

BBABIO 43970

Review

The light-harvesting chlorophyll *a* / *b*-binding proteins

Stefan Jansson *

Department of Plant Physiology, University of Umeå, S-901 87 Umeå (Sweden)

(Received 16 July 1993)

Key words: Light-harvesting chlorophyll *a* / *b*-binding protein; Antenna protein; LHC I; LHC II; Chlorophyll–protein complex

Contents

I. Introduction	2
II. The LHC proteins	2
A. Lhca1	3
B. Lhca2	4
C. Lhca3	4
D. Lhca4	5
E. Lhca1 + Lhca4 = LHCI-730	5
F. Lhca2 + Lhca3 = LHCI-680	5
G. Lhcb1	5
H. Lhcb2	6
I. Lhcb1 + Lhcb2 = LHC II	7
J. Lhcb3	8
K. Lhcb4	8
L. Lhcb5	9
M. Lhcb6	10
III. Are there additional LHC proteins?	11
A. A third LHCI-730 protein – A contaminant?	11
B. Psaf in LHC I?	11
C. A small LHC I component – Psae?	11
D. LHC Iie – An eleventh LHC protein and/or an ELIP?	11
IV. The LHC proteins of lower plants	12
V. The three-dimensional structure	12
VI. The light-harvesting antenna of higher plants	13
A. Photosystem I	13
B. Photosystem II	14
C. A consensus model	15
D. Can we correlate the protein complexes with freeze-fracture particles?	16
VII. Concluding remarks	16
Acknowledgements	17
References	17

* Corresponding author. Fax: +46 90 166676.

Abbreviations: DCCD, dicyclohexylcarbodiimide; Deriphat-PAGE, *N*-lauryl- β -iminodipropionate polyacrylamide gel electrophoresis; EFs, exoplasmic fracture face, stacked thylakoid region; EFu, exoplasmic fracture face, unstacked thylakoid region; ELIP, early light-induced protein; IEF, isoelectric focusing; ImL, intermittent light; LHC, light-harvesting chlorophyll *a* / *b*-binding; LHC I, light-harvesting complex of PS I; LHC II, light-harvesting complex of PS II; MSR, membrane spanning region; PAGE, polyacrylamide gel electrophoresis; PFu, protoplasmic fracture face, unstacked thylakoid region; PFs, protoplasmic fracture face, stacked thylakoid region; PS I, Photosystem I; PS II, Photosystem II; RC I, reaction centre of PS I; RC II, reaction centre of PS II. Gene nomenclature follows Refs. 1 and 2.

I. Introduction

The light-harvesting chlorophyll *a/b*-binding (LHC) proteins function in both Photosystem (PS) I and II as coordinators of antenna pigments. They constitute a family of abundant and important but still rather poorly characterised proteins, and the exact number of different LHC proteins in the thylakoid membrane of higher plants was unknown until recently. Their stoichiometry, specific functions and localisation within PS I and PS II remain to be established, and the well-known concepts 'PS II α ', 'PS II β ' and 'peripheral antenna' have not been defined in terms of LHC protein composition.

Many problems in this field can be attributed to four properties of the LHC protein family. Firstly, all members of the family are of a similar size, and many comigrate in polyacrylamide gels. Moreover, the apparent molecular masses (both absolute and relative) are dependent both on the gel system and the plant species used. Consequently, gel electrophoresis is not an ideal way of detecting the different LHC proteins. Secondly, since they all have sequence homology, antibodies raised against any of the proteins will typically cross-react with several other members of the family. Thirdly, the pigments are attached to their apoproteins by non-covalent bonds. Even with modern, low-denaturing separation techniques such as Deriphat-PAGE [3] and flat-bed isoelectric focusing (IEF) [4], the use of detergents to solubilise the complexes is inevitable and leads to significant pigment loss. Finally, many of the proteins can appear after isolation both as monomers and as multimeric complexes that may or may not exist *in vivo*.

The nomenclature for the LHC proteins has consequently been one of the more confusing areas in plant molecular biology, but recent studies at both the protein and DNA level have made it possible for a large number of research groups to finally agree on a new, comprehensive nomenclature for the genes encoding the LHC proteins (Table I) [2]. As yet, this classification only applies to the genes, but the corresponding names for the proteins will be used in this review to avoid ambiguities (e.g., Lhca1 for the protein encoded by the *Lhca1* gene). However, as some designations for chlorophyll-protein complexes are useful and widespread they will be used occasionally. LHC I is here defined as the four Lhca polypeptides, while CP29, CP26 and CP24 are the chlorophyll-protein complexes of the gene products of *Lhcb4*, *Lhcb5* and *Lhcb6*, respectively. The meaning of LHC II, on the other hand, is quite ambiguous. It was originally used to designate the chlorophyll *a/b*-binding pigment-protein complex of PS II that is reversibly phosphorylated, migrates as an oligomer in many polyacrylamide gel systems, and that precipitates from a Triton X-100-solubilised thylakoid solution at low concentrations of

TABLE I

A conversion table between the new nomenclature and the designations found in Green et al. [99], Thornber et al. [30], Bassi et al. [39] and Harrison and Melis [75]; adopted from [2]

Gene product	Green et al.	Thornber et al.	Bassi et al.	Harrison and Melis
Lhca1	Type I LHCI	LHC Ib	LHCI-730	–
Lhca2	Type II LHCI	?	LHCI-680	–
Lhca3	Type III LHCI	LHC Ia	LHCI-680	–
Lhca4	Type IV LHCI	LHC Ib	LHCI-730	–
Lhcb1	Type I LHCII	LHC IIb 28 kDa	LHCII	b
Lhcb2	Type II LHCII	LHC IIb 27 kDa	LHCII	c
Lhcb3	Type III LHCII	LHC IIb 25 kDa	LHCIIa	d?
Lhcb4	Type II CP29	LHC IIa	CP29	a
Lhcb5	Type I CP29	LHC IIc	CP26	?
Lhcb6	CP24	LHC IId	CP24	e

Mg²⁺. It is now known that LHC II corresponds to a mixture of Lhcb1 and Lhcb2. However, Lhcb3, Lhcb4, Lhcb5 and Lhcb6 are also LHC proteins of PS II which are found to various extents in 'LHC II' preparations. For this reason, the name 'LHC II' has deliberately been used very restrictively in this review.

Table I summarises the new nomenclature and makes reference to some old designations. It is important to note that the LHC proteins of ferns, mosses and algae can *not* be classified using the same system. This certainly does not mean that the light-harvesting antenna of lower plants is fundamentally different from that of angiosperms and gymnosperms, but its protein components are nevertheless sufficiently different to prevent us from *directly* applying this nomenclature to *all* genes encoding LHC proteins from lower plants. However, it is possible that in the future we will find *Lhc* genes in lower plants that are almost identical to their higher plant counterparts.

In this article, I will first review the data obtained for the ten different LHC proteins of higher plants, then briefly describe the LHC proteins of lower plants and discuss the structure of the higher plant light-harvesting antenna. This review will focus on the progress made after 1988, when a review article had the title 'The chlorophyll-protein complexes of higher plant photosynthetic membranes *or* Just what green band is that?' [5]. Today, this question has been answered, and we can instead specifically address more important and complex issues concerning the function of the higher plant light-harvesting antenna.

II. The LHC proteins

The LHC proteins are encoded by nuclear genes, synthesised in the cytosol and post-translationally assimilated by the chloroplast and inserted into the thylakoid membrane. The translocation process is medi-

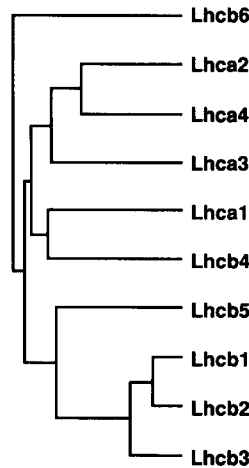


Fig. 1. Sequence homologies between the higher plant LHC proteins. The dendrogram has been created using the PILEUP program of the GCG package [81], and the length of the horizontal axis is proportional to the difference between the sequences of the mature LHC proteins.

ated by an N-terminal extension of the proteins, the transit peptide, which is cleaved from the mature LHC protein during the process. All amino acid sequences shown in this review are for mature polypeptides only. The structure of the transit peptides and the associated processes have been reviewed recently [6], and will not be covered here. The numerous studies of the regulation of these genes will also not be covered, and readers interested in this subject are recommended to read Refs. 7 and 8.

All higher plant LHC proteins have sequence homology, and have no doubt evolved from a common ancestor. Since we know their primary sequences, it is possible to calculate the similarities between all higher plant LHC proteins. In the dendrogram of Fig. 1, the lengths of the horizontal axes are proportional to the difference between the sequences of the mature LHC proteins. Some conclusions that can be drawn from this will be mentioned in the text.

Since there is no sequence similarity between the LHC proteins and any known pigment-binding protein

TABLE II

Comparison of the LHC proteins

Protein	Amino acids	M_r (kDa)	App. M_r (kDa)	Theor. pI	Actual pI	N-terminal blockage	No. of genes
Lhca1	201	22	20–22	5.1–5.7	?	No	1–2
Lhca2	211	23	19–23	4.6–4.9	?	No	1
Lhca3	233–234 ^a	25 ^a	23–25	5.7–6.3 ^a	?	Yes	1–2
Lhca4	199–200	22	18–21	4.8–4.9	?	No	1–2
Lhcb1	230–233	25	27–28	4.7–5.4	4.0–4.2	Yes	3–16
Lhcb2	228	25	25–27	4.7–5.2	4.0–4.2	Yes	1–4
Lhcb3	223	24	24–25	4.5–4.6	3.9–4.1	No	1–4
Lhcb4	257–258 ^a	28 ^a	29–31	4.9–5.0 ^a	4.7	Yes	?
Lhcb5	246–251 ^a	27 ^a	26–29	4.8–5.2 ^a	4.3	Yes	1
Lhcb6	210–211	23	20–22	4.8	4.4	No	2

The values of apparent molecular mass and pI are only approximate, the actual values vary with the experimental conditions.

^a Approximate values, N-terminus of the mature protein is not known.

of cyanobacteria, the evolutionary origin of the LHC polypeptides has previously not been understood. However, the recently characterised PsbS protein, a 22 kDa protein of PS II that has been hypothesised to have a role in the assembly of PS II [9], shows significant sequence homology to the LHC proteins [10,11]. The fact that this protein also is found in cyanobacteria makes it a strong candidate for the LHC ancestor [10].

The following section will describe the higher plant LHC polypeptides one by one, starting with a short summary in Table II. Three supramolecular complexes, LHCI-730, LHCI-680 and the Lhcb1/Lhcb2 complex, are described separately.

II.A. Lhca1

Lhca1 genes (or cDNAs) have so far been characterised from four species; tomato [12], *Arabidopsis thaliana* [13], tobacco [14] and Scots pine [15]. Lhca1 is part of the LHCI-730 complex (section II.E), and the N-terminus of the mature Lhca1 (as well as Lhca2 and Lhca4) protein has been experimentally determined in

	10	20	30	40	50	60
Tomato	SADWMPGQPRPSYLDGSAPGDFGFDPLGLGEV	PANLERYKESELIHCRWAMLA	VPGIIVPEALGLGN			
Tobacco
Arabid.	A.H.....E.....A.....
S. pine	..E.....PH.....
	70	80	90	100	110	120
Tomato	WVKAQEWAAIPGGQATYLGQFVPWGTLP	TILAEFLAIAFVEHQ	RSMEKDSKKYPGGAF	PLGYS		
Tobacco
Arabid.
S. pine	..E..K...I...S....F.V....Y..IV..V.....	S..NG.P.P..R.....F.				
	140	150	160	170	180	190
Tomato	KDPAKFEEELKVKEIKNGRLALLAFVGF	CVQSSAYPGTGPLENLATHLADPWHNNIGDVI	IPKGI	FPN		
Tobacco	...K.....
Arabid.	...K.L.....
S. pine	...V..K.Y.....	A..N.....	S....R..AEI...RSL..E			

Fig. 2. Amino acid sequence of Lhca1 from tomato [12], tobacco [14], *Arabidopsis thaliana* [13] and Scots pine [15].

	10	20	30	40	50	60	70
Tomato	AADPDRPLWFP	GSTPPWLDGSL	PGDFGFDPL	GLASDPESLR	WNQQAELVHCR	WAMLGAAGIF	IPELLTK
<i>Petunia</i>E.....G.....K.A.....S.....F.....
<i>S. pine</i>	.VQTE.....	.N.....G.....	.T.K.MV.....C.....
	80	90	100	110	120	130	140
Tomato	IGILNTPSWYT	AGEQEYFTDTT	TLFIVELVLIG	WAEGRRWADI	IKPGCVNTDPI	FPNNKLTGTD	VGYPGG
<i>Petunia</i>V.....VI.....
<i>S. pine</i>	L.....L.....K.....	.V.MIFL.....LE..S.....
	150	160	170	180	190	200	210
Tomato	LWFDPLGWGSG	SPAKIKELRTKE	IKNGRLAMLA	VMGAWFQHIY	TGTGPIDNLFA	HLADPGHATIF	AAFSPK
<i>Petunia</i>
<i>S. pine</i>E.....AE.....L.....

Fig. 3. Amino acid sequence of Lhca2 from tomato [21], *Petunia* [20] and Scots pine [15].

spinach, pea [16] and barley [17]. Lhca1 consists of 200 amino acids (Fig. 2) and is, along with Lhca4, the smallest of the LHC proteins. Fig. 1 shows that Lhca1 is more related to the largest, Lhcb4, than to the other Lhca proteins. Considerable sequence similarity to Lhcb4 also exists in the central region (residue 70–120 in Fig. 2) where Lhca1 shows a very limited similarity to the other Lhca sequences. This suggests that Lhca1 and Lhcb4 may have similar functions.

No reliable data are available on the pigment content of this, or any other, LHC I protein. Two *Lhca1* genes are present in tomato [18], one in *Arabidopsis thaliana* [19] and at least two in Scots pine (Jansson and Gustafsson, unpublished).

II.B. Lhca2

The *Lhca2* gene codes for a slightly larger LHC I polypeptide (Fig. 3) that is associated with the LHCI-680 complex (see subsection II.F). *Lhca2* sequences have been obtained from *Petunia* [20], tomato [21] and Scots pine [15], and the mature protein consists of 211 amino acids. *Lhca2* is a single-copy gene in both tomato [21], *Arabidopsis thaliana* [19] and Scots pine (Jansson and Gustafsson, unpublished). In hindsight, it is possible that the reported number of seven homologous genes in *Petunia* [20] is an over-estimation. It is quite likely that a number of hybridising bands in that report originate from *Lhca1*, *Lhca3* and *Lhca4*, rather than from additional *Lhca2* genes. In tomato, the

Lhca2 gene has the highest level of expression of all *Lhca* genes, measured on mRNA level [22]. Evidence for an association between Lhca2 and the PsaE protein has recently been presented [17]. Interestingly, Lhca2 has a higher similarity to Lhca4, which is part of LHCI-730, than to the other LHCI-680 protein, Lhca3 (Fig. 1).

II.C. Lhca3

The Lhca3 polypeptide is the largest of the LHC I proteins (Fig. 4), and corresponding genes or cDNAs have been isolated from tomato [23], pea (Thornber, personal communication) and Scots pine [15]. The N-terminal sequence of this protein is not known, due to N-terminal blocking of the protein. However, sequencing of proteolytic fragments has shown that this protein is the larger of the two LHCI-680 proteins [16]. Evidence has been presented that Lhca3 can be isolated in an oligomeric form after electrophoresis [17]. Only one *Lhca3* gene is present in tomato [23] and *Arabidopsis thaliana* [19], while Scots pine contains two (Jansson and Gustafsson, unpublished).

DCCD-binding experiments have suggested that Lhca3 and Lhca4 are important for proton funnelling from the water-splitting complex of PS II to the lumen [24]. It is hard to imagine how LHC I proteins could modulate the activity of PS II, but it is possible that future work will clarify this issue.

	10	20	30	40	50	60	70
Tomato	-ASTP-PVKQA	-NRQLWFASKQ	SLSYLDGSL	PGDFGFDPL	GLSDPEGTGG	FIEPKWLAYGE	VINGRFAMLGAAGAIAP
Pea	-AA.....	GVD.P.....Y.....R.....V.....
<i>S. pine</i>	Q..SNK.....	--D.....Y.....M.....	P...M...G..V...	I...Y...V.....
	80	90	100	110	120	130	140
Tomato	EILGKAGLIPQ	ETALAWFQTG	VIPAGTYNYW	ADNYTLFVLE	MALMGFAEHRR	FQDWAKPGSMG	KYFLGLEKLGGS
Pea	.Y...V.....F...
<i>S. pine</i>L.....P.....PF.....L.....	YRN.....F.....A...
<i>Arabid.</i>
	160	170	180	190	200	210	220
Tomato	GDPAYPGG	PLFNPLGFGK	DEKSMKELKL	KEIKNGRLAM	LAILGYFIQAL	VTGVGPYQNLL	DHLADPVNNNVLTSLK
Pea	.N.....	F.....L.....V.....G.....V.....
<i>S. pine</i>	.N.....V.....G.....F.....H..F..NM..L..
<i>Arabid.</i>	.N.....	F.....L.....V.....G.....

Fig. 4. Amino acid sequence of Lhca3 from tomato [23], pea (Thornber, personal communication), Scots pine [15] and a partial *Arabidopsis thaliana* sequence [EMBL Z17941]. Since the transit peptide cleavage site is not known, the putative N-terminus of the protein is shown.

	10	20	30	40	50	60
Tomato	KKGQWLPLGLASPDYLDGSLPGDNGFDPLGLVEDPENLKWFIQAELVNGRWAMLGVAGMLLPEVFTSI					
Arabid.	K..E.....T...A.....A.....V.....K.					
S. pine	-..E.....S..S..N.....A....S...YV.....I...L...					
	70	80	90	100	110	120
Tomato	GILNVPKWYDAGKSEYFASSTLFVIEFILFHYVEIRRWQDIKNPGSVNQDPIFKNYSLPNKCQGY					
Arabid.	..I...E.....EQ.....Q.....KGEV...					
S. pine	.LI.....V.....L.....Y.....L..Q.....EV...					
	140	150	160	170	180	190
Tomato	GGIFNPLNFAPTEEAKEKELANGLAMLAFLGFIQHNVTGKGFDFNLLQHLSDPWHNTIIQTLSN					
Arabid.Q.....V.....E.....V..F-N					
S. pineS.SM.....V.....V.....Q GK					

Fig. 5. Amino acid sequence of Lhca4 from tomato [25], *Arabidopsis thaliana* [19] and Scots pine (Jansson and Gustafsson, unpublished).

II.D. Lhca4

The Lhca4 protein is the second constituent of LHCI-730 [16], where it has been reported to form a dimeric complex with Lhca1 [17]. *Lhca4* genes and/or cDNAs have so far been characterised from tomato [25], *Arabidopsis thaliana* [26] and Scots pine (Jansson and Gustafsson, unpublished). Tomato contains two *Lhca4* genes [25], but *Arabidopsis thaliana* [19] and Scots pine (Jansson and Gustafsson, unpublished) contain only one. The mature Lhca4 protein consists of 199–200 amino acids (Fig. 5) and frequently co-migrates with Lhca1 during SDS-PAGE.

II.E. Lhca1 + Lhca4 = LHCI-730

Lhca1 and Lhca4 are thought to be associated in the pigment–protein complex LHCI-730 [16,17], which was previously described under the name LHCPIb [27]. Also named LHC Ib, this complex is characterised by a strong 77 K fluorescence emission peak at 730 nm, while the other LHC I subcomplex, LHCI-680, fluoresces at 680 nm. It has been hypothesised that all LHC I sub-complexes originally emit fluorescence at a long wavelength, but that absence of this fluorescence in LHCI-680 is a consequence of pigment loss [16]. LHCI-730 has been reported to contain only one polypeptide [27] but we now know that this is due to co-migration of Lhca1 and Lhca4. The presence of three polypeptides in LHCI-730 has also been reported [28] but this too is likely to be an error (section III.A).

One estimate of the amount of pigment bound to LHCI-730 is 0.5 neoxanthin, 11 violaxanthin, 17.3 lutein, 2.4 β -carotene and 40 Chl *b* per 100 Chl *a* molecules [29], and another gave a Chl *a/b*-ratio of 2.3 [30].

II.F. Lhca2 + Lhca3 = LHCI-680

The pigment–protein complex LHCI-680, or LHCPIa [27] consists of two polypeptides, Lhca2 and Lhca3 [16,17]. LHCI-680 can be further fractionated into LHCI-680A, containing Lhca3, and LHCI-680B, containing Lhca2 [17], suggesting that these two proteins are not intimately associated. A complex, named LHC Ia, corresponding to either LHCI-680 or LHCI-680A, has been identified in barley and is reported to have a Chl *a/b*-ratio of 1.4 [30].

Bassi and Simpson [28] suggested that excitation energy is transferred from 'LHC II' (Lhcb1 and Lhcb2) to the reaction centre (RC) of PS I via the path LHC II \rightarrow LHCI-680 \rightarrow LHCI-730 \rightarrow RC I. The observation that LHCI-730 is synthesised before LHCI-680 (at least Lhca2) when etiolated soybean seedlings are transferred to light [31] supports this proposal. Conversely, later studies have indicated that LHCI-680 might be a link between LHCI-730 and RC I [32]. It is also quite possible that LHCI-680 and LHCI-730 independently deliver excitation energy to the core antenna. Nevertheless, Lhcb1/Lhcb2 appears to associate with LHCI-680, not with LHCI-730 [32].

	10	20	30	40	50	60	70
Tomato	RKTATKAKPAS-SSSPWYGPDRVKYLGPFSGESPSYLTGEFP GDYGDWDTAGLSADPETFAKNRELEVIHCRWAMLGAL						
Rice	...A.P...AS.G...A...L.....P.....S.....						
S. pine	..AT..KLT.SA.T.....L.....P.....N.....						
	80	90	100	110	120	130	140
Tomato	GCVFPELLARNGVKFGEAVWFKAGSQIFSEGLDYLGNPSLVHAQSILAIWACQVILMGAVEGYRIAGGPLGEVTDPL						
RiceA.....I.....V.....						
S. pineA.....I.....V.....						
	160	170	180	190	200	210	220
Tomato	YPGGSFDPLGLADDPEAFELKVKEIKNGRLAMFSMGFFVQAIVTGKGPLENLADHLADP VNNNAWAFATNFVPGK						
Rice	...A.....C.....Y.....Y.....						
S. pineE.....I.....Y.....						

Fig. 6. Amino acid sequence of Lhcb1 from a dicot; tomato (Lhcb1 * 7, or cab3C) [125], a monocot; rice [126] and a gymnosperm; Scots pine (Lhcb1 * 2) [50].

It was believed for a long time that the precursor of Lhcb1 could be processed at multiple sites. In vitro studies showed that one *Lhcb1* gene product, when taken up by intact chloroplasts or isolated thylakoids, gave rise to polypeptides of different electrophoretic mobility [41,42]. The 'secondary site' was experimentally determined [43] and the processing protein partially purified [44]. However, other in vitro studies have shown that one precursor corresponds to one polypeptide only. It appears today that processing at the sec-

Lhcb1 is very similar to Lhcb2, but 14 distinctive amino acids (underlined in Fig. 7) have been identified [50]. Most protocols for preparation of 'LHC II' mainly yield Lhcb1 and Lhcb2, while various amounts of Lhcb3 and other Lhcb proteins are found as well. The native form of the protein is most likely a trimer, although dissociation into monomers and dimeric complexes is readily obtained upon solubilisation. The ratio between the monomeric, dimeric and trimeric forms differs from species to species and is also dependent on detergent concentration [51].

Lhcbl is reversibly phosphorylated at a Thr (or Ser) residue close to the N-terminus of the protein [40,52]. This phosphorylation is discussed further in section II.I.

II.H. *Lhcb2*

The *Lhcb2* genes code for the second-most abundant LHC polypeptide (Fig. 7), the '25 kDa LHC II protein' [45]. This polypeptide is slightly smaller than Lhcb1 (approx. 228 amino acids) and is found in amounts equivalent to roughly one third of Lhcb1 [53,54]. *Lhcb2* genes or cDNAs, formerly called Type II LHC II, have been characterised from 10 species [2]. In angiosperms, they are present in one or two copies per genome, while the copy number in Scots pine appears to be larger (Jansson and Gustafsson, unpublished).

	10	20	30	40	50	60	70		
Tomato	RRTVKSAPQSIWYGEDR	KYLGPFS	EQTPSYLTGE	FPDYGWD	TAGLSAD	PETFARN	RELEVIHCRWAMLGALGC		
RiceP.....L.....S.....							
B. pineR.....E.....P.....G.....				K.....		
	80	90	100	110	120	130	140	150	
Tomato	FPEILSK	NGVKVFGE	AVWFKAGSQ	IFSEGGLD	LYLGNPN	LVHAQSILAI	WACQVVL	MGFVEGVRVGGPLGEGGDKIY	
RiceV.....						V.....	
B. pineL.....A.....			I.....	LI.....PL.....	
	160	170	180	190	200	210	220		
Tomato	PGGA	FDPLGLADD	PEAF	AELKVKEIK	NGRLAMF	SMFGFFVQ	AIVTGKGP	IENTLSDHINDPVANN	AWAYATNFVPGK
RiceDT.....	L.....		R.....	F.....VT.....	
B. pineY.....LA.....		

Fig. 7. Amino acid sequence of Lhcb2 from a dicot; tomato (Lhcb2 * 1 or cab4) [127], a monocot; rice [126] and a gymnosperm; black pine [128]. 'Type-specific' amino acids, distinguishing the protein from Lhcb1 [50] are underlined.

Many studies have focused on this protein [53–56], which is believed to be a part of the outermost, peripheral antenna of PS II. As mentioned above, it is similar to Lhcb1 in many respects, although there are several reports that the phosphorylation kinetics of Lhcb2 is faster than that of Lhcb1 [53,57]. However, the final degree of phosphorylation is the same; one third on a molar basis [57]. The phosphorylation site is the residue Thr³, as in Lhcb1 [40]. It is also believed that this protein is involved in long-term acclimation to different light regimes. More specifically, the amount of this protein increases, relative to Lhcb1, when the PS II antenna size is large [53,54].

II.1. *Lhcb1 + Lhcb2 = LHC II*

In most of the work dealing with ‘LHC II’, no distinction is made between Lhcb1 and Lhcb2. It is believed that they exist in mixed trimers and all four possible trimers are assumed to be present in vivo [53,56]. Most of the differences between Lhcb1 and Lhcb2 are found in the N-terminal region protruding from the membrane, with the membrane-spanning regions being more or less identical. It is likely that these minor differences correspond to important variations in the function of the proteins, since the 14 distinctive amino acids are highly conserved [50]. These two proteins have almost identical molecular mass and amino acid sequence, but resolve clearly in most polyacrylamide gel systems. This, together with the report that LHC II could be palmitoylated [58], has led some authors to suggest that the size heterogeneity of Lhcb1/Lhcb2 could be a consequence of palmitoylation [59]. However, it has been shown that neither Lhcb1 nor Lhcb2 is acylated [45], therefore palmitoylation cannot be the reason for this difference in gel mobility.

The precise pigment composition of Lhcb1 and Lhcb2 is still not known, although crystallographic studies indicate that 14 (or 15) chlorophyll molecules, probably 8 Chl *a* and 6 Chl *b*, are bound to each monomer [60]. The carotenoid content is also not known with certainty. Peter and Thornber [3] report 1 neoxanthin, 0.5 violaxanthin and 2 lutein molecules per monomer, whereas the results of Bassi et al. [61] indicate 0.5 neoxanthin, trace amounts of violaxanthin and 2 lutein molecules per monomer. Reconstitution experiments have shown that all three xanthophylls are required for the creation of a stable, monomeric Lhcb1 pigment–protein complex [62]. It is also known that violaxanthin is more loosely bound than the other carotenoids and that it dissociates during isolation [63]. This suggests that there are specific binding sites for all three carotenoids on each monomer (1 neoxanthin, 1 violaxanthin and 2 lutein), and that if sub-stoichiometrical amounts of these carotenoids are detected, some

pigment molecules have dissociated. Violaxanthin can be photo-converted into zeaxanthin via antheraxanthin in the xanthophyll cycle [64]. Since zeaxanthin is present in ‘LHC II’ preparations, it has been assumed that the xanthophyll cycle takes place in the major LHC complex (i.e., Lhcb1 and Lhcb2). A model for the interaction between zeaxanthin formation and energy dissipation in the complex has been presented [65]. However, it has recently been found that the fraction of violaxanthin present in the Lhcb1/Lhcb2 complex can *not* be converted to zeaxanthin. Instead, the zeaxanthin formation in PS II appears to take place in Lhcb4, Lhcb5 and Lhcb6 [66].

The phosphorylation of Lhcb1 and Lhcb2 has attracted the attention of many researchers, since it is believed to regulate State 1/State 2 transitions [52,67]. Phosphorylation is triggered by the redox state of a component in the photosynthetic electron transport chain close to, or inside, the cytochrome *b₆/f* complex. The result is a lateral movement of Lhcb1 and Lhcb2 from PS II in the grana regions into the PS I-enriched stroma-exposed regions [67]. It is likely that this phosphorylation is a key element for short-term regulation of energy balance between PS I and PS II.

It is known that Lhcb2 is phosphorylated more rapidly than Lhcb1 [55,57]. On the other hand, synthetic peptides corresponding to the Lhcb1 and Lhcb2 N-termini are phosphorylated at the same rate in vitro [40]. Thus, it is possible that sequence difference around residue 30 (GES vs. EQT) affects the structure of the N-terminus of Lhcb1 and Lhcb2, resulting in different phosphorylation kinetics.

Experiments have shown that one-third of all Lhcb1 and Lhcb2 proteins can be phosphorylated in the light [57]. A serious drawback of these studies is that they only measure *de novo* phosphorylation. One-third of the proteins corresponds to one polypeptide in each trimer being phosphorylated in the light. Whether or not the other two polypeptides of one trimer are already phosphorylated in darkness could not be deduced from these results. Evidence for Lhcb1 and/or Lhcb2 phosphorylation in the dark is provided by trypsin treatment removing the ability of Lhcb1/Lhcb2 to be precipitated with Mg²⁺ [68]. Mg²⁺ ions are thought to interact with negative charges in the LHC II proteins [69] and many LHC II purification protocols include precipitation with 5 mM Mg²⁺. Trypsin cleaves after Arg and Lys residues, and reduces the apparent molecular mass of Lhcb1/Lhcb2 by roughly 2 kDa, corresponding to approximately 20 amino acids. The actual site for the trypsin cleavage has not been experimentally determined, but it is likely that cleavage occurs after the residue Arg²². If this assumption is correct, five positive amino acids are removed but only one acidic residue, Asp²¹. It seems unlikely that the cation-dependent aggregation should be due to this

	10	20	30	40	50	60	70
Tomato	SNDLWYGFDRVKYLGPFSAQTPSYLNGEFGDYGWDTAGLSADPEAFKNRALEVIHGRWAMLGALGCIFFPEVL						
Pea	G.....T.....T.....						
Barley	G.....R.....V.....						
	80	90	100	110	120	130	140
Tomato	EKWKVDFKEPVWFKAGSQIFSDGGLDYLGNPNLVHAQSILAVLGFQVVMGLVEGFRINGLPGVGEGNDLYPG						
Pea	Q...R.....E.....I.....						
Barley	Q...G.E.....E.....L.....D.....						
	150	160	170	180	190	200	210
Tomato	GQYFDPLGLADDPPTFAELKVKEIKNGRLAMFSMFGFFVQAIVTGKGPLENLLDHLDNFVANNWVYATKFEVPGA						
PeaV.....						
BarleyV.....F...D.....F...A...S						

Fig. 8. Amino acid sequences of Lhcb3 from tomato [70], pea (EMBL X69215) and barley [71].

single residue. An alternative explanation is that the critical residue is a phosphorylated amino acid close to the N-terminus, and the absence of cation-dependent precipitation after trypsin treatment is circumstantial evidence of phosphorylation of Lhcb1/Lhcb2 in the dark. An accurate estimation of the phosphorylation level of Lhcb1 and Lhcb2 from dark-adapted plants might confirm this proposal.

II.J. Lhcb3

Lhcb3 is the smallest constituent of 'LHC II'. It is quite similar to the *Lhcb1* and *Lhcb2* gene products, and is often found in 'LHC II' preparations. However, it is not clear if this protein forms heterotrimers with Lhcb1/Lhcb2 (see below). *Lhcb3* genes have been isolated from tomato [70], pea (EMBL X69215) and barley [71]. Protein sequencing has also identified the protein in *Arabidopsis thaliana* [72], maize [73] and wheat [74], although the latter sequence was originally reported to correspond to Lhcb2. At least four *Lhcb3* genes are present in the tomato genome [46], while barley appears to have only one [71]. Lhcb3 is, in contrast to Lhcb1 and Lhcb2, neither N-terminally blocked nor reversibly phosphorylated, consistent with the absence of the N-terminal 10 amino acids found in Lhcb1 and Lhcb2 (Fig. 8) where phosphorylation takes place. Furthermore, Lhcb3 is tightly bound to RC II under all conditions, might not be depleted in the thylakoids of the chlorophyll *b*-less barley mutant *chlorina-f2* [75] and appears to be assembled early in PS II development [76].

It is not known whether Lhcb3 is present in trimers, and if so, if it is found in mixed trimers with Lhcb1 and Lhcb2. During Deriphat- and mild SDS-PAGE, it migrates together with the Lhcb1/Lhcb2 trimers [3,77] but in a sucrose gradient Lhcb3 migrates like the minor Lhcb proteins, whereas Lhcb1/Lhcb2 forms a larger complex [4]. This indicates that Lhcb3 could form a separate, monomeric complex under some experimental conditions. The chlorophyll content of Lhcb3 seems to be similar to that of Lhcb1/Lhcb2 [4].

Webber and Gray [74] presented evidence for the binding of Ca^{2+} ions to Lhcb3. These experiments were mainly performed in vitro, so non-specific binding cannot be ruled out. Nevertheless, it has been speculated that this Ca^{2+} could be important for the functioning of the water-splitting complex [71], which might explain why Lhcb3 never dissociates from RC II.

II.K. Lhcb4

Lhcb4 genes code for the apoprotein of CP29, or LHC IIa. This protein is the largest of the LHC proteins, approx. 257 amino acids, although the actual N-terminus of the mature protein is not known because of N-terminal blockage. The primary sequence of this protein differs from the other LHC proteins in having a 42 amino acid insertion close to the first predicted α -helical region (see Fig. 14, below). *Lhcb4* cDNAs have so far only been isolated from barley [78] and *Arabidopsis thaliana* (Green and Pichersky, personal communication) (Fig. 9) but protein sequences have been obtained from spinach [79] and tomato [80].

	10	20	30	40	50	60
Arabid.	VFVGFKKKAAPKKSAAKTVTTDRPLWYPGAISPDWLDGSLVGDYGFDPFGLGKPAEYLQFDIDSL					
Barley	R.....P...AK-APP.....F...QA.EY...T.....Y.V....					
	70	80	90	100	110	120
Arabid.	DQNLAKNLAGDVIGTRTEAADAKSTPFQPYSEVFGIQRFRECELIHGRWAMLATLGALSVEWLT					
BarleyQ....EI...F.D..V.....A...L.....T.....					
	130	140	150	160	170	180
Arabid.	GVTWQDAGKVELVDGSSYLQPLPFSISTLIWIEVLVIGYIEFQRNAELDSEKRLYPGGKFFDP					
BarleyT.T.....P.R.....SY...					
	200	210	220	230	240	250
Arabid.	LGLAADPEKTAQLQLAEIKHARLAMVAF LGFAVQAAATGKGPLNNWATHLSDPLHTTIIDTFSSS					
BarleyKET.....R.....F...G...					

Fig. 9. Amino acid sequence of Lhcb4 from barley [78] and *Arabidopsis thaliana* (Green and Pichersky, personal communication). Since the transit peptide cleavage site is not known, the putative N-terminus of the protein is shown.

Lhcb4 often co-purifies with Lhcb5 (CP26). Despite this biochemical similarity, the sequence identity is only 41%, while Lhcb4 and Lhca1 are 49% identical (Fig. 1). Close to the N-terminus, Lhcb4 contains a sequence motif (WFPG) that has been proposed to be LHC I-specific [25]. Calculations of overall sequence identity can sometimes be misleading but by using the GAP and BESTFIT programs of the GCG package [81], more careful studies are possible. These analyses show that Lhcb4 is equally related to Lhcb5 and Lhca1 (relative value = 153), less related to the other Lhca sequences (134–141) and even less to the other Lhcb sequences (126–134). The sequence similarity between Lhcb4 and the Lhca polypeptides explains the findings of Høyer-Hansen et al. [82] that monoclonal antibodies raised against LHC I and CP29 cross-react to a considerable extent.

Concerning the amount of pigment bound to Lhcb4, data are conflicting. Henrysson et al. [79] report 9 Chl *a*, 3 Chl *b*, 1 neoxanthin, 1 violaxanthin and 1 lutein per monomer. The data of Peter and Thornber [3], on the other hand, suggests 9 Chl *a*, 4 Chl *b*, 1 neoxanthin, 2 violaxanthin and 2 lutein, while results presented by Bassi et al. indicate that 10 Chl *a*, 4 Chl *b*, 1 neoxanthin, 1.5 violaxanthin and 2 lutein molecules are bound to each monomer [61], assuming that Lhcb4 is present in one copy per RC II (see section VI.B). Recent studies indicate that violaxanthin bound to Lhcb4, as well as to Lhcb5 and Lhcb6, is photo-convertible into zeaxanthin [66].

Spinach Lhcb4 separates in high-resolution gels as two [38], and in an IEF flat-bed separation as three, closely migrating bands with slightly different molecular mass (Spangfort, personal communication). Amino acid sequencing of proteolytic fragments has shown that these proteins are all products of *Lhcb4* genes (Spangfort and Jansson, unpublished). It is not known if they originate from three different genes, or if they are the result of post-translational modifications.

There is considerable evidence showing that Lhcb4 is a part of the inner antenna that connects the major LHC II complex with RC II [39,83]. A recent report also indicated that even Lhcb4 can be phosphorylated, although to a much lower level than Lhcb1 and Lhcb2 [3].

Lhcb4 is most likely the 30 kDa Ca^{2+} -binding protein that has been isolated from spinach by Irrgang et al. [84]. Ca^{2+} appears to bind to a carboxyl group on the protein. These authors show that addition of Ca^{2+} changes the 77 K fluorescence emission peak from 680 to 679 nm. They observed a high tendency for dimerization of the protein in the presence of Ca^{2+} and also obtained two dimeric forms (60 and 66 kDa) which were stable during SDS-PAGE. Pigment loss can probably explain why this preparation has a peak absorption at 670, rather than 678 nm as reported by Henrysson et al. [79]. The preparation of Irrgang et al. contains only 4 Chl *a*, 1 Chl *b* and an undetermined number of carotenoids per polypeptide.

It has been shown that a 31 kDa polypeptide (CP31) accumulates in the thylakoids of maize leaves chilled in light [85]. This protein reacts strongly with Lhcb4 antibodies and is almost certainly an unprocessed, partially processed or covalently modified Lhcb4 protein [86]. This form of the protein binds pigments [86] and its distribution in the thylakoid membrane is identical to 'normal' Lhcb4 (D. Campbell, personal communication). Its formation causes a major shift in the 77 K fluorescence emission spectrum of purified 'LHC II', which in this case contains Lhcb4. The illuminated 'chilled LHC II' has a strong fluorescence emission peak at 740 nm, which probably results from emission from a Chl *b* species. Returned to normal temperatures, these plants lose the 31 kDa protein, but the abnormal fluorescence of the LHC II complex remains [86].

The molecular mechanism leading to these changes is unknown, but two interesting similarities with other

	10	20	30	40	50	60	
Tomato	L--FKKKAAAP----	AKAKAAVSPADDELAKWYGPDRRIFLPEGLDRSEIPEYLNQVPG					
Barley	.FDRFN.K--.APKPKPAPV.TSSAGI.....Y..N.....V.....						
S. pine	.FG.....PP.PPPKS.....-A..TE.....DD.....						
	70	80	90	100	110	120	
Tomato	DYGYDPFGLSKKPEDFAKYQAYELIHARWAMLGAAGFI IPEAFNKFGANCPEAVWFKTGALL						
BarleyG.....F.....L.....						
S. pineGR...N.D.....Y.....						
	130	140	150	160	170	180	
Tomato	LDGNTLNYFGKNIPINLILAVVAEVVLVGGAEEYRIINGLDLEDKLFHGGPFDFPLGLAKDPDQ						
BarleyNS.....T...EFD.....T....						
S. pine	...G...S...AS.....A.....AT...NFD.....						
	190	200	210	220	230	240	250
Tomato	AAILKVKEIKNGRLAMFSLGFFIQAYVTGQGPVENLAHLSDPFGNNLLTVIGGASERVPTL						
Barley	..L.....E..F..C.....S..A....S..						
S. pine	F..L.....L.....E.....I..L..SL..A...						

Fig. 10. Amino acid sequence of Lhcb5 from tomato [80], barley [88] and Scots pine (Jansson and Gustafsson, unpublished). Since the transit peptide cleavage site is not known, the putative N-terminus of the protein is shown.

results should be mentioned. Firstly, given that Lhcb4 and Lhca1 have sequence homology, it is possible that the 77 K fluorescence emission at 740 nm in the 'chilled LHC II' could have some analogy to the 730 nm emission of LHCI-730. Secondly, there are many similarities between the set-up used by Hayden and co-workers and the studies performed by groups interested in the xanthophyll cycle. Lhcb4 binds large amounts of photo-convertible violaxanthin, and it is possible that the observed perturbation in the 77 K emission spectrum of 'chilled LHC II' could be related to zeaxanthin formation in Lhcb4.

II.L. *Lhcb5*

The location of the *Lhcb5* gene product has been a matter of confusion. When the gene was first isolated from tomato [80], the corresponding protein was found in the CP29 complex [80] as defined by White and Green [87]. However, other groups have detected the protein in CP26 as defined by Bassi et al. [39] or in the equivalent complex, LHC IIc [2]. This confusion is probably a result of different isolation procedures and therefore different empirical definitions of the complexes. *Lhcb5* genes or cDNAs have been isolated from tomato [80], barley [88] and Scots pine (Jansson and Gustafsson, unpublished). Tomato and Scots pine have only one *Lhcb5* gene (ref. 80, and Jansson and Gustafsson, unpublished), as has *Arabidopsis thaliana* [19].

Lhcb5 is also N-terminally blocked, which has prevented determination of the cleavage site for the transit peptide. The mature protein might start close to a conserved Leu residue (Fig. 10), corresponding to a polypeptide of 245–251 amino acids, but it is also quite possible that the protein is 10–20 amino acids shorter. The data of Peter and Thornber [3] indicate that 7 Chl *a*, 4 Chl *b*, 1 neoxanthin, 0.5 violaxanthin and 2 lutein are associated with each monomer, although the authors themselves hesitate to use their data for such exact calculations. Bassi et al. [61] provide evidence that 9 Chl *a*, 5 Chl *b*, 0.5 neoxanthin, 1 violaxanthin and 2 lutein are bound to each polypeptide.

Preparations of Lhcb5 often yield two polypeptides [3,38,39]. The origin of these polypeptides is not clear. They might be products of two homologous genes, but only one *Lhcb5* gene is present in the plant species

studied so far (Refs. 19,80; and Jansson and Gustafsson, unpublished). There is a possibility that one of the polypeptides might be a product of a post-translational modification, or the smaller protein, appearing in varying amounts, may represent a Lhcb5 degradation product.

Lhcb5 binds a Cu^{2+} ion [89], and a 28 kDa polypeptide, depleted under conditions of Cu-deficiency [90], could possibly be Lhcb5. Cu^{2+} ions often have Cys residues as ligands. The only Cys residue in Lhcb5, Cys¹¹³ in Fig. 10, is conserved; no other LHC protein has a cysteine at that position so it might be a Cu^{2+} -ligand. Cu^{2+} ions are normally found in oxidases [90], so the speculation that Lhcb5 could be an oxidase in the xanthophyll cycle [89] seems reasonable. The observation that most of the copper in PS II is recovered in a 'LHC II-fraction' [90] is probably just a consequence of Lhcb5 co-purifying with Lhcb1 and Lhcb2.

II.M. *Lhcb6*

Lhcb6, the apoprotein of CP24 or LHC IId, is the smallest of the Lhcb proteins (Fig. 11). *Lhcb6* DNA sequences have been characterised from tomato [91] and spinach [92], and protein sequencing has identified the protein in barley [93]. Tomato has two *Lhcb6* genes [91] but *Arabidopsis thaliana* has only one [19]. Lhcb6 consists of 210 amino acids and its primary sequence is the most divergent of all the LHC sequences (Fig. 1). Four closely migrating Lhcb6 bands can be obtained when maize PS II membranes are run on a non-denaturing IEF gel [39]. The number of Lhcb6 genes in the maize genome is unknown, so it cannot be determined yet whether these proteins are different gene products or results of post-translational modifications. There are no consistent data on the pigment content of this protein, Chl *a/b* ratios between 0.9 [3] and 1.6 [4] have recently been reported. Peter and Thornber [3] report on a stoichiometry of 20:22:4:1:1 for Chl *a*, Chl *b*, lutein, violaxanthin and neoxanthin, and the corresponding figures from Bassi et al. are 11:9:5:2:0 [61].

III. Are there additional LHC proteins?

Most evidence implies that there are only 10 LHC proteins present in the thylakoids of higher plants,

	10	20	30	40	50	60	70
Tomato	AAAVAPKKSWIPAVKSGGNLVDPEWLDGSLPGDFGFDPLGLGKDP AFLK WYREAE LIHGRWAMA AAVLGIF						
SpinachG...FL.....L.....						
	80	90	100	110	120	130	140
Tomato	VGQAWSGIPWFEEAGADPGAIA PF SFGSL LGTQLLLMGWVESKRWVDFDND S QSIDWATPWSKTAENFAN						
SpinachT.....V.....T.....P.....VE.....R.....S.						
	150	160	170	180	190	200	210
Tomato	FTGEQGYPGGKFFDPLALAGTLNNGVYVPDTEKLERLKLAEIKHSRLAML LIF YFEAGQGKTP LGALGL						
Spinach	S.....S.....N...D.....A.....						

Fig. 11. Amino acid sequence of Lhcb6 from tomato [91] and spinach [92].

although there have been several reports of additional LHC proteins. These 'putative LHC proteins' are presented here one by one.

III.A. A third LHCI-730 polypeptide – A contaminant?

The presence of three polypeptides in barley LHCI-730 has been reported [28]. This third polypeptide is most probably a contaminant, since subsequent studies indicate that only two LHCI-730 proteins exist [16,17].

III.B. PsaF in LHC I?

The product of the *psaF* gene ('subunit III') has been found in a LHC I preparation [94]. This led the authors to the conclusion that PsaF might be a part of the LHC I complex (named LHC Ic), and not RC I. This has been a matter of discussion, but today there is hardly any reason to believe that PsaF is part of the light-harvesting antenna or binds chlorophyll.

III.C. A small LHC I component – PsaE?

There was an early report that LHC I contains a 10 kDa polypeptide [95], and similar results have been reported recently [30]. This protein (called LHC Id) has not been found in other LHC I preparations [16,27,28], but since it is known that the PsaE protein often co-purifies with LHC I [17], there are reasons to believe that this protein is PsaE.

III.D. LHC Iie – An eleventh LHC protein and / or an ELIP?

Another putative LHC polypeptide is the ≈ 14 kDa polypeptide of LHC II. It has been named LHC Iie, shown to have a Chl *a/b*-ratio of 1.4 and to contain high levels of carotenoids, especially violaxanthin, but also some lutein and neoxanthin [3]. Although the presence of such a protein has not been confirmed by work in other laboratories, the characterisation of such a pigment-protein complex with current techniques [3,96] is intriguing. It is possible that LHC Iie could be a proteolytic fragment of one of the 10 LHC proteins already fully characterised. Prolonged storage of Lhcb4 can cause proteolytic cleavage and fragments with the apparent molecular mass of 14–15 kDa that retain pigments and show a standard absorption spectrum have been found (Spangfort, personal communication). However, a more likely explanation is that LHC Iie is actually an early light-induced protein (ELIP).

The ELIPs are synthesised transiently when dark-grown plants are transferred to light [97] and also after photoinhibition [98], presumably during the repair process. There are at least two types of ELIPs in barley, one of 13–14 kDa and the other of 18–19 kDa [97].

These proteins show sequence homology to the LHC proteins and the topology in the membrane is probably similar as well [99]. Interestingly, they also have sequence similarity to both a β -carotene binding protein from the unicellular alga *Dunaliella baradawil* [100], and a protein that is induced during desiccation in the resurrection plant *Craterostigma plantagineum* [101]. Altogether these observations imply that the ELIPs might be replacing the LHC proteins under 'unusual' conditions. The 12–15 kDa polypeptides that are the most abundant thylakoid proteins in a Chl *b*-deficient soybean mutant under conditions where Lhcb1 and Lhcb2 are strongly depleted [102] might also be ELIPs, although it has recently been claimed that these proteins are contaminating histones [103].

The ELIPs are abundant during early thylakoid development, and LHC Iie has been characterised from young leaves of barley [14]. In such material, many chloroplasts are in an early developmental state and could thus be expected to contain ELIPs. There is a possibility that LHC Iie is a homologue to the smaller type of ELIPs. This would explain why this protein has not been detected in other studies, since most other studies are performed on material in a later developmental state. The discovery that the Chl *b*-less barley mutant *chlorina-f2* is enriched in LHC Iie, while the other Lhcb proteins are depleted [3], further corroborates the assignment of LHC Iie as an ELIP.

There is no direct evidence that a higher plant ELIP binds pigments. If the assumption that LHC Iie is an ELIP is correct, it means that they bind both Chl *a* and Chl *b* and a large amount of carotenoids. Should ELIPs, then, be considered to be LHC proteins? If future work will show that ELIPs/LHC Iie functions in light-harvesting, the answer to this question must be yes. An alternative possibility is that ELIPs could have a function in carotenoid biosynthesis or assembly, as suggested by Adamska et al. [98].

To conclude this section, there is no good reason to believe that there are genes encoding LHC proteins functioning under 'normal' conditions beyond the ten types already characterised. If on the other hand ELIPs turn out to be LHC proteins, then they constitute a class of LHC proteins found only during early seedling development and under stress conditions, such as high light and/or drought stress. On current evidence, the possibility that other LHC proteins could be present under other special conditions cannot be excluded.

IV. The LHC proteins of lower plants

We still have very fragmentary information concerning the primary structure of LHC proteins from phylogenetic groups other than angiosperms and gymnosperms. One gene encoding an LHC protein has been isolated from each of four algal species; *Chlamy-*

domonas reinhardtii [104], *Chlamydomonas moewusii* [105], *Dunaliella tertiolecta* [106] and *Dunaliella salina* [107], and one from the moss *Physcomitrella patens* [107], although these species undoubtedly have additional *Lhc* genes (see below). From the homosporous fern *Polystichum munitum*, one potentially functional *Lhcb* gene and a couple of pseudogenes have been characterised [108].

Deduced amino acid sequences of the lower plant LHC proteins are compared to each other and to the higher plant proteins in Fig. 12. It should be kept in mind that the outcome of such an analysis is dependent on the method used, and the results should be interpreted with care. Nevertheless, the proteins from *C. reinhardtii*, *C. moewusii* and *D. tertiolecta* represent one type ('Type I') and the protein from *D. salina* another ('Type II') of green algal *Lhcb* proteins. There are reasons to expect that the LHC protein family of green algae to be larger than that of higher plants. There appear to be many additional LHC polypeptides in *C. reinhardtii* [8]; seven out of probably ten different LHC I polypeptides have been characterised by N-terminal sequencing [109]. Four different LHC II sequences have also been detected at the protein level in *D. tertiolecta* [106]. The LHC II proteins of *C. reinhardtii* differ in size and phosphorylation from the ones found in higher plants, but some show cross-reactivity with antibodies raised against different LHC proteins from angiosperms [110]. Immunological and spectroscopic data have identified three of these polypeptides as possible functional equivalents of *Lhcb4*, *Lhcb5* and *Lhcb6* [110].

The precursors of the LHC I [111] and LHC II [112] proteins of *Euglena gracilis* have a structure different from their higher plant counterparts. They are synthesised as large polypeptides, and might be transported to the chloroplast through Golgi vesicles and cleaved to their mature size either in the endomembrane or in the chloroplast [112]. The individual mature LHC II proteins of *Euglena* are quite similar to the *Lhcb1*/*Lhcb2*/*Lhcb3* proteins, whereas its LHC I proteins have approximately the same similarity to both the *Lhca* and *Lhcb* proteins of higher plants (Fig. 12). Note that the characterised fern, and possibly also the moss, LHC II protein appears to be more similar to *Lhcb1* than to *Lhcb2* (Fig. 12).

V. The three-dimensional structure

By using electron crystallography, the structure of the major LHC II complex has been determined at 6 Å resolution [60]. In the unit cell, three LHC II monomers form a trimeric complex. The monomer is composed of two symmetrical, tilted α -helical membrane spanning regions (MSR), 1 and 3 in Fig. 13, or helix A and B in Ref. 60. The helices, 31 and ≈ 33 amino acids long,

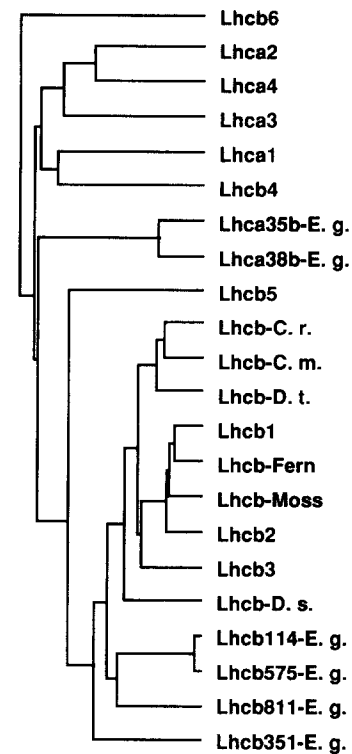


Fig. 12. Sequence homologies between the higher and lower plant LHC proteins. The length of the horizontal branches is proportional to the sequence divergence. The lower plant sequences are mentioned in the text, the designations of the different *E. gracilis* proteins correspond to the ones found in Refs. 111 and 112.

protrude into the stroma and correspond to the first and third MSR that have been predicted from the primary sequence [113]. These regions are homologous to each other [12], but it is notable that the C-terminal parts of the helices are not symmetrical. On the stromal side there are hook-like structures attached to the two helices. A third MSR (2 in Fig. 13, or C in the original reference) corresponds to the second predicted helix, which shows much lower sequence conservation, see Fig. 14. Clearly, the four-helix model that has been proposed [114] must be wrong.

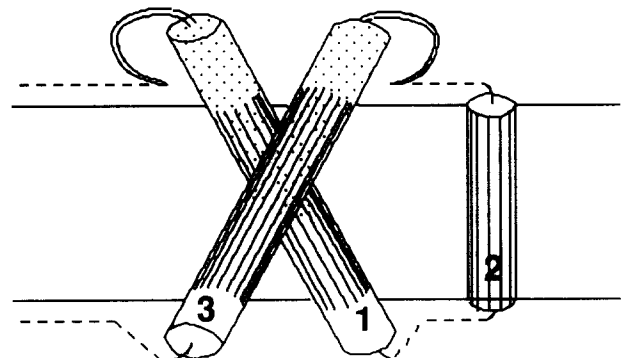


Fig. 13. Schematic representation of the three-dimensional structure of *Lhcb1*. Re-drawn from Kühlbrandt and Wang [60].

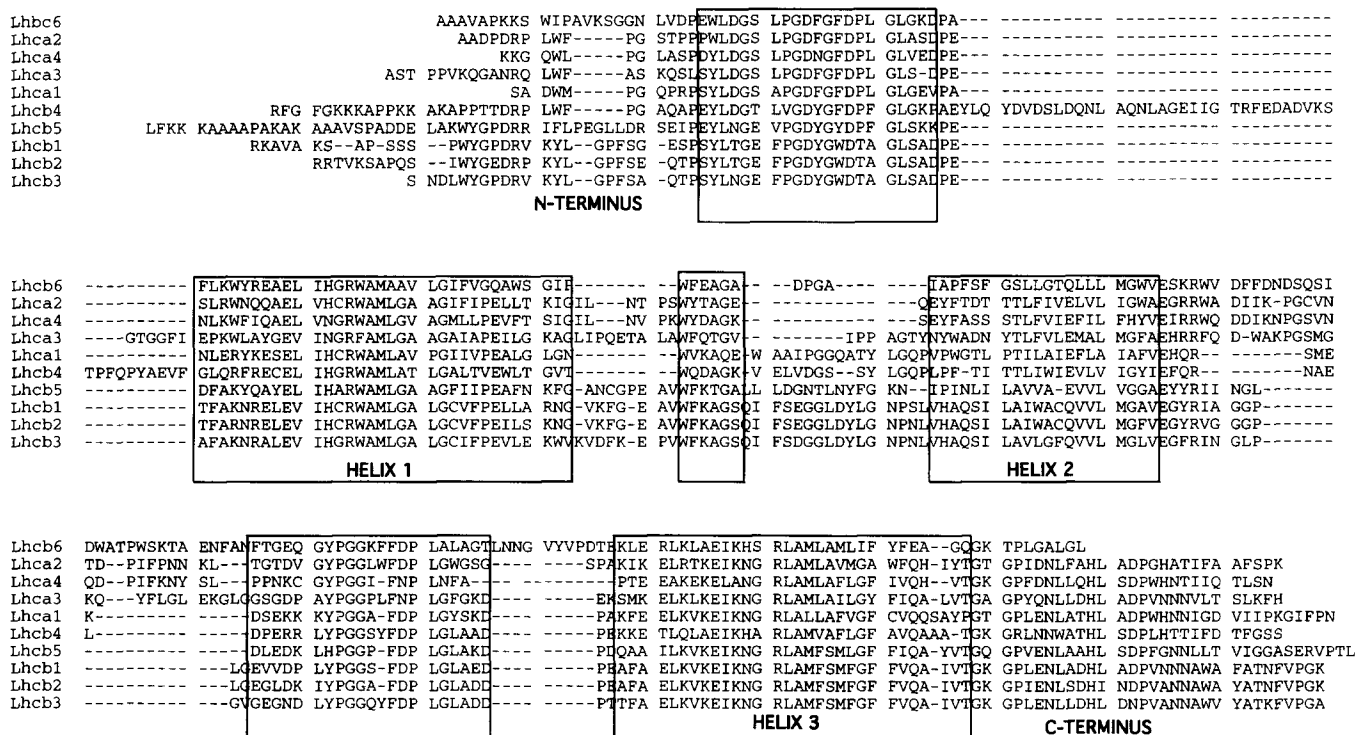


Fig. 14. The domain structure of the LHC polypeptides. The tomato proteins are taken as representatives for the different proteins, except in the case of Lhcb4, where the barley protein sequence is shown. Sequences were initially aligned with the PILEUP program of the GCG package [81], and then manually corrected.

The chlorophyll molecules are organised roughly perpendicular to the membrane plane and are found in two layers, with seven (or eight) molecules at the stromal side of the membrane and seven at the luminal side [60].

Theoretical considerations have predicted that the folding of all LHC proteins will be very similar in many parts [80]. Thus, it is possible to compare crystallographic data with all LHC protein sequences to predict structural elements that all LHC proteins have in common. An alignment the LHC protein sequences and a tentative assignment of different domains are shown in Fig. 14. The predicted membrane-spanning α -helices,

the hydrophilic regions, attached to the N-side of the first and third helices, and a conserved motif in the connector between helix 2 and helix 3, are boxed.

VI. The light-harvesting antenna of higher plants

A few of the models for the light-harvesting antenna of higher plants that have been presented will be discussed here. These models not only differ from each other in their graphical presentation, they also show differences in the position and stoichiometries of the different proteins.

VI.A. PS I

Electron microscopy data suggest that the PS I core is surrounded by a monolayer of eight LHC polypeptides (Fig. 15) [115]. Since the four Lhca proteins seem to be present in approximately equal amounts, this implies that there are probably two Lhca proteins of each kind coordinated with each PS I. Bruce and Malkin [116] estimate one of the LHC I proteins to be present in two copies per RC I. These eight Lhca molecules bind approximately 120 chlorophyll molecules, in addition to the 80–100 molecules bound to the PS I core [28].

Knoetzel et al. [17] have proposed a model of LHC I (Fig. 16), where both LHCI-680 and LHCI-730 are

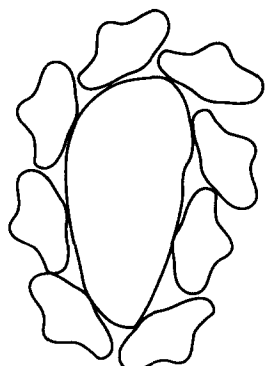


Fig. 15. Structure of PS I core (inner circle) with eight LHC I proteins attached. Re-drawn from Boekema et al. [115].

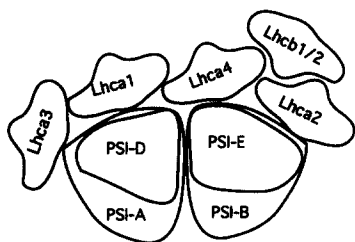


Fig. 16. Proposed organisation of the PS I light-harvesting antenna. Re-drawn from Knoetzel et al. [17].

recognised. The Lhcb1/Lhcb2 trimers are connected to LHCI-680 (Lhca2) in this model. It is assumed that energy is transferred from LHC II or LHCI-730 over LHCI-680 to the reaction centre.

PS I has recently been found to be heterogeneous like PS II; PS I α appears to have an antenna $\approx 40\%$ larger than that of PS I β [117]. PS I α is present in the margins of the grana stacks whereas PS I β is found in the stromal regions. PS I α (probably equivalent to PS I-300 [28]) differs from PS I β in having Lhcb1 and Lhcb2 polypeptides functionally attached [118], while the stoichiometry between the Lhca proteins is identical. The PS I complex used for the electron microscopy studies mentioned above (PS I-200) was derived from stromal regions and corresponds to PS I β .

It has been assumed that Lhcb1-Lhcb2 complexes in the stromal regions deliver their energy to PS I. One implication of the data presented above is that this process takes place in the margins of the grana stacks and not in the stromal regions [117]. This might require that models of the regulation of the energy balance between the two photosystems by phosphorylation of Lhcb1-Lhcb2 are revised.

II.B. PS II

In the case of the light-harvesting antenna of PS II, three models have recently been published. Two of them are based on the recent finding that PS II appears to be dimeric in the grana membranes, albeit monomeric in the stroma exposed thylakoids [119,120].

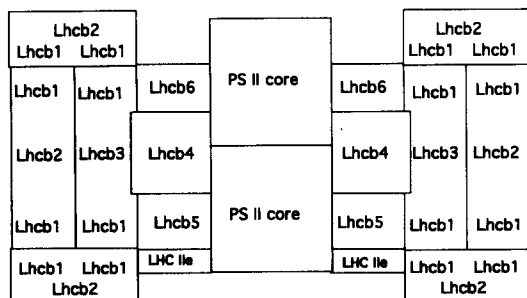


Fig. 17. Proposed organisation of the PS II light-harvesting antenna. Re-drawn from Peter and Thornber [3].

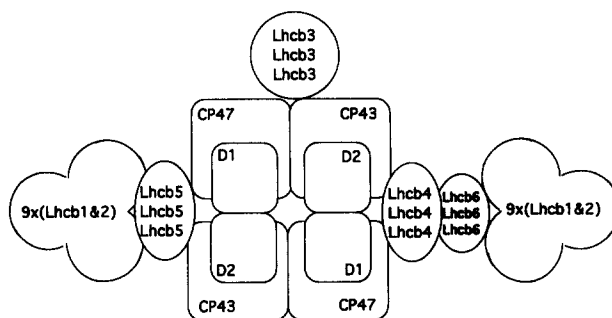


Fig. 18. Proposed organisation of the PS II light-harvesting antenna. Re-drawn from Dainese et al. [120].

The major difference between the three models is that the stoichiometries between the polypeptides are different. It is also notable that none of the models explain the difference between PS II α and PS II β in terms of antenna polypeptides.

Peter and Thornber [3] present a model of the organisation of LHC II in which the two PS II units in the dimer have the same antenna structure (Fig. 17). Lhcb4, Lhcb5 and Lhcb6 (and LHC IIe) are found close to the core and a trimer containing Lhcb3 is the innermost part of the Lhcb1/Lhcb2/Lhcb3 antenna. A different model, based on a presumed stoichiometry of 1.5 minor Lhcb protein/PS II, where one Lhcb3, Lhcb4, Lhcb5 and Lhcb6 homotrimer is found in each dimeric, asymmetric PS II complex (Fig. 18), has been proposed by Dainese et al. [120]. In this model Lhcb3 is proposed to form one branch of the antenna, Lhcb5 together with a fraction of Lhcb1/Lhcb2 another, and Lhcb4, Lhcb6 and Lhcb1/Lhcb2 a third. Finally, Harrison and Melis [75] suggest a model (Fig. 19) which differs from the others by having an inner antenna consisting of two copies of Lhcb3 and Lhcb4, but only one Lhcb6 polypeptide per reaction centre. Lhcb5 is not included in this model.

Data from many groups indicate that the minor Lhcb proteins (i.e., Lhcb3, Lhcb4, Lhcb5 and Lhcb6) are found in equimolar amounts [3,75,76,121], but it is not known whether the number 1, 1.5 or 2 copies per

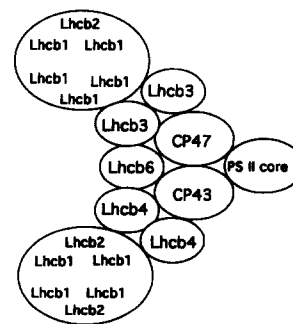


Fig. 19. Proposed organisation of the PS II light-harvesting antenna. Re-drawn from Harrison and Melis [75].

reaction centre is correct. I believe that the most probable number is one, for the following reasons. The measured pigment stoichiometries and the ratio between the minor and the major (Lhcb1/Lhcb2) polypeptides make the number two quite unlikely. Firstly, there are reasons to believe that the minor Lhcb proteins, like Lhcb1, Lhcb2 and the Lhca polypeptides, bind approximately 14 chlorophyll molecules. This assumption is supported by the isolation of pure fractions with a chlorophyll/protein ratio of 12–13 [3,79]. The lower chlorophyll/protein ratios often found in purified minor polypeptides do not contradict this, since estimations based on purified, detergent-solubilized pigment–protein complexes are not fully reliable.

Secondly, the stoichiometries of Lhcb1 and Lhcb2, treated together here, compared to the minor polypeptides must be considered. This ratio is probably variable, since huge differences in Chl *a/b*-ratios and the amount of Lhcb1/Lhcb2 ('LHC II') are found under different growth conditions. However, the stoichiometries within PS II preparations have been determined by a couple of authors. Dainese and Bassi [121] used Coomassie-binding to make an accurate determination of the amount of LHC polypeptides in PS II membranes. The molar ratio of Lhcb1 + Lhcb2 + Lhcb3/a minor protein (i.e., Lhcb4, Lhcb5 and Lhcb6) was between 7.6:1 and 9.4:1, making the ratio of (Lhcb1/Lhcb2) over any of the minor Lhcb proteins between 7:1 and 8:1. In contrast, Peter and Thornber [3] suggest, based on the chlorophyll content in Deriphat-PAGE separated complexes, a stoichiometry of 8:3:1 between Lhcb1, Lhcb2 and the minor proteins corresponding to 11 Lhcb1/Lhcb2 proteins for each of the minor proteins. Harrison and Melis [75] propose, primarily based on the less accurate method of densitometric scanning of stained gels, a 4.5/1.5/1 stoichiometry of Lhcb1, Lhcb2 and the minor proteins corresponding to a 6:1 ratio. The discrepancies might, at least partially, be due to different light conditions during plant growth, but there appear to be at least six, maybe up to twelve, Lhcb1 + Lhcb2 proteins per each minor Lhcb protein in PS II.

If there are two copies of each minor polypeptide associated with one PS II, it does not match the measured antenna size of PS II. In that case there would be 20–32 LHC associated with each PS II. The PS II antenna, including the 40–50 Chl molecules bound to the CP43 and CP47 proteins, would have a size of 320–500 Chl, which is well beyond most measurements. A consensus value for the antenna size of PS II is 230 to 250 chlorophyll molecules [75].

The figure of 1.5 minor protein/reaction centre cannot be excluded based on stoichiometrical considerations, but this figure implies that the actual configuration of the minor Lhcb proteins is trimeric. There is, at least in my opinion, a number of reasons to believe

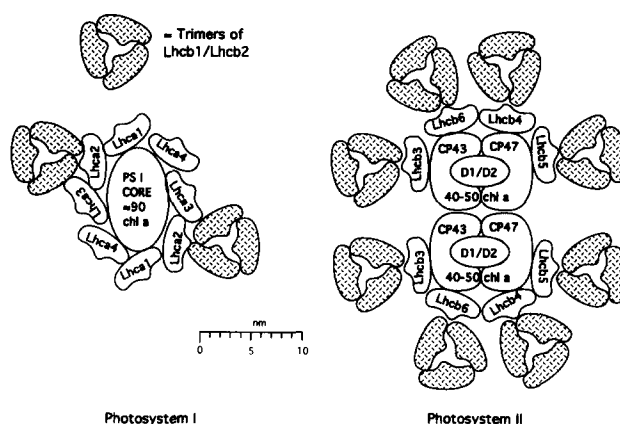


Fig. 20. A novel model for the higher plant light-harvesting antenna.

that this is not the case. Recent data indicate that the Lhcb proteins closely connected to the reaction centre are monomeric [122] as are the Lhca proteins [115]. They also appear to migrate as monomers in various non-denaturing gel systems [3,77]. The finding that oligomerisation of Lhcb4, for example, can be observed under certain conditions (e.g. Ref. 84) is not in conflict with this, since the association between monomeric Lhca1 and Lhca4 can also be quite strong in vitro [16,17].

From this presentation one can conclude that the light-harvesting antenna of a one RC II probably consists of one each of Lhcb3, Lhcb4, Lhcb5 and Lhcb6, and 2 to 4 trimers of Lhcb1/Lhcb2.

VI.C. A consensus model

I would like to propose a model, which has features in common with all those presented above. In this model (Fig. 20) trimers of Lhcb1/Lhcb2 are the most peripheral antenna to both PS I and PS II, analogous to the cyanobacterial phycobilisome, and these trimers could associate with either PS I and PS II [123]. An inner antenna, composed of 2 + 2 + 2 + 2 Lhca polypeptides in PS I and one each of Lhcb3, Lhcb4, Lhcb5 and Lhcb6 in PS II, connects the peripheral antenna to the core, although the positions of the proteins within the inner antenna are highly hypothetical. The inner antenna may regulate the efficiency of energy transfer from the outer antenna to the core, and the localisation of the xanthophyll cycle to the polypeptides in the inner antennas might be related to this function.

The difference between both PS I α and PS I β and PS II α /PS II β is likely to be solely a difference in the number of Lhcb1/Lhcb2 trimers associated with the inner antenna. Without Lhcb1/Lhcb2 trimers, the antenna sizes of PS I and PS II would be \approx 210 and \approx 100 chlorophyll molecules, respectively. PS I α has two Lhcb1/Lhcb2 trimers attached, corresponding to the 40% larger antenna size [117]. The number of

trimers attached to PS II might be even more variable, so further work is needed to clarify the difference between PS II α and PS II β in terms of LHC protein composition.

This model assumes that Lhcb3 is monomeric and does not form trimers together with Lhcb1 and Lhcb2. This is controversial, it is possible that Lhcb3, Lhcb1 and/or Lhcb2 heterotrimers are formed *in vivo*. In that case, the model could be modified by replacing a Lhcb3 monomer and its Lhcb1/Lhcb2 trimer with a heterotrimer.

VI.D. Can we correlate the protein complexes with freeze-fracture particles?

Freeze-fracture electron microscopy has been used to study the protein complexes of the thylakoid membrane, and some of the particles found have been correlated with various supramolecular complexes. Lhcb1/Lhcb2 trimers form the 8 nm PF particles present both in the stacked (PFs particles) and unstacked (PFu small particles) regions [124]. PS I complexes (PS I-200) correspond to the 11–13 nm PFu particles and PS II forms another type of structure, EF particles, with ≈ 9 nm EFu and ≈ 16 nm EFs particles found in the unstacked and stacked regions, respectively [124]. If the spatial localisation of monomeric and dimeric PS II is considered, it seems likely that the EFu particles correspond to the monomeric and the EFs particles to the dimeric PS II. The increase in size of PS II particles when plants grown under ImL (which lack chlorophyll *b* and LHC polypeptides) are transferred to light has been suggested to be a consequence of the various amounts of LHC proteins added [124].

Using the present knowledge of antenna structures, it should be possible to refine these assignments. I hypothesise that the different supramolecular complexes are related to various freeze-fracture particles according to Table III. It must be pointed out that this

a tentative correlation, which has to be substantiated by experimental studies.

VII. Concluding remarks

Some important questions concerning the LHC proteins remain to be answered. Without knowledge of the stoichiometries of the different proteins, any attempt to build a model will have a great deal of uncertainty, thus it is necessary for the number of LHC proteins bound to each photosystem to be definitely established. The difference in polypeptide content between the different types of PS I and PS II centres is also an important area for future research. A combination of different techniques, such as reconstitution experiments, structure determinations, mutant studies and spectroscopy will be necessary to gain a more detailed and precise understanding of the structure of the higher plant light-harvesting antennas and the relation between their components. It is also possible that the characterisation of LHC proteins of lower plants will give further insight to this problem.

Two issues, mentioned in the Introduction, deserve to be addressed finally. First, is it impossible to purify LHC proteins without pigment loss, or is there no fixed pigment content for the complexes? The same pigment–protein complex, prepared in different ways, can have vastly different pigment contents. Even in different ‘LHC II’ fractions, obtained from non-denaturing IEF (which is believed to result in no or little release of pigments), Chl *a/b*-ratios can vary by about 20% [61]. I believe that the pigment content is *not* as variable as it appears from experimental data; improved techniques have to be found in order to determine the accurate number of pigments bound to each polypeptide. Reconstitution experiments also indicate that binding sites for different carotenoids are specific, at least in Lhcb1 [63]. On the other hand, the presence of sub-stoichiometric amounts of some carotenoids in some fairly pure preparations is nevertheless intriguing, and could perhaps mean that the number of bound carotenoids may vary *in vivo*.

Second, exactly what is ‘LHC II’? In the Introduction, I defined ‘LHC II’ very strictly, and my reasons are hopefully clear to all readers by now. The so-called ‘LHC II’-proteins (i.e., Lhcb1 and Lhcb2) are *not* pure LHC II proteins, instead they function as antenna for both PS I and PS II. In contrast, Lhcb3, Lhcb4, Lhcb5 and Lhcb6 are restricted to PS II, although to name these proteins (and not Lhcb1 and Lhcb2) ‘LHC II’ would result in total confusion. Moreover, since it is uncertain if Lhcb3 is located in ‘LHC II’-trimers, we do not know whether or not Lhcb3 should be called a ‘LHC II’-protein. To conclude, any scientist using the designation ‘LHC II’ should do so with care and only after giving a clear definition of the term.

TABLE III

Correlation between various supramolecular complexes and different freeze-fracture particles

Complex	Size (nm)	Freeze-fracture particle
Lhcb1/Lhcb2 trimers	7.3×7.3	8 nm PF
PS I-100	8×12	PFu 6–9 nm (only in ImL)
(core)		
PS I-200	11×16	PFu large (11–13 nm)
(core + LHC I)		
PS II core	7×9	8 nm EF (only in ImL)
(monomer)		
PS II core + inner	9×14	9–10 nm EFu
(monomer)		
PS II core + inner	14×18	EFs 16 nm
(dimer)		

Acknowledgements

I am indebted to Vaughan Hurry, Lottie Eriksson and Adrian Clarke for comments on the manuscript that significantly improved the quality of the language, and to Petter Gustafsson for comments and general support. This work was supported by grants from the Swedish Council for Forestry and Agricultural Research.

References

- Hallick, R.B. (1989) *Plant Mol. Biol. Rep.* 7, 266–275.
- Jansson, S., Pichersky, E., Bassi, R., Green, B.R., Ikeuchi, M., Melis, A., Simpson, D.J., Spangfort, M., Staehelin, L.A. and Thornber, J.P. (1992) *Plant Mol. Biol. Rep.* 10, 242–253.
- Peter, G.F. and Thornber, J.P. (1991) *J. Biol. Chem.* 266, 16745–16754.
- Dainese, P., Hoyer-Hansen, G. and Bassi, R. (1990) *Photochem. Photobiol.* 51, 693–703.
- Green, B.R. (1988) *Photosynth. Res.* 15, 3–32.
- Keegstra, K. and Olsen, L.J. (1989) *Annu. Rev. Plant. Physiol. Plant Mol. Biol.* 40, 471–501.
- Buetow, D.E., Chen, H. and Erdős, G. (1988) *Photosynth. Res.* 18, 61–97.
- Chitnis, P.R. and Thornber, J.P. (1988) *Photosynth. Res.* 16, 41–63.
- Ljungberg, U., Åkerlund H.E. and Andersson, B. (1986) *Eur. J. Biochem.* 158, 477–482.
- Kim, S., Sandusky, P., Bowlby, N.R., Aebersold, R., Green, B.R., Vlahakis, S., Yocum, C.F. and Pichersky, E. (1992) *FEBS Lett.* 314, 67–71.
- Wedel, N., Klein, R., Ljungberg, U., Andersson, B. and Herrmann, R.G. (1992) *FEBS Lett.* 314, 61–66.
- Hoffman, N.E., Pichersky, E., Malik, V.S., Castresana, C., Ko, K., Darr, S.C. and Cashmore, A.R. (1987) *Proc. Natl. Acad. Sci. USA* 84, 8844–8848.
- Jensen, P.E., Kristensen, M., Hoff, T., Lehnbeck, J., Stummann, B.M. and Henningsen, K.W. (1992) *Physiol. Plant.* 84, 561–567.
- Palomares, R., Herrmann, R.G. and Oelmüller, R. (1991) *J. Photochem. Photobiol. B: Biol.* 11, 151–162.
- Jansson, S. and Gustafsson, P. (1991) *Mol. Gen. Genet.* 229, 67–76.
- Ikeuchi, M., Hirano, A. and Inoue, Y. (1991) *Plant Cell Physiol.* 32, 103–112.
- Knoetzel, J., Svendsen, I. and Simpson, D.J. (1992) *Eur. J. Biochem.* 206, 209–215.
- Pichersky, E., Hoffman, N.E., Bernatzky, R., Piechulla, B., Tanksley, S.D. and Cashmore, A.R. (1987) *Plant Mol. Biol.* 9, 205–216.
- McGrath, J.M., Terzaghi, W.B., Sridhar, P., Cashmore, A.R. and Pichersky, E. (1992) *Plant Mol. Biol.* 19, 725–733.
- Stayton, M.M., Brosio, P. and Dunsmuir, P. (1987) *Plant Mol. Biol.* 10, 127–137.
- Pichersky, E., Tanksley, S.D., Piechulla, B., Stayton, M.M. and Dunsmuir, P. (1988) *Plant Mol. Biol.* 11, 69–71.
- Piechulla, B., Kellmann, J.-W., Pichersky, E., Schwartz, E. and Förster, H. (1991) *Mol. Gen. Genet.* 230, 413–422.
- Pichersky, E., Brock, T.G., Nguyen, D., Hoffman, N.E., Piechulla, B., Tanksley, S.D. and Green, B.R. (1989) *Plant Mol. Biol.* 12, 257–270.
- Jahns, P. and Junge, W. (1990) *Eur. J. Biochem.* 193, 731–736.
- Schwartz, E., Shen, D., Aebersold, R., McGrath, J.M., Pichersky, E. and Green, B.R. (1991) *FEBS Lett.* 280, 229–234.
- Zhang, H., Hanley, S. and Goodman, H.M. (1991) *Plant Physiol.* 96, 1387–1388.
- Lam, E., Ortiz, W., Mayfield, S. and Malkin, R. (1984) *Plant Physiol.* 74, 650–655.
- Bassi, R. and Simpson, D. (1987) *Eur. J. Biochem.* 163, 221–230.
- Damm, I., Steinmetz, D. and Grimme, L.H. (1990) in *Current Research in Photosynthesis* (Baltscheffsky, M., ed.), Vol. 2, pp. 607–610, Kluwer, Dordrecht.
- Thornber, J.P., Morishige, D.T., Anandan, S. and Peter, G.F. (1991) in *Chlorophylls* (Scheer, H., ed.), pp. 549–585, CRC Press, Boca Raton.
- Henry, R.L., Armbrust, T., Gallegos, G. and Guikema, J.A. (1992) *Plant Physiol. Biochem.* 30, 347–356.
- Knoetzel, J. and Simpson, D. (1991) *Planta* 185, 111–123.
- Broglie, R., Bellemare, G., Bartlett, S.G., Chua N. H. and Cashmore, A.R. (1981) *Proc. Natl. Acad. Sci. USA* 78, 7304–7308.
- Coruzzi, G., Broglie, R., Cashmore, A. and Chua, N.H. (1983) *J. Biol. Chem.* 258, 13399–13402.
- Dunsmuir, P. (1985) *Nucleic Acids Res.* 13, 2503–2518.
- Greenland, A.J., Thomas, M.V. and Walden, R.M. (1987) *Planta* 170, 99–110.
- Leutweiler, L.S., Meyerowitz, E.M. and Tobin, E.M. (1986) *Nucleic Acids Res.* 14, 4051–4064.
- Allen, K.D. and Staehelin, L.A. (1992) *Plant Physiol.* 100, 1517–1526.
- Bassi, R., Rigoni, F. and Giacometti, G.M. (1990) *Photochem. Photobiol.* 52, 1187–1206.
- Michel, H.P., Buvinger, W.E. and Bennett, J. (1990) in *Current Research in Photosynthesis* (Baltscheffsky, M., ed.), Kluwer, Dordrecht, pp. 747–753.
- Lamppa, G.K. and Abad, M.S. (1987) *J. Cell Biol.* 105, 2641–2648.
- Kohorn, B.D., Harel, E., Chitnis, P.R., Thornber, J.P. and Tobin, E.M. (1986) *J. Cell Biol.* 102, 972–981.
- Abad, M.S., Oblong, J.E. and Lamppa, G.K. (1991) *Plant Physiol.* 96, 1220–1227.
- Abad, M.S., Clark, S.E. and Lamppa, G.K. (1989) *Plant Physiol.* 90, 117–124.
- Jansson, S., Selstam, E. and Gustafsson, P. (1990) *Biochim. Biophys. Acta* 1019, 110–114.
- Green, B.R., Shen, D., Aebersold, R. and Pichersky, E. (1992) *FEBS Lett.* 305, 18–22.
- Sigrist, M. and Staehelin, L.A. (1992) *Biochim. Biophys. Acta* 1098, 191–200.
- Lubben, T.H., Donaldson, G.K., Viitanen, P.V. and Gatenby, A.A. (1989) *Plant Cell* 1, 1223–1230.
- Yalovsky, S., Paulsen, H., Michaeli, D., Chitnis, P.R. and Nechushtai, R. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5616–5619.
- Jansson, S. and Gustafsson, P. (1990) *Plant Mol. Biol.* 14, 287–296.
- Huner, N.P.A., Campbell, D., Krol, M., Hayden, D.B., Myscich E.M., Basalyga, S. and Williams, J.P. (1992) *Plant Physiol.* 99, 830–836.
- Mullet, J.E. (1983) *J. Biol. Chem.* 258, 9941–9948.
- Larsson, U.K., Anderson, J.M. and Andersson, B. (1987) *Biochim. Biophys. Acta* 894, 69–75.
- Mäenpää, P. and Andersson, B. (1989) *Z. Naturforsch.* 44C, 403–406.
- Larsson, U.K. and Andersson, B. (1985) *Biochim. Biophys. Acta* 809, 396–402.
- Larsson, U.K., Sundby, C. and Andersson, B. (1987) *Biochim. Biophys. Acta* 894, 59–68.
- Islam, K. (1987) *Biochim. Biophys. Acta* 893, 333–341.

- 58 Mattoo, A.K. and Edelman, M. (1987) *Proc. Natl. Acad. Sci. USA* 84, 1497–1501.
- 59 Peter, G. and Thornber, J.P. (1988) in *Photosynthetic Light-Harvesting Systems* (Scheer, H. and Schneider, S., eds.), pp. 175–186, Walter de Gruyter, Berlin.
- 60 Kühlbrandt, W. and Wang, D.N. (1991) *Nature* 350, 130–134.
- 61 Bassi, R., Pineau, B., Dainese, P. and Marquardt, J. (1993) *Eur. J. Biochem.* 212, 297–303.
- 62 Plumley, F.G. and Schmidt, G.W. (1987) *Proc. Natl. Acad. Sci. USA* 84, 146–150.
- 63 Cammarata, K.V. and Schmidt, G.W. (1992) *Biochemistry* 31, 2779–2789.
- 64 Demmig-Adams, B. (1990) *Biochim. Biophys. Acta* 1020, 1–24.
- 65 Horton, P., Ruban, A.V., Rees, D., Pascal, A.A., Noctor, G. and Young, A.J. (1991) *FEBS Lett.* 292, 1–4.
- 66 Dainese, P., Marquardt, J., Pineau, B. and Bassi, R. (1992) *Proceedings of the IX International Congress on Photosynthesis*, Nagoya, Japan.
- 67 Allen, J.F. (1992) *Biochim. Biophys. Acta* 1098, 275–335.
- 68 Steinback, K.E., Burke, J.J. and Arntzen, C.J. (1979) *Arch. Biochem. Biophys.* 195, 546–557.
- 69 Berg, S., Dodge, S., Krogmann, D.W. and Dilley, R.A. (1974) *Plant Physiol.* 53, 619–27.
- 70 Schwartz, E., Stasys, R., Aebersold, R., McGrath, J.M., Pichersky, E. and Green, B.R. (1991) *Plant Mol. Biol.* 17, 923–925.
- 71 Brandt, J., Skovgaard Nielsen, V., Thordal-Christensen, H., Simpson, D. and Okkels, S. (1992) *Plant Mol. Biol.* 19, 699–703.
- 72 Morishige, D.T. and Thornber, J.P. (1991) *FEBS Lett.* 293, 183–187.
- 73 Bassi, R. and Dainese, P. (1990) in *Current Research in Photosynthesis* (Baltscheffsky, M., ed.), Vol. 2, Kluwer, Dordrecht, pp. 209–216.
- 74 Webber, A.N. and Gray, J.C. (1989) *FEBS Lett.* 249, 79–82.
- 75 Harrison, M.A. and Melis, A. (1992) *Plant Cell Physiol.* 33, 627–637.
- 76 Morrissey, P.J., Glick, R.E. and Melis, A. (1989) *Plant Cell Physiol.* 30, 335–344.
- 77 Falbel, T.G. and Staehelin, L.A. (1992) *Photosynth. Res.* 34, 249–262.
- 78 Morishige, D.T. and Thornber, J.P. (1992) *Plant Physiol.* 98, 238–245.
- 79 Henrysson, T., Schröder, W.P., Spangfort, M. and Åkerlund, H.-E. (1989) *Biochim. Biophys. Acta* 977, 301–308.
- 80 Pichersky, E., Subramaniam, R., White, M.J., Reid, J., Aebersold, R. and Green, B.R. (1991) *Mol. Gen. Genet.* 227, 277–284.
- 81 Devereux, J., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* 12, 387–395.
- 82 Høyer-Hansen, G., Bassi, R., Hønerberg, L.S. and Simpson, D.J. (1988) *Planta* 173, 13–21.
- 83 Camm, E.L. and Green, B.R. (1989) *Biochim. Biophys. Acta* 974, 180–184.
- 84 Irrgang, K.-D., Renger, G. and Vater, J. (1991) *Eur. J. Biochem.* 201, 515–522.
- 85 Hayden, D.B., Baker, N.R., Percival, M.P. and Beckwith, P.B. (1986) *Biochim. Biophys. Acta* 851, 86–92.
- 86 Hayden, D.B., Covello, P.S. and Baker, N.R. (1988) *Photosynth. Res.* 15, 257–270.
- 87 White, M.J. and Green, B.R. (1987) *Eur. J. Biochem.* 165, 531–535.
- 88 Sørensen, A.B., Lauridsen, B.F. and Gausing, K. (1992) *Plant Physiol.* 98, 1538–1540.
- 89 Arvidsson, P., Bratt, C.E., Andréasson, L. and Åkerlund, H. (1992) *Proceedings of the IX International Congress on Photosynthesis*, Nagoya, Japan.
- 90 Droppa, M. and Horváth, G. (1990) *Crit. Rev. Plant Sci.* 9, 111–123.
- 91 Schwartz, E. and Pichersky, E. (1990) *Plant Mol. Biol.* 15, 157–160.
- 92 Spangfort, M., Larsson, U.K., Ljungberg, U., Ryberg, M. and Andersson, B. (1990) in *Current Research in Photosynthesis* (Baltscheffsky, M., ed.), Vol. 2, Kluwer, Dordrecht, pp. 253–256.
- 93 Morishige, D.T., Anandan, S., Jaing, J.T. and Thornber, J.P. (1990) *FEBS Lett.* 264, 239–242.
- 94 Anandan, S., Vainstein, A. and Thornber, J.P. (1989) *FEBS Lett.* 256, 150–154.
- 95 Haworth, P., Watson, J.L. and Arntzen, C.J. (1983) *Biochim. Biophys. Acta* 724, 151–158.
- 96 Irrgang, K.-D., Bechtel, C., Vater, J. and Renger, G. (1990) in *Current Research in Photosynthesis* (Baltscheffsky, M., ed.), Vol. 1, Kluwer, Dordrecht, pp. 355–358.
- 97 Grimm, B., Kruse, E. and Koppstech, K. (1989) *Plant Mol. Biol.* 13, 583–593.
- 98 Adamska, I., Ohad, I. and Koppstech, K. (1992) *Proc. Natl. Acad. Sci. USA* 89, 2610–2613.
- 99 Green, B.R., Pichersky, E. and Koppstech, K. (1991) *Trends Biochem. Sci.* 16, 181–186.
- 100 Lers, A., Levy, H. and Zamir, A. (1991) *J. Biol. Chem.* 266, 13698–13705.
- 101 Bartels, D., Hanke, C., Schneider, K., Michel, D. and Salamini, F. (1992) *EMBO J.* 11, 2771–2778.
- 102 Droppa, M., Ghirardi, M.L., Horvath, G. and Melis, A. (1988) *Biochim. Biophys. Acta* 932, 138–145.
- 103 Welty, B.A. and Morishige, D.T., Thornber, J.P. (1992) *Plant Cell Physiol.* 33, 1049–1055.
- 104 Imbault, P., Wittemer, C., Johanningmeier, U., Jacobs, J.D. and Howell, S.H. (1988) *Gene* 73, 397–407.
- 105 Larouche, L., Tremblay, C., Simard, C. and Bellemare, G. (1991) *Curr. Genet.* 19, 285–288.
- 106 LaRoche, J., Bennett, J. and Falkowski, P.G. (1990) *Gene* 95, 165–171.
- 107 Long, Z., Wang, S.Y. and Nelson, N. (1989) *Gene* 76, 299–312.
- 108 Pichersky, E., Soltis, D. and Soltis, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 195–199.
- 109 Bassi, R., Soen, S.Y., F.G., Zuber, H. and Rochaix, J.-D. (1992) *J. Biol. Chem.* 267, 25714–25736.
- 110 Bassi, R. and Wollman, F. (1991) *Planta* 183, 423–433.
- 111 Houlné, G. and Schantz, R. (1988) *Mol. Gen. Genet.* 213, 479–486.
- 112 Muchhal, U.S. and Schwartzbach, S.D. (1992) *Plant Mol. Biol.* 18, 287–299.
- 113 Karlin-Neumann, G.A., Kohorn, B.D., Thornber, J.P. and Tobin, E.M. (1985) *J. Mol. Appl. Genet.* 3, 45–61.
- 114 Anderson, J.M. and Goodchild, D.J. (1987) *FEBS Lett.* 213, 29–33.
- 115 Boekema, E.J., Wynn, R.M. and Malkin, R. (1990) *Biochim. Biophys. Acta* 1017, 49–56.
- 116 Bruce, B.D. and Malkin, R. (1988) *J. Biol. Chem.* 263, 7302–7308.
- 117 Svensson, P., Andreasson, E. and Albertsson, P.-Å. (1991) *Biochim. Biophys. Acta* 1060, 45–50.
- 118 Andreasson, E. and Albertsson, P.-Å. (1993) *Biochim. Biophys. Acta* 1141, 175–182.
- 119 Peter, G.F. and Thornber, J.P. (1991) *Plant Cell Physiol.* 32, 1237–1250.
- 120 Dainese, P., Santini, C., Ghiretti-Magaldi, A., Marquardt, J., Tidu, V., Mauro, S., Bergantino, E. and Bassi, R. (1992) *Proceedings of the IX International Congress on Photosynthesis*, Nagoya, Japan.
- 121 Dainese, P. and Bassi, R. (1991) *J. Biol. Chem.* 266, 8136–8142.
- 122 Holzenberg, A., Bewley, M.C., Wilson, F.H., Nicholson, W.V. and Ford, R.C. (1993) *Nature* 363, 470–472.
- 123 Mustardy, L., Cunningham, F.X. and Gantt, E. (1992) *Proc. Natl. Acad. Sci. USA* 89, 10021–10025.

- 124 Staehelin, L.A. (1986) *Encyclopedia of Plant Physiology – New Series*, Vol. 19, pp. 1–84.
- 125 Pichersky, E., Bernatzky, R., Tanksley, S.D., Breidenbach, B., Kausch, A.P. and Cashmore, A.R. (1985) *Gene* 40, 247–258.
- 126 Matsuoka, M. (1990) *Plant Cell Physiol.* 31, 519–526.
- 127 Pichersky, E., Hoffman, N.E., Malik, V.S., Bernatzky, R., Tanksley, S.D., Szabo, L. and Cashmore, A.R. (1987) *Plant Mol. Biol.* 9, 109–120.
- 128 Kojima, K., Yamamoto, N. and Sasaki, S. (1992) *Plant Mol. Biol.* 19, 405–410.