Volume 13 Number 7 1985 Nucleic Acids Research

The petunia chlorophyll *a/b* binding protein genes: a comparison of *Cab* genes from different gene families

Pamela Dunsmuir[1]

CSIRO Division of Plant Industry, Canberra, Australia

Received 18 January 1985; Accepted 5 March 1985

ABSTRACT

In Petunia (Mitchell) there are at least 16 genes which encode the chlorophyll a/b binding proteins; these genes have been classified into small multigene families based upon nucleotide sequence homology (1). A gene from each of five distinct Cab gene families is compared here. These genes have uninterrupted open reading frames of 266 or 267 amino acids corresponding to the Cab precursor proteins of sizes around 32000 daltons. A comparison of the amino acid sequences deduced here with published information from direct NH_2-terminal analysis of a mature Cab protein in pea (10) suggests that a 34-36 amino acid transit peptide is cleaved from the NH_2-terminal of the petunia precursor proteins. The proposed transit peptide sequences are more divergent than the mature peptide sequences between the Cab genes from different gene families. There are two regions within the mature Cab proteins which are conserved between all genes - a sequence of 28 amino acids near the NH_2 terminal, and another sequence of 26 amino acids in the middle of the protein. The DNA sequences proximal to the Cab coding regions contain typical eukaryote promoter elements - TATA and CCAAT boxes, and in addition those genes which are known to be expressed in petunia leaf tissue also have an extensive region of homology (48 nucleotides) centered at approximately 130 nucleotides from the proposed transcription start sites.

INTRODUCTION

In higher plants primary light energy capture is performed by a membrane bound complex consisting of chlorophylls a and b and the major thylakoid proteins - the chlorophyll a/b binding proteins (Cab proteins). The complex is termed the light harvesting complex (LHC) and it is in close association with the photosystem to which it delivers light energy. The LHC associated with photosystem II (LHC-II, which is the major LHC) can be isolated from the thylakoid membranes and it contains approximately 50% of the chlorophyll in mature chloroplasts (2).

The complex is composed of several distinct Cab polypeptides and the chlorophylls. The nature of the interaction between the proteins and chlorophylls is ill-defined; it is unknown whether both chlorophylls bind to a single peptide or whether specific peptides bind only chlorophyll a or chlorophyll b.

The Cab proteins of the major light harvesting complex, LHC-II, are known to be specified by multiple nuclear genes in a range of higher plant species, [petunia, barley, tobacco (1), pea, wheat (3,4) and lemna (5)]. The Cab proteins are synthesized in the cytoplasm as precursor polypeptides of about 30-32 kD. These precursor proteins undergo post-translational cleavage upon, or after, transport into the chloroplast where they bind chlorophylls a and b and are assembled as light harvesting structures in the photosynthetic membranes. The mature Cab proteins which are isolated from the chloroplast range in size from 26-28 kD. This transition between the cytoplasmic Cab precursor protein to the chloroplast Cab mature protein is associated with the removal of a fragment around 4 kD which has been termed the transit sequence (6).

In Petunia (Mitchell) we have estimated that there are approximately sixteen nuclear genes encoding the Cab proteins of the LHC-II (1). Analyses of a number of distinct Cab cDNA clones led us to classify these genes into several small multigene families based upon nucleotide sequence homology in the protein coding regions, and also in the adjacent 3' untranslated regions of the expressed genes. Further, we reported that at least one Cab gene from each of four separate gene families is expressed in green leaf tissue (1).

Here we report nucleotide sequence analyses for five genes from different Cab gene families. These nucleotide sequence data enable us to deduce the complete amino acid sequences of the Cab precursor proteins in petunia and hence provide information about the transit peptide sequences as well as the mature Cab protein sequences. In addition to the comparison of nucleotide and amino acid sequences within the Cab gene protein coding regions, we present a comparison of the 5' and 3' flanking regions of the genes.

METHODS

The growth of Petunia (Mitchell), extraction of nucleic acids, and the construction of the cDNA clones, and the petunia-Charon 28 clone library has been described in detail (1). The sequence information presented here was assembled from analyses by the chemical method of Maxam and Gilbert (7) and/or the chain-termination method of Sanger et al (8). The sequences of both DNA strands were determined with either one of the two methods.

RESULTS AND DISCUSSION

The construction of the genomic library of petunia EcoRI fragments in the vector Charon 28, and the isolation and localization of the Cab protein genes on the EcoRI fragments have been described (1). The Cab gene regions were subcloned from the genomic fragments into pUC plasmids and M13-mp phage (9) for nucleotide sequence analyses. The Figure 1 diagrams the restriction maps within the regions which have been sequenced for

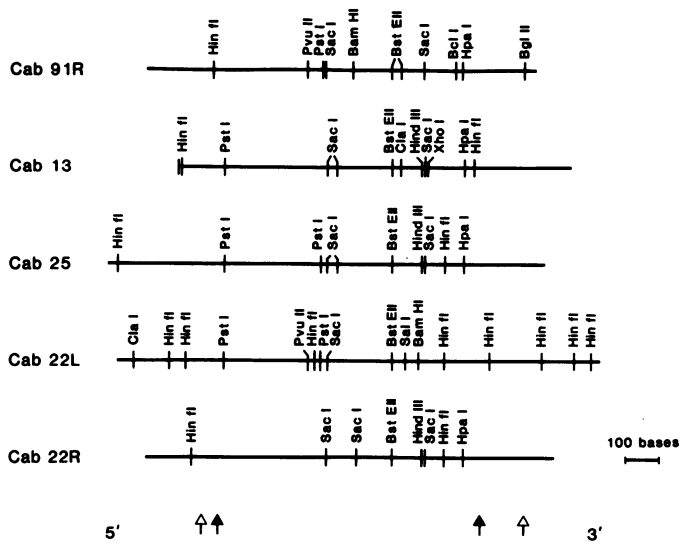

Figure 1 Restriction maps of the Petunia (Mitchell) Cab genes. The regions diagrammed correspond to the regions of the Cab phage (1) which have been sequenced. The proposed gene structure is also diagrammed: ⇧ - approximate transcript length, ⬆ - precuror protein length.

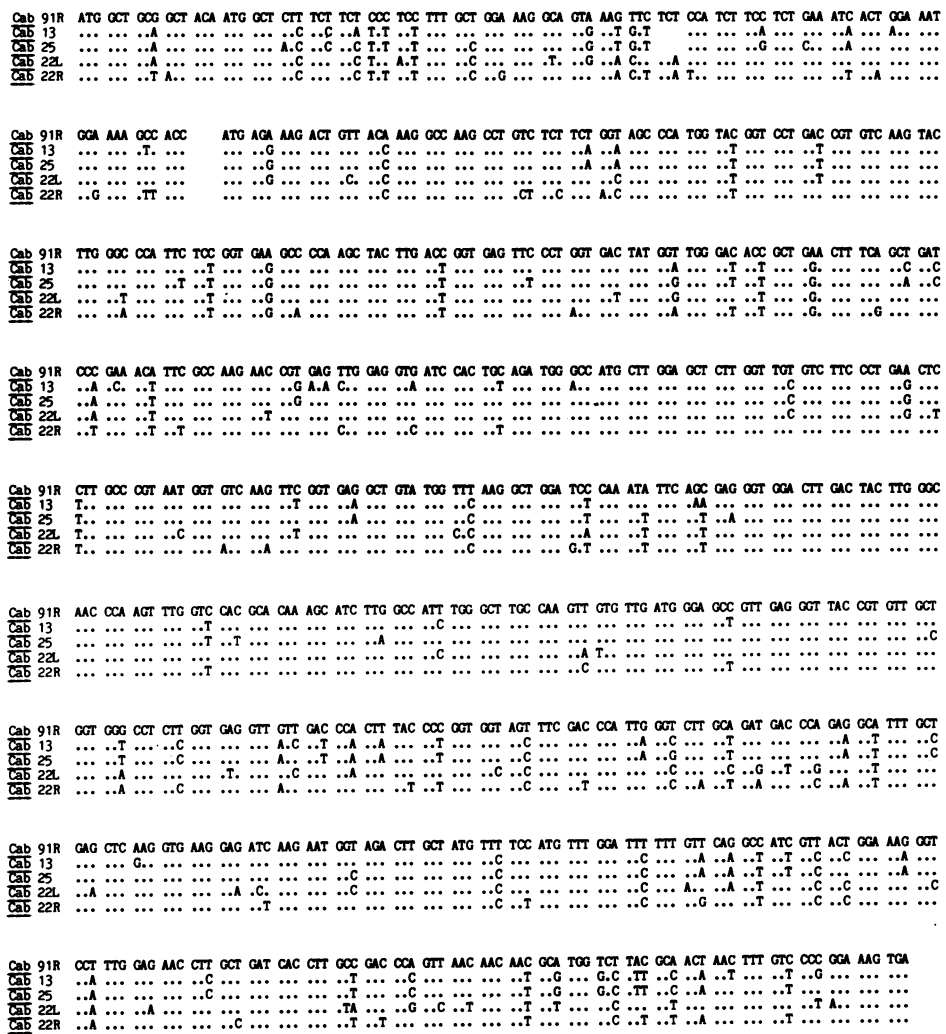

Figure 2 Complete nucleotide sequences for the coding region of each of the Cab genes. Nucleotide replacements are shown, or dots indicate the same nucleotides. A space which corresponds to three nucleotides has been inserted into the Cab 13 and Cab 25 sequences in order to achieve the best alignment with the other Cab genes.

the five Cab genes [Cab 25: 1340 nucleotides, Cab 13: 1200 nucleotides, Cab 22L: 1490 nucleotides, Cab 22R: 1200 nucleotides and Cab 91R: 1160 nucleotides]. In petunia genomic DNA, the Cab genes 22L and 22R are closely linked with an

inverted orientation, and the Cab gene 91R is closely linked in tandem to the gene Cab 91L which is not discussed here (1).

(1) The Cab protein sequence

For each of the Cab genes which have been sequenced, there is only one open reading frame that could correspond to the expected Cab precursor protein length. The nucleotide sequences and the deduced amino acid sequences for each of these different Cab gene coding regions are shown in Figures 2 and 3 respectively. In Cab 13 and Cab 25 the open reading frame extends for two hundred and sixty-six amino acids from the first ATG to the translation termination signal TGA. In Cab 22L, Cab 22R and Cab 91R there are two hundred and sixty-seven amino acids from the first ATG to the translation stop codon. The length difference results from there being an additional codon at amino acid position 21 in these genes (the amino acids are numbered from the NH_2-terminal of the precursor protein). Apart from this insertion/deletion of one nucleotide triplet, the five Cab protein coding regions are very similar.

These precursor proteins specified by the genes which were sequenced have estimated molecular weights which range in size from 32,743 to 33,029. There are no data that establish unequivocally the position of the transit polypeptide nor the NH_2-terminal residue of the mature Cab proteins, however the circumstantial evidence which follows suggests that the transit peptide is cleaved from the NH_2-terminal of the Cab precursor proteins and that the NH_2-terminal of the mature proteins is either the Met (M) residue at position 35 (position 34 in Cab genes 13 and 25) or the Arg (R) residue at position 36 (position 35 in Cab genes 13 and 25). There is some amino acid sequence information available about a mature Cab protein from pea; Mullet (10) has reported that the carboxyl terminal is lysine (K) and that the sequence (R,K)-S-A-T-T-K-K is located at the NH_2 terminal of the mature Cab polypeptides. Our sequencing data from petunia are consistent with these observations in pea. The carboxyl terminus of the Cab precursor proteins is lysine (K), and an amino acid sequence R-K-T-$\frac{V}{A}$-T-K-A-K, which is similar to that reported at the NH_2-terminal in pea, occurs in the precursor proteins at residue 36 (35 in Cab 13 and Cab 25). There are no

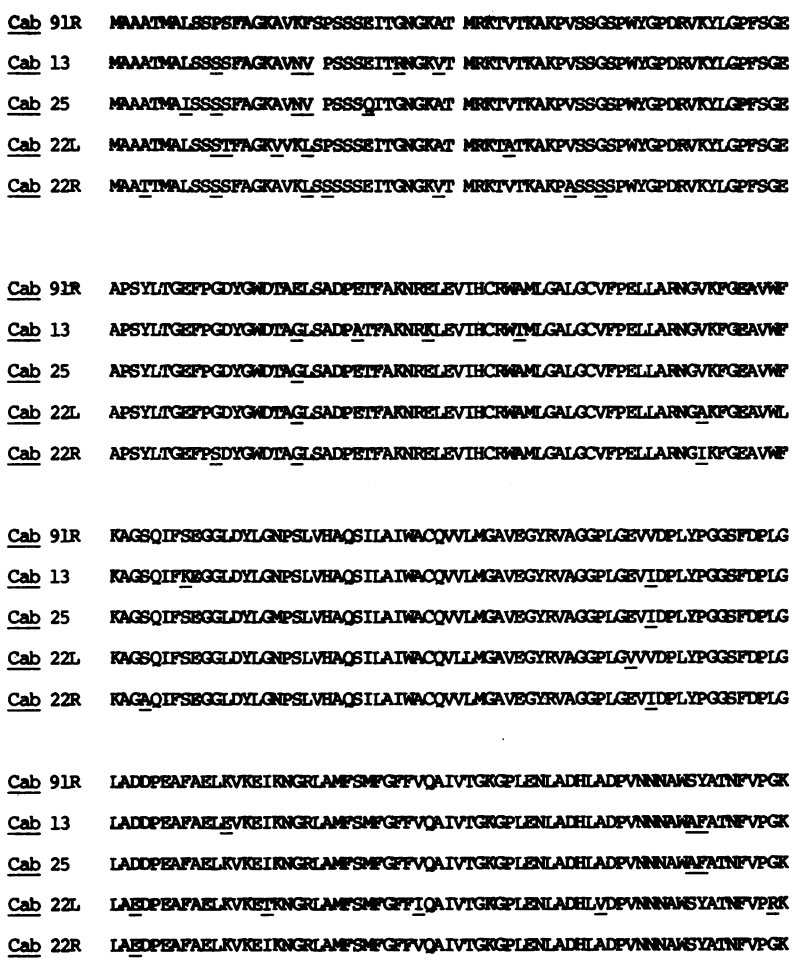

Figure 3 Complete amino acid sequences of the Cab polypeptides predicted from nucleotide sequence analyses. Amino acid replacements are indicated by underlining. A space which corresponds to one amino acid has been inserted into the Cab 13 and Cab 25 protein sequence.

other regions in the precursor protein where any similar sequences occur. If this sequence were at the NH_2-terminal of the mature Cab proteins in petunia then transit peptides of 35 or 34 residues (approximately 4100kD) would be cleaved from the NH_2-terminal of the precursor proteins to generate mature Cab proteins which range in size from 28700 to 28800 kD.

The small subunit polypeptides of the chloroplast enzyme ribulose bisphosphate carboxylase, like the Cab proteins, are encoded by nuclear genes which are translated into precursor proteins in the cytoplasm then post translationally transported into the chloroplast with cleavage of the transit peptide to produce the mature small subunit protein (11,12). Amino acid sequence data are available for this protein from a variety of plant species and it is established that the transit peptide is cleaved from the NH_2-terminal of the precursor protein to produce a mature protein with methionine (M) at the NH_2-terminal (12). It has been found that this NH_2-terminal methionine is frequently chemically modified in a variety of species, the effect of which is that the penultimate residue-glutamine (Q) is often incorrectly determined to be the mature peptide NH_2-terminal (13). Highfield and Ellis (12) have proposed that this NH_2-terminal methionine modification is associated with the removal of the transit peptide.

It is possible that the mature Cab proteins, like the mature small subunit proteins, have a modified NH_2-terminal methionine since the sequence in petunia which is similar to that reported as the NH_2-terminal in pea (9) i.e. R-K-T-$\frac{V}{A}$-T-K-A-K, is preceded by methionine. We do have some evidence that the NH_2-terminal of the petunia Cab proteins are blocked (Wolber and Dunsmuir, unpublished results). While it is not possible at present to state unequivocally the NH_2-terminal of the petunia Cab proteins, circumstantial evidence indicates that it is the methionine (M) or arginine (R) residue at position 35 or 36 in the precursor protein, (position 34 or 35 in Cab 13 and Cab 25).

(11) Comparison of the proposed transit peptides from different Cab genes

It is clear from the sequence data of Figure 2 and 3 that the proposed part of the Cab protein coding region which specifies the transit peptide [amino acid residues 1 to 34 for Cab 22L, Cab 22R and Cab 91R, and residues 1 to 33 for Cab 13 and Cab 25] shows greater nucleotide and amino acid sequence divergence between the genes than that which specifies the mature peptide. A summary of the nucleotide, and predicted amino acid alterations which occur within the proposed transit peptide

Table 1:

Gene	Nucleotide Replacements A	B	Total	Amino Acid Replacements C	D	Total
Cab 13	10	5	15	3	2	5
Cab 25	11	5	16	2	3	5
Cab 22L	8	4	12	3	1	4
Cab 22R	14	5	19	2	3	5

A comparison of the nucleotide an amino acid sequence divergence between the proposed transit peptides of Petunia (Mitchell) Cab genes. (The transit sequences are assumed to extend to the methionine residue at position 34/35). Each of the Cab genes is compared with the gene Cab 91R which has been chosen as the prototype since it is known to be transcribed in leaf tissue. A-nucleotide replacements which do not alter the encoded amino acid, B-nucleotide replacements which effect an amino acid replacement, C-amino acid changes within the same class of amino acid, D-amino acid changes between classes of amino acids.

region of the genes is shown in Table 1. For the purpose of these sequence comparisons, the gene Cab 91R has been chosen as a gene against which other sequences are compared. This gene has been chosen as the 'prototype' since it is known to be transcribed in petunia leaf tissue, (1).

The average nucleotide divergence within the proposed transit peptide coding regions of the different Cab genes is 15%; the predicted amino acid divergence within this region is about 8%. There are four examples of amino acid substitutions which effect the class of encoded amino acid, two of these occur at the same residue in both Cab 13 and Cab 25 (K to N, residue 19) the third in Cab 25 (E to Q), and the fourth in Cab 13 at position 28 (G to R).

There are at present no data from other plant species for the Cab protein transit peptide sequences, however there are data for the transit peptides of the small subunit of ribulose bisphosphate carboxylase from a range of species. For the small subunit gene that portion which encodes the transit peptide is far more divergent between species, than the portion which codes for the mature peptide (4,5); comparisons of different small subunit gene sequences within any one species are not available.

It is of interest to compare the transit peptides between the small subunit proteins and the Cab precursor proteins since

Table 2:

Gene	Nucleotide Replacements A	B	Total	Amino Acid Replacements A	B	Total
Cab 13	68	9	77	2	7	9
Cab 25	65	4	69	2	2	4
Cab 22L	61	9	70	6	3	9
Cab 22R	59	8	67	6	2	8

A comparison of the nucleotide and amino acid sequence divergence between the proposed mature Cab proteins of Petunia (Mitchell). Each of the Cab genes is compared with the gene Cab 91R. A, B, C and D are the same categories as in Table 1.

certain features of these sequences appear to be conserved. The small subunit transit sequences are quite basic, a feature proposed to facilitate interaction of the precursor protein and the highly acidic chloroplast envelope (12). The petunia Cab transit peptides are also positively charged having two or three basic residues and no acidic residues. There are only two regions of the small subunit transit peptide which are invariant between the all species so far analysed, these are the residues $SNGGR^+$ at -8 to -4 and the pair of residues around the processing site C,M at -1 and +1 [the numbering used here is with respect to the NH_2-terminal of the mature peptide, i.e., methionine is +1], (4). In the Cab transit peptides, the sequence NGK^+ at -5 to -3 is conserved between all genes, and the residues T,M occur in each gene at the proposed processing sites.

(111) Comparison of the mature peptide sequence between different Cab genes

The number of nucleotide sequence changes and the predicted amino acid replacements within the mature peptide coding sequence calculated from pairwise comparisons of each of the Cab genes with Cab 91R are summarized in Table 2. The nucleotide changes which occur are distributed evenly throughout the mature peptide coding regions except for one region of 35 bases (nucleotide residues 429 to 464) which is invariant between all genes.

On the basis of the number of different amino acid replacements which occur throughout the mature protein length it is expected that, on average, there would be an amino acid replacement in at least one of the genes each 11 residues.

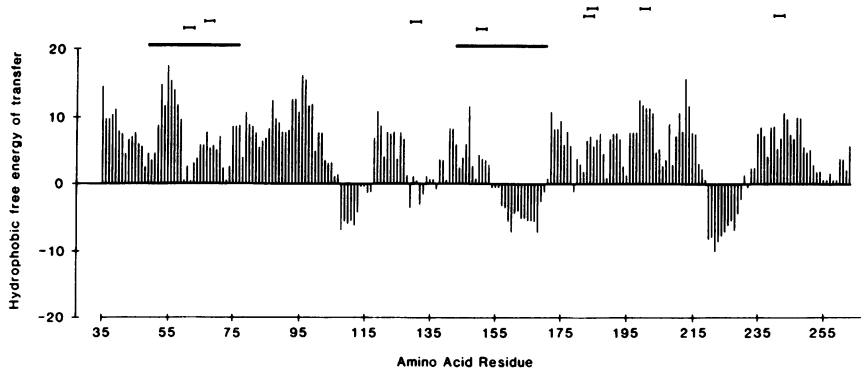

Figure 4 An hydropathy plot of the amino acid sequence of the Cab 91R mature protein. The long regions within the mature protein which are conserved between all Cab genes analysed here are designated by the heavy bar. The repeating amino acid triplets which occur in the protein are indiacted by thin lines.

However there are several segments of the mature protein which do not exhibit this expected distribution of replacements. A twenty-eight amino acid length near the NH_2 terminal (amino acid residues 49 to 77), and a twenty-six amino acid length in the middle of the sequence, (amino acid residues 140 to 166), are invariant between all of the genes. These invariant regions are indicated in Figure 4 which diagrams an hydropathy plot for a mature Cab protein (see below). The conserved region towards the NH_2 terminal of the protein coincides with a portion of the peptide which is external to the membrane on the basis of the hydropathy plot (Figure 4). The other conserved amino acid sequence in the middle of the protein overlaps the region which appears to be membrane-spanning from the hydropathy plot. In a Chou and Fasman (15) analysis of the Cab proteins (data not presented), the NH_2-terminal conserved segment appears to comprise largely turns, and the conserved segment in the middle of the protein is an α helix surrounded by short regions of β sheets [which from the hydropathy plot (Figure 4) appear to be outside the membrane].

There are ten different positions in the mature protein amino acid sequences where changes in the chemical nature of the residue occur. These residues (at positions 86, 92, 99, 109,

138, 141, 186, 212, 217, 258, 259) are not evenly distributed throughout the mature peptide coding sequence. The significance of these changes is not yet clear, they may indicate regions of the Cab protein which are structurally and functionally unimportant, or they may represent critical alterations which are important for distinct functions for the different Cab proteins.

A hydropathy plot of the proposed mature Cab protein amino-acid sequence for Cab 91R is shown in Figure 4. Assuming that those regions of the protein which span the thylakoid membrane should have at least 16 residues and a negative "hydrophobic free energy" of partition (14), then the segment between residues 157 and 174 may be one such membrane spanning region and the stretch between residues 206 and 222 may also be membrane spanning. However it may be unrealistic to predict the membrane traversing properties of these Cab protein sequences from hydropathy diagrams since we do not know how the associated chlorophylls a and b, which are also highly hydrophobic, will affect the net hydrophobicity of these Cab protein segments.

The structure of the light harvesting complex is not yet well defined, however from examination of isolated LHC the number of chlorophylls (a and b) per Cab polypeptide has been estimated as approximately seven in spinach (16) and 12 in pea (10). In an attempt to identify possible sites of interaction of the chlorophylls with the Cab polypeptides I have searched for repeating sequences of amino acids in the mature proteins. There are no long stretches of amino acids which are repeated in direct or inverted orientation in the Cab proteins, however there are several pairs of amino acid triplets distributed throughout the protein. Five pairs occur which are invariant between all of the petunia Cab genes, these are YLG (residues 60 and 147), GEA (residues 66 and 129), GPL (residues 182 and 240), PLG (residues 183 and 199) and DPL (residues 189 and 198): these triplets are indicated on Figure 3. Four of these amino acid sequences pairs are conserved in the Cab gene of pea which has been sequenced (3). It is at present unclear whether these sequences function in chlorophyll-Cab protein interactions however it will be interesting to see whether these triplets occur in other thylakoid proteins which interact with chlorophylls when the

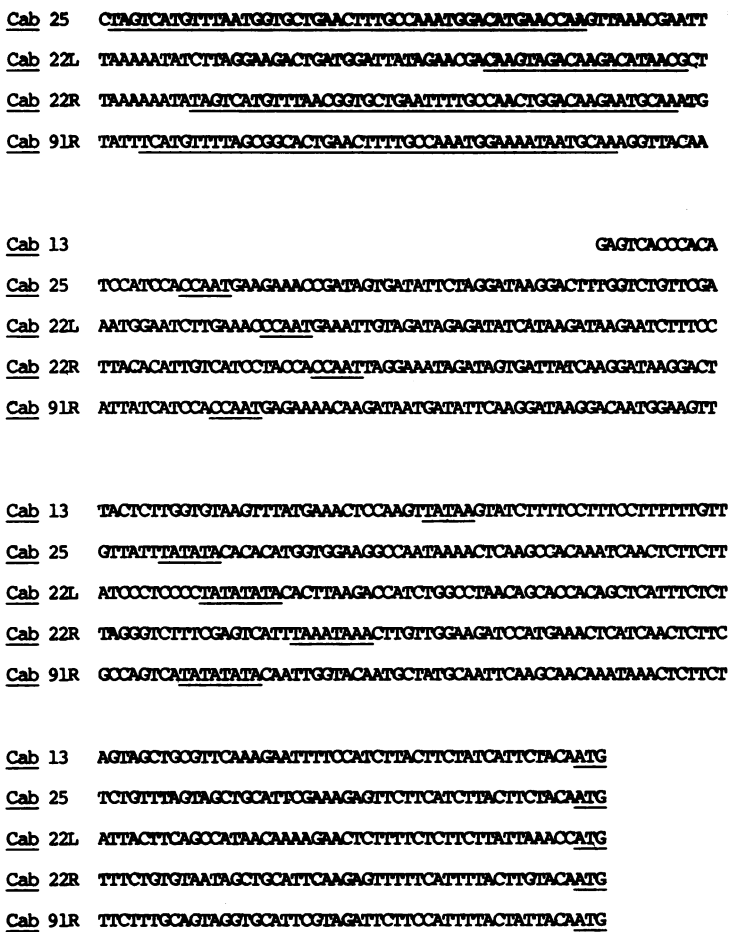

Figure 5 The nucleotide sequences proximal to the Cab protein coding regions. Areas of extensive homology are underlined. The TATA and CCAAT sequences in the genes are indicated and the proposed transcription sites depicted with an asterisk.

amino acid sequences of these proteins become available.

(IV) Comparison of the nucleotide sequences at the 5' ends of the different Cab genes

The nucleotide sequences proximal to the Cab protein coding sequences are summarized in Figure 5. At least two hundred and twenty nucleotides are presented for Cab 25, 22L, 22R, 91R, and for Cab 13, one hundred and twenty nucleotides. On the basis of the cDNA clones which we have isolated, we know that Cab 22R and

Cab 91R are expressed in green leaf tissue (1). We have not yet investigated whether the other genes are expressed, however the sequence information presented here shows that there are no nonsense mutations in any of these genes, thus at the level of the encoded polypeptide there is no reason to suspect that these genes are non-functional.

Generally in eukaryotic genes which are transcribed by RNA polymerase II there are several 'concensus' nucleotide sequences positioned at the 5' end of a gene which function in the regulation of transcription of the gene. There is an AT-rich region, the TATA box (17) centered about 25 to 30 nucleotides upstream from the start site of transcription. Every eukaryotic gene which has been analysed to date has some form of TATA sequence (18). A second region which occurs in many eukaryotic genes at position -70 to -80 from the transcription start, is the sequence CCAAT (18) which is believed to be involved in regulation of the level of transcription. All of the Cab genes which we have analysed have both of these sequence motives - the CCAAT site, and the TATA box separated by 50 to 55 nucleotides at the 5' side of the coding regions. The precise transcription start sites are not yet known for the Cab genes however there are sequences which occur at around -30 nucleotides from the TATA box in each gene which are similar to the concensus transcription start, that is TCAT - these possible transcription starts are at about 50 nucleotides from the translation start for each of the Cab genes.

In the addition to these typical 'promoter' sequences which are associated with the 5' ends of the Cab genes, there is an extensive region of homology of 48 nucleotides, centered at around -130 from the proposed transcription start in Cab 91R, 22R, and 25. Only part of the sequence occurs in the Cab 22L gene (~15 nucleotides), and we do not yet have the data for Cab 13. This extensive region of homology is not perfect, however it appears significant in view of the absence of extensive homology in the preceding 100 nucleotides adjacent to the transcription start of the genes. It is at this position with respect to the protein coding region that sequences which are involved with the

Cab 13 TGAAGTAGTCCCTAAAATAAAAGTACTCTAGTATCAGATTATTTCTTGGCCTTGTAAACTGATGT

Cab 25 TGAAGTGTTCCCTAAAAGAAGAGCCCTCTAGTATCAGATTGTTTCTTGGCCTTGTAAAACTGATG

Cab 22L TGAGCTTAACAAATGATGAATCTTTAAATCTTTCAATTAGTGTGAGATGAGTTTGTAGCTTGTGA

Cab 22R TGAAGTTTTTAGAATTGAGTTTTCACTAATTATCGGGTTGTTTGATGGCCTTGTAAATTTGGCTA

Cab 91R TGAATTTCTTAAAAATTAGTCTCTTCTAGTGTTCGATTGTTCGGTGGCTCTATATAACTAGCTAT

Cab 13 ATATTACCAGAGATTACATGTGGAATTTTTTTGAC

Cab 25 TATATTACCCGAGAATACATGTGAATTTTGTTTAATTGTGTAGTTGTCAATTACACTTTTCCGTG

Cab 22L GTGATGAACCCAAAGAAGGGATCAGAGTTTTTCTTTTAGCATTCTGGGTCATGGGTTCAGTAAAG

Cab 22R TTGCAAATTATGGTAATCATATATGAAACTTTGTTTGGTCTTCAATAATTTTGAATGGCCATAAA

Cab 91R TGTAAATTACAGTGGGTTATATATGAAATATTTGTTTGATCTCCAATAATTGAGTAGTAACATTT

Cab 25 CTCAGAGATACAACTATCAAAATAAA

Cab 22L TGACTGTGGTGTGGTGTAAATTAGTG

Cab 22R ATTTAAAATCCTCTACGTCGGCGCTC

Cab 91R AGCTCCAGATCT

Figure 6 The nucleotide sequences distal to the Cab protein coding regions. The conserved sequences which are present in all Cab cDNA clones are underlined.

regulation of transcription, have been located in other eukaryote systems (17).

(V) Comparison of the regions flanking the Cab genes at the 3' end

We have already classified the multiple Cab genes in petunia into small gene families based upon divergence in the 3' untranslated regions and also upon the nucleotide sequence divergence in the protein coding regions (1); the Cab genes which we are analysing here fall into different gene families on the basis of these criteria. A summary of the Cab gene 3' flanking sequences is presented in Figure 6. The 3' adjacent regions of Cab 13 and Cab 25 are similar and the protein coding regions are also more closely related to each other than any other 'interfamily' pair of Cab genes which we have compared (Fig. 2); however the 5' adjacent regions are unrelated (Fig. 5) thus we still consider that these two genes belong to separate families.

We know that genes belonging to the same gene family have long conserved flanking regions at both the 5' and 3' ends (Dunsmuir et al., in preparation). Aside from the Cab 13, Cab 25 3' flanking region similarly the other genes have 3' flanking regions which are divergent. There is one long conserved sequence of 20 nucleotides (indicated in Fig. 5) which occurs in all genes except Cab 22L. This sequence is also present in all five different cDNA clones which we have analysed (1) and appears to be correlated with gene expression. It is interesting that Cab 22L is also without the long conserved region at the 5' end of the genes which occurs in Cab 25, 22R and 91R. Perhaps it is not expressed, or regulated differently from the other Cab genes. The short sequence TTTGTTT (Fig. 5) which we have found to occur in all cDNA clones (1) is absent in both Cab 13, and 22L. We are investigating whether either of these genes are expressed.

CONCLUSION

I have summarized nucleotide sequence data describing five different chlorophyll a/b binding protein genes in Petunia (Mitchell). These genes have been ascribed previously to distinct Cab gene families based upon their sequence relatedness (1). There is no evidence for intervening sequences in any of these genes, however heterogeneity in both length and amino acid sequence is predicted by the nucleotide sequencing information reported here. The genes Cab 13 and Cab 25 encode precursor proteins of 266 amino acids while Cab 22L, Cab 22R and Cab 91R specify proteins of 267 amino acids. This length difference results from a deletion/insertion in the region of the protein which is proposed to correspond to the transit peptide at the NH_2-terminal. A comparison of the predicted petunia Cab precursor sequences with the amino acid sequence information for the mature Cab protein in pea suggests that the mature peptides in petunia initiate at methionine (position 34/35) or arginine (position 35/36). Hence the Cab protein transit peptides would be between 33 and 34 amino acids long. These proposed transit peptide regions are more divergent between the Cab genes than the regions which correspond to the mature Cab peptides (nucleotide divergences - 15% and 10% respectively, amino acid divergence -

8% and 4%).

At present it is unknown whether the amino acid replacements which occur between the different Cab proteins have any functional significance. Experiments are in progress to determine if different proteins are expressed in response to varying environmental stimuli. These studies should also provide some insight into the significance of the differences in the 5' and 3' flanking regions to the regulation of expression of distinct Cab genes.

ACKNOWLEDGEMENTS

I would like to thank Bryan Clarke for excellent technical assistance and John Bedbrook for helpful discussions.

[1] Current address: Advanced Genetic Sciences, Inc., 6701 San Pablo Avenue, Oakland, California 94608

REFERENCES

1. Dunsmuir, P., Smith, S.M., Bedbrook, J.R. J. Mol. Appl. Genet. 1983;2:285-300.
2. Boardman, N.K., Anderson, J.M. In: Akoyunoglou, G., Argyroundi-Akoyunoglou, J.M., eds, Chloroplast Development. Amsterdam:Elsevier 1978.
3. Coruzzi, G., Broglie, R., Cashmore, A.R., Chua, N.H. J. Biol. Chem. 1983;1399-1402.
4. Broglie, R., Coruzzi, G., Lamppa, G., Keith, B., Chua, N.H. Biotechnology 1983;1:55-61.
5. Stiekema, W.J., Wimpee, C.F., Tobin, E.M. Nucl. Acid Res. 1983;11:8051-8061.
6. Schmidt, G.W., Bartlett, S.G., Grossman, A.R., Cashmore, A.R., Chua, N.H. J. Cell Biol. 1981;91:468-478.
7. Maxam, A.M, Gilbert, W. In: Grossman, L., Moldave, K., eds, Methods Enzymol 1980;65:499-560.
8. Sanger, F., Nicklen, S., Coulson, A.R. Proc. Natl. Acad. Sci. USA 1977;74:5463-5467.
9. Vieira, J., Messing, J. Gene 1982;19:269-276.
10. Mullet, J.E. J. Biol. Chem. 1983;258:9941-9948.
11. Chua, N.H., Schmidt, G.W. Proc. Natl. Acad. Sci. 1978;75:6110-6114.
12. Highfield, P.E., Ellis, R.J. Nature 1978;271:420-424.
13. Martin, P.G. Aust. J. Plant Physiol. 1979;6:401-8.
14. Von Heijne, G. Eur. J. Biochem. 1981;116:419-422.
15. Chou, P.Y., Fasman, G.D. Ann. Rev. Biochem. 1978;47:251-276.
16. Ryrie, I., Anderson, J.M., Goodchild, D.J. Eur. J. Biochem. 1980;107:345-354.
17. Goldberg, M. 1979 Ph.D. Thesis. Stanford University.
18. Nevins, J.R. Ann. Rev. Biochem. 1983;52:441-466.
19. Berry-Lowe, S.L., McKnight, T.D., Shah, D.M., Meagher, R.B. J. Mol. App. Genet. 1982;483-498.