypeptides in the Swiss-Prot + TrEMBL non-redundant data base

For each cab protein the first line in the sequence motif column indicates two potential pigment-binding sites (shaded) and specific amino acid residues. The other lines show variations in residues observed in the multiple alignment of 250 sequences used in the study. Conserved residues are shown in boldface.

| Chl a/b binding protein | Sequence motif |
|------------------------|----------------|
| 1 Lhca1                | EPLAIAFVBEHRQS |
|                        | A S G Q N G T  |
| 2 Lhca2                | ELVLIQSWAEGRRW |
|                        | Q M I FL M    |
| 3 Lhca3                | EMALMGFAEBRRF |
|                        | L               |
| 4 Lhca4                | EFILFHYVEIRRW  |
|                        | F S W L        |
|                        | M V            |
| 5 Lhca5                | QFLLMGFAETRKW |
| 6 Lhcb1, Lhcb2, Lhcb3  | QVVLGMGAVEGYRV |
|                        | GFI FF LAF I   |
|                        | LVI LISA I G A |
| 7 Lhcb4 (CP29)         | EPLAIAFVBEQRN |
|                        | V V GYI        |
| 8 Lhcb5 (CP26)         | EYVVLGGABBTRYR |
|                        | L A            |
| 9 Lhcb6 (CP24)         | QLMHGWVESKRW   |
|                        | L               |

a good agreement with the assignments in the LHCII model (19).

Consensus on prediction data for all sequences used in the study showed that the most significant differences among LHCI and LHCII are observed in the luminal part of the helix A, loop ab, and the helix B (Fig. 3). According to the prediction, the conserved motif (WF)AML(AG) in helix A has an α-helix conformation (the letters in parentheses indicate the residues occurred at this position). However, the luminal part of the helix has regions with coil conformation suggesting possibilities of helix kinks. The length of the region with the coil conformation varies among studied LHCs; however, it is the longest in Lhca1.

The TM helix B in Lhca1, Lhca3, and Lhca4 is characterized by a uniform α-helical structure with 3.6 residues per turn of the helix and side chains projected outward from the helix. Such a structure is optimal for binding pigments in the interhelical space. The stromal half of helix B has conserved motifs with arginines. Our results indicate a significantly different prediction for the secondary structure of the luminal part of helix B in Lhca2. The predicted conformation for this region is the extended β-sheet. In contrast to the α-helix conformation, the extended β-sheet is characterized by significant van der Waals interactions between atoms of neighboring residues and by the absence of intrasegment hydrogen bonds.

Building a Three-dimensional Structural Model of Lhca4

Fig. 4 presents the three-dimensional structural model of the Lhca4 based on the structural similarity with LHCII. The model resolves protein ligands to nine Chls, whose structure was determined with 3.4-A resolution in the LHCII model (a1, a2, a3, a4, a6, b1, b3, b5, and b6) and three potential His residues to extra Chls. Some of the pigments identified by the LHCII model (a7 and b2) were removed from the Lhca4 model because of the absence of apparent ligands in the protein environment.

The consensus prediction methods systematically predict that helices A and C are longer than helix B because of the transmembrane topology of the protein. Axes of helices A and C form a cross in one projection and are nearly parallel to each other and to the membrane normal in another projection. A shorter helix B is overall perpendicular to the membrane plane. The helix D at C terminus is parallel to the membrane plane; however, a short cd loop would allow for lateral translations in the membrane plane.

Pigment Binding to Helices A and C (Fig. 5A and Table II)—The model indicates that Chls a1, a2, a4, and a5 and two lutein molecules are bound to the protein region in helices A and C with conserved motifs and rigid α-helix conformation. The oxygen atom of the carbonyl group in Glu-A44 is directly involved in binding of a central magnesium of a4 with a distance of ~2 Å. On the contrary, the distance from Asn-A47 to the central magnesium in a5 of ~4.5 Å suggests an involvement of a water molecule in the binding. Oxygen atoms of Glu-C153 and Asn-C156 are ligands to a1 (~2.5 Å distance to magnesium) and a2 (3 Å) in the helix C.

To obtain a better overlap with the backbone conformation of LHCII, helix kinks in the luminal parts of the helices A and C were imposed. The secondary structure prediction (Fig. 2) predicts regions of coil conformation enabling changes of the angles in the rigid helix structure. The possible helix kink could explain the involvement of Gly-A57 into binding of a6. The α-carbon atom of the conserved Gly-A57 (see alignment in Fig. 1) is the only candidate for the ligands to the magnesium atom in a6 in the local environment. A symmetrical luminal part of the helix C binds a3 by a direct liganding from the oxygen atom in Gln-C170, which is also conserved in multiple alignments. His-C171 is an example of potential Chl ligands in Lhca4. In a majority of Lhca4 and Lhca2, this residue is conserved as histidine, the strongest ligand to Chl, whereas this position is occupied by Gln in Lhca1 and Ala in Lhcb4, Lhcb1, and Lhca3.

Overall, the helices A and C and the binding of six Chls a1—a6 reveal a pseudo-C2 symmetry in the central domain of Lhca4 (Fig. 5A) suggesting a duplication of these two helices in evolution (17). Binding of four Chls (a1, a2, a4, and a5) in close proximity to the luteins enables ultrafast excitation energy transfer from Car to Chl a observed in LHCII (26) and LHCI polypeptides.2 The model also visualizes pairs of charged residues between helices A and C that are most probably involved in structural stabilization of this central region via electrostatic interactions. For example, Lys-A38, a positive residue, has a negatively charged counterpart in helix C (Glu-C151).

Pigment Binding to Helix D (Fig. 5B and Table II)—Helix D is located in the luminal part of the complex. In Lhca4 it has a length of 8—9 amino acid residues with a conformation of α-he-
lix. The model identifies His-D185, a 100% conserved residue, as a direct ligand to the magnesium atom of Chl b3. An amino acid residue at position 191 is conserved as His in Lhca4, Lhca2, Lhca1, and Lhcb4. This is another potential candidate for binding extra Chl in Lhca4, although the position might be unoccupied in monomers.

Our model shows that the surface of the helix D is predominantly negatively charged (Asp-D180, Gln-D184, Asn-D192, and Gln-D196).

Pigment Binding to Helix B (Fig. 5C and Table II)—The main structural motif of the helix B for Lhca4 is two residues of glutamate, Glu-B94 and Glu-B102, with seven residues between them. Structurally, these two conserved Glu residues fit perfectly as ligands to b6 and b5. The distances from the oxygen atoms of Glu-B94 and Glu-B102 to the central magnesium atoms of these pigments are 2 and 2.5 Å, respectively.

The structural model further supports our observation based on the analysis of the specific motifs in the region (Fig. 1 and Table I) that His-B99 is a residue, which is unique for Lhca4. In the structural model this residue directs toward the helix AB interface and is sterically able to ligate an extra pigment (Fig. 5C). Surprisingly, this pigment locates in close proximity to Chl b5. Taking into account the rotamer conformations of the Glu-B102 (b5) and His-B99 (extra Chl), the interpigment distances in the dimer might be as close as 3 Å with a substantial overlap of the tetrapyrrole planes. The closest neighbor to this dimer is Chl a5 liganded by Asn-A47. All three pigments have direct ligands to the central magnesium atom. Some adjustments in the pigment orientations are possible including decrease of the distance between the oxygen atoms in Asn-A47 to magnesium of a5. We mentioned above that a water molecule might mediate pigment binding with Asn-A47. The shorter distance between the ligand (Asn-A47) and the pigment (a5) would result in stronger pigment-pigment interactions in the trimer a5-b5-extra Chl. Figs. 2 and 3 illustrate that in Lhca4 helix B has a defined secondary structure of /H-turn suggesting that rigid conformations in this region would favor binding of the pigment trimer.

Earlier it was suggested that strong excitonic interactions in the dimer (or trimer) of Chl a and Chl b molecules as well as specific Chl b-protein interactions could be responsible for a significant red spectral shift of the absorption band of the red pigment (11, 27). In the PSI core, which is also characterized by red-absorbing species, the strong pigment-pigment interactions as well as influence of the protein charges are thought to cause the red spectral shifts (6–8). The stromal part of the helix B includes the double arginine motif (Table I). Our three-dimensional model helps to visualize...
the relative position of these residues in the helix (Fig. 5C). The first residue in this motif, Arg-B104, locates on the outer surface of the helix B. In different LHCs, this position is occupied by different residues (Fig. 1 and Table I and see below). The second Arg-B105 is directed toward the Chl b1, which suggests an involvement of this residue in binding of b1. For example, in PSI structure (code 1JB0) arginine is commonly bound to the Chl a via the H-bond of a water molecule and a keto group in ring E of the tetrapyrrole. However, our three-dimensional model does not visualize any direct ligands to b1 suggesting that this pigment might be missing in Lhca4. In this case the conserved Arg might participate in structure-stabilizing electrostatic interactions. This residue is 100% conserved in all available sequences of cab proteins.

Ten Chls observed in the three-dimensional structural model of Lhca4 (Chls a1-a6, b3, b5, b6, and the extra Chl bound to His-B99) are in good agreement with the estimates of the number of Chls based on reconstitution analysis of the LHCI polypeptides (13, 14). With the observed Chl a/b ratio of ~4, this accounts for 2 or 3 Chl b molecules. Time-resolved spectroscopy studies suggest a peripheral location of Chl b in Lhca4 (27, 28) and involvement of one of the Chl b molecules into the dimer (trimer) formation (11, 27). According to our structural model (Fig. 5C), molecules b3 and b6 might form a pool of Chl b in the Lhca4. Center to center distance from b6 to b5 is about 15 Å. Insertion of the extra Chl bound to His-B99 between b6 and b5 could explain the influence of the excited b6 on the transient spectra of the Chl trimer. The mechanism of this effect needs to be resolved.

The extra pigment-binding sites (His-C171 and His-D191) could result in variability of the Chl a/b ratio in Lhca4. The binding of Chls to these residues might be ensured by Lhca1-Lhca4 heterodimerization (see below).

Superposition of LHCI and LHCII Structures

To support further the suggested conformation for helix B in the proposed Lhca4 structural model, we performed a superposition of the modeled three-dimensional structures of helix B in different LHCI polypeptides. Fig. 6 identifies the structural differences among the LHCI polypeptides (Fig. 6, A–D) and LHCII (Fig. 6E). These differences are observed in the region between the conserved ligands to b6 and b5 (Fig. 1 and Table I). We used this structural pattern as a molecular ruler in our adjustments of the structure (see also Table II).

The results of the overlap for the three-dimensional structural models of helices B in the studied LHCs further confirm that in all LHCs, except Lhca4, there are no residues to extra Chls in the region between ligands b6 and b5 (Figs. 5C and 6A). However, in Lhca1 (Fig. 6B) and Lhca3 (Fig. 6C), the models identify His residues in the region adjacent to the conserved specific motifs with arginines (Table I). Based on the conformation of helix B, these histidine residues, which are potential
ligands to Chls in Lhca1 and Lhca3, are directed outside the interface between helices A and B. This suggests that in the monomeric Lhca1 and Lhca3, these extra pigments might be absent. Importantly, they are relatively far (~11 Å) from the nearest pigment b5. However, based on the structural model, we can suggest that some rotamer conformations of this His residue might result in a closer location of the extra pigment to b5. It should be noted that these His residues are not conserved between Lhca1 and Lhca3 (Table I). This implies that the suggested red spectral shifts due to pigment-pigment interactions in Lhca1 and Lhca3 are species-dependent even within a group of Lhca.

The side chains of conserved Arg in the specific motifs shown in Table I are directed toward b1. There are no potential direct ligands to this pigment in the protein environment in all LHCs; therefore, we suggest that this pigment if it is present in the structure might be stabilized by hydrophobic interactions and by the H-bond from the conserved Arg to the keto group in ring E of the Chl b1. The pigment-binding sites in all LHCs for which we obtained multiple alignments (Fig. 1), PHI-Blast results (Table I), and secondary structure predictions (Fig. 3) are summarized in Table II. For a series of LHCl and LHCII polypeptides, the conserved pigment-binding sites were probed by mutational analysis (29–31).

The residues preceding the conserved arginine in all LHCs are located on the outer surface of the helix B. Our data show that this position is occupied by Arg in Lhca4, Lhca2, and Lhca3, Gln in Lhca1 and Lhcb4, and Tyr in Lhcb1. We note as a very important structural pattern that this residue is always charged or polar for all studied LHC except Lhcb1 and Lhcb5 (CP26). We suggest that the outer surface of helix B partici-
Pigments \(a_7\) and \(b_2\) identified in the LHCII model are not shown because of the absence of apparent binding residues in vicinity of pigments. Binding of putative extra Chls is suggested based on the presence of strong ligands.

### Table II

Comparison of pigment-binding sites in studied Lhc polypeptides (see data bank accession numbers) based on the multiple sequence alignments (Fig. 1) and structural modeling (Figs. 4–6).

| Protein region | Ch1 | Lhca4 (S14305) | Lhca2 (P10708) | Lhca3 (P27522) | Lhca1 (P12360) | Lhcb4 (Q07473) | Lhcb1 (Q39141) |
|----------------|-----|----------------|----------------|----------------|----------------|----------------|----------------|
| Helix A        | a4  | Glu-A44        | Glu-A43        | Glu-A48        | Glu-A43        | Glu-A86        | Glu-A45        |
|                | a5  | Asn-A47        | His-A46        | Asn-A51        | His-A46        | His-A89        | His-A48        |
|                | a6  | Gly-A57        | Gly-A56        | Gly-A61        | Gly-A57        | Gly-A99        | Gly-A58        |
| Extra          |     |                |                |                |                |                | His-B100       |
| Helix B        | b6  | Glu-B94        | Glu-B93        | Glu-B107       | Glu-B101       | Glu-B141       | Gln-B111       |
| Extra          |     | His-B99        |                |                |                |                |                |
| Helix C        | a1  | Glu-C153       | Glu-C158       | Glu-C173       | Glu-C147       | Glu-C188       | Glu-C161       |
|                | a2  | Asn-C156       | Asn-C161       | Asn-C178       | Asn-C150       | Asn-C191       | Asn-C164       |
|                | a3  | Gln-C170       | Gln-C175       | Gln-C190       | Gln-C164       | Gln-C205       | Gln-C178       |
| Extra          |     | His-C171       | His-C176       |                |                |                |                |
| Helix D        | b3  | His-D185       | His-D190       | His-D205       | His-D180       | His-D220       | His-D193       |
| Extra          |     | His-D191       | His-D196       | Asn-D211       | His-D186       |                |                |

Fig. 6. Comparison of three-dimensional structural models of the second transmembrane helix in Lhca4 (A), Lhca1 (B), Lhca3 (C), Lhca4 (D), and Lhcb1 (E). Structures are obtained by three-dimensional fitting of the primary sequences of this region to the structural model of Lhca4 (Fig. 4). Scale and orientation are similar for all models. Superposition of the structures is shown in F. See also Table II for comparison of pigment-binding sites.

Pigments in the docking of LHCl polypeptides to the surface of the PSI core and minor LHClII (for example, Lhcb6 (CP24) and Lhcb4 (CP29) to the surface of the PSII core).

PsaA, PsaB, PsaJ, and PsaF subunits of the PSI were suggested to be involved in the docking of LHCl (18, 32, 33). The conserved hydrophobic residues on the outer surface of the helix B indicate the involvement of hydrophobic interactions into the docking process. It is interesting that this surface in Lhcb1, Lhcb2, Lhcb3, and probably in Lhcb5 is completely hydrophobic.
Three-dimensional Model of the LHCI-730 Heterodimer

Earlier it was demonstrated that Lhca4 and Lhca1 polypeptides are assembled in the LHCI peripheral antenna as heterodimers (16). Fig. 7 shows a suggested three-dimensional model of the Lhca1-Lhca4 heterodimer (LHCI-730). A structural model of Lhca1 was obtained using the model of Lhca4 as a three-dimensional template. Docking of two subunits of the heterodimer was performed manually based on the structural information retrieved from the close inspection of the three-dimensional models of the monomers.

Formation of the cross-structure by helices A and C in the center of the complex (Fig. 5A) imposes conformational constraint for docking the subunits Lhca1 and Lhca4 side by side. Therefore, we performed first a transformation of one structure relative to another along the C2 symmetry axis, which is parallel to the normal of the membrane and one of the projections of the helices.

An important structural feature of all cab proteins is the long N and C termini and interhelical loops that probably participate both in LHC oligomerization (LHCI dimerization and LHCI trimerization) and docking to the surface of photosystem core complexes. Based on the structural models of Lhca1 and Lhca4, we suggest that the lateral conformational mobility of termini and interhelical loops ab and bc in both subunits is essential for dimerization. Earlier mutagenesis studies of the reconstituted Lhca1 and Lhca4 (34) suggested that the N and C termini of Lhca1 are specifically important for the heterodimerization. The structural models visualize for Lhca1 and Lhca4 a series of conserved hydrophobic and charged residues in N termini that might be involved in the heterodimerization. For example, the N terminus adjacent to the stromal part of helix A is overall negatively charged in Lhca4. This is achieved by positioning of Asn-N23 and Glu-N32 that are conserved only among Lhca4 close to Asp residues (Asp-N16, Asp-N22, Asp-N26, and Asp-N33) that are conserved in all sequences. Although we do not know the real conformation of the termini and loops, the proposed model of the LHCI-730 heterodimer in Fig. 7 visualizes the relative position of the loops that might be involved in dimerization.

In the predicted heterodimer the surface of helix D in Lhca1 is doomed to the luminal part of the helix B in Lhca4 and vice versa. Hydrophobic and electrostatic interactions between the surfaces could stabilize the complex. Secondary structure prediction data (Fig. 3) indicate for Lhca1 the presence of a short α-helix in the luminal loop ab (see Fig. 7). The charged residues on this extra helix (particularly, conserved lysine) could interact with the negative charges on the surface of helix D in Lhca4. Alternatively, amino acid residues on the outer surfaces of extra helix in loop ab in Lhca1 and helices D in both subunits might participate in docking to the surface of the PSI. Specific conformation of these regions in Lhca1 and Lhca4 needs to be resolved in further structural modeling or crystallography studies.

The suggested conformational change of helix D upon heterodimerization would result in a translation of b3 bound to His-D191 into the interhelical space in the membrane. Possible changes in the interpigment distances could result in appearance of additional red pigments in the heterodimers. Furthermore, heterodimerization might ensure binding of extra Chls to His-C171 (Fig. 5A) and His-D191 (Fig. 5B) in Lhca4 and His-D186 in Lhca1. This would result in additional binding of the pigments upon heterodimerization, which is experimentally observed for the LHCI-730 heterodimers (13, 14, 16). The extra pigment binding sites in the luminal part of helix C (His-C171), and on the surface of helix D (His-D191), conserved for dimer-forming polypeptides might be vacant in the monomeric form. This explains variability in pigment/protein ratios in LHCI monomers and dimers.

The model of the LHCI-730 heterodimer allows predicting another source for the red spectral shifts in the PSI. If the outer surface of the helix B participates in docking with the PSI surface, then extra Chls bound to His on this surface in Lhca1 (Fig. 6B) and Lhca3 (Fig. 6C) might be involved in pigment-pigment interactions with Chls on the surface of PSI. Earlier reports on Lhca1 and Lhca4 mutants in vivo (35) suggested that specific interactions between the LHCI-730 heterodimer and the PSI core might result in additional red spectral shift in fluorescence spectra. This is also supported by observations of the red spectral forms in Lhca2 and Lhca3 in vivo (12).

In conclusion, inspection of the homology-based three-dimensional structural models of the Lhca4 and LHCI-730 heterodimers indicate at least three causes of the red spectral shifts in the peripheral LHCI antenna that could be based on the specifics of structure. (i) Strong interpigment interactions result from binding of extra Chls in close proximity to b5 and a5 in the helix AB interface. (ii) Conformational lateral movement of the helix D toward the luminal surfaces of the helix B upon heterodimerization results in translation of Chl b3 toward pigments in the interhelical space including those bound to His-C171 in the luminal part of the helix C. (iii) There are strong interpigment interactions between the Chls bound to the surface of helix B (in Lhca1 and Lhca3) and the Chls on the surface of PSI upon docking of LHCI to PSI.

Finally we would like to emphasize that future development of the presented structural models of the LHCI-730 heterodimer will include global energy minimizations of the structure based on extended computer simulations, conformational analysis of the N and C termini and loop regions, and analysis of docking of LHCI to the surfaces of PSI core complex.

REFERENCES
1. Blankenship, R. E. (2002) Molecular Mechanisms of Photosynthesis, Blackwell Scientific, Oxford
2. Haldrup, A., Jensen, P. E., Lunde, C., and Scheller, H. V. (2001) Trends Plant
