Functions and Evolution of Lhcb Isoforms Composing LHCII, the Major Light Harvesting Complex of Photosystem II of Green Eukaryotic Organisms

Aurelie Crepin, Stefano Caffarri

Functions and evolution of Lhcb isoforms composing LHCII, the major light harvesting complex of Photosystem II of green eukaryotic organisms.

Running title: LHCII isoforms in green organisms

Aurélie Crepin[1,*], Stefano Caffarri[1,*]

[1]*Aix Marseille Université, CEA, CNRS, Biosciences and Biotechnologies Institute of Aix-Marseille (BIAM), Laboratoire de Génétique et Biophysique des Plantes, 13009 Marseille, France.*

*Address correspondence to these authors at the Laboratoire de Génétique et de Biophysique des Plantes (LGBP), Aix-Marseille Université, Faculté des Sciences de Luminy, 163 Avenue de Luminy, 13009, Marseille, France ; Tel: +33-4-91829562 ; Fax: +33-4-91829566,

E-mail: stefano.caffarri@univ-amu.fr or aurelie.crepin@univ-amu.fr

Abstract: Oxygenic photosynthesis provides energy and oxygen for almost all forms of life on earth. This process is based on the energy of photons, which is used to split water and use its electrons to reduce carbon atoms to create organic molecules and thus fix the light energy into a chemical form. Two photosystems working in series are involved in light harvesting and conversion. Both are multi-protein supercomplexes composed of a core complex, where the photochemical reaction takes place, and an antenna system involved in light harvesting. In plants and green algae, the antenna of photosystem II, the photosynthetic complex involved in water splitting, comprises the Light Harvesting Complex II (LHCII) trimers, the most abundant membrane protein on earth. LHCII is composed of highly conserved Lhcb isoforms and all green organisms count a high number of Lhcb. In vascular plants they are classified in three distinct subclasses, Lhcb1, 2 and 3, while in algae and non-vascular plants, these isoforms are less differentiated and called Lhcbm proteins. In this review, we compare the LHCII proteins of different organisms, from green algae to angiosperms, and discuss the role of the modifications that occurred through evolution. We highlight the various functions of the different isoforms in photosynthesis, ranging from light harvesting, a common role to all these proteins, to regulations of photosynthesis that rely on specific isoforms.

Keywords: Photosynthesis, Light harvesting complexes in plants and green algae, LHCII isoforms, Lhcb3, State transitions

1. INTRODUCTION

Living organisms require energy to sustain the whole metabolism. All this energy comes from the sun and is made available for life by photosynthetic organisms through its transformation into a chemical form. The very first and indispensable event necessary for the photosynthetic process is the harvesting of the light. This process is performed principally by the so called "antenna systems", which are pigment-protein complexes peripherally located around photosystem II (PSII) or photosystem I (PSI), the multiprotein complexes performing the first reactions necessary to transform photons into chemical energy. While the photosystem cores (the sites where the catalytic activity is performed) are well conserved in all photosynthetic organisms despite billions of years of separate evolution (see [1-5]), the antenna systems have evolved very differently in the various photosynthetic clades (see [1]). In eukaryotic green organisms (plants and green algae), the antenna system is composed by the Light harvesting complexes (Lhc proteins), which belong to a large family of homologous proteins [6, 7].
All Lhcs have a similar structure: they are membrane proteins with three transmembrane α-helices located in the thylakoid membranes inside the chloroplast [4, 5, 8-11]. To be able to harvest light, these proteins bind an extraordinary number of pigments as cofactors: 11-14 chlorophylls (Chls) *a* or *b* and 2-4 carotenoids (Cars), depending on the specific Lhc [4, 5, 11, 12]. Lhcs are present in a monomeric form or organized as trimers. In particular, trimeric complexes, called LHCII (the major light harvesting complexes of photosystem II) are the most abundant Lhc proteins, being able to coordinate up to 40% of the total Chl present in the chloroplast. They also represent the most abundant membrane protein on earth.
LHCII are heterotrimers composed by highly similar isoforms encoded by an elevated number of genes in all green organisms. The similarity is extremely conserved even between very distant green organisms, such as vascular plants and unicellular green algae, which are separate by about 1 billion years of evolution [13]. It is likely that all sequences evolved from a single ancestral LHCII gene that differentiated into several isoforms in the various green clades. Despite a remarkable homology, LHCII isoforms have also evolved specific sequence differences that allow, at least in plants, separating them in subclasses (Lhcb1, 2 and 3), each one represented by one or several proteins. When these specific sequence markers are not evident (as in green algae and non-vascular plants as the moss *Physcomitrella patens*), LHCII isoform are named Lhcbmx (where "x" is a progressive number).

In plants as in green algae, LHCII trimers are present in several copies per PSII. In the case of plants, up to four LHCII per monomeric PSII core has been estimated [14, 15]. The LHCII trimers are located at various positions around the PSII core (Fig. 1A) [5, 11, 16, 17]: the most strongly bound LHCII trimer, called S-LHCII, is located between the monomeric antenna CP26 (Lhcb5) and CP29 (Lhcb4) and in direct contact with the core antenna CP43 (PsbC); the moderately bound M-LHCII is in contact with S-LHCII and the monomeric antennas CP29 and CP24 (Lhcb6). Note that CP24 is absent in green algae and at this position next to M-LHCII a trimer called N-LHCII has been identified in *Chlamydomonas reinhardtii* [18, 19]. Additional loosely bound L-LHCII trimers, which easily detach from purified PSII supercomplexes, cannot be easily located. However, in spinach, a position of a L-LHCII has been proposed in contact with the core complex near to CP26 [20]. Very recent work on PSII megacomplexes, where the binding of loosely bound LHCII is retained between two adjacent PSII, also provided important information about the localization of these additional L-trimers (Fig. 1A) [21].

It should be noted that the distinct type of LHCII have a different isoform content: it has been clearly demonstrated that plant S-LHCII contains only Lhcb1 and 2 isoforms [22], but not Lhcb3, which is a specific isoform of M-LHCII, which is also enriched in Lhcb1 [17, 20, 23, 24]. Proteomic analysis on different trimers of Arabidopsis also showed that the sub-class isoforms are also differently represented in the S, M and L-LHCII trimers [25]. From these proteomic data, it can be estimated that the molar ratio of Lhcb1:Lhcb2:Lhcb3 in Arabidopsis is about 7:4:1, which is similar to the 8:3:1 estimated for barley [14]. It is however important to highlight that this ratio depends also on growth conditions, since LHCII amount is regulated by light intensity. Interestingly, the subunits proportions seems also be correlated to the number of gene copies for each isoform: in all genomes studied so far, the number of Lhcb1 isoforms has always been found higher than the number of Lhcb2 isoforms, and Lhcb3 is usually present just in one or two copies. In *Arabidopsis thaliana*, for instance, there are 5 isoforms for Lhcb1, 3 for Lhcb2, and only one for Lhcb3.

Recent findings by using biochemical, genetic and spectroscopic approaches have provided important information about the role of the different LHCII isoforms, indicating some reason for their remarkably high number.
In this review we summarize the present knowledge about LHCII proteins, with a particular focus on the role of specific isoforms and their evolution.

2. LHCII ISOFORMS IN GREEN ORGANISMS

2.1 LHCII evolution over a billion years

LHC sequences seem to have appeared more than a billion years ago, after the divergence of glaucophytes from the red and green algae lines [7, 26, 27]. It is considered that they evolved from bacterial high light inducible proteins, involved in energy dissipation [26, 28], which may have functioned as a dimer binding chlorophylls and carotenoids [29]. Alternatively, LHC might derive from the two helix stress-enhanced proteins also present in glaucophytes [27].
The apparition of these LHC sequences and their specialization in light harvesting gave rise to the successful and lasting transformation of photosystems' antennae, and thus LHCII trimers of similar organization and structure can be found from green algae to flowering plants. For a better comparison, we retrieved Lhcb sequences from representative organisms over the green line whose genomes have now been sequenced: a green alga (*Chlamydomonas reinhardtii*), a charophyte (*Klebsormidium flaccidum*), two bryophytes (the liverwort *Marchantia polymorpha* and the moss *Physcomitrella patens*), a lycophyte (*Selaginella moellendorffii*), two gymnosperms (*Gingko biloba* and the more recent *Picea glauca*) and two angiosperms (a monocot, *Zea mays*, and a dicot, the well-studied *Arabidopsis thaliana*). The full list is provided in Supplemental Table 1, including an incomplete Lhcb3 sequence from *Gingko biloba* partially completed from [30].
As indicated before, neither *C. reinhardtii* [31] nor *K. flaccidum* present LHCII sequences that are clearly assignable to the Lhcb1, 2 or 3 sub-classes identified in vascular plants. They are therefore labeled "Lhcbm". This is also the case for the bryophytes, with a notable exception: if most of their Lhcb sequences cannot be assigned to the Lhcb1 or Lhcb2 classes, they do contain a clear Lhcb3 sequence, showing that this isoform was the first to differentiate and, as suggested before [32], a marker of land plant evolution. Lhcb1 and Lhcb2 only differentiated from lycophytes (here from *S. moellendorffii*), one of the oldest lineages of vascular plants. All three

Lhcb types are thus present in almost all vascular plants studied so far, with the notable exception of Lhcb3 in a particular clade of the gymnosperms, where the gene has been lost [33].
The LHCII sequences of *Chlamydomonas,* even though they are not as clearly differentiated as in vascular plants, have been classified in four types depending on sequence similarities [34]: type I (Lhcbm3, Lhcbm4, Lhcbm6, Lhcbm8 and Lhcbm9), type II (Lhcbm5), type III (Lhcbm2 and Lhcbm7), and type IV (Lhcbm1). Despite similar spectroscopic properties [35], reverse genetics and functional studies have determined some specific role for different isoforms: Lhcbm9 is considered to have a role in energy quenching during stress, especially sulfur or nitrogen deprivation, which induce increase of lhcbm9 gene expression [36, 37]. Lhcbm2/7, Lhcbm5, and Lhcbm4/5/6 were proposed to be involved in state transitions [38-40], a fast mechanism of energy balancing between the two photosystems that will be discussed later in this review; in addition, depletion of Lhcbm4/5/6 causes a reduction in Non-Photochemical Quenching (NPQ) and photoprotection [40]. Finally, the absence of Lhcbm1 does not affect state transitions, but causes a reduction of NPQ instead [41].
As in *Chlamydomonas*, there are now evidences that plant Lhcb1-2-3 isoforms also have specific roles in photosynthesis and its regulation. However, the reason for which each class is represented by several sub-isoforms in every plant studied is still under discussion, especially as most of the gene duplications appear to be recent events, leading to very different numbers of copies in evolutionary close organisms: *Arabidopsis*, for instance, presents 5 genes for Lhcb1, while another angiosperm, *Zea mays* contains 9 genes for this isoform and *Hordeum vulgare* at least 14 [24]. All these sequences segregate on organism-specific branches in a phylogenetic tree (Fig. 2) [42], showing that the numerous duplications occurred after the divergence of monocots and eudicots about 140 Myr ago [43]. A reason sometimes mentioned to explain this high number of gene copies is the need for a very high rate of transcription and translation, to answer for the heavy requirements in protein quantity, as well as the compensation for possible mutations and loss of function of these genes. Another explanation would be the optimization of the regulation of these genes: instead of a single gene with a limited number of regulatory motifs, these duplications allow a better integration of signals and possibly a different regulation for each of them [44, 45].
This duplication and specification process shows that, despite its high conservation through the green line, the LHCII complex is still a flexible antenna system, easily adapted in each organism in response to the different requirements in light harvesting regulation.

2.2 Sequence conservation and divergence, clues for regulation mechanisms

Given the remarkable similarity of the structure, organization and function of the Lhcb proteins, it is not surprising that all the sequences analyzed present very high sequence conservation, especially at the level of the helices, which contain most of the chlorophyll binding sites [8, 9, 24, 46]. All sequences identified so far, from algae to flower plants, also possess a characteristic motif (WYGPDR) that has been shown to be important for trimerization [47]. It should be noted however that this part of the protein does not make a direct contact between monomers (Fig. 1B); rather it stabilizes a domain (containing Chl 601, the xanthophyll in site V1 and a phosphatidylglycerol lipid) that makes interactions with the adjacent monomer. However, the conservation of this motif suggest that trimerization of these proteins was an early adaptation and an essential property of these antennas [26].
The N-terminal sequences of plant Lhcb1 and Lhcb2 are also rather well-conserved in length as well as in sequence. In Lhcbm sequences, however, the N-terminal peptide is more variable.

Various mature sequences were reported for the same proteins of *Chlamydomonas reinhardtii* (e.g. [31, 35]), and it was even proposed that some Lhcb might undergo a different processing, leading to two alternative mature proteins ([48, 49]; this is also discussed in a later section). Thus the precise sequences of mature Lhcbm proteins in *Chlamydomonas* are still unknown. On the contrary, in non-vascular plants as *Physcomitrella patens*, the mature Lhcbm sequences can be deduced based on the similarity of their N-terminus with Lhcb1-2 isoforms of vascular plants [32] (Supplementary Table 2).

It is interesting to point that the N-terminus and the stromal region of LHCII were suggested to be involved to grana formation in vascular plants: the positively charged N-terminal peptide could interact with the negatively charged stromal surface of a facing LHCII to promote membrane stacking [50]. We can observe that the positive charge of the first N-terminus residues is well conserved amongst Lhcb1-2-m isoforms. This is the same for most of the charges exposed on the stromal side: Glu38, Glu62, Glu150, Asp153, Asp169, Glu171, Lys177 (as numbered according to AtLhcb1.1 mature sequence). Interestingly, Asp19, a charged residue very conserved amongst almost all LHCII sequences, is replaced by a Asn in the five Lhcbm of type I of *Chlamydomonas*. Glu30, a residue well conserved in Lhcb1 sequences (Fig. 3) and several Lhcbm sequences, is also replaced by Ala or Ser in type I Lhcbm proteins and in Lhcbm5 of *Chlamydomonas*. These differences could have a biological significance and it cannot be excluded that they are involved in electrostatic interactions important for grana stacking, in which case their absence could be partially responsible for the loosened grana in *Chlamydomonas*. However, considering the conservation of most of the other charges, additional mechanisms are likely necessary for grana formation.

Sequence alignments also reveal isoform-specific motifs. Lhcb3 divergence, for instance, is marked by several sequence specificities, which will be discussed in section 1.3.

Lhcb1 and Lhcb2 are known to be highly similar; however already longtime ago 14 distinctive amino acids differentiating them were proposed by S. Jansson [51]. With the additional sequences analyzed here, we propose just 10 Lhcb2-specific amino acids (Fig. 3). As expected, most of the differences are present in the N-terminal domain of the proteins. The first is the presence of an Arg in Lhcb2 instead of a Lys in Lhcb1 before the phosphorylatable Thr3 (numbered according to the *Arabidopsis* Lhcb1.1 mature sequence) in almost all sequences, which seems to affect the affinity of the STN7 kinase for the different isoforms ([52] and discussion below). This phosphorylation motif is followed by the deletion of 3-4 amino acids in Lhcb2 compared to Lhcb1, making its N-terminus shorter. The second Lhcb2-specific residue is the Pro at position 8 (Fig. 3), which might be important for the N-terminal structure of Lhcb2. This effect may be associated to the presence of another difference, highly distinctive of the Lhcb1-2-3 isoforms: the residue immediately preceding the trimerization motif. This residue is a Leu in most of Lhcb3 sequences and a Ile in all Lhcb2 proteins analyzed here; however, it is a Pro in all Lhcb1 sequences, probably inducing a more rigid turn in this isoform that may directly impact its structure and specific role. The trimerization motif is then followed by another Pro in all Lhcb2 sequences, replacing a Val in Lhcb1 and inducing a further divergence in the structure. Further away in the N-terminus, a three amino acid motif clearly differentiates the sequences, though part of the Lhcb2 motif is also conserved in Lhcb3. The other divergences are found in the C/A loop. These differences are distant from the pigment binding sites (Fig. 1C), which therefore are very likely unaffected, but they might slightly modify the protein structure and its interactions with other proteins and thus the assembly of the photosystem supercomplexes.

The phosphorylatable Thr3 involved in state transitions is present in several isoforms in each organism studied here, underlying the importance of this photosynthetic regulation. The conservation and isoform specificities of this site will be discussed in section 1.4.
The Lhcb proteins also bear other phosphorylation sites, most of which are organism-specific. Some of these sites were only reported once and not confirmed in other phosphoproteomic studies (reviewed in [53]), and they will not be discussed here. However, one site (other than the Thr involved in state transitions) has been highly conserved during evolution: the Ser residue preceding the trimerization site (Fig. 3), immediately before the isoform-specific P/L/I residues (for ease of discussion, this residue will be called hereafter Ser13, based on the *A. thaliana* Lhcb1.1 mature sequence). This residue is conserved in all Lhcb1-2 sequences from land plants, but not in the Lhcb3 isoforms, as well as in some of the sequences of *Chlamydomonas* and *Klebsormidium*. However, if Ser13 was found phosphorylated in several studies (e.g. [54, 55] for *Arabidopsis thaliana*; [56] for *Selaginella moellendorffii*), only Lhcb1 isoforms were found phosphorylated. The kinase involved, which has yet to be identified, has thus a high specificity for Lhcb1, probably relying on the isoform-specific residues in close proximity of the Ser13 and/or the specific structure of its N-terminus. It was shown that the kinase is neither STN7 nor STN8 [57], the two kinases usually responsible for the better-known phosphorylation of LHCII and of other PSII proteins.
Despite the recurrent finding of phosphorylated Ser13 in phosphoproteomic studies, this phosphorylation site has yet to be studied in depth. P-Ser13 was found in *A. thaliana* cells in suspension [54] and wild-growing *S. moellendorffii* [56]; a study found that the amount of Ser13 phosphorylation decreased after a one-minute treatment to ethylene, a hormone involved in several physiological processes in plants [55]. Two works investigated LHCII phosphorylation using Phos-tag [58, 59], a small molecule able to slow down phosphorylated proteins on gels, regardless of the phosphorylation site. In both cases, Lhcb1 was present as two bands: one non-phosphorylated, and the other bearing a single phosphorylation, recognized by antibodies specifically designed against the P-Thr involved in state transitions [60]. This suggests that in the conditions necessary to induce state transitions, Ser13 is not phosphorylated or, alternatively, phosphorylations on Thr3 and Ser13 are mutually exclusive. Additional studies will thus be necessary to determine the conditions needed for Ser13 phosphorylation, as well as its potential role in regulating photosynthesis.

Concerning the light harvesting properties, the alignment of the Lhcb1-2-3-m sequences shows that the pigment binding sites are highly conserved, as confirmed by the studies on recombinant LHC of plants [24, 61] and green algae [35]. It is thus likely that the Lhcbm, Lhcb1 and Lhcb2 proteins have a highly similar pigment organization. Lhcb3, though, presents important modifications compared to the other sequences around some pigment binding site, probably responsible for its slightly different spectroscopic properties [24, 61]. These divergences and their possible consequences are discussed in section 2.3.
Other isoform-specific light harvesting properties and photosynthetic regulations could be due to their localization, which at the moment is well known only for Lhcb3 (see section 2.3). Few works have investigated this point, in part because it is generally difficult to discriminate isoforms (and sub-isoforms) by biochemical methods. The exact localization of Lhcb1-2 isoforms and sub-isoforms is thus little known. However, an investigation by mass spectrometry on the composition of S-, M- and L-LHCII in Arabidopsis [25] revealed that the M-trimer contains very little Lhcb2 and is enriched in the Lhcb1.4 isoform and poor for the Lhcb1.5

isoform, that S-LHCII is enriched in Lhcb1.1-3 isoforms (which are identical), while the loosely bound L-LHCII trimers are enriched in Lhcb1.5. In the case of *Chlamydomonas*, a mass spectrometry analysis of PSII supercomplexes [19] showed that specific isoforms are well associated to PSII (mainly Lhcbm1, Lhcbm2/7, Lhcbm3) and other Lhcbm are mainly present in the "extra LHCII" population (especially Lhcbm5).
Concerning a second important function of LHCII, that is the dissipation of the excess energy for photoprotection, it should be noted that the analysis of the sequence divergences gives little clue on the molecular mechanisms involved. Energy quenching in LHCII is activated under high light, for instance during the operation of the xanthophyll cycle, which induces the pH-dependent de-epoxydation of violaxanthin into zeaxanthin and implicates xanthophyll molecules present in Lhcb complexes. However the precise mechanism involved in the quencher formation remains to be determined and so far the binding of zeaxanthin has been proposed to favor a conformational change of LHCII by inducing new chlorophyll-chlorophyll or xanthophyll-chlorophyll interactions that would open an energy dissipation channel [62-64]. This mechanism however seems to be common to multiple, if not all Lhcb sequences, as the pigment binding properties seem shared between complexes.
Aggregation of LHCII trimers, discussed below, can also lead to the formation of excitation energy traps [65]. In *Chlamydomonas*, the Lhcbm1 isoform, which is essential for energy dissipation [41], might be involved in this phenomenon during state transitions. Nevertheless, not enough information is available on the precise mechanism involved and thus so far it is not possible to differentiate LHCII isoforms for their possible specific roles in energy quenching based on sequence comparison.
Interestingly, in vascular plants during a long term adaptation to high light inducing photosystem remodeling, the content of a specific isoform, Lhcb3, has been shown to decrease, in line with the observed decrease of M-LHCII trimers [15]. Thus, the roles in photosynthesis regulation and photosystem architecture of this specific isoform are discussed in depth in the following section.

2.3 Lhcb3, a special LHCII isoform of plant PSII

The Lhcb3 isoform is considered a marker of land plants evolution [26, 32]. It is indeed absent from green algae, but present in the moss *Physcomitrella patens* [32] and the liverwort *Marchantia polymorpha*, even though Lhcb1 and Lhcb2 are not differentiated in these species.
Lhcb3 is characterized by a shorter N-terminal sequence, which lacks the phosphorylation site necessary for state transitions (Fig. 4). It also presents several specific amino-acid substitutions, especially in a short region between helix B and helix E, in helix C and in the loop between helix C and helix A (Fig. 4 and Fig. 1C). These and other amino acid differences between Lhcb3 and Lhcb1-2 may explain some specific properties that distinguish this isoform from Lhcb1 and Lhcb2: in vitro studies have shown differences in pigment binding, as well as a slightly red-shifted fluorescence emission maximum at 683.5 nm, compared to 681 nm in Lhcb1-2 [24, 61]. Moreover and contrarily to the other isoforms, Lhcb3 is also unable to form homotrimers [24] and is thus only found as heterotrimers, mainly with Lhcb1 in the M-LHCII trimer [23, 25, 66].
The changes differentiating Lhcb3 from other Lhcb appeared progressively during evolution. Most of the modifications implying changes of amino acids properties occurred early and are already present in bryophytes, as the shorter N-terminus and the absence of the phosphorylatable threonine. These mutations are thus probably the most important ones for the differentiation of the Lhcb3 isoform and for its specific properties. It has to be noted that they are also the only Lhcb3-specific features displayed by the Lhcb3.2 and Lhcb3.3 sequences in *Marchantia*

polymorpha. Interestingly, the genes for these two proteins seem to be little expressed (if at all) to the profit of lhcb3.1 (which codes for a more standard Lhcb3 isoform), as estimated by the number of sequences found in the EST databases. The different spectroscopic properties of Lhcb3, likely due to other modifications downstream the N-terminus, thus seem to be important for photosynthesis in plants.

Most of the specific Lhcb3 residues (Fig. 4) are localized in the B/E loop and in the C/A loop (Fig. 1C).Two particularly important changes are the insertion, occurred early in evolution, of an aromatic residue at the end of helix B next to the luminal side (Fig. 1C) and of an aromatic residue in the C/A loop on the stromal side. The probable role of these changes is discussed below.

Other changes seem unique to vascular plants, including a characteristic "DL" motif that precedes the trimerization motif in the N-terminal region. Some appeared even later in evolution and are only found in angiosperms and in the gymnosperms that retained Lhcb3, such as the change of a "WAx" motif in helix C for a "LGF" one. This last mutation is close to Chls b605, b606 and b607 (Fig. 1C) [24], and may modify their environment and properties, possibly explaining the observed shift of Chls b absorption from 650 to 655 nm in Lhcb3 [24]. At the end of the same helix, the Lhcb3 characteristic "NGLP" sequence was further modified in angiosperms by the mutation of the highly conserved Pro residue in a Asp.

It is possible that the progressive modification of Lhcb3 occurred during evolution has modulated its specific properties and therefore its role could be not exactly the same in bryophytes and angiosperms. To our knowledge, Lhcb3 properties have only been studied in angiosperms this far [17, 24, 61, 67]. Studies of Lhcb3 pigment binding and absorption properties in bryophytes, as well as its impact on PSII structure, may give further insights into land plant evolution and the role that Lhcb3 played in light harvesting adaptation.

In addition to its specific biochemical and biophysical properties, Lhcb3 has a specific localization: it has been shown to be part of the M-LHCII trimer [11, 17, 23-25] (Fig. 1A). The concomitant presence of this isoform and the CP24 (Lhcb6) monomeric antenna, also specific of land plants [26, 32, 33], has been shown to be responsible for the particular binding and orientation of the M trimer in plants. Indeed the binding of this trimer is affected in kolhcb6 mutant plants and Lhcb3 is depleted, showing a strong link between the two proteins [68]. On the contrary, in Lhcb3 mutant plants, CP24 is not depleted; in this case the only changes are an increase in Lhcb1 and Lhcb2 isoforms, which likely replace Lhcb3 in the M trimers, and the different orientation and strength of binding of the M-LHCII trimer to PSII [17, 67, 69]. Lack of Lhcb3 induces indeed a rotation of about 21° of the M trimer compared to its position in WT plants [67]. A shift of this trimer was also observed in *Picea abies*, which lost both Lhcb3 and CP24 during evolution [33]; in this case, however, the rotation was of 52°, closer to the 45° shift observed in *Chlamydomonas reinhardtii*, which also lacks both genes [18, 19]. These results suggested a close interaction between Lhcb3 and CP24; this proposition has been very recently confirmed by cryo-electron microscopy data [5, 11]. Indeed, according to the recent PSII structures from *Arabidopsis* and Pea [5, 11], the B/E loop and the C/A loop stand very close to CP24 on the luminal and stromal side respectively (Fig. 1). In particular, the pea PSII model (at 3.6 Å resolution for this part of the supercomplex), shows that specific amino acid changes in Lhcb3 (in particular the insertion of Trp77, as numbered on AtLhcb3 sequence) allow a direct contact of this subunit with CP24. It is therefore likely that the mutations occurred during evolution responsible for the apparition of Lhcb3 and CP24 have been fundamental for the structural changes of PSII. It is interesting to observe that CP24 is the PSII antenna most similar

to the Lhca proteins (Light harvesting complex of PSI) [6] and, in a similar manner as Lhca complex, it binds a β-carotene at the interface with CP29 [11, 70]. It is therefore possible to speculate that first an ancient Lhca became capable of binding to PSII and evolved in CP24, and in parallel an ancient Lhcbm evolved in Lhcb3, and both interacting proteins were responsible for the specific binding of LHCII-M to PSII.

It is also interesting to note that very recent structural data [11] also indicate that the different structure of the C/A loop in Lhcb3 likely prevent homotrimerisation of this isoform by steric hindrance with the N-terminus of the adjacent monomer, in accordance with biochemical data [24, 61].

The specific orientation of the M trimer observed in the presence (for instance in *Arabidopsis*) and in the absence of Lhcb3 (as in green algae) has likely an impact on energy transfer: indeed, the model of the *Chlamydomonas* complex [19] suggests that the energy transfer from the M trimer to the core involves Chl 610, instead of Chls 611 and 612 as in the case of *Arabidopsis* [5].

As a further hint of its specific role in light harvesting, Lhcb3 presents a differential regulation compared to Lhcb1 and 2: if Lhcb3 content decreases in high light, its expression is not affected by low light, while Lhcb1 and Lhcb2 expression increase [15, 71]. This behavior could be explained by the fact the Lhcb3 is specifically located in the M-trimer, whose content at low light is at most one trimer per monomeric PSII, while in high light is decreased. On the contrary, Lhcb1 and Lhcb2 also participate in the formation of the loosely bound LHCII L-trimers (Fig. 1A), whose content is modulated in response to the light intensity with a substantial increase at low light and decrease at high light.

Lhcb3 may have a role in high light adaptation: koLhcb3 Arabidopsis plants have indeed been shown to experience higher photoinhibition after a transition to high light [67]. Trimer M has been proposed to be an important site for NPQ in land plants: it was suggested to detach from the core upon high light exposure [72], and two independent cross-linking studies in moss [73] and *Arabidopsis* [74] have pointed it out as the possible site of interaction of PsbS, the protein involved in qE activation in plants (the fastest component of NPQ).

In addition to its potential role in photosynthesis regulation in high light, it has been suggested that Lhcb3 has a role in state transitions: in kolhcb3 mutant plants, LHCII was shown to be more phosphorylated than in the WT, and state transitions were faster [67]. This point will be further discussed in the next section.

Concerning the role of Lhcb3 in the light harvesting process in the PSII supercomplex, partially conflicting reports have been published. If Damkjaer and co-workers found no depletion of PSII activity in koLhcb3 plants [67], Adamiec and coworkers, by using a biochemical and time resolved fluorescence approach, found that PSII antenna size was slightly increased and excitation energy transfer slightly decreased in the koLhcb3 mutant. This is likely due to a less efficient transfer between the Lhcb3-depleted M trimer and the neighboring CP24 and CP29 [69], which could be explained by the positional shift of the M trimer. Additional studies of the energy transfer in both WT and Lhcb3-depleted plants, as well as in plants that have lost Lhcb3 and/or CP24, such as the *Pinus* and *Picea* lines [33], are necessary to fully understand the role of Lhcb3.

As a conclusion, more studies will be necessary to fully understand the role of Lhcb3 in PSII architecture and light-harvesting regulation. However, it is becoming clear that, as mentioned by Drop et al. [19] it is unlikely that Lhcb3 (and CP24) evolved to provide PSII with a bigger antenna; their absence in *Chlamydomonas* leaves place for an additional, “naked” (N) LHCII

trimer. Instead, it seems that its apparition in land plants answers for the need of specific regulations necessary for the terrestrial environment, such as the ability to rapidly modify PSII antenna to face fluctuating light conditions.

3. STATE TRANSITIONS: THE IMPACT OF LHCII EVOLUTION ON A WIDESPREAD REGULATION MECHANISM

LHCII is at the heart of another conserved regulation of light harvesting: the so-called state transitions [75-79]. These transitions are involved in the equilibration of excitation energy between the two photosystems. The current model asserts that under non-saturating light exciting preferentially PSII (i.e. blue light around 450-500 nm and red light in the 630-660 nm range [4, 5, 11, 12]), the N-terminus of a part of the major Lhcb proteins is phosphorylated by a thylakoid-embedded kinase. The enzymes involved are Stt7 in *Chlamydomonas* [80] and the homologous STN7 in *Arabidopsis* [81], which are activated at the Cytochrome b_6f by an over reduction of the plastoquinone (PQ) pool. This phosphorylation induces the movement of part of LHCII and its binding to PSI, allowing direct energy transfer from LHCII to PSI and thus rebalancing the excitation between photosystems [82, 83]. The process is reversed when the PQ pool is significantly oxidized, as in the dark or under a light preferentially absorbed by PSI (i.e. far red light). Under these conditions the Stt7/STN7 kinases are inactive and a phosphatase leads to LHCII dephosphorylation and its movement back to PSII. The phosphatase involved in *Arabidopsis* is the PPH1 protein (also called TAP38) [84, 85]. It has yet to be characterized in green algae (Fig. 5).

The mechanism is thus essential to maintain the redox poise of the PQ pool in order to optimize the electron transport chain. This regulation has been shown to be particularly important under non saturating fluctuating light as plants impaired for state transitions show a strongly reduced growth in these conditions [86].

In the green alga *Chlamydomonas*, state transitions can also been induced by anaerobiosis, when NADH cannot be oxidized by respiration and electrons are channeled to the PQ pool via the chloroplast NADH dehydrogenase. Moreover in *Chlamydomonas* a much greater number of trimers is involved in state transitions: around 80% of the LHCII trimers, compared to about 20% in *Arabidopsis* [75, 87-89]. However, recent results suggest that in *Chlamydomonas*, even if 80% of LHCII can detach from PSII, only a fraction binds to PSI and the rest become protected via aggregation-induced shortening of the excited-state lifetime [90]. This difference might point to a second physiological role of state transitions in algae. The fact that in *Chlamydomonas* a good part of the energy of the detached LHCII trimers might be quenched instead of being redistributed to PSI suggests that state transitions might be an important regulation to excess light in this organism. Indeed, it has been shown that several hours of high light treatment induce a transition to state 2 in *Chlamydomonas* [91]. In *Arabidopsis*, on the contrary, high light induces a return to state 1, along with the dephosphorylation of LHCII [92], pointing to a main role of regulations acting at the level of PSII under excess light. The necessity for the intervention of state transitions as a high light regulation in *Chlamydomonas* has been attributed to the differences in the activation of NPQ [90, 91]: in *Arabidopsis*, where PsbS is constitutively expressed, NPQ is activated in few seconds. By contrast, in *Chlamydomonas*, LhcSR3, the main protein involved in NPQ, as well as PsbS, are expressed upon exposition to high light [93]. In the time needed for the expression of these proteins, state transitions could provide fast-induced energy dissipation.

Besides some specific physiological function of state transitions, other differences can be noted between these organisms: in algae, two LHCII trimers, as well as the minor antenna protein CP29 and to a smaller extent CP26, bind to PSI on the opposite site of LHCI [94] (Fig. 5). In plants, only one LHCII trimer has been observed at this site [25, 95], though there are now evidences that other trimers can attach on the other side of PSI and transfer their energy through LHCI [96-98]. This divergence might be explained by the small differences of algae and plant photosystems. On PSI side, the PsaH, L, O proteins, which were shown to be important for LHCII docking [99, 100], show little changes between *Chlamydomonas* and *Arabidopsis*. However, the PSI supercomplex of *Chlamydomonas* has a different absorption compared to *Arabidopsis* PSI, with a spectrum less shifted toward far red wavelengths [101, 102]. The thylakoid membranes also show a different structure, with less stacked membranes in *Chlamydomonas* [103, 104]. This last fact could be a possible explanation for the large quantity of trimers involved in state transitions in this organism: with less thylakoid stacking, the trimers are more accessible for the phosphorylation. In *Arabidopsis*, the bulky stromal domain of STN7 should prevent its access to the grana cores, which indeed show a decrease in LHCII phosphorylation [58].

Nevertheless, the Lhcb isoform themselves likely play a major role in state transitions and in the PSI-LHCII supercomplex structure. Interestingly, the phosphorylation site is remarkably conserved among green organisms and all organisms studied here present the phosphorylation site on several Lhcb proteins (Supplemental Table 3), which answers to the characteristic basic/basic/T/X/basic residues motif described before [52, 105]. For vascular plants, it has been shown that N-terminus in Lhcb2 is even more conserved than in Lhcb1 [52]: in all these proteins, the motif corresponds to a RRTV(K/R) sequence. By contrast, Lhcb1 proteins show more variability around the phosphorylated Thr, with sometimes a lengthening of the sequence preceding the phosphorylation site. In both Lhcb1 sequences of *Picea glauca*, for instance, two amino acids were inserted after the Lys residues of the motif. Both Lhcb1-2 motifs are also present in the Lhcbm sequences of bryophytes and charophytes. In these isoforms, the N-terminus of the proteins present mixed characteristics of Lhcb1 and Lhcb2, though some of them are clearly closer of Lhcb2. Such is the case, for instance, of the Lhcbm2 isoform in *Physcomitrella patens* (Supplementary Table 3).

The better conservation of the Lhcb2-specific N-terminus is interesting considering that this isoform is more rapidly phosphorylated than Lhcb1 during state transitions [60]. This fact has been recently confirmed by Liu and coworkers [52], who showed that this specific and highly conserved sequence is indeed responsible for the faster phosphorylation kinetic: a substitution of the N-terminus of Lhcb2 with the Lhcb1 sequence, and reciprocally, did invert their phosphorylation properties. The kinase activity was deeply affected not only by the length of the peptide that preceded the phosphorylation site, but also by the actual sequence: the “RR” residues, specific of Lhcb2 in plants seem, alone, to be responsible for the faster kinetics.

It is important to note that the LHCII kinases of all organisms studied so far, from *Chlamydomonas* to *Arabidopsis*, are orthologs [81], and present a similar specificity [52]: the *Chlamydomonas* Stt7 kinase also induces a faster phosphorylation of *Arabidopsis* Lhcb2 and of all peptides presenting the RRTV(K/R) sequence. This sequence is indeed present (Supplemental Table 3) and was found to be phosphorylated [105] in one of the 9 Lhcbm proteins of *C. reinhardtii*: Lhcbm1. However, puzzlingly, it has been shown that this protein seems to have a minor role in state transitions: if it is phosphorylated by Stt7 in state 2 [105], and even though it was found associated to the PSI complex in state 2 [94], the Lhcbm1-mutant strain is not affected

for state transitions, but for NPQ instead [38, 41]. As discussed above, however, it is now proposed that a large pool of LHCII, after detachment from PSII, does not bind to PSI and rather aggregates and switches in a quenched state [90]. With the additional evidence that state transitions in algae are also involved in high light regulation [91], it is likely that Lhcbm1 is important for such a regulation (Fig. 5). By contrast, Lhcb2, which shares the same phosphorylation motif and fast phosphorylation kinetics, has been found to be essential for state transitions in plants [58, 59, 106].

Besides Lhcbm1, the *Chlamydomonas* isoforms that were found to be essential for state transitions also present a specific Stt7 phosphorylation motif, with 2 basic residues preceding the phosphorylation site. However, none of them show a particular similarity of their N-terminus with Lhcb1-2 sequences. In addition, type I sequences (Lhcbm 3/4/6/8/9) present a repeated motif (Supplemental Table 3), in which both Thr residues were reported to be phosphorylated [105].

These differences in the N-terminus around the phosphorylation site suggest that Lhcb1-2-3 and specific Lhcbm isoforms have different roles in states transitions. Such a specific role can for instance be seen by analyzing the phenotype of koLhcb3 plants. As indicated in the previous section, Damkjaer and coworkers showed that these plants present a higher LHCII phosphorylation and an increased rate of state transitions [67]. The authors suggest that the composition of the LHCII antennae permits to control the availability of phosphorylation sites and, together with the different kinetics of phosphorylation of the isoforms, the rate of state transitions. It is thus likely that Lhcb3 early differentiation in land plants had an immediate effect on state transitions and phosphorylation rates. In addition, together with CP24, it has a direct impact on the structure of PSII supercomplexes and its presence might limit CP29 mobility and availability for state transitions. Once again, the *Pinus* and *Picea* species, both lacking Lhcb3 and CP24, might be interesting models to study the effect of these proteins on state transitions in plants. To our knowledge, this regulation mechanism has been little studied in gymnosperms so far, although recent studies found a relatively high phosphorylation of LHCII in high light in conifers compared to angiosperms [107, 108], suggesting a different regulation of energy dissipation and possibly also of state transitions. It is nonetheless possible that the absence of Lhcb3 in pine species facilitates this higher phosphorylation.

As for the role of the other Lhcb proteins in state transitions, reverse genetics, structural and biochemical studies of the PSI-LHCII supercomplexes have brought a lot of information. In *Chlamydomonas*, for instance, the most abundant supercomplex found counts two LHCII trimers and one monomeric antenna. Another supercomplex shows the binding of a single LHCII with two monomeric Lhcb [94]. The monomer present in both supercomplexes is likely CP29, suggested to be essential for LHCII binding to PSI [109, 110]; the second one might be CP26, which has sometimes been found attached to PSI in state 2 [39, 94]. The Lhcbm isoforms themselves seem not sufficient for binding to PSI, as no LHCII trimer was found associated to PSI in the absence of CP29 [110]. However, a significant reduction of state transitions was observed after silencing of Lhcbm2/7 genes, underlying their involvement in the formation of the PSI-LHCII supercomplex [38]. Lhcbm5 is also thought to be an important isoform for state transitions, as it is considered a part of the “extra” LHCII pool [19] and was found associated to PSI in state 2 [39, 94, 111]. The silencing of Lhcbm4, 6 and 8 also reduced the amplitude of state transitions [40]; such was not the case for Lhcbm1, even though it was also found associated to the PSI-LHCII supercomplex [94]. All these subunits were found phosphorylated in state 2, but for Lhcbm2/7 [94, 105, 111], which is puzzling considering the reduction of state transitions in

their absence: either they are present in the complex in association with other subunits, though it was showed that they are able to form homotrimers [35], or their state 2-specific binding to PSI involves another protein, possibly CP29.
In flowering plants, by contrast, neither CP29 nor CP26 are present in the PSI-LHCII supercomplex [25, 95] and LHCII subunits are directly involved in the binding to PSI (Fig. 5). The purification and biochemical characterization of the PSI-LHCII complex (with a single LHCII bound on the PsaH side of PSI) has permitted to determine that the trimer involved is composed of Lhcb1 and Lhcb2 isoforms and belongs mainly to the "extra" pool of L-LHCII [25]. Only about a monomer per LHCII trimer is phosphorylated in this complex [58]. Moreover this phosphorylation was shown to be almost exclusively on the Lhcb2 isoform [58, 59] and protected at the interface between LHCII and PSI, suggesting the direct involvement of the phosphorylated N-terminus of Lhcb2 in the binding to PSI [58]. Phosphorylated Lhcb1, by contrast, is retained in the grana membranes where the quantity of phosphorylated Lhcb2 is decreased, and might thus have a role at the level of PSII during state transitions [58] (Fig. 5).
This specificity, both for the isoform and phosphorylation, highlights the importance of the differentiation of Lhcb1 and Lhcb2 N-terminus: it is likely that the structural change that has been proposed to occur upon phosphorylation [112] leads to a different conformation in each isoform. Unfortunately, the first amino-acids of these proteins, which bear the phosphorylation site, are absent from all structural models of LHCII, either isolated [8, 50] or bound to PSII [4, 5, 11]. This may be due to the fact that the structural studies were performed on heterogeneous particles, in which the trimers are composed of both Lhcb1 and Lhcb2. As the N-terminus of these proteins differs both in length and in amino acid composition, the electronic density is lost at this level. Another possible cause of the lack of available structure for the N-terminus of Lhcb1/2 is their high flexibility when not phosphorylated and in solution. Therefore the structure and dynamic of this peptide have only been studied by spectroscopy [112] or EPR [113, 114]. These studies confirmed that the N-terminal segment of LHCII proteins has a high mobility and can adopt diverse, but not random, conformations. It also seems that phosphorylation does affect the conformation of the N-terminal segment, if not of the entire LHCII protein [112, 115]. It is also interesting to observe that according to a pseudo-atomic model of the PSI-LHCII supercomplex of Arabidopsis [25, 58], the first seven Lhcb2-specific residues (Fig. 3) are at the interface between Lhcb2 and PSI and likely important for LHCII binding, thus providing a reason for the differentiation of the Lhcb2 isoform. Further structural data will have to wait for future studies, possibly by resolving the high resolution structure of the PSI-LHCII supercomplex.
Finally, some Lhcb proteins have been shown to undergo a different N-terminus processing upon their import in the chloroplast, leading to two alternative proteins of slightly different molecular weights. This phenomenon has been described in *Chlamydomonas*, at least for the Lhcbm3 and Lhcbm6 isoforms [48], but seems to be also present in plants [49]. In both cases, the second processing site is after the phosphorylatable Thr. If the larger peptide represents the majority of the mature proteins in the conditions studied [49], it is possible that this ratio is modified in other light conditions, leading to another, possibly long term regulation of state transitions by reducing the amount of Lhcb isoforms available for phosphorylation.
In brief, despite the overall conservation of the phosphorylation site and pattern, state transitions have gone through profound transformations between green algae and flower plants, probably due to the evolution of PSI and PSII antennae and the differentiation of the Lhcbm isoforms in Lhcb1, Lhcb2 and Lhcb3. In parallel, CP29 mobility and availability may have been limited by

Lhcb3 and CP24 apparition. Together, these adaptations could explain the differential structures of the PSII and PSI-LHCII supercomplexes in *Chlamydomonas* and *Arabidopsis*. Future studies will have to further explore the role of each Lhcb subunit in this sophisticated mechanism, which we are only beginning to decipher.

CONCLUSIONS

Despite a billion year of evolution, the proteins composing the LHCII antenna in the green line show an impressive similarity. However, if the pigments binding sites and properties are almost unchanged between green algae and plants, with the exception of the Lhcb3 isoform, LHCII isoforms multiplied during evolution and underwent a progressive differentiation into the Lhcb1-2-3 subclasses in vascular plants, while several Lhcbm isoforms evolved in green algae and non-vascular plants. Both in the case of Lhcb1-2-3 and Lhcbm isoforms, most of the differences are found in the N-terminus and in the loop regions. This suggests that the specific residues of each isoform might have an impact on their regulation and interactions, which are important for the photosystem architecture, as demonstrated in the case of Lhcb3. Moreover, comparative studies on angiosperms and green algae also point out to an effect of this evolution on light harvesting regulation, through a modification of various acclimation mechanisms, at the first rank of which are state transitions. The increasing number of studies on other organisms, such as mosses and gymnosperms, should allow uncovering the missing information on the properties of these proteins and of their role on the structure, function and regulation of photosystems.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

SC received support by the French National Research Agency Grant ANR-12-JSV8-0001-01.

ABBREVIATIONS

Car = Carotenoid
Chl = Chlorophyll
Lhc = Light harvesting complex
Lhcb = Light harvesting complexes of PSII
Lhca = Light harvesting complexes of PSI
CP29 (Lhcb4), CP26 (Lhcb5), CP24 (Lhcb6) = Monomeric Lhc of PSII
LHCII = Major Light Harvesting Complex of PSII
Lhcb1, 2 and 3: nomenclature of the isoforms composing LHCII in vascular plants
Lhcbm(x): nomenclature of LHCII isoforms when they are not classifiable as Lhcb1, 2 and 3 (generally in non-vascular plants and green algae)
NPQ = Non Photochemical (energy) Quenching
PSI, PSII = Photosystem I and II, respectively

REFERENCES

[1] Hohmann-Marriott, M.F.; Blankenship, R.E. Evolution of photosynthesis. *Annu. Rev. Plant Biol.*, **2011**, *62*, 515-548.

[2] Umena, Y.; Kawakami, K.; Shen, J.R.; Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 A. *Nature*, **2011**, *473*, 55-60.

[3] Ago, H.; Adachi, H.; Umena, Y.; Tashiro, T.; Kawakami, K.; Kamiya, N.; Tian, L.R.; Han, G.Y.; Kuang, T.Y.; Liu, Z.Y.; Wang, F.J.; Zou, H.F.; Enami, I.; Miyano, M.; Shen, J.R. Novel Features of Eukaryotic Photosystem II Revealed by Its Crystal Structure Analysis from a Red Alga. *J. Biol. Chem.*, **2016**, *291*, 5676-5687.

[4] Wei, X.; Su, X.; Cao, P.; Liu, X.; Chang, W.; Li, M.; Zhang, X.; Liu, Z. Structure of spinach photosystem II-LHCII supercomplex at 3.2 A resolution. *Nature*, **2016**, *534*, 69-74.

[5] van Bezouwen, L.S.; Caffarri, S.; Kale, R.S.; Kouril, R.; Thunnissen, A.W.H.; Oostergetel, G.T.; Boekema, E.J. Subunit and chlorophyll organization of the plant photosystem II supercomplex. *Nat. Plants*, **2017**, *3*, 17080, doi: 10.1038/nplants.2017.80.

[6] Jansson, S. A guide to the Lhc genes and their relatives in Arabidopsis. *Trends Plant Sci.*, **1999**, *4*, 236-240.

[7] Buchel, C. Evolution and function of light harvesting proteins. *J. Plant Physiol.*, **2015**, *172*, 62-75.

[8] Liu, Z.; Yan, H.; Wang, K.; Kuang, T.; Zhang, J.; Gui, L.; An, X.; Chang, W. Crystal structure of spinach major light-harvesting complex at 2.72 A resolution. *Nature*, **2004**, *428*, 287-292.

[9] Pan, X.; Liu, Z.; Li, M.; Chang, W. Architecture and function of plant light-harvesting complexes II. *Curr. Opin. Struct. Biol.*, **2013**, *23*, 515-525.

[10] Ballottari, M.; Girardon, J.; Dall'osto, L.; Bassi, R. Evolution and functional properties of photosystem II light harvesting complexes in eukaryotes. *Biochim. Biophys. Acta- Bioenerg.*, **2012**, *1817*, 143-157.

[11] Su, X.; Ma, J.; Wei, X.; Cao, P.; Zhu, D.; Chang, W.; Liu, Z.; Zhang, X.; Li, M. Structure and assembly mechanism of plant C2S2M2-type PSII-LHCII supercomplex. *Science*, **2017**, *357*, 815-820.

[12] Caffarri, S.; Tibiletti, T.; Jennings, R.C.; Santabarbara, S. A comparison between plant photosystem I and photosystem II architecture and functioning. *Curr. Protein Pept. Sci.*, **2014**, *15*, 296-331.

[13] Bowman, J.L.; Floyd, S.K.; Sakakibara, K. Green genes-comparative genomics of the green branch of life. *Cell*, **2007**, *129*, 229-234.

[14] Peter, G.F.; Thornber, J.P. Biochemical composition and organization of higher plant photosystem II LIGHT-HARVESTING PIGMENT-PROTEINS. *J. Biol. Chem.*, **1991**, *266*, 16745-16754.

[15] Kouril, R.; Wientjes, E.; Bultema, J.B.; Croce, R.; Boekema, E.J. High-light vs. low-light: Effect of light acclimation on photosystem II composition and organization in Arabidopsis thaliana. *Biochim. Biophys. Acta- Bioenerg.*, **2013**, *1827*, 411-419.

[16] Dekker, J.P.; Boekema, E.J. Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim. Biophys. Acta- Bioenerg.*, **2005**, *1706*, 12-39.

[17] Caffarri, S.; Kouril, R.; Kereiche, S.; Boekema, E.J.; Croce, R. Functional architecture of higher plant photosystem II supercomplexes. *Embo J.*, **2009**, *28*, 3052-3063.

[18] Tokutsu, R.; Kato, N.; Bui, K.H.; Ishikawa, T.; Minagawa, J. Revisiting the supramolecular organization of photosystem II in Chlamydomonas reinhardtii. *J. Biol. Chem.*, **2012**, *287*, 31574-31581.

[19] Drop, B.; Webber-Birungi, M.; Yadav, S.K.; Filipowicz-Szymanska, A.; Fusetti, F.; Boekema, E.J.; Croce, R. Light-harvesting complex II (LHCII) and its supramolecular organization in Chlamydomonas reinhardtii. *Biochim. Biophys. Acta- Bioenerg.*, **2014**, *1837*, 63-72.

[20] Boekema, E.J.; van Roon, H.; van Breemen, J.F.; Dekker, J.P. Supramolecular organization of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. *Eur. J. Biochem.*, **1999**, *266*, 444-452.

[21] Nosek, L.; Semchonok, D.; Boekema, E.J.; Ilik, P.; Kouril, R. Structural variability of plant photosystem II megacomplexes in thylakoid membranes. *Plant J.*, **2017**, *89*, 104-111.

[22] Hankamer, B.; Nield, J.; Zheleva, D.; Boekema, E.; Jansson, S.; Barber, J. Isolation and biochemical characterisation of monomeric and dimeric photosystem II complexes from spinach and their relevance to the organisation of photosystem II in vivo. *Eur. J. Biochem.*, **1997**, *243*, 422-429.

[23] Dainese, P.; Bassi, R. Subunit Stoichiometry of the Chloroplast Photosystem- II Antenna System and Aggregation State of the Component Chlorophyll-a/b Binding Proteins. *J. Biol. Chem.*, **1991**, *266*, 8136-8142.

[24] Caffarri, S.; Croce, R.; Cattivelli, L.; Bassi, R. A look within LHCII: Differential analysis of the Lhcbl-3 complexes building the major trimeric antenna complex of higher-plant photosynthesis. *Biochemistry*, **2004**, *43*, 9467-9476.
[25] Galka, P.; Santabarbara, S.; Khuong, T.T.; Degand, H.; Morsomme, P.; Jennings, R.C.; Boekema, E.J.; Caffarri, S. Functional analyses of the plant photosystem I-light-harvesting complex II supercomplex reveal that light-harvesting complex II loosely bound to photosystem II is a very efficient antenna for photosystem I in state II. *Plant Cell*, **2012**, *24*, 2963-2978.
[26] Koziol, A.G.; Borza, T.; Ishida, K.; Keeling, P.; Lee, R.W.; Durnford, D.G. Tracing the evolution of the light-harvesting antennae in chlorophyll a/b-containing organisms. *Plant Physiol.*, **2007**, *143*, 1802-1816.
[27] Engelken, J.; Brinkmann, H.; Adamska, I. Taxonomic distribution and origins of the extended LHC (light-harvesting complex) antenna protein superfamily. *BMC Evol. Biol.*, **2010**, *10:233* doi: 10.1186/1471-2148-10-233.
[28] Montane, M.H.; Kloppstech, K. The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function? *Gene*, **2000**, *258*, 1-8.
[29] Staleva, H.; Komenda, J.; Shukla, M.K.; Slouf, V.; Kana, R.; Polivka, T.; Sobotka, R. Mechanism of photoprotection in the cyanobacterial ancestor of plant antenna proteins. *Nat. Chem. Biol.*, **2015**, *11*, 287-291.
[30] Chinn, E.; Silverthorne, J.; Hohtola, A. Light-regulated and organ-specific expression of types 1, 2, and 3 light-harvesting complex b mRNAs in Ginkgo biloba. *Plant Physiol.*, **1995**, *107*, 593-602.
[31] Elrad, D.; Grossman, A.R. A genome's-eye view of the light-harvesting polypeptides of Chlamydomonas reinhardtii. *Curr Genet.*, **2004**, *45*, 61-75.
[32] Alboresi, A.; Caffarri, S.; Nogue, F.; Bassi, R.; Morosinotto, T. In silico and biochemical analysis of Physcomitrella patens photosynthetic antenna: identification of subunits which evolved upon land adaptation. *PLoS ONE*, **2008**, *3*, e2033, doi: 10.1371/journal.pone.0002033.
[33] Kouril, R.; Nosek, L.; Bartos, J.; Boekema, E.J.; Ilik, P. Evolutionary loss of light-harvesting proteins Lhcb6 and Lhcb3 in major land plant groups - break-up of current dogma. *New Phytol.*, **2016**, *210*, 808-814.
[34] Minagawa, J.; Takahashi, Y. Structure, function and assembly of Photosystem II and its light-harvesting proteins. *Photosynth. Res.*, **2004**, *82*, 241-263.
[35] Natali, A.; Croce, R. Characterization of the major light-harvesting complexes (LHCBM) of the green alga Chlamydomonas reinhardtii. *PLoS One*, **2015**, *10*, e0119211, doi: 10.1371/journal.pone.0119211.
[36] Nguyen, A.V.; Thomas-Hall, S.R.; Malnoe, A.; Timmins, M.; Mussgnug, J.H.; Rupprecht, J.; Kruse, O.; Hankamer, B.; Schenk, P.M. Transcriptome for Photobiological Hydrogen Production Induced by Sulfur Deprivation in the Green Alga Chlamydomonas reinhardtii. *Eukaryot. Cell*, **2008**, *7*, 1965-1979.
[37] Grewe, S.; Ballottari, M.; Alcocer, M.; D'Andrea, C.; Blifernez-Klassen, O.; Hankamer, B.; Mussgnug, J.H.; Bassi, R.; Kruse, O. Light-Harvesting Complex Protein LHCBM9 Is Critical for Photosystem II Activity and Hydrogen Production in Chlamydomonas reinhardtii. *Plant Cell*, **2014**, *26*, 1598-1611.
[38] Ferrante, P.; Ballottari, M.; Bonente, G.; Giuliano, G.; Bassi, R. LHCBM1 and LHCBM2/7 polypeptides, components of major LHCII complex, have distinct functional roles in photosynthetic antenna system of Chlamydomonas reinhardtii. *J. Biol. Chem.*, **2012**, *287*, 16276-16288.
[39] Takahashi, H.; Iwai, M.; Takahashi, Y.; Minagawa, J. Identification of the mobile light-harvesting complex II polypeptides for state transitions in Chlamydomonas reinhardtii. *Proc. Natl. Acad. Sci. USA*, **2006**, *103*, 477-482.
[40] Girolomoni, L.; Ferrante, P.; Berteotti, S.; Giuliano, G.; Bassi, R.; Ballottari, M. The function of LHCBM4/6/8 antenna proteins in Chlamydomonas reinhardtii. *J. Exp. Bot.*, **2017**, *68*, 627-641.
[41] Elrad, D.; Niyogi, K.K.; Grossman, A.R. A major light-harvesting polypeptide of photosystem II functions in thermal dissipation. *Plant Cell*, **2002**, *14*, 1801-1816.
[42] Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J.F.; Guindon, S.; Lefort, V.; Lescot, M.; Claverie, J.M.; Gascuel, O. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.*, **2008**, *36*, W465-9, doi: 10.1093/nar/gkn180.
[43] Chaw, S.M.; Chang, C.C.; Chen, H.L.; Li, W.H. Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *J. Mol. Evol.*, **2004**, *58*, 424-441.
[44] Caffarri, S.; Frigerio, S.; Olivieri, E.; Righetti, P.G.; Bassi, R. Differential accumulation of Lhcb gene products in thylakoid membranes of Zea mays plants grown under contrasting light and temperature conditions. *Proteomics*, **2005**, *5*, 758-768.
[45] Frigerio, S.; Campoli, C.; Zorzan, S.; Fantoni, L.I.; Crosatti, C.; Drepper, F.; Haehnel, W.; Cattivelli, L.; Morosinotto, T.; Bassi, R. Photosynthetic antenna size in higher plants is controlled by the plastoquinone

redox state at the post-transcriptional rather than transcriptional level. *J. Biol. Chem.*, **2007**, *282*, 29457-29469.
[46] Remelli, R.; Varotto, C.; Sandona, D.; Croce, R.; Bassi, R. Chlorophyll binding to monomeric light-harvesting complex. A mutation analysis of chromophore-binding residues. *J. Biol. Chem.*, **1999**, *274*, 33510-33521.
[47] Hobe, S.; Förster, R.; Klingler, J.; Paulsen, H. N-proximal sequence motif in light-harvesting chlorophyll a/b-binding protein is essential for the trimerization of light- harvesting chlorophyll a/b complex. *Biochemistry*, **1995**, *34*, 10224-10228.
[48] Stauber, E.J.; Fink, A.; Markert, C.; Kruse, O.; Johanningmeier, U.; Hippler, M. Proteomics of Chlamydomonas reinhardtii light-harvesting proteins. *Eukaryot. Cell*, **2003**, *2*, 978-994.
[49] Clark, S.E.; Lamppa, G.K. Determinants for cleavage of the chlorophyll a/b binding protein precursor: a requirement for a basic residue that is not universal for chloroplast imported proteins. *J. Cell Biol.*, **1991**, *114*, 681-688.
[50] Standfuss R, van Scheltinga ACT, Lamborghini M, Kuhlbrandt W. Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5A resolution. *Embo J*, **2005**, 24, 919-928.
[51] Jansson, S. The light-harvesting chlorophyll a/b-binding proteins. *Biochim. Biophys. Acta*, **1994**, *1184*, 1-19.
[52] Liu, W.; Tu, W.; Liu, Y.; Sun, R.; Liu, C.; Yang, C. The N-terminal domain of Lhcb proteins is critical for recognition of the LHCII kinase. *Biochim. Biophys. Acta- Bioenerg.*, **2016**, *1857*, 79-88.
[53] Grieco, M.; Jain, A.; Ebersberger, I.; Teige, M. An evolutionary view on thylakoid protein phosphorylation uncovers novel phosphorylation hotspots with potential functional implications. *J. Exp. Bot.*, **2016**, *67*, 3883-3896.
[54] Sugiyama, N.; Nakagami, H.; Mochida, K.; Daudi, A.; Tomita, M.; Shirasu, K.; Ishihama, Y. Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in Arabidopsis. *Mol. Syst. Biol.*, **2008**, *4*, 193, doi: 10.1038/msb.2008.32.
[55] Yang, Z.; Guo, G.; Zhang, M.; Liu, C.Y.; Hu, Q.; Lam, H.; Cheng, H.; Xue, Y.; Li, J.; Li, N. Stable isotope metabolic labeling-based quantitative phosphoproteomic analysis of Arabidopsis mutants reveals ethylene-regulated time-dependent phosphoproteins and putative substrates of constitutive triple response 1 kinase. *Mol. Cell. Proteomics*, **2013**, *12*, 3559-3582.
[56] Chen, X.; Chan, W.L.; Zhu, F.Y.; Lo, C. Phosphoproteomic analysis of the non-seed vascular plant model Selaginella moellendorffii. *Proteome Sci.*, **2014**, *12*, 16, doi: 10.1186/1477-5956-12-16.
[57] Ingelsson, B.; Vener, A.V. Phosphoproteomics of Arabidopsis chloroplasts reveals involvement of the STN7 kinase in phosphorylation of nucleoid protein pTAC16. *FEBS Lett.*, **2012**, *586*, 1265-1271.
[58] Crepin, A.; Caffarri, S. The specific localizations of phosphorylated Lhcb1 and Lhcb2 isoforms reveal the role of Lhcb2 in the formation of the PSI-LHCII supercomplex in Arabidopsis during state transitions. *Biochim. Biophys. Acta- Bioenerg.*, **2015**, *1847*, 1539-1548.
[59] Longoni, P.; Douchi, D.; Cariti, F.; Fucile, G.; Goldschmidt-Clermont, M. Phosphorylation of the Light-Harvesting Complex II Isoform Lhcb2 Is Central to State Transitions. *Plant Physiol.*, **2015**, *169*, 2874-2883.
[60] Leoni, C.; Pietrzykowska, M.; Kiss, A.Z.; Suorsa, M.; Ceci, L.R.; Aro, E.M.; Jansson, S. Very rapid phosphorylation kinetics suggest a unique role for Lhcb2 during state transitions in Arabidopsis. *Plant J.*, **2013**, *76*, 236-246.
[61] Standfuss, J.; Kuhlbrandt, W. The three isoforms of the light-harvesting complex II - Spectroscopic features, trimer formation, and functional roles. *J. Biol. Chem.*, **2004**, *279*, 36884-36891.
[62] Dall'Osto L, Cazzaniga S, Bressan M, Palecek D, Zidek K, Niyogi KK, Fleming GR, Zigmantas D, Bassi R. Two mechanisms for dissipation of excess light in monomeric and trimeric light-harvesting complexes. *Nat. Plants*, **2017**, *3*, 17033.
[63] Ruban AV, Berera R, Ilioaia C, van Stokkum IHM, Kennis JTM, Pascal AA, van Amerongen H, Robert B, Horton P, van Grondelle R. Identification of a mechanism of photoprotective energy dissipation in higher plants. *Nature*, **2007**, *450*, 575-578.
[64] Papadatos S, Charalambous AC, Daskalakis V. A pathway for protective quenching in antenna proteins of Photosystem II. *Sci. Rep.*, **2017**, *7*, 2523.
[65] van Oort B, van Hoek A, Ruban AV, van Amerongen H. Aggregation of light-harvesting complex II leads to formation of efficient excitation energy traps in monomeric and trimeric complexes. *FEBS Lett.*, **2007**, *581*, 3528-32.
[66] Jackowski, G.; Jansson, S. Characterization of photosystem II antenna complexes separated by non-denaturing isoelectric focusing. *Z. Naturforsch. C*, **1998**, *53*, 841-848.

[67] Damkjaer, J.T.; Kereiche, S.; Johnson, M.P.; Kovacs, L.; Kiss, A.Z.; Boekema, E.J.; Ruban, A.V.; Horton, P.; Jansson, S. The photosystem II light-harvesting protein Lhcb3 affects the macrostructure of photosystem II and the rate of state transitions in Arabidopsis. *Plant Cell*, **2009**, *21*, 3245-3256.
[68] Kovacs, L.; Damkjaer, J.; Kereiche, S.; Ilioaia, C.; Ruban, A.V.; Boekema, E.J.; Jansson, S.; Horton, P. Lack of the light-harvesting complex CP24 affects the structure and function of the grana membranes of higher plant chloroplasts. *Plant Cell*, **2006**, *18*, 3106-3120.
[69] Adamiec, M.; Gibasiewicz, K.; Lucinski, R.; Giera, W.; Chelminiak, P.; Szewczyk, S.; Sipinska, W.; van Grondelle, R.; Jackowski, G. Excitation energy transfer and charge separation are affected in Arabidopsis thaliana mutants lacking light-harvesting chlorophyll a/b binding protein Lhcb3. *J. Photochem. Photobiol. B*, **2015**, *153*, 423-428.
[70] Xu, P.; Roy, L.M.; Croce, R. Functional organization of photosystem II antenna complexes: CP29 under the spotlight. *Biochim. Biophys. Acta- Bioenerg.*, **2017**, *1858*, 815-822.
[71] Ballottari, M.; Dall'Osto, L.; Morosinotto, T.; Bassi, R. Contrasting behavior of higher plant photosystem I and II antenna systems during acclimation. *J. Biol. Chem.*, **2007**, *282*, 8947-8958.
[72] Betterle, N.; Ballottari, M.; Zorzan, S.; de Bianchi, S.; Cazzaniga, S.; Dall'osto, L.; Morosinotto, T.; Bassi, R. Light-induced dissociation of an antenna hetero-oligomer is needed for non-photochemical quenching induction. *J. Biol. Chem.*, **2009**, *284*, 15255-15266.
[73] Gerotto, C.; Franchin, C.; Arrigoni, G.; Morosinotto, T. In Vivo Identification of Photosystem II Light Harvesting Complexes Interacting with PHOTOSYSTEM II SUBUNIT S. *Plant Physiol.*, **2015**, *168*, 1747-1761.
[74] Correa-Galvis, V.; Poschmann, G.; Melzer, M.; Stuhler, K.; Jahns, P. PsbS interactions involved in the activation of energy dissipation in Arabidopsis. *Nat. Plants*, **2016**, *2*, 15225, doi: 10.1038/nplants.2015.225.
[75] Allen, J.F. Protein phosphorylation in regulation of photosynthesis. *Biochim. Biophys. Acta*, **1992**, *1098*, 275-335.
[76] Wollman, F.A. State transitions reveal the dynamics and flexibility of the photosynthetic apparatus. *Embo J.*, **2001**, *20*, 3623-3630.
[77] Tikkanen, M.; Piippo, M.; Suorsa, M.; Sirpio, S.; Mulo, P.; Vainonen, J.; Vener, A.V.; Allahverdiyeva, Y.; Aro, E.M. State transitions revisited-a buffering system for dynamic low light acclimation of Arabidopsis. *Plant Mol. Biol.*, **2006**, *62*, 779-793.
[78] Rochaix, J.D. Regulation and dynamics of the light-harvesting system. *Annu. Rev. Plant Biol.*, **2014**, *65*, 287-309.
[79] Goldschmidt-Clermont, M.; Bassi, R. Sharing light between two photosystems: mechanism of state transitions. *Curr. Opin. Plant Biol.*, **2015**, *25*, 71-78.
[80] Depege N, Bellafiore S, Rochaix JD. Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in Chlamydomonas. *Science*, **2003**, *299*, 1572-5.
[81] Bellafiore, S.; Barneche, F.; Peltier, G.; Rochaix, J.D. State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. *Nature*, **2005**, *433*, 892-895.
[82] Le Quiniou, C.; van Oort, B.; Drop, B.; van Stokkum, I.H.; Croce, R. The High Efficiency of Photosystem I in the Green Alga Chlamydomonas reinhardtii Is Maintained after the Antenna Size Is Substantially Increased by the Association of Light-harvesting Complexes II. *J. Biol. Chem.*, **2015**, *290*, 30587-30595.
[83] Santabarbara, S.; Tibiletti, T.; Remelli, W.; Caffarri, S. Kinetics and heterogeneity of energy transfer from light harvesting complex II to photosystem I in the supercomplex isolated from Arabidopsis. *Phys. Chem. Chem. Phys.*, **2017**, *19*, 9210-9222.
[84] Shapiguzov A, Ingelsson B, Samol I, Andres C, Kessler F, Rochaix JD, Vener AV, Goldschmidt-Clermont M. The PPH1 phosphatase is specifically involved in LHCII dephosphorylation and state transitions in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **2010**, *107*, 4782-7.
[85] Pribil M, Pesaresi P, Hertle A, Barbato R, Leister D. Role of Plastid Protein Phosphatase TAP38 in LHCII Dephosphorylation and Thylakoid Electron Flow. *PLoS Biol.*, **2010**, *8*, e1000288.
[86] Tikkanen M, Grieco M, Kangasjarvi S, Aro EM. Thylakoid protein phosphorylation in higher plant chloroplasts optimizes electron transfer under fluctuating light. *Plant Physiol.*, **2010**, *152*, 723-35.
[87] Vallon, O.; Wollman, F.A.; Olive, J. Lateral distribution ofthe main protein complexes of the photosynthetic apparatus in Chlamydomonas reinhardtii and in spinach: And immunocytochemical study using intact thylakoid membranes and a PS II enriched membrane preparation. *Photobiochem. Photobiophys.*, **1986**, *12*, 203-220.
[88] Iwai, M.; Takahashi, Y.; Minagawa, J. Molecular remodeling of photosystem II during state transitions in Chlamydomonas reinhardtii. *Plant Cell*, **2008**, *20*, 2177-2189.

[89] Bassi, R.; Rigoni, F.; Barbato, R.; Giacometti, G.M. Light-harvesting chlorophyll a/b proteins (LHCII) populations in phosphorylated membranes. *Biochim. Biophys. Acta*, **1988**, *936*, 29-38.
[90] Unlu, C.; Drop, B.; Croce, R.; van Amerongen, H. State transitions in Chlamydomonas reinhardtii strongly modulate the functional size of photosystem II but not of photosystem I. *Proc. Natl. Acad. Sci. USA*, **2014**, *111*, 3460-3465.
[91] Allorent, G.; Tokutsu, R.; Roach, T.; Peers, G.; Cardol, P.; Girard-Bascou, J.; Seigneurin-Berny, D.; Petroutsos, D.; Kuntz, M.; Breyton, C.; Franck, F.; Wollman, F.A.; Niyogi, K.K.; Krieger-Liszkay, A.; Minagawa, J.; Finazzi, G. A dual strategy to cope with high light in Chlamydomonas reinhardtii. *Plant Cell*, **2013**, *25*, 545-557.
[92] Mekala NR, Suorsa M, Rantala M, Aro EM, Tikkanen M. Plants Actively Avoid State Transitions upon Changes in Light Intensity: Role of Light-Harvesting Complex II Protein Dephosphorylation in High Light. *Plant Physiol.*, **2015**, 168, 721-34.
[93] Tibiletti T, Auroy P, Peltier G, Caffarri S. Chlamydomonas reinhardtii PsbS protein is functional and accumulates rapidly and transiently under high light. *Plant Physiol.*, **2016**, 10.1104/pp.16.00572.
[94] Drop, B.; Yadav, K.N.S.; Boekema, E.J.; Croce, R. Consequences of state transitions on the structural and functional organization of photosystem I in the green alga Chlamydomonas reinhardtii. *Plant J.*, **2014**, *78*, 181-191.
[95] Kouril, R.; Zygadlo, A.; Arteni, A.A.; de Wit, C.D.; Dekker, J.P.; Jensen, P.E.; Scheller, H.V.; Boekema, E.J. Structural characterization of a complex of photosystem I and light-harvesting complex II of Arabidopsis thaliana. *Biochemistry*, **2005**, *44*, 10935-10940.
[96] Bell, A.J.; Frankel, L.K.; Bricker, T.M. High Yield Non-detergent Isolation of Photosystem I-Light-harvesting Chlorophyll II Membranes from Spinach Thylakoids: IMPLICATIONS FOR THE ORGANIZATION OF THE PS I ANTENNAE IN HIGHER PLANTS. *J. Biol. Chem.*, **2015**, *290*, 18429-18437.
[97] Benson, S.L.; Maheswaran, P.; Ware, M.A.; Hunter, C.N.; Horton, P.; Jansson, S.; Ruban, A.V.; Johnson, M.P. An intact light harvesting complex I antenna system is required for complete state transitions in Arabidopsis. *Nat. Plants*, **2015**, *1*, 15176, doi: 10.1038/nplants.2015.176.
[98] Yadav, K.N.; Semchonok, D.A.; Nosek, L.; Kouril, R.; Fucile, G.; Boekema, E.J.; Eichacker, L.A. Supercomplexes of plant photosystem I with cytochrome b6f, light-harvesting complex II and NDH. *Biochim. Biophys. Acta- Bioenerg.*, **2017**, *1858*, 12-20.
[99] Lunde, C.; Jensen, P.E.; Haldrup, A.; Knoetzel, J.; Scheller, H.V. The PSI-H subunit of photosystem I is essential for state transitions in plant photosynthesis. *Nature*, **2000**, *408*, 613-615.
[100] Zhang, S.; Scheller, H.V. Light-harvesting complex II binds to several small subunits of photosystem I. *J. Biol. Chem.*, **2004**, *279*, 3180-3187.
[101] Melkozernov, A.N.; Kargul, J.; Lin, S.; Barber, J.; Blankenship, R.E. Spectral and kinetic analysis of the energy coupling in the PSI-LHC I supercomplex from the green alga Chlamydomonas reinhardtii at 77 K. *Photosynth. Res.*, **2005**, *86*, 203-215.
[102] Gibasiewicz, K.; Szrajner, A.; Ihalainen, J.A.; Germano, M.; Dekker, J.P.; van Grondelle, R. Characterization of low-energy chlorophylls in the PSI-LHCI supercomplex from Chlamydomonas reinhardtii. A site-selective fluorescence study. *J. Phys. Chem. B*, **2005**, *109*, 21180-21186.
[103] Mullineaux, C.W. Function and evolution of grana. *Trends Plant Sci.*, **2005**, *10*, 521-525.
[104] Anderson, J.M.; Chow, W.S.; De Las Rivas, J. Dynamic flexibility in the structure and function of photosystem II in higher plant thylakoid membranes: the grana enigma. *Photosynth. Res.*, **2008**, *98*, 575-587.
[105] Lemeille, S.; Turkina, M.V.; Vener, A.V.; Rochaix, J.D. Stt7-dependent phosphorylation during state transitions in the green alga Chlamydomonas reinhardtii. *Mol. Cell. Proteomics*, **2010**, *9*, 1281-1295.
[106] Pietrzykowska, M.; Suorsa, M.; Semchonok, D.A.; Tikkanen, M.; Boekema, E.J.; Aro, E.M.; Jansson, S. The light-harvesting chlorophyll a/b binding proteins Lhcb1 and Lhcb2 play complementary roles during state transitions in Arabidopsis. *Plant Cell*, **2014**, *26*, 3646-3660.
[107] Verhoeven, A.; Osmolak, A.; Morales, P.; Crow, J. Seasonal changes in abundance and phosphorylation status of photosynthetic proteins in eastern white pine and balsam fir. *Tree Physiol.*, **2009**, *29*, 361-374.
[108] Verhoeven, A.S.; Kertho, A.; Nguyen, M. Characterization of light-dependent regulation of state transitions in gymnosperms. *Tree Physiol.*, **2016**, *36*, 325-334.
[109] Kargul, J.; Turkina, M.V.; Nield, J.; Benson, S.; Vener, A.V.; Barber, J. Light-harvesting complex II protein CP29 binds to photosystem I of Chlamydomonas reinhardtii under State 2 conditions. *Febs J.*, **2005**, *272*, 4797-4806.

[110] Tokutsu, R.; Iwai, M.; Minagawa, J. CP29, a monomeric light-harvesting complex II protein, is essential for state transitions in Chlamydomonas reinhardtii. *J. Biol. Chem.*, **2009**, *284*, 7777-7782.
[111] Takahashi, H.; Okamuro, A.; Minagawa, J.; Takahashi, Y. Biochemical characterization of photosystem I-associated light-harvesting complexes I and II isolated from state 2 cells of Chlamydomonas reinhardtii. *Plant Cell Physiol.*, **2014**, *55*, 1437-1449.
[112] Nilsson, A.; Stys, D.; Drakenberg, T.; Spangfort, M.D.; Forsen, S.; Allen, J.F. Phosphorylation controls the three-dimensional structure of plant light harvesting complex II. *J. Biol. Chem.*, **1997**, *272*, 18350-18357.
[113] Dockter, C.; Muller, A.H.; Dietz, C.; Volkov, A.; Polyhach, Y.; Jeschke, G.; Paulsen, H. Rigid core and flexible terminus: structure of solubilized light-harvesting chlorophyll a/b complex (LHCII) measured by EPR. *J. Biol. Chem.*, **2012**, *287*, 2915-2925.
[114] Fehr, N.; Dietz, C.; Polyhach, Y.; von Hagens, T.; Jeschke, G.; Paulsen, H. Modeling of the N-terminal Section and the Lumenal Loop of Trimeric Light Harvesting Complex II (LHCII) by Using EPR. *J. Biol. Chem.*, **2015**, *290*, 26007-26020.
[115] Ding, J.H.; Li, N.; Wang, M.L.; Zhang, Y.; Lu, S.Q.; Long, M. The impact of N-terminal phosphorylation on LHCII conformation in state transition. *Acta Mech. Sinica*, **2014**, *30*, 447-456.

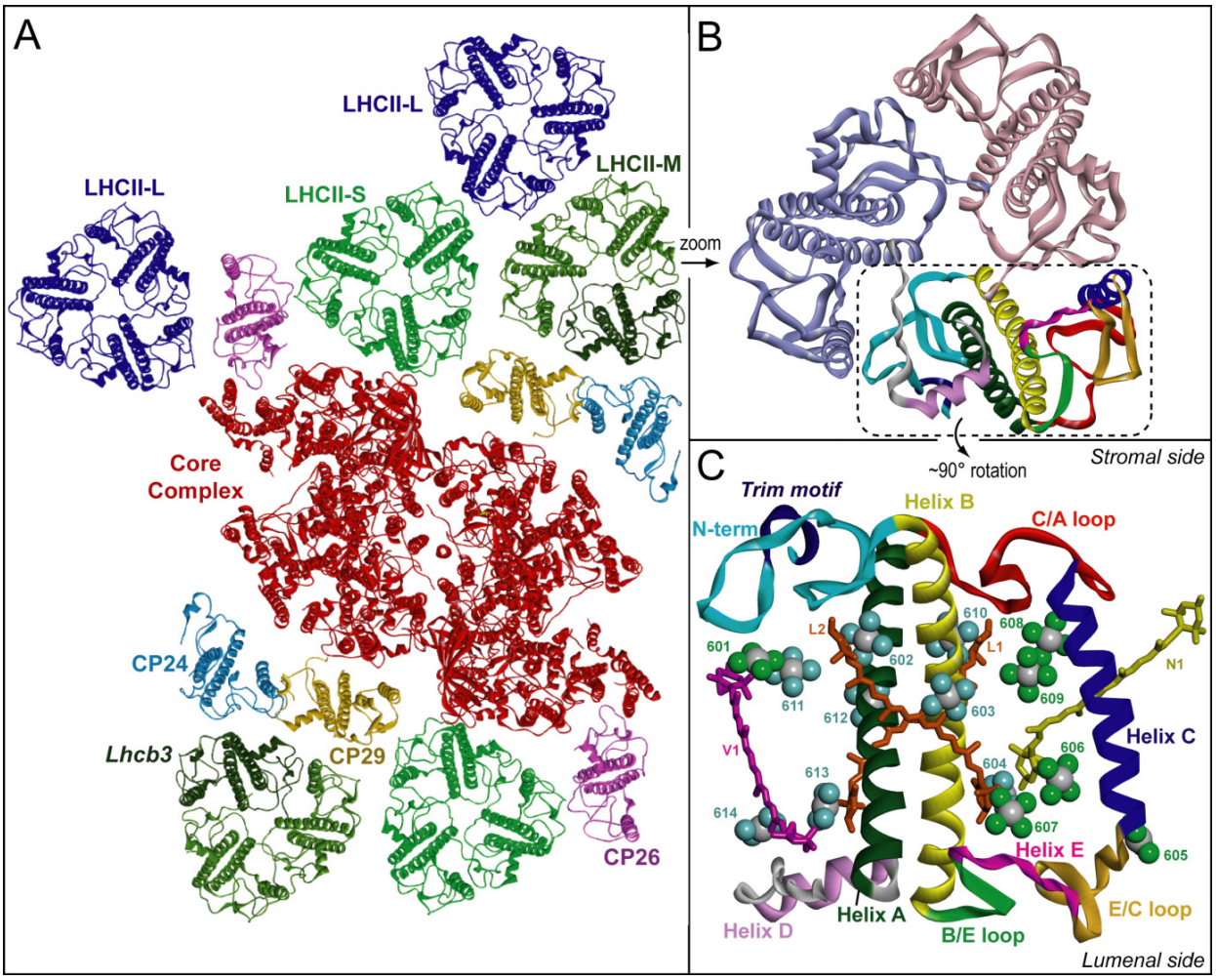

Fig. (1). A) Top view of dimeric PSII from the lumenal side, normal to the membrane plane. Positions of LHCII-S (strongly bound), LHCII-M (moderately bound) and monomeric Lhc (CP24, CP26 and CP29) are based on PDB model 5MDX [5]; Lhcb3 is the dark green monomer in LHCII-M. Two positions of LHCII-L (loosely bound trimers) is based on electron microscopy data on PSII megacomplexes)[21].The core complex is shown in red. **B)** Zoom on a LHCII trimer in the same orientation as LHCII-M of panel A. Two monomers are coloured in violet and pink, while the third monomer is coloured based on specific regions which are labelled in panel C. **C)** LHCII monomer of panel B rotated of ~90°, based on the PDB model 1RWT [8]. Specific regions of the protein are shown in different colours: N-terminus, light blue (with the trimerization motif in dark blue); helix B, yellow; B/E loop, light green; helix E, magenta; E/C loop, orange; helix C, blue; C/A loop, red; helix A, dark green; helix D, pink. Chlorophylls *a* (light blue) and *b* (green) are shown for simplicity by only four atoms: the central magnesium and the 4 nitrogen atoms (NA, NB, NC, ND). Lutein in site L1 and L2, neoxanthin is site N1 and violaxanthin in site V1 are also shown.

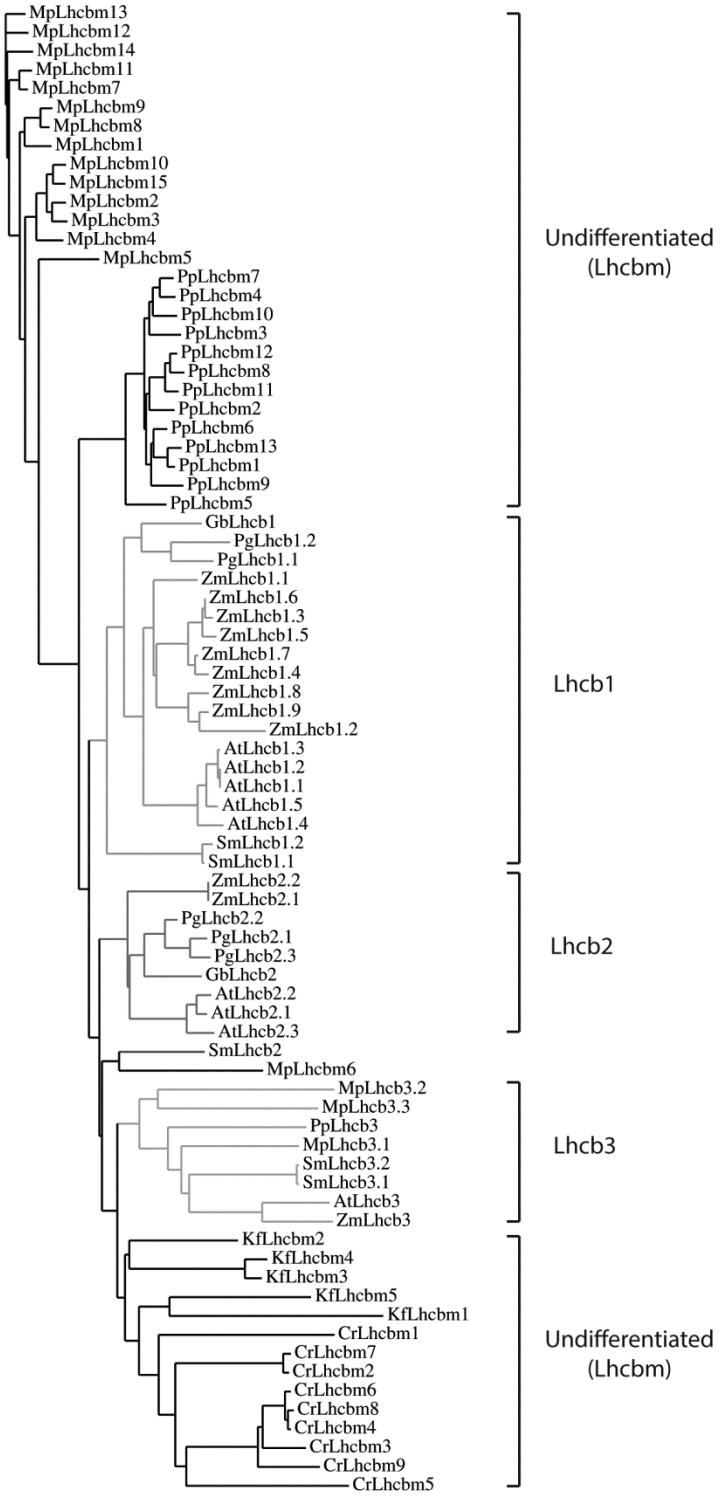

Fig. (2). Phylogenetic tree of the retrieved LHCII sequences. Branch length is proportional to the number of substitutions per site. Cr: *Chlamydomonas reinhardtii*; Kf: *Klebsormidium flaccidum*; Mp: *Marchantia polymorpha*; Pp: *Physcomitrella patens*; Sm: *Selaginella moellendorffii*; Gb: *Gingko biloba*; Pg: *Picea glauca*; Zm: *Zea mays*; At: *Arabidopsis thaliana*. The tree was realized using Phylogeny.fr [42].

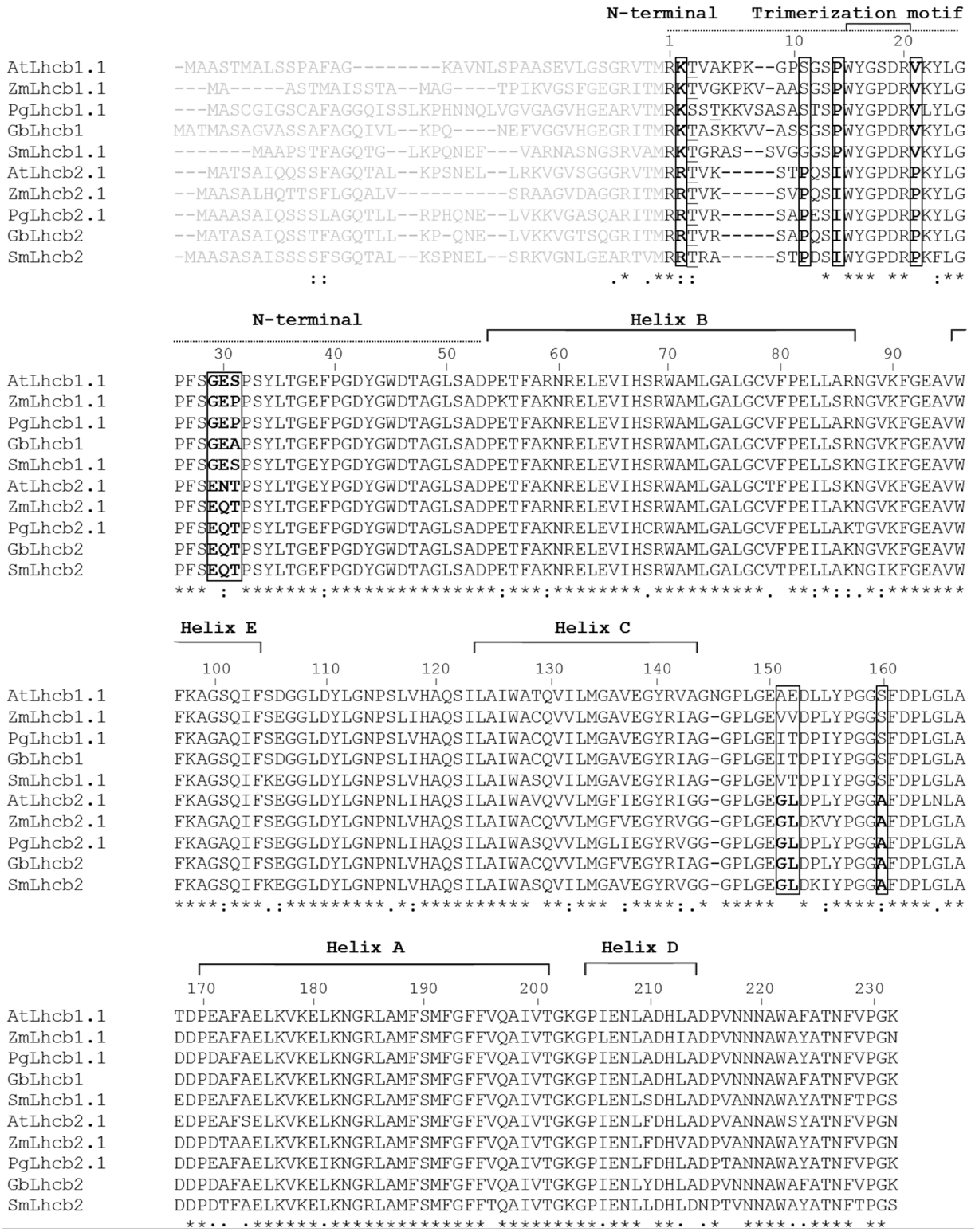

Fig (3). Identification of Lhcb1- and Lhcb2-specific amino acids. One Lhcb1 and one Lhcb2 sequence from each vascular plant analyzed were aligned. The specificity of the considered amino acids was verified on all sequences. Numbering is indicated for the AtLhcb1.1 sequence. Light grey: signal peptide. Underlined: phosphorylatable threonine (Thr3). Specific amino acids of Lhcb1 or Lhcb2 are in bold and in boxes. Abbreviations of plant species are as in Fig. 2.

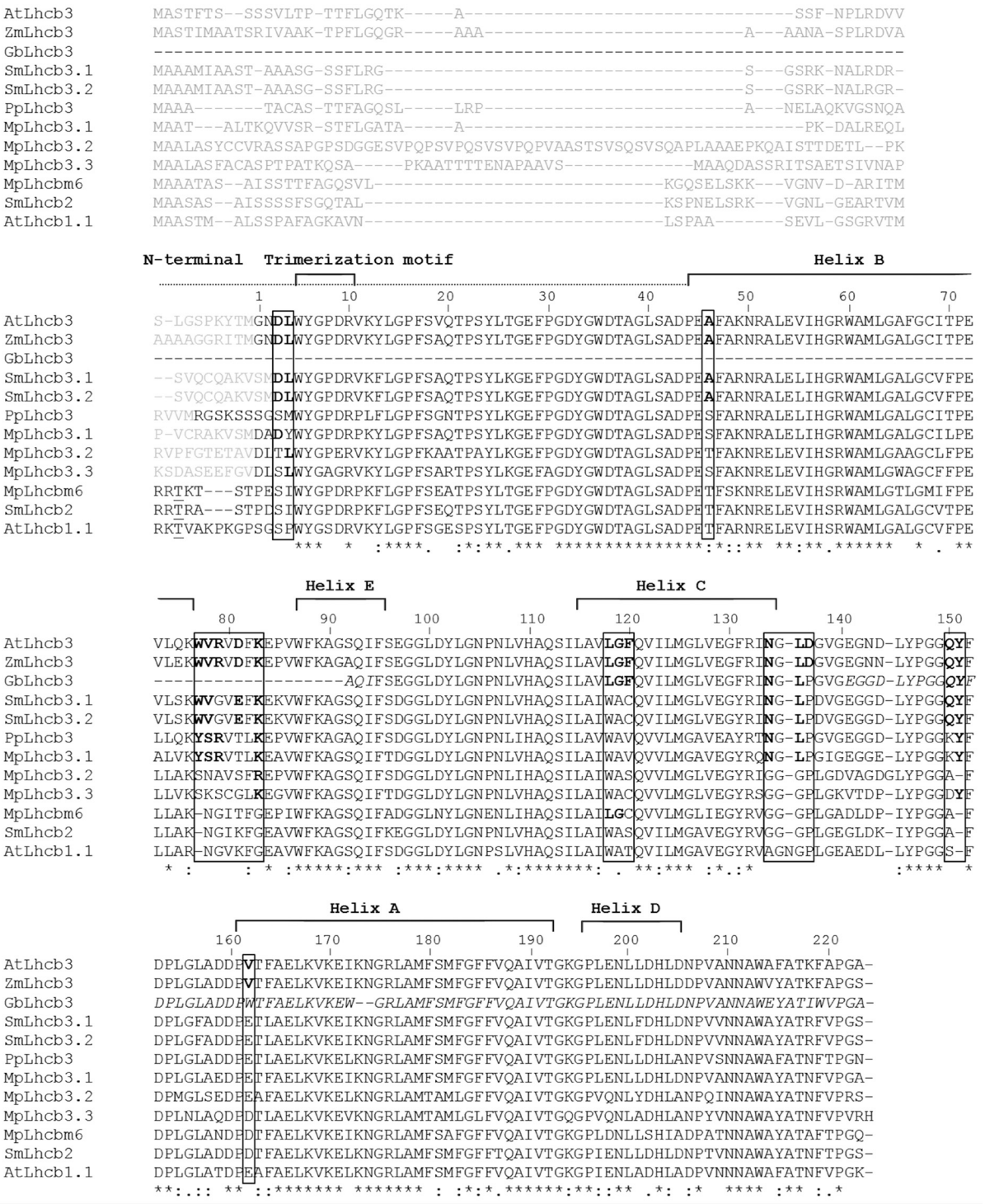

Fig. (4). Determination of Lhcb3-specific amino acids. All Lhcb3 sequences were aligned; two phylogenetically closest sequences, MpLhcbm6 and SmLhcb2, as well as AtLhcb1.1 were also added for comparison. Amino acids numbers are based on AtLhcb3 sequence. Light grey: signal peptide. Underlined: phosphorylatable threonine. Bold and in boxes: specific amino acids of Lhcb3, based on angiosperms sequences. GbLhcb3 is incomplete in the EST database, and was partially completed (residues in italics) with the sequence determined in [30]. Abbreviations of plant species are as in Fig. 2.

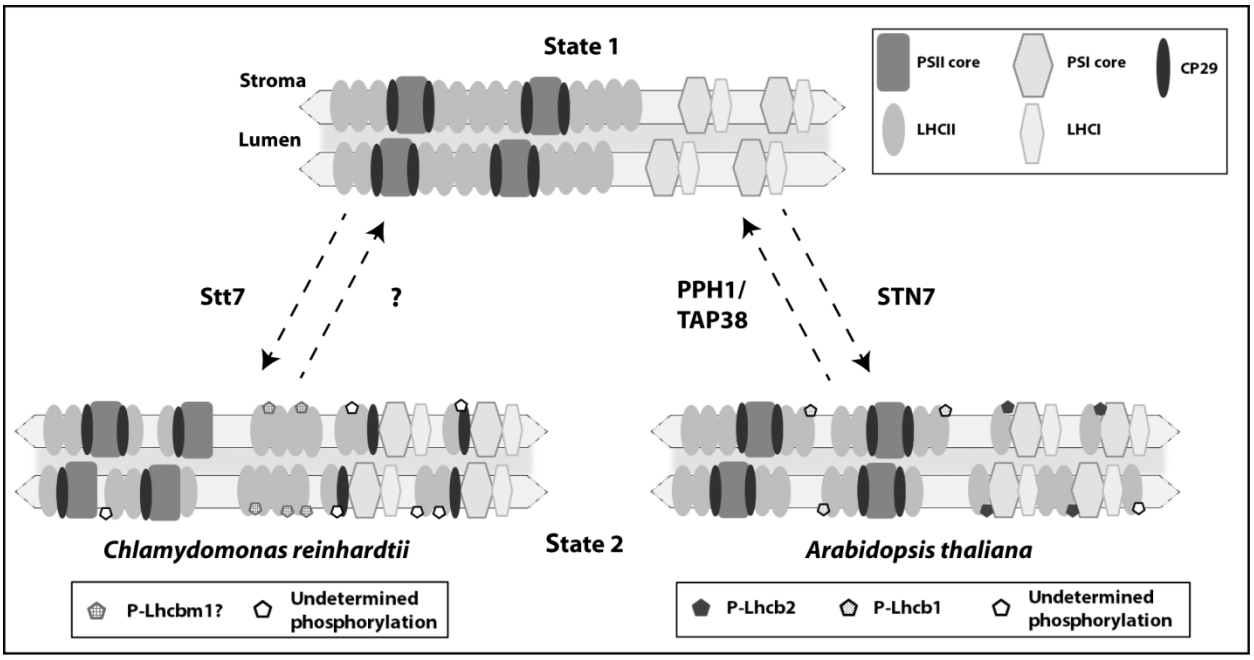

Fig. (5). Model of state transitions in *Chlamydomonas reinhardtii* and *Arabidopsis thaliana*. In state 1, in both organisms, LHCII is dephosphorylated and transfers energy to PSII. Upon transition to state 2, LHCII trimers are phosphorylated by the Stt7/STN7 kinase and part of them detaches from PSII and binds to PSI. In *Chlamydomonas*, PSI binds one to two LHCII trimers, as well as CP29. In this organism, a part of the LHCII trimers aggregates and switches to a quenched state; the Lhcbm1 isoform might be involved in this process. In *Arabidopsis*, PSI binds one LHCII trimer on its core side via a phosphorylated Lhcb2 isoform. At least one other trimer has been suggested to bind on the LHCI antenna side, but no information is available concerning the phosphorylation of this additional trimer. In *Arabidopsis*, phosphorylated Lhcb1 isoforms are mainly found in PSII supercomplexes, but their specific role remains to be determined. The state transitions process is reversed in *Arabidopsis* by the PPH1 (TAP38) phosphatase and by a still unknown phosphatase in *Chlamydomonas*.