Plant Physiol. (1993) 102: 1047–1048

Plant Gene Register

Chlorophyll *a*/*b*-Binding Protein Gene Expression in Cotton[1]

David M. Anderson*, Richard L. Hudspeth, Susan L. Hobbs, and John W. Grula
Phytogen, 101 Waverly Drive, Pasadena, California 91105

To investigate the regulation of genes encoding CAB proteins in cotton (*Gossypium hirsutum*) and identify potentially useful promoters, we determined the structure and expression of two genes (*Lhcb1*1* and *Lhcb3*1*) encoding type I and type III proteins of the LHCII subfamily (Table I).

Northern analyses of the gene expression pattern in cotton demonstrated substantial LHCII transcript accumulation in the following photosynthetic organs and tissues: leaves (greatest accumulation), bracts, seedling stems, pericarp, and immature petals (least accumulation). No LHCII transcripts were detected in roots, mature petals, seeds, anthers, ovaries and pistils, or immature fiber and ovules. Although the pattern of expression was the same with either an *Lhcb1* or an *Lhcb3* gene probe, differences in the intensity of hybridization suggested that *Lhcb1* transcripts may be as much as 10-fold more abundant than *Lhcb3* transcripts.

To test the level of expression conferred by the *Lhcb1*1* gene promoter, we linked its upstream and downstream regions (approximately 700 and 500 bp of sequence, respectively) to the *GUS* gene (Jefferson, 1987). In transient expression assays performed with cotton cotyledons (Hamilton et al., 1992), this construction exhibited only 5-fold less activity than a very highly expressed GUS construction containing a double CaMV 35S enhancer sequence and the tobacco mosaic virus omega sequence (Skuzeski et al., 1990). The relatively high expression conferred by the *Lhcb1*1* gene promoter may be due in part to three GATA motifs that reside between the TATA and CAAT box sequences. These GATA motifs are highly conserved in *Lhcb1* gene promoters from several different plants (reviewed in Castresana et al., 1987) and are also found in the CaMV 35S promoter (Lam and Chua, 1989).

Table I. *Characteristics of CAB genes Lhcb1*1 and Lhcb3*1 from cotton*

Organism:
Gossypium hirsutum cv Acala SJ5 (*Lhcb1*1*) and cv Coker 201 (*Lhcb3*1*).

Techniques:
Acala SJ5 genomic library was made with DNA isolated from delinted seed and required an *mcr*⁻ host strain and packaging extract (Graham et al., 1990). Transient expression assay employed cotton cotyledons and a biolistic device (Hamilton et al., 1992).

Sequence Identification and Gene Nomenclature:
Sequence and structural homologies with other *CAB* genes in the LHCII subfamily (Green et al., 1991; Schwartz et al., 1991). Nomenclature according to Jansson et al. (1992).

Structural Features of Genes:
*Lhcb1*1* has no introns; *Lhcb3*1* has two introns of 210 and 330 bp.

Structural Features of Deduced Amino Acid Sequence:
The *Lhcb1*1* gene sequence predicts a transit peptide of 34 amino acids and mature polypeptide of 230 amino acids (M_r 24,804). The *Lhcb3*1* sequence predicts a transit peptide of 39 amino acids and a mature polypeptide of 223 amino acids (M_r 24,330).

Gene Copy Number:
Genome blots and cDNA clone analyses indicate cotton (Acala SJ5) has at least 15 members in the *LHCII* gene subfamily, with 7 to 10 *Lhcb1* genes and 6 or more *Lhcb3* genes.

ACKNOWLEDGMENTS

We thank Drs. Joe Kamalay and Leon Dure for their Coker 201 genomic library, and Dr. Linda Walling for a soybean CAB clone used to screen the library.

Received December 28, 1992; accepted January 9, 1993.
Copyright Clearance Center: 0032-0889/93/102/1047/02.
The GenBank/EMBL accession numbers for the sequences reported in this article are L07119 for *Lhcb1*1* and L07120 for *Lhcb3*1*.

LITERATURE CITED

Castresana C, Staneloni R, Malik VS, Cashmore AR (1987) Molecular characterization of two clusters of genes encoding the type I CAB polypeptides of PSII in *Nicotiana plumbaginifolia*. Plant Mol Biol **10:** 117–126

Graham MW, Doherty JP, Woodcock DM (1990) Efficient construction of plant genomic libraries requires the use of *mcr*– host strains and packaging mixes. Plant Mol Biol Rep **8:** 18–27

Green BR, Pichersky E, Kloppstech K (1991) Chlorophyll *a*/*b*-binding proteins: an extended family. Trends Biochem Sci **16:** 181–186

Hamilton DA, Roy M, Rueda J, Sindhu RK, Sandford J, Mascarenhas JP (1992) Dissection of a pollen-specific promoter from maize by transient transformation assays. Plant Mol Biol **18:** 211–218

Jansson S, Pichersky E, Bassi R, Green BR, Ikeuchi M, Melis A, Simpson DJ, Spangfort M, Staehelin LA, Thornber JP (1992) A

[1] Supported by the J.G. Boswell Company (Los Angeles).
* Corresponding author; fax 1–818–792–0823.

Abbreviations: CAB, chlorophyll *a*/*b*-binding; CaMV, cauliflower mosaic virus; GUS, glucuronidase; LHCII, light-harvesting complex II.

nomenclature for the genes encoding the chlorophyll *a/b*-binding proteins of higher plants. Plant Mol Biol Rep **10:** 242–253

Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol Biol Rep **5:** 387–405

Lam E, Chua N-H (1989) ASF-2: a factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in *Cab* promoters. Plant Cell **1:** 1147–1156

Schwartz E, Stasys R, Aebersold R, McGrath JM, Green BR, Pichersky E (1991) Sequence of a tomato gene encoding a third type of LHCII chlorophyll *a/b*-binding polypeptide. Plant Mol Biol **17:** 923–925

Skuzeski JM, Nichols LM, Gesteland RF (1990) Analysis of leaky viral translation termination codons *in vivo* by transient expression of improved β-glucuronidase vectors. Plant Mol Biol **15:** 65–79