

Update section

Sequence

Sequence of a tomato gene encoding a third type of LHCII chlorophyll *a/b*-binding polypeptide

Egbert Schwartz,¹ Robert Stasys,¹ Ruedi Aebersold,³ J. Mitchel McGrath,¹ Beverley R. Green² and Eran Pichersky^{1*}

¹Biology Department, University of Michigan, Ann Arbor, MI 48109, USA (*author for correspondence);

²Botany Department, University of British Columbia, Vancouver, B.C., Canada; ³Biomedical Research Centre and Department of Biochemistry, University of British Columbia, Vancouver, B.C., Canada

Received 31 May 1991; accepted 4 June 1991

Genes encoding two types of chlorophyll *a/b*-binding (CAB) polypeptides localized in the light-harvesting complex (LHCII) of Photosystem II (PSII) have been characterized in a number of plant species (reviewed in [3, 4]). Type I genes contain no introns and encode polypeptides whose mature forms (232–234 residues) migrate on SDS-PAGE with an apparent molecular mass of 26–28 kDa. Type II genes contain a single intron and encode polypeptides 2–4 residues shorter than the Type I LHCII polypeptides. The mature Type I and Type II LHCII CAB proteins are approximately 85% identical in their primary amino acid sequence. In tomato, Type I and Type II correspond to the two major LHCII polypeptides [4, 8; B.R. Green *et al.*, in prep.]. Both types are believed to be phosphorylatable at Thr or Ser residues near the N-terminus [2]. LHCII has a third minor polypeptide of about 25 kDa [6] which is not phosphorylated [1] and is not N-terminally blocked [7]. Morishige and Thornber [7] obtained a 59 residue N-terminal sequence of this protein from barley, which showed that it was distinct from both Type I and II sequences but more related to them than to other CAB polypeptides. They therefore designated it LHCII Type III CAB [7]. Shorter N-terminal sequences

identical or almost identical to the barley sequence were also obtained from wheat [10] and maize (R. Bassi, pers. comm.). Comparisons of the Type III N-terminal sequences with the corresponding Type I and II sequences indicate that the N-terminal sequence of the Type III protein is shorter by 7–11 residues [7, 10].

We have isolated and characterized tomato genomic and cDNA clones encoding an LHCII Type III CAB polypeptide (Fig. 1). It is designated *cab13* following our earlier CAB gene nomenclature [9] or *lhbc1* according to the nomenclature proposed by Jansson and Gustafsson [5]. The gene contains two introns (Fig. 1) and resides on chromosome 12 (data not shown). It encodes a precursor polypeptide of 265 residues. N-terminal sequencing of two tomato Type III polypeptides gave the sequence XNDLWYGP-DXVXYL.. (identity of X uncertain) which is compatible with Ser as the first residue of the mature protein (Fig. 1). This means that the mature polypeptide encoded by *cab13* has 223 residues and commences with a Ser residue (arrowhead in Fig. 1) not a Gly as reported for the corresponding protein in the other three species. The calculated molecular mass of the mature polypeptide is 24.3 kDa. In addition to *cab13*

```

TTTTAAAAATCAAGATAGGTGATATAAAGTGGAGATATAGTTCATATCTCTTCAAGCCAA 60
GAGATTGGTTAATCAACACTACACAAAACATCTCCAACCACAATTTCTACTTGTGCAATA 120
TTAGCCACCCACGTAGGATTCTCATTTTCTCTATACAAATAACAACCAATTCCCTCTTTA 180
GAACCATATCCACATTTGCCCAAGCACAAAATATATTTTGGAGAAGTTGAAGAAA ATGGC 239
M A
ATCAATGGCAGCAACAGCTAGCTCCACCACAGTTGTAAAGCAACACCATTTTTGGGCCA 299
S M A A T A S S T T V V K A T P F L G Q
AACCAAAAATGCTAACCTCTTAGAGATGTTGTTGCCATGGGCTCTGCCAGATTACCATT 359
T K N A N P L R D V V A M G S A R F T M
G GTATATATACATATATTCTCTTTCTTTTCTCCATGTTTTAGTTTTTATGCTAGTTT 418
ATCTTGGTATTGCCAAGCTATGAAAACAGTTTCTTGAGTATCCTTCTCGGACCTCAGAC 478
GCGAGAGCTTGTGCATACCTTTTTTTTCTTTTCTTCTACTGGATTGATGTGACTGTAT 538
TTTTTTTCTCTCATCTTATTAG AGTAATGATTGTGGTACGGACCTGACCGTGTAAAG 597
↓ S N D L W Y G P D R V K
TACTTGGGACCCTTTTCTGCTCAAACCTCTTACTTGAACGGAGAGTTCCCTGGTGAC 657
Y L G P F S A Q T P S Y L N G E F P G D
TACGGATGGGACACTGCTGGTTTTATCCGCTGATCCCGAGGCCCTTGGCCAAGAACAGAGCT 717
Y G W D T A G L S A D P E A F A K N R A
CTTGAG GTCAGTTATTTGTTCACTCACACTTACCAAAAAGAGAACTAAGATTTGCATAC 776
L E
GTCCTACCTTCCCTCAGACGCCACTCATTTTGCTTAGTGACATTGAACTATATAG GTTAT 835
V I
CCATGGAAGATGGGCAATGCTTGGTGCACACTAGGTTGCATTTTTCCCGAAGTTCTTGAAAA 895
H G R W A M L G A L G C I F P E V L E K
ATGGGTGAAGGTAGACTTCAAAGAACAGTATGGTCAAAGCAGGATCACAAATCTTCAG 955
W V K V D F K E P V W F K A G S Q I F S
TGACGGTGGACTTGACTACTTGGGAAACCCTAACCTTGTCCATGCTCAGAGTATTCTAGC 1015
D G G L D Y L G N P N L V H A Q S I L A
TGTGCTAGGTTTCCAAGTGGTGCCTTATGGGACTTGTGAAGGTTTCAAGAAATCAACGGACT 1075
V L G F Q V V L M G L V E G F R I N G L
TCCAGGCGTTGGAGAAGCAACGACTTATACCCCGGTGGCCAATATTTGACCCACTCGG 1135
P G V G E G N D L Y P G G Q Y F D P L G
CCTTGGCGATGACCCTACCACATTTGCTGAACTCAAGGTAAAGGAAATCAAAAACGGAAG 1195
L A D D P T T F A E L K V K E I K N G R
ACTAGCTATGTTCTCCATGTTTGGATTCTTCTGTTCAAGCTATTGTTACCGGTAAGGCC 1255
L A M F S M F G F F V Q A I V T G K G P
TCTTGAAAACCTATTAGATCACCTTGACAACCTGTGGCTAACAAATGCTTGGGTTTACGC 1315
L E N L L D H L D N P V A N N A W V Y A
CACTAAGTTTGTCTTGGAGCT TAGATTTATGTTCAATTCATTGGAAGTACTACCATT 1374
T K F V P G A
TGTGTAACCTGAAATCTTCAGAAATGCAAAATATATGTTAGATGCAATCGTTTTCAACCTTGT 1434
TTACTAACATTTGTAGAGATAAAGAGGTCATGAGCACAAAATGAACTATTAGAATAGAAA 1494
ATTTGGATATAACAATTTTCAATTTACTGAACTCAATCAAATCTATTTCTATATTCAATCTT 1554
CTGTTCTTACACCAAATCTAATCGAAAATCACAAATCCAACAGAAAATCGACAGAAAAC 1614
AAACAAATCTAAATCTAAGAGCTAGTGAATTTGGTAACAGCTTTAGTACCCTCAGAGACG 1674
GCGTGCTTAGCCAAATCACCCGGGAAGGACCAATCGAACAGCAGTTTGAATCTCCCGAGAA 1734
GTGATGGTAGGCTTTTGTGTTGTTACCTAGCAAGCCTGGATGATTCCTGAGCAAGCTTCTCG 1794
AAGATATC 1802

```

Fig. 1. Nucleotide sequence of the tomato *cab13* (*lhbc1*) gene. Intron positions were determined by comparing the genomic sequence with the sequence of a full-length cDNA clone. The amino acid sequence is shown below the nucleotide sequence. The peptide sequence obtained from the N-terminus is underlined. The arrow indicates the position of cleavage of the precursor.

(*lhbc1*), the tomato genome contains at least two other LHCII Type III CAB genes, and all three genes encode almost identical proteins (B.R. Green *et al.*, in prep.). The mature tomato Type III polypeptides are approximately 80% identical to either the Type I or Type II LHCII CAB polypeptides.

References

1. Bassi R, Rigoni F, Barbato R, Giacometti GM: Light-harvesting chlorophyll a/b proteins (LHCII) populations in phosphorylated membranes. *Biochim Biophys Acta* 936: 29–38 (1988).
2. Michel HP, Buvinger WE, Bennet J: Redox control and sequence specificity of a thylakoid kinase. In: Baltscheff-

- sky (ed) *Current Research in Photosynthesis*, vol 2, pp. 747–753. Kluwer Academic Publishers, Dordrecht (1990).
3. Chitnis PR, Thornber JP: The major light-harvesting complex of Photosystem II: aspects of its molecular and cell biology. *Photosyn Res* 16: 41–63 (1988).
 4. Green BR, Pichersky E, Kloppstech K: Chlorophyll *a/b*-binding proteins: an extended family. *Trends Biochem Sci* 16: 181–186 (1991).
 5. Jansson S, Gustafsson P: Evolutionary conservation of the chlorophyll *a/b*-binding proteins: cDNA encoding Type I, II and III LHCI polypeptides from the gymnosperm Scots pine. *Mol Gen Genet*, in press (1991).
 6. Machold O: Polypeptide composition of the photosystem I chlorophyll *a/b*-protein from three different plant species. *Carlsberg Res Commun* 51: 227–238 (1986).
 7. Morishigie DT, Thornber JP: The major light-harvesting chlorophyll *a/b* protein (LHC IIb): The smallest subunit is a novel CAB gene product. In: Baltscheffsky (ed) *Current Research in Photosynthesis*, vol 2, pp. 261–264. Kluwer Academic Publishers, Dordrecht (1990).
 8. Pichersky E, Hoffman NE, Malik VS, Bernatzky R, Tanksley SD, Szabo L, Cashmore AR. The tomato *Cab-4* and *Cab-5* genes encode a second type of CAB polypeptides localized in PS II. *Plant Mol Biol* 9: 109–120 (1987).
 9. Pichersky E, Green BR: The extended family of chlorophyll *a/b*-binding proteins of PSI and PSII. In: Baltscheffsky (ed), *Current Research in Photosynthesis*, vol 3, pp. 553–556. Kluwer Academic Publishers, Dordrecht (1990).
 10. Webber AN, Gray JC: Detection of calcium binding by photosystem II polypeptides immobilised onto nitrocellulose membrane. *FEBS Lett* 249: 79–82 (1989).