AN INVESTIGATION OF THE LHC FAMILY OF GENES AND PROTEINS IN THE HOMOSPOROUS FERN *Ceratopteris richardii.*

O. Oberschmidt [1], S. Kim[2], J.M. McGrath[2], M.M. Jansco[2], R. Bassi[3,4], B. Piechulla[1] & E. Pichersky[2]
[1]Institute of Plant Biochemistry, University of Gottingen, 37073 Gottingen Germany, [2]Biology Department, University of Michigan, Ann Arbor MI 48109 USA, [3]Biology Department, University of Padova, Padova Italy, [4]Present address: Faculty of Science, University of Verona, Verona Italy

1. Introduction

At least ten different types of chlorophyll *a/b*-binding (CAB) polypeptides per plant have been found in angiosperms and gymnosperms (1). The expression characteristics of the *Lhc* genes, which encode the CAB proteins, have been examined in great detail in tomato and other plants (2). Their transcription in angiosperms follows a distinct pattern; maximal transcription occurs at midday, whereas no transcription occurs during the night. Control of this pattern is attributed to a circadian 'clock' (3).

In contrast to the extensive investigation of the *Lhc* gene family in seed plants, the corresponding proteins of non-seed plants, such as the ferns and fern allies, have been little characterized. Several *Lhcb* genes (encoding LHCII proteins) have been previously isolated from a homosporous fern, *Polystichum munitum*, but most of these were defective and unlikely to be expressed (4). The single potentially functional gene appeared to encode an LHCII CAB polypeptide, but its mode of expression was not investigated. Here we report the isolation of several *Lhcb* cDNA clones from another homosporous fern species, *Ceratopteris richardii*. We show that there are several distinct LHCII and LHCI CAB proteins in this fern, and that the *Lhcb* genes examined in this study are under the control of a circadian clock. We have also found that these genes are expressed in both gametophytic and sporophytic tissues.

2. Results and Discussion

2.1 Isolation and characterization of LHCI and LHCII: Unstacked *C. richardii* thylakoids (chl *a/b* ratio=2.767) solubilized with 0.06% (w/v) dodecylmaltoside and fractionated by flat bed IEF yielded twelve green fractions. Analysis of each of these fractions on SDS-PAGE revealed that in many cases, some of the proteins were present in more than one IEF fraction. For example, IEF fractions 1 through 3 contain LHCII proteins in the 28-30 kDa range in different relative amounts. To resolve the IEF fractions into smaller, non-denatured chlorophyll-binding complexes, each green band (or a small group of closely spaced bands) were subjected to a sucrose gradient ultracentrifugation. Each sucrose gradient fractionation yielded several green bands, which were analyzed with respect to their absorption spectra and cross-reactivity of their polypeptides (in Western blots) with antibodies against specific CAB and PS proteins. These results are summarized in Table I. The heaviest sucrose gradient band in fractions 4-10 was identified as a PSI-LHCI complex, which had a chl absorption maximum at 679.2 nm. The PSI reaction center was isolated in IEF fraction 4, whereas the PSII reaction center was found in IEF fractions 5-10.

10. Photosynthesis Cong.
20.-25. Aug. 1995
Montpellier

An additional 26 kDa polypeptide was present only in fraction 3 and it was tentatively identified as CP24. Its chl *a* absorption maximum is 673.4 nm, which is red-shifted when compared with maize LHCII (674.8 nm). IEF fractions 11 and 12, which appear enriched in polypeptides of 28 and 29 kDa, have low chl *b* content and their chl *a* absorption maximum is at 676 nm; these characteristics are indicative of the presence of the CP26 complex (5).

Protein sequence determination of several electroeluted proteins from LHCI and LHCII was attempted. All were found to be N-terminally blocked, except for one protein in LHCII. The position of this N-terminal sequence, about a dozen residues shorter than in LHCII Type I and II CAB proteins (Fig. 1), strongly suggests that it is derived from an LHCII Type III CAB protein, encoded by an *Lhcb*3 gene. In angiosperms too this is the only LHCII CAB protein that is not N-terminally blocked (6).

Table 1

IEF green band No.	Sucrose gradient band No.	Chl a/b ratio	Chl a Red Max (nm)	Immunolgoical Identification
1	1	1.15	673.80	LHCII
	2	1.15	673.70	LHCII
2	1	1.15	673.50	LHCII
	2	1.37	673.00	LHCII
3	1	1.11	673.10	LHCII
	2	1.22	673.60	LHCII + CP24(?)
	3	1.51	672.90	LHCII + CP24(?)
4	1	9.53	679.00	PSI-LHCI
	2	>20	677.30	PSI RC
	3	3.17	676.30	?
	4	3.32	673.10	LHCII
5 - 7	1	8.48	679.30	PSI-LHCI
	2	>20	675.10	PSII RC
	3	>10	673.90	?
	4	4.41	675.50	?
	5	4.76	673.90	?
8 - 10	1	8.52	679.10	PSI-LHCI
	2	>20	674.80	PSII RC
	3	>10	674.20	PSII RC
	4	4.15	675.20	?
11 - 12	1	3.61	675.90	?
	2	>15	676.10	CP29/CP26

2.2 Isolation and characterization of Lhcb cDNA clones: A cDNA library (7) made from mRNA isolated from light-germinated green gametophytes (*C. richardii* gametophytes require light for germination) was screened with *Lhcb* gene probes from tomato and *Polystichum munitum*. Two full-length and one partial cDNA clones (*Lhcb**1, *Lhcb**2, *Lhcb**3) were obtained and characterized. The amino acid sequences of the proteins encoded by *Lhcb**1 and *Lhcb**2 (Fig. 1) indicate that they are very similar to the previously characterized *P. munitum Lhcb* gene (92-97% amino acid identity) and also to LHCB1 and LHCB2 sequences from angiosperms (89%-91% amino acid identity with the tomato *Lhcb*1*7 gene product). As was shown earlier (4), it is not possible to distinguish LHCB1 and LHCB2 sequences in ferns since the gene duplication that gave rise to the two types may have occurred in the angiosperm lineage after it diverged from the fern lineage. The incomplete cDNA clone of Lhcb*3 encodes a protein that is more divergent from any of these fern and angiosperm sequences (76-80% amino acid sequence identity), but the shortness of the sequence precludes a definite assessment of its character. Southern blots of *C. richardii* genomic DNA with cDNA clone Lhcb*1 (Fig. 1 in ref. [8]) indicate that the number of *Lhcb* genes in this species is >5, and possibly >10, consistent with the data showing multiple LHCII proteins (Table 1). The molecular data

show that *C. richardii* has at least one *Lhcb*3-like gene encoding an LHCII Type III CAB protein, and several *Lhcb* genes encoding CAB polypeptides of intermediate LHCII Type I - Type II nature.

```
                                          ↓
1.  ..TS.MA...STFA.KAVKLSPSSSEIT--..NG..T.RKTAT.AKPAS...P......VK.......ES..
2.  ..TS..A.T.TFS-....KPVN-...............-....SK.--........S...............
3.  MAAATASLSSASVIGQQLN----ELSRKVGAGEARVQM-MAPKKSA--PSGSIWYGPDRPLYLGPFSGSPPS
4.  ..........SVFS....KPAAS...............RTSS..P.--........................
5.  ------------------------------------------------------------------------
6.                                                     DY.......K...
7.                                                   SNDL......VK...

1.  ... ..............................C...........F....R...............S...S
2.  ..S...............................S................................S....
3.  YLTGEFPGDYGWDTAGLSADPETFAKNRELEVIHARWAMLGALGCVTPELLAKNGVKFGEAVWFKAGAQIFA
4.  ..............................................I.D.......................
5.  ------------------------------------------------...................S....

1.  ..........................V..................V..........................
2.  ................................................I.......................
3.  EGGLDYLGNPSLVHAQSILAIWACQVILMGAVEGYRVAGGPLGEVEDPLYPGGSFDPLGLADDPEAFAELKV
4.  ............I.......................I...................................
5.............N...............V..........G......GL..I...A...........D.......

1.  ..I........................L...A......VN.....F..N.V...
2.  ......................................A.........N.....
3.  KELKNGRLAMFSMFGFFVQAIVTGKGPIENLSDHLADPTVNNAWAYATQFTPGK
4.  ................................................N.V...
5.  ..I.........................K..........E........N.....
```

Fig. 1: Comparison of sequences of protein encoded by *Lhcb* genes. 1) tomato *Lhcb*1*7, 2) *P. munitum Lhcb**1 (formerly "F3"), 3) *C. richardii Lhcb**1, 4) *C. richardii Lhcb**2, 5) *C. richardii Lhcb**3, 6) N-terminal sequence of a *C. richardii* LHCII protein, 7) N-terminal sequence of tomato LHCB3*1. Dots indicate identity with sequence shown in full. Dashes indicate missing sequence or introduced gaps.

2.3 *Expression of Lhcb genes in C. richardii*: Northern blots of mRNA extracted from green gametophytes show substantial expression of *Lhcb* genes (data not shown). Northern blots from sporophytic leaves (Fig. 2) show that steady-state levels of total *Lhcb*

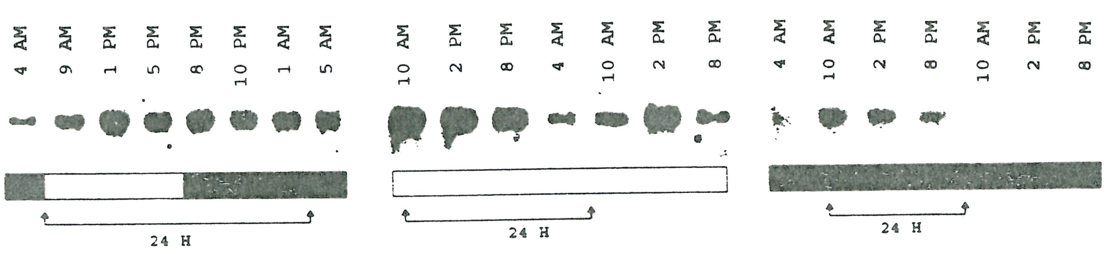

Fig. 2: Northern blots of mRNA extracted from plants grown under different light regimes. mRNAs were extracted at times indicated.

mRNA of plants grown under 12 h light/12 h dark period peak in early afternoon and then decline (see also ref. 9). Some mRNA is seen throughout the dark period, which may indicate that *Lhcb* mRNA in *C. richardii* is rather stable in the dark. In plants subjected to continuous illumination or darkness, the levels of *Lhcb* mRNA continue to oscillate (Fig. 2), indicating that the expression of the *Lhcb* genes in *C. richardii* is under a circadian

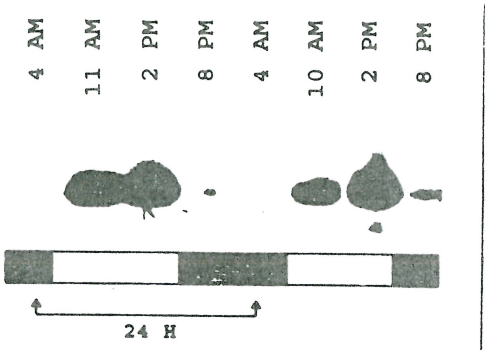

Fig. 3: Primer extension experiment on mRNAs extracted from plants growing under light/dark regime, using a primer specific for *Lhcb**1.

control. Primer extension experiments using gene-specific oligonucleotides show that the amplitudes of oscillation of some *Lhcb* mRNAs is more pronounced, *i.e.*, there is almost no mRNA in the dark period (*e.g.*, *Lhcb**1, Fig. 3), which might indicate that such mRNAs are least stable in the dark (2). The primer extension experiments with both *Lhcb**1 and *Lhcb**2 also showed that both genes are expressed in gametophytic and sporophytic tissues, with Lhcb*2 being more abundant in both tissues (data not shown).

3. REFERENCES

1 Jansson S (1994) Biochim Biophys Acta 1184:1-19.

2 Kellmann J, M Wiesse M, Pichersky E, Piechulla B (1993) Mol Gen Genet 237:439-448

3 Piechulla (1993) Plant Mol Biol 22:533-542.

4 Pichersky E, Soltis P, Soltis D (1990) Proc Natl Acad Sci 87: 195-199.

5 Bassi R, and Simpson D. (1987) In Progress in Photosynthesis Research (edited by J. Biggins), Vol. II, pp 81-88. Martinus Nijhof, Dordrecht.

6 Schwartz E, Stasys R, Aebersold R, McGrath JM, Green BR and Pichersky E (1991) Plant Mol Biol 17:923-925.

7 McGrath JM, Hickok LG, Pichersky E (1994) Plant Syst Evol 189: 193-202.

8 McGrath JM, Hickok LG, Pichersky E (1994) Plant Syst Evol 189: 203-210.

9 Oberschmidt O, Hucking C, Piechulla B (1995) Plant Mol Biol 27: 147-153.