Lil3 dimerization and chlorophyll binding in *Arabidopsis thaliana*

Astrid Elisabeth Mork-Jansson[1], Daniela Gargano[1], Karol Kmiec[1], Clemens Furnes[1], Dmitriy Shevela[1,2] and Lutz Andreas Eichacker[1]

[1]Center for Organelle Research, University of Stavanger, Stavanger, Norway

[2]Department of Chemistry, Chemical Biological Centre (KBC), Umeå University, Sweden

*Corresponding author

Lutz Eichacker, University of Stavanger, CORE, Richard-Johnsensgt. 4, 4036 Stavanger, +4751831896,

E-mail: lutz.eichacker@uis.no (LE)

Abstract

The two-helix light harvesting like (Lil) protein Lil3 belongs to the family of chlorophyll binding light harvesting proteins of photosynthetic membranes. A function in tetrapyrrol synthesis and stabilization of geranylgeraniol reductase has been shown. Lil proteins contain the chlorophyll a/b-binding motif; however, binding of chlorophyll has not been demonstrated. We find that Lil3.2 from Arabidopsis thaliana forms heterodimers with Lil3.1 and binds chlorophyll. Lil3.2 heterodimerization (25 ± 7.8 nM) is favored relative to homodimerization (431 ± 59 nM). Interaction of Lil3.2 with chlorophyll a (231 ± 49 nM) shows that heterodimerization precedes binding of chlorophyll in *Arabidopsis thaliana*.

Keywords

Lil3, Chl a, microscale thermophoresis, HCA plot

Abbreviations

LHC – light harvesting complex, Chl – Chlorophyll,

Highlights (**3-5 bullet points, max 85 characters pr. point)**

- Lil3.1 and Lil3.2 can form homodimers in *Arabidopsis thaliana*
- Lowest binding constant is found for heterodimerization
- Lil3 heterodimerization precedes chlorophyll binding

Introduction

The LHC motif was originally described as an overall hydrophobic amino acid sequence composed of 22 amino acids with two charged amino acids: glutamic acid (E), arginine (R) and three conserved glycine (G) residue with the consensus sequence ELINGRLAMLGFLGFLVPELIT (Jansson, 1999). Two 16-mers peptides synthetically designed in the N-terminus of this motif were shown to bind Chlorophyll (Chl) (Eggink and Hoober, 2000). A motif was defined and shown to be required for Chl binding: E-X-X-H/N-X-R or R-X-N/H-X-X-E (Eggink and Hoober, 2000, Green and Pichersky, 1994, Jansson, 1994, Jansson, 1999) in which residues E and H/N were found responsible for coordination of the central Mg^{2+} ion in Chl. The anion carbonyl group in E- and the guanidinium group in R-residues were also shown to play an important role for salt ion paring in E139-R142, E65-R185 and E180-R70 (Kuhlbrandt et al., 1994). The LHC motif has been established as a foundation for classification of the LHC-like protein family. The family is divided according to the number of transmembrane helices: three-helix early light-induced proteins (ELIP), two-helix stress-enhanced proteins (SEP) and one-helix proteins (OHP) (Andersson et al., 2003, Engelken et al., 2010, Heddad and Adamska, 2000).

The light harvesting like protein 3 (Lil3) belongs to the two-helix stress-enhanced proteins (SEPs) in higher plants. However, the proteins' role in

protecting the plant against stress has been discussed the last decade (Craigon et al., 2004, Reisinger et al., 2008, Staleva et al., 2015, Tanaka et al., 2010). Lil3 has been shown to associate with Chl and tocopherol synthesis in *Arabidopsis thaliana* (Tanaka et al., 2010). Recently, the transmembrane amino acids of the LHC-motif in Lil3 were reported to structurally anchor geranylgeranyl reductase (GGR) to the membrane, and to be responsible for the oligomerization of GGR (Takahashi et al., 2013). It has been shown that a recombinant GGR protein functionally did not require Lil3 for the reduction of geranylgeranyl pyrophosphate (GGPP) to phytyl pyrophosphate (PYPP) in plant (Takahashi et al., 2013). GGR catalyzes the NADPH dependent three-step reduction of the pyrophosphate (PP) form of GG or its esterified form in Chl_{GG}. The products of this enzymatic reaction is PYPP or Chl_{PY} (Tanaka et al., 1999).

Cyanobacterial high-light inducible proteins (Hlips) are the ancestors of LHC antennae and other members of the LHC super family. A recent study showed that HliD purified from Synechocystis bound Chl a and β-carotene and exhibited an energy dissipative function (Staleva et al., 2015). The above-mentioned study suggested that the quenching mechanism works *via* a direct energy transfer from a Chl a Q_y state to the β-carotene S_1 state in the LHC superfamily. In barley (Hordeum vulgare L.), Lil3 was identified as the first protein to bind Chl during deetiolation (Reisinger et al., 2008). In this study, we determined the dissociation constants for interaction of Lil3 proteins with chlorophyll a with microscale thermophoresis. We find that the dissociation

constant for heterodimerization of proteins Lil3.1 and Lil3.2 from Arabidopsis thaliana is lower than for binding of chlorophyll a by Lil3.2.

Methods

Split Ubiquitin

The split Ubiquitin assay was carried out according to the DUALmembrane starter kit (Dualsystems Biotech Inc. Schlieren, Switzerland). Genes were amplified from cDNA using *Pwo* polymerase (Roche, Basel, Switzerland), cloned into pPCR-Script (Stratagene, California, USA). The coding sequence of *Lil3:1* (At4g17600), *Lil3:2* (At5g47110), were cloned into the bait (pBT-C) and prey (pPR-N and pPR-C) vectors (Dualsystems Biotech AG, Schlieren, Switzerland). Yeast NMY51 cells were co-transformed with the resulting plasmids according to the manufacturer's instructions (Dualsystems Biotech AG, Schlieren, Switzerland). In the split ubiquitin assay with Lil3:1 as bait, the selective media (SD–WLAH) was supplied with 5mM *of 3-aminotriazole (3-AT).* Coexpression analysis experiments were repeated four times, two times in each direction.

Protein expression

The Lil3 genes, (Lil3.1 – AT4G17600 and Lil3.2 – AT5G47110), of *A. thaliana* were PCR amplified from pUNI51 plasmids containing the Lil3 sequences obtained from TAIR. The amplified sequences were cloned into the pET151d expression vector (Invitrogen, Carlsbad, CA, USA). The Lil3_pET151d plasmids were transformed to BL21 *E. coli* chemically competent cells for expression of recombinant protein. Cultures were induced at OD_{600} = 400 nm

with a final concentration of 1 mM IPTG and incubated at 16 °C on prior to harvest (6000 x g, 15 min, 4°C).

Cell lysis and Protein Purification

Cells were lysed in 20x volume lysis buffer with 0,8 mg/mg lysosome and complete protease inhibitor cocktail tablet (Roche), incubated on ice 1 hour prior to sonication 6x 10 seconds at 30 amplitudes. The filtrated crude protein extract was purified on a 1 mL His-Trap HP column (GE Healthcare Life Sciences) according to the manufacturers instructions, and 1 mL fractions collected. Fractions were separated on a 12 % SDS-PAGE (Laemmli, 1970), stained by Coomassie Brilliant Blue (CBB) and transferred to nitrocellulose (NC) membranes (Towbin et al., 1979, Towbin et al., 1992) for subsequent immunological identification of recombinant Lil3_his (monoclonal anti-polyHistidine Antibody produced in mouse, Sigma Aldrich). Lil3_his fractions were desalted on a HiTrap desalting column (GE healthcare, Buckinghamshire, United Kingdom) using desalting buffer (25 mM Hepes pH7.5, 30 mM NaCl) and protein concentrations determined by the BCA method. Purified Lil3 was verified by mass spectrometry (Waters Corporation, Milford, MA, USA).

Surface plasmon resonance

The SPR measurements were carried out on a Biacore® T-100/T-200 instrument (GE-Healthcare) using CM5 chips. Purified Lil3.1 and Lil3.2 were either immobilized (ligand) on a CM5 chip or in the mobile phase (analyte) and run using single channel analysis. The CM5 chips used were

preconditioned with 3x 10 s injections of running buffer (1x HBS-EP: 10 mM Hepes pH7.4, 0.15 M NaCl, 3.4 mM EDTA and 0.005% surfactant P20), 2x 10s injections of 100 mM HCl, 2x 10 s. injections of 50 mM NaOH and finally 2x 10 s. injection of sodium monododecyl sulfate with the flow rate set to 100 µl/min. Reactive succinimide esters were created by injecting a 1:1 mixture of *N*-ethyl-*N*`-(3-dimethylaminopropyl) carbodiimide hydrochloride (Life technologies, Oslo, Norway) and *N*-hydroxysuccinimide (Life Technologies, Oslo Norway) for 420 s with the flow rate set to 5 µl/min. Ligands where then covalently immobilized (1 min., 10 µl/min) and unreacted esters blocked by ethanolamine-HCl (240s., 5 µl/min). The ligands were covalently linked to the CM5 chip surface at densities of 200 RU for multiple channel runs and 500 RU for single channel runs. Flow channel one (FC1) was treated as FC2-FC4 except no ligand was coupled to the sensor chip channel. Analyte flow rate was set to 100 µl/min to avoid mass transport. Regeneration of chip surfaces was obtained by injecting flowing buffer (30 µl/min, 10-30 min).

Pigment isolation

Etioplasts were isolated from 4.5 days old H. vulgare seedlings as described in (Eichacker et al., 1996, Klein et al., 1987, Muller and Eichacker, 1999). Acetone was added at 80 % (v/v) final concentration in the dark and extracts incubated 20 min prior to transfer and over night incubation at -80 °C. Aggregated proteins were removed by centrifugation (20800 xg, 10 min, 0°C). The concentration of Chl a standard (Sigma, St. Louis, USA) was determined in 100 % aceton (Porra et al., 1989) and a respective volume added to the supernatant prior to evaporation of acetone in a vacuum centrifuge.

Thin Layer Chromatography

For pigment characterization, pigment/lipids frozen on dry ice were thawed into an organic phase composed of 100% (v/v) acetone on ice. Samples were loaded on reversed phase (C18) high-pressure thin-layer chromatography (HPTLC) plates (Merck, Darmstadt, DE) and plates were developed in a solvent composed of 58.8 % (v/v) acetone, 39.2 % (v/v) methanol, and 2 % (v/v) water. HPTLC plates were scanned for fluorescence emission (excitation 633 nm/670 BP30 emission filter) in a Typhoon scanner (GE Healthcare, Buckingham, GB).

Fluorescence spectroscopy assays

Pigment/lipid extracts were characterized by emission (740 nm) and excitation (440nm) spectra measured at 77K in a Horiba Yvonne Florolog-3 Spectrophotometer (Fluorolog®, HORIBA, France). Fluorescence emission (600 to 800 nm) at room temperature was measured upon excitation of HPTLC spots at 440 nm using a y-scale optical light cable.

Microscale thermophoresis, MST

MST experiments were performed on a Monolith NT.115 system using 20 % LED and 20 % IR laser power and consumables (NanoTemper Technologies, München, Germany). The intrinsic fluorescence of the externally added Chl a in the pigment/lipid mix was monitored and applied at a final concentration of 4 µM Chl a diluted in MST buffer with 1 mg/mL BSA and 0,025 % Tween 20. A two-fold dilution series starting at 39 µM

was prepared for the unlabelled Lil3.2 in 25 mM Hepes pH7.5 and 30 mM NaCl. Samples were filled into Premium coated capillaries for measurement. The negative control was conducted by substituting Lil3.2 with (Glu1)-Fibrinopeptid B human (Sigma Aldrich, St. Louis, USA).

MST measurements of protein dimerization was performed by labelling Lil3.1 and Lil3.2 with NT647 using 20 % LED and 20 % IR laser power. The fluorophore was diluted in MST buffer with 0,025 % Tween 20 and monitored at a concentration of 500 nM for the heterodimerization and homodimerization . A twofold dilution series ranging from 6 nM to 25 µM and 3 nM to 100 µM was prepared in MST buffer with 0,025 % Tween 20 for the heterodimerization and homodimerization of unlabelled Lil3.2 respectively . Samples were filled into Premium coated capillaries for measurement.

Hydrophobic cluster analysis

Hydrophobic cluster analysis was carried out using a HCA plot (http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py?form=HCA) (Callebaut et al., 1997, Gaboriaud et al., 1987). The amino acid sequence (i+1, amino acid number 10-260 labeled on abscissa) was virtually rolled along the helical axis, defined by amino acids i+18 (HCA plot) (Callebaut et al., 1997, Gaboriaud et al., 1987), on the basis of an α-helical arrangement with i+3.6 to equal 360°, and with the N-terminus left and C-terminus right. The helical axis was placed horizontal and was characterized by the axis parallel orientation of the H-bonds between amino acids i+4. The 0° to 360° helical outline was plotted on the ordinate and the progression of the amino

acid sequence was defined by an angle of 100° between amino acids i+1 and a positioning of amino acids i+3 at an angle of -60° and of amino acids i+4 at an angle of +40° relative to the helix axis (insert). Amino acid side chains of residue pairs, i and i +/- 3, and i and i +/- 4, are positioned for intra-helix interactions. Amino acid side chains of residue pairs, i and i +/- 1 (100°) (dashed lines, insert) and of i +/- 2 (200°), are disfavored to interact.

Protein identification by mass spectrometry (MS)

Lil3_pET151d plasmids (Lil3.1 – AT4G17600 and Lil3.2 – AT5G47110) were purified by affinity chromatography and proteins were separated by 12 % SDS-PAGE. Preparation of proteins for MS analysis was done according to the in-gel digestion OMX-S protocol (OMX, Seefeld, Germany) (Granvogl et al., 2007). MS and MS/MS analysis of peptides was performed using a Waters Q-Tof mass spectrometer (Waters Corporation, Milford, MA, USA) according to the procedure described earlier (Bryan et al., 2014).

Results

Dimerization of Lil3

The dimerization of Lil3 was investigated using the split ubiquitin system in yeast. Genes *Lil3:1* and *Lil3:2* were fused to the N-terminus of the Cub moiety protein and to the C-terminus and N-terminus of the NubG moiety protein to generate BTC-*lil3:1,* BTC-*lil3:2,* PRN-*lil3:1,* PRN-*lil3:2, PRC-lil3:1 and PRC-lil3:2* respectively. The PRN and PRC constructs were co-expressed in yeast NMY51 cells with BTC-*lil3:1 and* BTC-*lil3:2*. Growth of yeast NMY51 cells, in

the absence of histidine, was readily determined upon coexpression of Lil3:1 and Lil3:2 indicating a direct interaction of both proteins as a heterodimer (Fig. 1). Also, coexpression of Lil3:2 constructs resulted in a strong growth induction. However, growth induction was found low in case of a coexpression of the Li3:1 constructs indicating a lower efficiency for direct interaction of Lil3.1 proteins relative to Lil3.2 proteins in homo-dimer interactions (Fig. 1).

Determination of the binding constant for Lil3 dimerization by Biacore analysis

Recombinant Lil3.1 and Lil3.2 of *A. thaliana* were investigated using surface plasmon resonance analysis (SPR, Biacore®) to validate a Lil3 dimerization and to determine the dissociation constants (K_D) (Fig 2 and Fig S1). Single channel analysis was performed for the homo-dimer interaction of Lil3.1 and Lil3.2 and K_D value of 2,8 x 10^{-5} M (Fig 2A) and 2,2 x 10^{-6} M (Fig 2B) were determined, respectively. In contrast, hetero-dimerization analysis of Lil3 by single channel analysis resulted in a K_D of 9,5 x 10^{-7} M (Fig 2C). Interestingly, the K_D value for hetero-dimerization was about 29-fold (Lil3.1) and 2.4-fold (Lil3.2) lower than for homo-dimerization of the proteins. Therefore, a 10-fold lower dissociation constant for Lil3.2 was determined relative to Lil3.1. Data strengthened the finding from the Y2H assays that Lil3.2 has a higher affinity for homodimerization then Lil3.1 and indicate that the assembly of hetero-dimers is favored in Arabidopsis thaliana.

Determination of the binding constant for Lil3 dimerization using Microscale thermophoresis analysis

Lil3 is a partially membrane- associated protein, and may loose it`s activity for interaction fast if not in a partial lipid environment. In the Biacore studies, immobilization of the *in vitro* expressed protein on the surface of a chip may have altered the proteins interaction capabilities. Therefore, the proteins were also investigated in liquid form using microscale thermophoresis (MST).

Serial dilution experiments were run in triplicate. For hetero-dimerization analysis, Lil3.1 was labeled with fluorescent dye NT647 and a concentration of 500 nM was incubated with unlabeled Lil3.2 ranging from 6 nM to 25 µM. A K_D value of 25 ± 7,8 was determined for the hetero-dimer interaction of both Lil3 proteins (Fig. 3A).

Homo-dimerization of Lil3 was also analyzed in triplicate via MST. Fluorescently labeled Lil3.2 at a concentration of 500 nM was added to a serial dilution of unlabeled Lil3.2 ranging from 3 nM to 100 µM. For the homo-dimerization of Lil3:2 a K_D of 431 ± 59 nM was determined (Fig. 3B). MST results confirmed the SPR and Y2H based results for heterodimerization of Lil3.1 and Lil3.2 and homodimerization of Lil3.2.

Lil3.2 binds Chl a in a microscale thremophoresis analysis

A binding of Chl a to Lil3.2 was investigated via MST using the intrinsic fluorescence of Chl a to enable a label-free analysis. The natural membrane environment for Lil3 was imitated using lipids, and pigments isolated from barley etioplasts were supplemented with a defined concentration of Chl a. The composition of the lipid/Chl a mixture used for the analysis was analyzed by spectroscopy and HPTLC (Fig S2). In the MST analysis, the Chl a/lipid

mixture was kept constant at a Chl a concentration of 100 nM and the solution was added to a serial dilution of unlabeled Lil3.2 ranging from 3 nM to 75 µM. A K_D of 231 ± 49 was determined for the Chl a interaction with Lil3.2 (Fig. 4A).

The interaction of Chl with the protein stabilizing agent bovine serum albumin (BSA) was tested in a negative control experiment (Gorza et al., 2014). The fluorescent Chl a/lipid mixture was maintained constant, and Lil3.2 was exchanged with Glufibrinogen, a peptide standard used in mass spectrometry. As for the binding experiment, the fluorescent mixture was kept at a constant concentration and a serial dilution of glufibrinogen was added ranging from 3.7 nM to 30 µM; however, no binding of Chl to either BSA or Glufibrinogen could be determined (Fig. 4B). Interestingly, the K_D value for binding of Chl a to Lil3.2 was about 2-fold lower than for homodimerization of the Lil3.2 protein, but K_D values for heterodimerization of both Lil3 proteins were found 9-fold lower indicating that heterodimerization of Lil3.2 and Lil3.1 precedes a binding of Chl a.

Two Chl binding motifs are conserved within the 2D structures of Lil3 and LHCP

To visualize the arrangement of hydrophobic regions, and alpha helical positions of conserved residues a two-dimensional (2D) hydrophobic cluster analysis (HCA) was performed for the Chl binding proteins LHCP AB 65 (LHCP) and Lil3 (Figure 5, (Callebaut et al., 1997)). In the HCA plot comparison, LHCP was found less hydrophobic than Lil3.1 and Lil3.2 in the region of amino acids (AA) 168-226 (Figure 5A and Fig S3). In this region,

MPEx (Snider et al., 2009) predicted two transmembrane helices for Lil3.1, but only one for LHCP. In order to correlate the conservation of AAs with the binding of Chl, conserved AAs were highlighted in the HCA plot (Figure 5B). Interestingly, AAs binding Chl in the LHCP crystal structure of *Pisum sativum* that were conserved in Lil3 revealed an overlapping pattern in the unfolded two dimensions of the HCA plot (Figure 5B, Supplemental Table 1).

Amino acids E174, N177, and L170, were found to establish one pattern for binding Chl 601, and were specified as i=E174, i+3 and i-4. For binding of Chl 604, amino acids M182, F186, and R179, established the same triangular pattern turned 180° as i=M182, i+4 and i-3. The metal binding atom of Chl 601, E174, had the same amino acids in the +/- 3/4 orientations for LHCP and for Lil3, except for the i-3 direction where the positively charged amino acid lysine was positioned in LHCP and the polar uncharged amino acid proline was positioned in Lil3. The distribution of proline in the region preceding the LHC motif and the predicted transmembrane regions of Lil3 strongly suggests that the sequence region between AA 140-170 in Lil3 generates a more flexible region before entering the predicted transmembrane helix region between amino acids 175-225. Therefore, the proline in the i-3 position is most likely marking the helix start, and is not interfering with Chl binding in position (i=E174) i+3 and i-4. For Chl 604 binding, AA M182 (i) and positions i+/-4 and i-3 are conserved, whereas position i+3 is taken by amino acid M in LHC and F in Lil3. The similar properties of AAs M and F indicate that Chl binding should not be affected in Lil3.

In LHC, molecules Chl 601 and Chl 604 are bound by the AAs in two distant Chl binding motives. This type of binding is not possible for the Lil3 sequence,

since the N-terminal region binding Chl in LHC is missing in Lil3 (Figure 6). To compensate for the lack of the N-terminal Chl binding motif in Lil3, ion pairing between E and R on opposing monomers and formation of a Lil3 dimer would create a binding pocket for Chl equivalent to the Lhcp structure (Figure 6).

Discussion

Binding of Chl to Lil3

The LHC II from Spinach (*Spinacia oleracea)* and Pea (*Pisum sativum)* has been crystallized and resolved at 2.7 Å and 2.5 Å respectively (Liu et al., 2004, Standfuss et al., 2005). Data show that the majority of the pigment molecules are bound to amino acids in helix one and three (Kuhlbrandt et al., 1994). Helix three contains the LHC motif which may not only provide an effective spacing of amino acids for binding Chl, but as a result of Chl binding, the hydrophobicity of Chl increases the hydrophobicity of the motif area, which may contribute to anchor the protein in the membrane (Granvogl, 2008). A coordination of Chl604, and Chl601 to the polar amino acids E180, and N183 of the LHC motif may in addition change their capacity for interaction with other amino acids and water molecules and further reduce the polarity of the protein (Hoober and Eggink, 1999, Hoober et al., 2007).

Further, hydrophobic amino acids of the LHC motif bind to the hydrocarbon backbone of the tetrapyrrol ring system orienting and anchoring the phytol chain towards the non-polar lipid environment of the membrane (Granvogl, 2008). In Lil3, full conservation of the N-terminal part of LHC motif, conservation of many amino acids in the n+3 and n+4 vicinity of the motif, and the transmembrane localization of the motif are strong indications for a

conservation of the Chl binding properties in Lil3 (Figure 5 and 6). For LHC, E174 has been identified as the metal binding atom for binding Chl 601 to Lhcp and N177, R64 and M67 to stabilize the pigment by hydrophobic interaction (Liu et al., 2004). For Chl 604, E59 was identified as the metal binding amino acid and H62, R179 and M182 to stabilize the structure by hydrophobic interaction (Liu et al., 2004).

For reconstitution, studies between LHCP and Chl had shown that the interaction between both binding partners was influence by the presence of detergents, and lipids (Horn and Paulsen, 2002, Paulsen et al., 1990, Plumley and Schmidt, 1987, Reinsberg et al., 2000). Our MST analysis was accordingly conducted in the presence of lipids extracted from etioplast membranes and supplemented with Chl a. Data fully support an interaction of Lil3 with Chl a (Figure 4). For Chl 604 (Liu et al., 2004), all amino acids conserved for Chl binding in helix 3 of LHCP remain structurally competent for binding of Chl in helix 1 in Lil3 (Figure 5). Recently, the one-helix LHC-like protein HliD was described to bind Chl a and β-carotene in *Synechocystis sp*. PCC 6803 (Staleva et al., 2015). Previous findings suggested that HliD binds Chl a as a dimer (Staleva et al., 2015). In our study, the HliD “homolog” Lil3 dimerizes and interacts with chlorophyll which is in agreement and extends the finding of Staleva et al 2015 (Fig 1, 2, 3 and 4) (Staleva et al., 2015). Split ubiquitin, SPR and MST analysis show that Lil3 is forming homo- and hetero-dimers, and the K_D values indicate that hetero-dimerization is favored (Figure 1, 2 and 3) The high affinity for hetero-dimerization of Lil3, K_D 25 ± 7,8 nM, strongly suggests that a dimerization event takes place prior to Chl a binding to the dimer (Fig 4, 5 and 6). Lil3 does not have the LHC motif in helix 1 of

LHCP where binding to Chl604 and Chl601 is achieved between helix 1 and 3 by two R/E amino acid pairs R70/E180 and E65/R185.

It is therefore proposed that ion pairing between the glutamic acid residues of each monomer with the arginines of the opposing monomer could substitute for the missing sequence conservation of LHC helix 1 in Lil3 and provide the two R/E pairs required for Chl binding in Lhcp (Figure 6). The sequence positions of the R/E pairs relative to the membrane plane and dimerization capability of the Lil3 protein therefore could reflect the core motif for evolution of the Chl binding capabilities around the LHC motif.

Acknowledgements

The study received funding from the Norwegian Forskingsrådet, grant number 192436, for DG and DS and grant number 197119, for AMJ, and CF. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

ANDERSSON, U., HEDDAD, M. & ADAMSKA, I. 2003. Light stress-induced one-helix protein of the chlorophyll a/b-binding family associated with photosystem I. *Plant Physiology,* 132**,** 811-20.

BRYAN, S. J., BURROUGHS, N. J., SHEVELA, D., YU, J., RUPPRECHT, E., LIU, L.-N., MASTROIANNI, G., XUE, Q., LLORENTE-GARCIA, I., LEAKE, M. C., EICHACKER, L. A., SCHNEIDER, D., NIXON, P. J. & MULLINEAUX, C. W. 2014. Localisation and interactions of the Vipp1 protein in cyanobacteria. *Molecular Microbiology,* 94**,** 1179-1195.

CALLEBAUT, I., LABESSE, G., DURAND, P., POUPON, A., CANARD, L., CHOMILIER, J., HENRISSAT, B. & MORNON, J. P. 1997. Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cellular and Molecular Life Sciences,* 53**,** 621-645.

CRAIGON, D. J., JAMES, N., OKYERE, J., HIGGINS, J., JOTHAM, J. & MAY, S. 2004. NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. *Nucleic Acids Research,* 32**,** D575-D577.

EGGINK, L. L. & HOOBER, J. K. 2000. Chlorophyll binding to peptide maquettes containing a retention motif. *J Biol Chem,* 275**,** 9087-90.

EICHACKER, L. A., MULLER, B. & HELFRICH, M. 1996. Stabilization of the chlorophyll binding apoproteins, P700, CP47, CP43, D2, and D1, by synthesis of Zn-pheophytin a in intact etioplasts from barley. *Febs Letters,* 395**,** 251-256.

ENGELKEN, J., BRINKMANN, H. & ADAMSKA, I. 2010. Taxonomic distribution and origins of the extended LHC (light-harvesting complex) antenna protein superfamily. *Bmc Evolutionary Biology,* 10**,** 233.

GABORIAUD, C., BISSERY, V., BENCHETRIT, T. & MORNON, J. P. 1987. Hydrophobic cluster-analysis - An efficient new way to compare and analyze amino-acid-sequences. *Febs Letters,* 224**,** 149-155.

GORZA, F. D., PEDRO, G. C., TRESCHER, T. F., DA SILVA, R. J., SILVA, J. R. & DE SOUZA, N. C. 2014. Morphological analysis and interaction of chlorophyll and BSA. *Biomed Res Int,* 2014**,** 872701.

GRANVOGL, B. 2008. *Entwicklung neuer Methoden zur massenspektrometrischen Charakterisierung von Membranproteinen.* des Doktorgrades der Naturwissenschaften, der Ludwig-Maximilians-Universität München.

GRANVOGL, B., GRUBER, P. & EICHACKER, L. A. 2007. Standardisation of rapid in-gel digestion by mass spectrometry. *Proteomics,* 7**,** 642-654.

GREEN, R. R. & PICHERSKY, E. 1994. Hypothesis for the evolution of three-helix Chl a/b and Chl a/c light-harvesting antenna proteins from two-helix and four-helix ancestors. *Photosynth Res,* 39**,** 149-62.

HEDDAD, M. & ADAMSKA, I. 2000. Light stress-regulated two-helix proteins in Arabidopsis thaliana related to the chlorophyll a/b-binding gene family. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 3741-6.

HOOBER, J. K. & EGGINK, L. L. 1999. Assembly of light-harvesting complex II and biogenesis of thylakoid membranes in chloroplasts. *Photosynthesis Research,* 61**,** 197-215.

HOOBER, J. K., EGGINK, L. L. & CHEN, M. 2007. Chlorophylls, ligands and assembly of light-harvesting complexes in chloroplasts. *Photosynthesis Research,* 94**,** 387-400.

HORN, R. & PAULSEN, H. 2002. Folding in vitro of light-harvesting chlorophyll a/b protein is coupled with pigment binding. *J Mol Biol,* 318**,** 547-56.

JANSSON, S. 1994. The light-harvesting chlorophyll a/b-binding proteins. *Biochim Biophys Acta,* 1184**,** 1-19.

JANSSON, S. 1999. A guide to the Lhc genes and their relatives in Arabidopsis. *Trends in Plant Science,* 4**,** 236-240.

KLEIN, R. R., MASON, H. S., GAMBLE, P. E. & MULLET, J. E. 1987. Regulation of Chloroplast-Encoded Chlorophyll-Binding Protein Translation. *Journal of Cellular Biochemistry***,** 84-84.

KUHLBRANDT, W., WANG, D. N. & FUJIYOSHI, Y. 1994. Atomic model of plant light-harvesting complex by electron crystallography. *Nature,* 367**,** 614-621.

LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227**,** 680-5.

LIU, Z. F., YAN, H. C., WANG, K. B., KUANG, T. Y., ZHANG, J. P., GUI, L. L., AN, X. M. & CHANG, W. R. 2004. Crystal structure of spinach major light-harvesting complex at 2.72 angstrom resolution. *Nature,* 428**,** 287-292.

MORK-JANSSON, A., BUE, A. K., GARGANO, D., FURNES, C., REISINGER, V., ARNOLD, J., KMIEC, K. & EICHACKER, L. A. 2015. Lil3 Assembles with Proteins Regulating Chlorophyll Synthesis in Barley. *PLoS One,* 10**,** e0133145.

MULLER, B. & EICHACKER, L. A. 1999. Assembly of the D1 precursor in monomeric photosystem II reaction center precomplexes precedes chlorophyll a-triggered accumulation of reaction center II in barley etioplasts. *Plant Cell,* 11**,** 2365-2377.

PAULSEN, H., RUMLER, U. & RUDIGER, W. 1990. Reconstitution of Pigment-Containing Complexes From Light-Harvesting Chlorophyll-A/B-Binding Protein Overexpressed in Escherichia-coli. *Planta,* 181**,** 204-211.

PLUMLEY, F. G. & SCHMIDT, G. W. 1987. Reconstitution of chlorophyll a/b light-harvesting complexes: Xanthophyll-dependent assembly and energy transfer. *Proc Natl Acad Sci U S A,* 84**,** 146-50.

PORRA, R. J., THOMPSON, W. A. & KRIEDEMANN, P. E. 1989. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* 975**,** 384-394.

REINSBERG, D., BOOTH, P. J., JEGERSCHOLD, C., KHOO, B. J. & PAULSEN, H. 2000. Folding, assembly, and stability of the major light-harvesting complex of higher plants, LHCII, in the presence of native lipids. *Biochemistry,* 39**,** 14305-13.

REISINGER, V., PLOSCHER, M. & EICHACKER, L. A. 2008. Lil3 assembles as chlorophyll-binding protein complex during deetiolation. *Febs Letters,* 582**,** 1547-1551.

SNIDER, C., JAYASINGHE, S., HRISTOVA, K. & WHITE, S. H. 2009. MPEx: a tool for exploring membrane proteins. *Protein science : a publication of the Protein Society,* 18**,** 2624-8.

STALEVA, H., KOMENDA, J., SHUKLA, M. K., SLOUF, V., KANA, R., POLIVKA, T. & SOBOTKA, R. 2015. Mechanism of photoprotection in the cyanobacterial ancestor of plant antenna proteins. *Nat Chem Biol*.

STANDFUSS, R., VAN SCHELTINGA, A. C. T., LAMBORGHINI, M. & KUHLBRANDT, W. 2005. Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5A resolution. *Embo Journal,* 24**,** 919-928.

TAKAHASHI, K., TAKABAYASHI, A., TANAKA, A. & TANAKA, R. 2013. Functional analysis of light-harvesting-like protein 3 (LIL3) and its light-harvesting chlorophyll-binding motif in Arabidopsis. *J Biol Chem***,** 987-999.

TANAKA, R., OSTER, U., KRUSE, E., RUDIGER, W. & GRIMM, B. 1999. Reduced activity of geranylgeranyl reductase leads to loss of chlorophyll and tocopherol and to partially geranylgeranylated chlorophyll in transgenic tobacco plants expressing antisense RNA for geranylgeranyl reductase. *Plant Physiology,* 120**,** 695-704.

TANAKA, R., ROTHBART, M., OKA, S., TAKABAYASHI, A., TAKAHASHI, K., SHIBATA, M., MYOUGA, F., MOTOHASHI, R., SHINOZAKI, K., GRIMM, B. & TANAKA, A. 2010. LIL3, a light-harvesting-like protein, plays an essential role in chlorophyll and tocopherol biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 16721-16725.

TOWBIN, H., STAEHELIN, T. & GORDON, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A,* 76**,** 4350-4.

TOWBIN, H., STAEHELIN, T. & GORDON, J. 1992. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. *Biotechnology,* 24**,** 145-9.

Legends

Figure 1. **Interaction of Lil3:1 and Lil3:2 in yeast.**

NMY51cells were co-transformed with *Lil3:1* and *Lil3:2* cloned into the prey (PRN) vector and *Lil3:1* or *Lil3:2* cloned into the bait (BTC) vector. The transformants were grown on SD medium lacking tryptophane (W) and leucine (L) with (-WL), positive control, or without histidine and adenine (-WLAH), negative control at 28 °C for 2 days. Serial dilutions of yeast strains were made to evaluate the specificity of the interaction. Growth under conditions suppressing growth was observed for co-expression of Lil3.1 and Lil3.2, Lil3.1 and Lil3.1 and Lil3.2 and Lil3.2.

Figure 2. Dimerization analysis of Lil3:1 and Lil3:2 using surface plasmon resonance.

Recombinant Lil3.1 and Lil3.2 of *A. thaliana* were subjected to surface plasmon resonance analysis (Biacore®, coloured line) determination of the interaction analysis (black line) and determination of the dissociation constant (K_D, value shown). Single channel analysis was performed and resulted in K_D values of 2,8 x 10^{-5} M (A) and 2,2 x 10^{-6} M (B) for the homodimerization of Lil3.1 and Lil3.2, respectively and a K_D of 9,5 x 10^{-7} M (C) for the heterodimer interaction study.

Figure 3. Thermophoretic quantification of Lil3 dimerization.

For heterodimerization analysis, unlabeled Lil3.2 was titrated against a constant amount of fluorescently labeled Lil3.1-NT647. Analysis of the MST traces and fitting of the data gave a dissociation constant K_D value of 25 ± 7,8 nM derived from the thermophoresis induced heterodimerization (A). For

homodimerization analysis, unlabeled Lil3.2 was titrated against a constant amount of fluorescently labeled Lil3.2-NT647. Thermophoretic changes show a K_D of 431± 59 nM for homodimerization of Lil3:2 (B).

Figure 4. Quantification of Lil3 interaction with Chl a

Binding of Chl a to Lil3 (K_D = 231 ± 49 nM) was investigated by thermophoresis (MST) in the presence of etioplast pigment/lipids extracts (A). Interaction of Chl a with protein, and with buffer components were tested by exchange of Lil3 against glufibrinogen (B).

Figure 5. Hydrophobic cluster analysis of the conserved region binding chlorin in LHCP and Lil3.

The amino acid sequences of LHC AB 65, and of Lil3.1 were analyzed by hydrophobic cluster analysis (Gaboriaud et al., 1987a; Callebaut et al., 1997b). Structural comparison of sequence conservation and hydrophobic cluster arrangements in both proteins was achieved by replacement of the α-helical net duplicate of LHCP (LHC) by the corresponding sequence of Lil3 (Lil3) (A and B). Black outline highlights the hydrophobic clusters (A and B). The sequence motifs and conserved Chl binding amino acids in the combined HCA plots are highlighted (B). The LHC sequence motif (LHC-motif) in LHCP (dashed line, B, upper HCA plot), and Lil3 Chl-motives (motif 1, i-4, i+3; motif 2, i-3, i+4) are outlined (dashed line, B, lower HCA plot). Conserved chlorin binding amino acids are outlined in red (B, upper and lower HCA plot). Amino acids in one letter code are colored or are replaced by symbols (Polar AA`s (N, D, E, Q), red; AA`s charged positively at neutral pH (K, R), blue; hydrophobic AA`s (F, I, V, L, W, M, Y), green). Amino acids replaced by

symbols are P, red star; T, empty square; S, square with internal black square; G, black diamond. Black outline highlights the hydrophobic clusters (A and B).

Figure 6. Generation of two Chl binding sites by E-R ion pairing in Lil3 dimers.

The predicted transmembrane helix region 1 of Lil3.1 is displayed as HCA plot. The dimeric state of Lil3 complexes is depicted by two helices plotted as mirror images to span the thylakoid membrane from the stroma to the lumen in the direction from N- to C- terminus (N1 to C1 and N2 to C2). The graphical sketch in the center explains the proposal that stabilization of the Lil3 dimer by E-R ion-pairing establishes the two LHC-Chl binding sites for Chl601, and Chl604 that link helix 1 and 3 in LHCP. Ion pairing between the glutamic acid residue (E) of each Lil3 monomer with the arginine (R) of the opposing Lil3 monomer results in two Chl binding sites on the opposing sides of the transmembrane helices 1 in the Lil3-dimer.

Figure 1

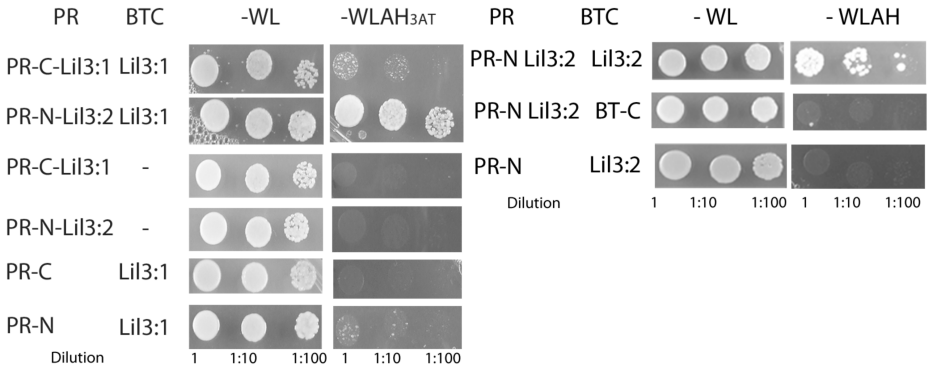


Figure 1. **Interaction of Lil3:1 and Lil3:2 in yeast.**

NMY51cells were co-transformed with *Lil3:1* and *Lil3:2* cloned into the prey (PRN) vector and *Lil3:1* or *Lil3:2* cloned into the bait (BTC) vector. The transformants were grown on SD medium lacking tryptophane (W) and leucine (L) with (-WL), positive control, or without histidine and adenine (-WLAH), negative control at 28 °C for 2 days. Serial dilutions of yeast strains were made to evaluate the specificity of the interaction. Growth under conditions suppressing growth was observed for co-expression of Lil3.1 and Lil3.2, Lil3.1 and Lil3.1 and Lil3.2 and Lil3.2.

Figure 2.

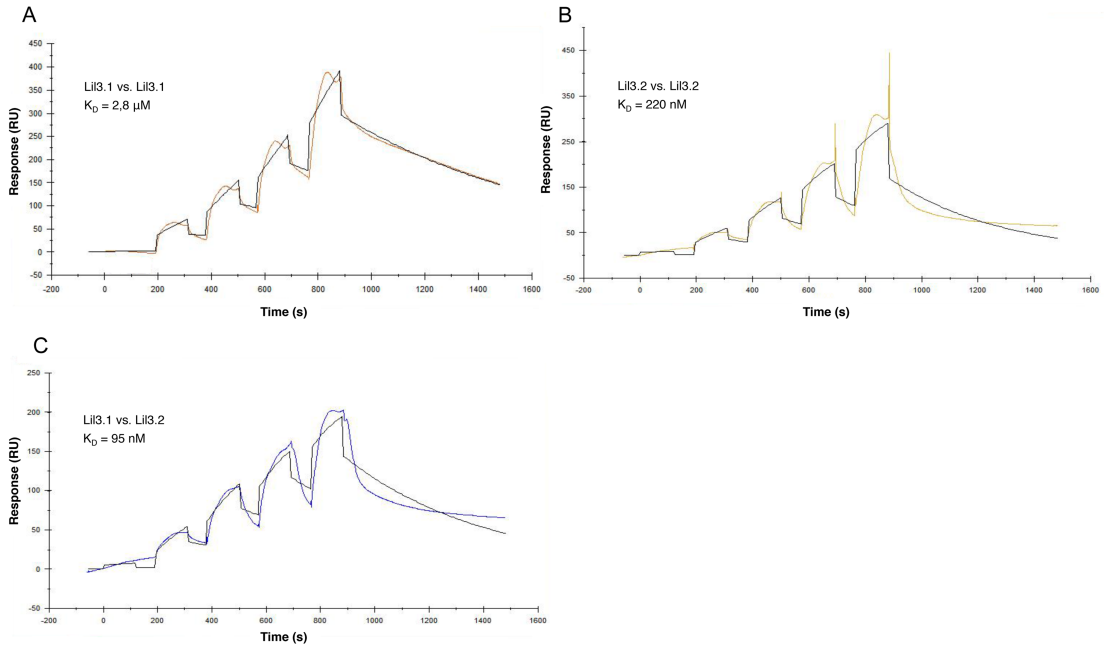


Figure 2. Dimerization analysis of Lil3:1 and Lil3:2 using surface plasmon resonance.

Recombinant Lil3.1 and Lil3.2 of *A. thaliana* were subjected to surface plasmon resonance analysis (Biacore®, coloured line) determination of the interaction analysis (black line) and determination of the dissociation constant (K_D, value shown). Single channel analysis was performed and resulted in K_D values of 2,8 x 10^{-5} M (A) and 2,2 x 10^{-6} M (B) for the homodimerization of Lil3.1 and Lil3.2, respectively and a K_D of 9,5 x 10^{-7} M (C) for the heterodimer interaction study.

Figure 3

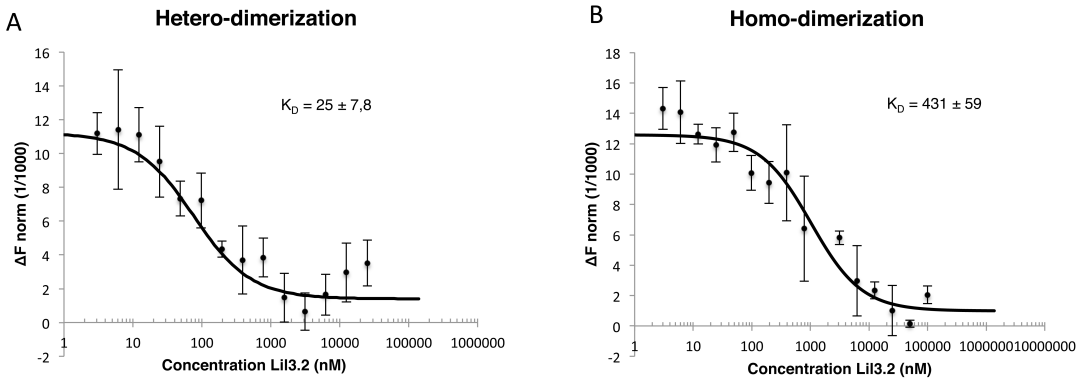


Figure 3. Thermophoretic quantification of Lil3 dimerization.

For heterodimerization analysis, unlabeled Lil3.2 was titrated against a constant amount of fluorescently labeled Lil3.1-NT647. Analysis of the MST traces and fitting of the data gave a dissociation constant K_D value of 25 ± 7,8 nM derived from the thermophoresis induced heterodimerization (A). For homodimerization analysis, unlabeled Lil3.2 was titrated against a constant amount of fluorescently labeled Lil3.2-NT647. Thermophoretic changes show a K_D of 431± 59 nM for homodimerization of Lil3:2 (B).

Figure 4

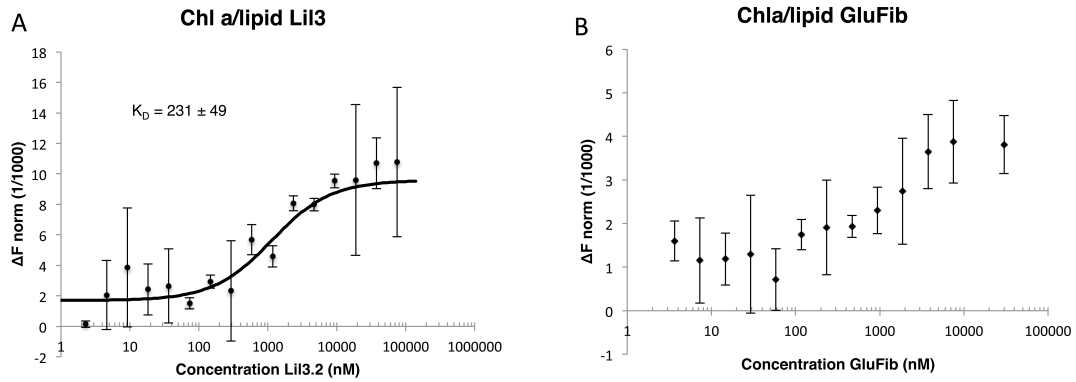


Figure 4. Quantification of Lil3 interaction with Chl a

Binding of Chl a to Lil3 (K_D = 231 ± 49 nM) was investigated by thermophoresis (MST) in the presence of etioplast pigment/lipids extracts (A). Interaction of Chl a with protein, and with buffer components were tested by exchange of Lil3 against glufibrinogen (B).

Figure 5

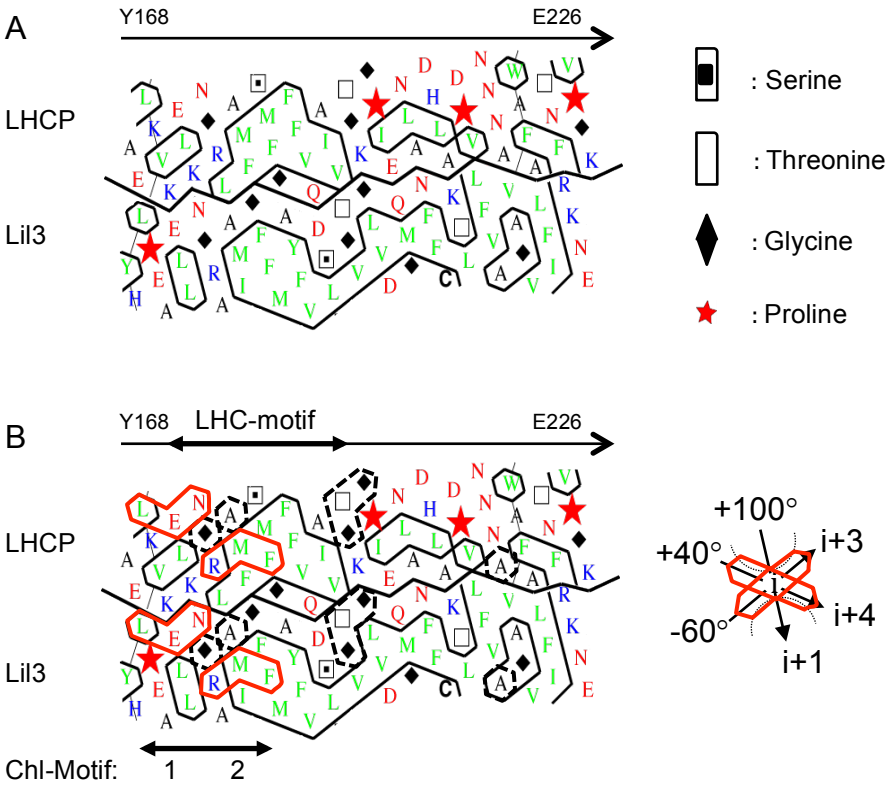


Figure 5. Hydrophobic cluster analysis of the conserved region binding chlorin in LHCP and Lil3.

The amino acid sequences of LHC AB 65, and of Lil3.1 were analyzed by hydrophobic cluster analysis (Gaboriaud et al., 1987a; Callebaut et al., 1997b). Structural comparison of sequence conservation and hydrophobic cluster arrangements in both proteins was achieved by replacement of the α-helical net duplicate of LHCP (LHC) by the corresponding sequence of Lil3 (Lil3) (A and B). Black outline highlights the hydrophobic clusters (A and B). The sequence motifs and conserved Chl binding amino acids in the combined HCA plots are highlighted (B). The LHC sequence motif (LHC-motif) in LHCP (dashed line, B, upper HCA plot), and Lil3 Chl-motives (motif 1, i-4, i+3; motif 2, i-3, i+4) are outlined (dashed line, B, lower HCA plot). Conserved chlorin

binding amino acids are outlined in red (B, upper and lower HCA plot). Amino acids in one letter code are colored or are replaced by symbols (Polar AA`s (N, D, E, Q), red; AA`s charged positively at neutral pH (K, R), blue; hydrophobic AA`s (F, I, V, L, W, M, Y), green). Amino acids replaced by symbols are P, red star; T, empty square; S, square with internal black square; G, black diamond. Black outline highlights the hydrophobic clusters (A and B).

Figure 6

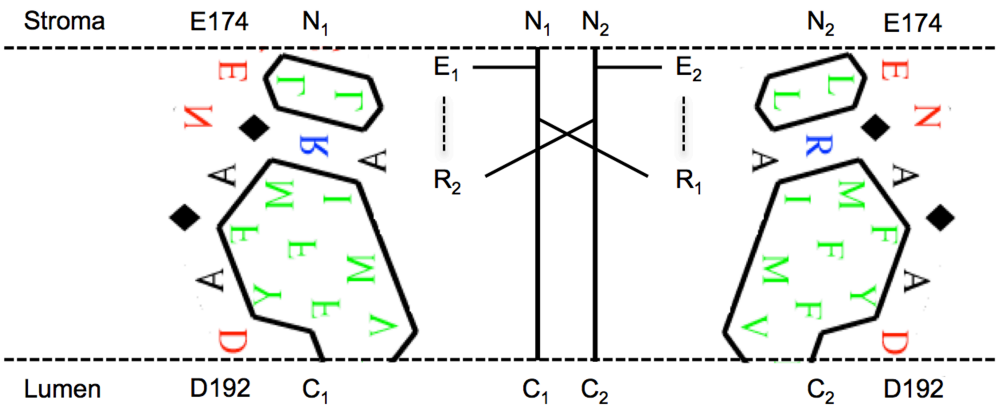


Figure 6. Generation of two Chl binding sites by E-R ion pairing in Lil3 dimers.

The predicted transmembrane helix region 1 of Lil3.1 is displayed as HCA plot. The dimeric state of Lil3 complexes is depicted by two helices plotted as mirror images to span the thylakoid membrane from the stroma to the lumen in the direction from N- to C- terminus (N1 to C1 and N2 to C2). The graphical sketch in the center explains the proposal that stabilization of the Lil3 dimer by E-R ion-pairing establishes the two LHC-Chl binding sites for Chl601, and Chl604 that link helix 1 and 3 in LHCP. Ion pairing between the glutamic acid residue (E) of each Lil3 monomer with the arginine (R) of the opposing Lil3 monomer results in two Chl binding sites on the opposing sides of the transmembrane helices 1 in the Lil3-dimer.

Supplementary material

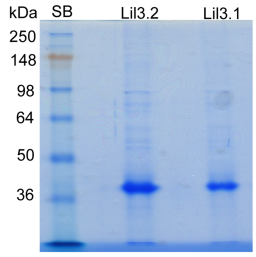
Table S1.

Conservation and capacity of amino acids for binding chlorophyll in Pea LHCP (2BHW) and Lil3.

LHC specific Chl number (60X)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
#<H>AA Chl interaction in LHC	19	0	9	19	2	0	2	0	0	3	3	2	0	0
# AA binding Chl in LHC	13	3	8	13	3	8	6	5	9	9	13	11	6	8
Potential AA binding Chl in Lil3	**6**	**1**	**2**	**4**	**1**	**0**	**1**	**0**	**0**	**1**	**2**	**2**	**0**	**0**
# Identical AA	4	1	1	3	0	0	1	0	0	0	2	2	0	0
# Conserved substitutions	1	0	1	1	1	0	0	0	0	1	0	1	0	0
# Semiconserved substitutions	1	0	0	0	0	0	0	0	0	0	0	0	0	0
AA binding Chl in LHC, conserved in Lil3	F196	N218	F224	R220	F230		N218			F230	R176	R176		
	L201		F227	M223							F196	F227		
	A202			F224								F230		
	L211			F227										
	E215													
	N218													
HCA, conservation of Chl motif in Lil3	+			+										
	L211			R220										
	E215			M223										
	N218			F227										

Figure S1

A



B

Identified protein	Accession number[a]	Detected m/z signals and charge state	*De novo* sequenced protein fragment	Total sequence coverage, %
Lil3	Q9LTB7	602.66 $[M+3H]^{3+}$ 821.37 $[M+4H]^{4+}$ 860.39 $[M+4H]^{4+}$	DGKTDWDSVIVSEAKR KDGKTDSVIVSEAKR GTWDLKQFEKDGKTDWDSVIVSE	10.2

[a] UniProtKB/SwissProt database entry number from *Arabidopsis thaliana*

Figure S1. Characterization of recombinant Lil3 protein by SDS-PAGE and by mass spectrometry.

(A) Lil3_pET151d plasmids (Lil3.1 – AT4G17600 and Lil3.2 – AT5G47110) were constructed and expressed in *E. coli*, then purified by affinity chromatography and separated by 12 % SDS-PAGE. (B) Mass spectrometric identification of Lil3 protein after SDS-PAGE. *De novo* sequence analysis was performed as described earlier (Bryan et al., 2014). Sequence coverage was calculated as the ratio of the number of amino acids in identified peptides divided by the number of amino acids in the complete proteinsequence.

Figure S2

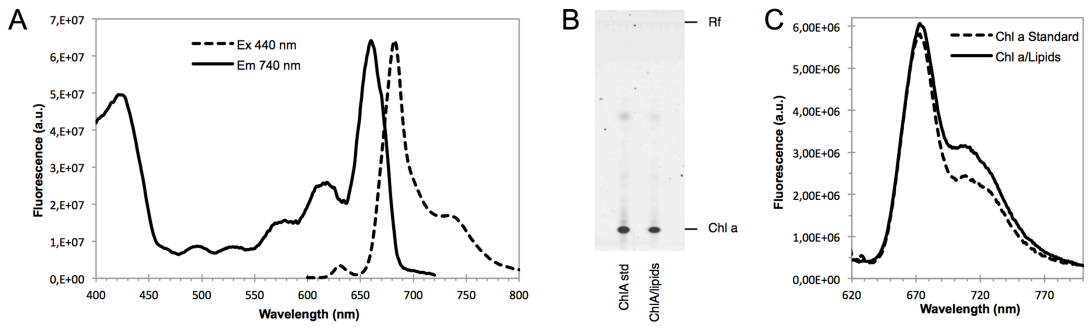


Figure S2. Characterization of the etioplast extract by spectroscopy and HPTLC.

Etioplasts were isolated from barley seedlings and proteins were precipitated by 80 % (v/v) acetone in darkness and the organic phase clarified by centrifugation. Excitation (E_m 740 nm) and emission spectra (E_x 440 nm) were recorded from the supernatant (A). An aliquot of the diluted supernatant was separated on a HPTLC plate and compared to the mobility of a Chl a standard (B). Chl a spots were picked from the HPTLC plate and emission spectra were recorded upon excitation at 440 nm (C).

Figure S3

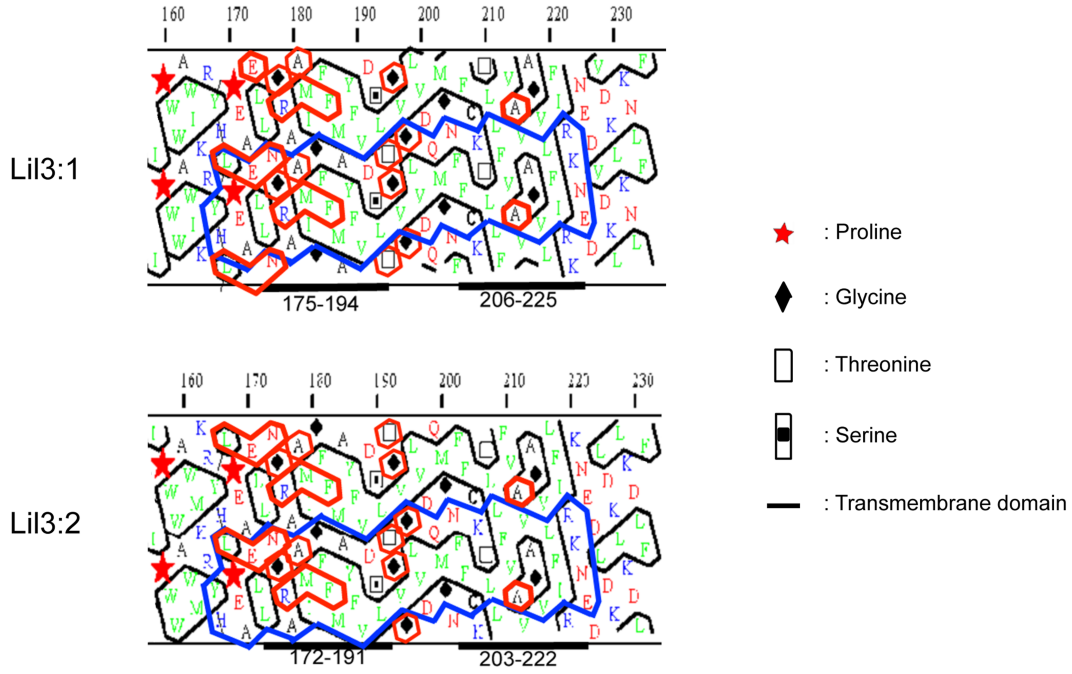


Figure S3. HCA analysis of the LHC motif and Chl binding region in Lil3.1 and Lil3.2.

The amino acid sequences of Lil3.1 and of Lil3.2 were analyzed by hydrophobic cluster analysis (Gaboriaud et al., 1987a; Callebaut et al., 1997b). Conserved amino acids and motives associated with chlorophyll binding are outlined in red according to crystal structures of LHCP (1RWT, 2BHW). Black outline highlights the hydrophobic clusters. The LHC-motif is outlined in blue. Amino acids in one letter code are colored or are replaced by symbols (Polar AA`s (N, D, E, Q), red; AA`s charged positively at neutral pH (K, R), blue; hydrophobic AA`s (F, I, V, L, W, M, Y), green). Amino acids replaced by symbols are P, red star; T, empty square; S, square with smaller black square; G, black diamond. Black bars under the HCA plot with amino acid numbers for the start and stop of the region refer to the trans-membrane regions predicted by MPEx http://blanco.biomol.uci.edu/mpex/.