Occurrence and expression of members of the ferritin gene family in cowpeas

Alison J. WARDROP, Ronald E. WICKS and Barrie ENTSCH[1]

Division of Molecular and Cellular Biology, School of Biological Sciences, University of New England, Armidale, NSW 2351, Australia

Ferritin gene expression has been demonstrated in a variety of plants including maize, *Arabidopsis*, cowpeas, soybeans, beans and peas. Most available evidence shows that the mature protein is located in plastids and its production is under gene transcriptional control. In maize, two different ferritin genes have been identified; they were found to express protein under different physiological conditions. Only single gene products have been found until now in the other plants, with the exception of cowpeas (*Vigna unguiculata*). Our previous work with cowpeas [Wicks and Entsch (1993) Biochem. Biophys. Res. Commun. **192**, 813–819] showed the existence of a family of at least three ferritin genes, each coding for a protein subunit with a unique amino acid sequence. Here we report the discovery of a fourth active gene in cowpeas and present the full cDNA sequences for two of the four known members of the cowpea gene family. We also provide preliminary evidence for a family of ferritin genes in soybeans (*Glycine max*) related to that in cowpeas. We conclude that a family of genes is probably present in all higher plants. We have used quantitative reverse transcriptase-mediated PCR to show that each of the four members of the cowpea ferritin gene family expresses mRNA in leaves and roots under normal growth with a complete nutrient supply. The results clearly show a marked differential pattern of mRNA levels formed during development from the four genes. We conclude that the composition of plant ferritin molecules from plant leaf extracts is probably a complex mixture of subunits, which might be different in roots and in leaves.

Key words: gene family, gene expression, plant ferritins, reverse transcriptase-PCR.

INTRODUCTION

Ferritins are multimeric proteins (24 subunits) that are present in plants, animals and micro-organisms. In all cases the protein is responsible for the storage and distribution of iron within cells [1]. Mammalian ferritins have been extensively characterized. Two active genes code for the functionally different subunit types, H and L, which assemble in various ratios in the final protein depending on the tissue in which the protein is found [2]. Until recently, only one functional type of plant ferritin subunit has been described; this subunit displays characteristics of both H-type and L-type animal ferritin subunits [2]. On the basis of sequence comparisons, plant and animal ferritins seem to be derived from a common evolutionary ancestor [3]. Although ferritin probably performs the same function in plants and animals, most experimental evidence supports the regulation of ferritin synthesis in plants at the level of transcription, as opposed to translation, in animals, and plant ferritins are located in plastids as opposed to the cytoplasm in animals [4–6]. In animals, iron regulates ferritin expression through a non-coding region of mRNA (the iron response element) that has not been detected in all reported plant ferritin sequences [7].

Coding regions for plant ferritin subunits are longer than those for animal ferritins; an extra sequence at the N-terminus of the plant protein subunit encodes a transit peptide responsible for plastid targeting, and also an extension peptide of approx. 24 amino acid residues that is specific to plant ferritins and found in the mature peptide [3,8]. In maize, two different ferritin genes have been identified that seem to be differentially expressed in response to abscisic acid or iron-mediated oxidative stress [9,10]. These genes are almost identical (95–96 %) in the protein-coding region. In dicotyledons, single gene products have been reported for soybeans [6], beans [11], peas [8,12] and *Arabidopsis thaliana* [13]. In *A. thaliana* [13] the iron-mediated oxidative stress transcriptional response described for maize [10] has been observed, but there is no evidence for the existence of an abscisic acid-regulated ferritin gene. In previous studies we have established that at least three cowpea (*Vigna unguiculata*) ferritin genes are expressed in developing leaves. Two of these genes code for protein products with substantial sequence divergence from the third product, which is closely related to the other legume subunits [14]. The possibility exists that members of this gene family in cowpeas might be differentially expressed in a manner analogous to that for maize or animal ferritins, and that all plants have a complex family of ferritin genes.

Here we present evidence for four distinct transcribed members of the cowpea ferritin gene family, and also provide preliminary evidence from PCR-amplifcation of soybean DNA that this gene family is not unique to cowpeas. For each of four known genes in cowpeas we have used quantitative reverse transcriptase-mediated PCR (RT–PCR) techniques to estimate the levels of mRNA during the growth and expansion of cowpea leaves as well as in roots. Results show clear patterns of differential accumulation of mRNA species.

MATERIALS AND METHODS

Materials

Agarose for gel electrophoresis was obtained from Pharmacia Biotech (Uppsala, Sweden). Restriction enzymes, reverse transcriptase (MMLV RT, RNase H Minus) and DNase (RQ1 RNase-free DNase) were all obtained from Promega Corporation (Madison, WI, U.S.A.). Radioactive [α-^{32}P]dATP was supplied

Abbreviations used: MMLV, Moloney murine leukaemia virus; RT–PCR, reverse transcriptase-mediated PCR.

[1] To whom correspondence should be addressed (e-mail bentsch@metz.une.edu.au).

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession numbers AF052057, AF052058, AF052512, AF052513 and AF052511.

© 1999 Biochemical Society

by Bresatec (Adelaide, SA, Australia). All other reagents were of A.R. grade. *Escherichia coli* strains JM105 and DH5α [15] were used for routine cloning; *E. coli* LE392 [15] was used for the construction of cDNA libraries in phage λgt10 from Pharmacia (Uppsala, Sweden). Cowpea (*V. unguiculata* var N41) and soybean (*G. max* var Valder) seeds were obtained from the NSW Department of Agriculture (Tamworth, NSW, Australia).

Plant growth

Cowpea seedlings were grown in moist vermiculite in a greenhouse under conditions of natural light and at ambient temperature. Growth was supported by a complete mineral mixture supplemented with extra soluble iron as FeEDTA. Leaf samples were taken at specified lamina lengths (see Table 3) and placed on ice before being transferred to liquid nitrogen for storage. Soybean seedlings were grown under the same conditions, with leaf samples taken at 5–7 cm lamina length for the preparation of DNA or RNA.

Cloning and sequencing of cowpea ferritin cDNA

Total RNA was prepared from expanding cowpea leaves (5–7 cm in lamina length) by the method of Chomczynski and Sacchi [16]. Poly(A)$^+$ RNA was isolated from total RNA by using a poly(dT) magnetic bead capture system (Promega PolyATract® mRNA Isolation System; Promega Corporation). Clones for the two reported cowpea ferritin cDNA species (CP2 and CP3) were isolated from two independent lambda cDNA libraries. One library was constructed from 4 μg of poly(A)$^+$ RNA with the cDNA Synthesis System Plus kit (Amersham International, UK) – first-strand synthesis primed with oligo-dT, and the cDNA was packaged into lambda gt10 phage particles using the Promega Packagene® Lambda DNA Packaging System. For the second library, cDNA was synthesized from 3 μg of poly(A)$^+$ RNA (oligo-dT priming) using Moloney murine leukaemia virus (MMLV) reverse transcriptase, and *Eco*RI/*Not*I adaptors were added (Pharmacia Biotech cDNA Synthesis Kit; Pharmacia, Sweden). The cDNA was cut with *Eco*RI, inserted into lambda gt10 and packaged using the Promega Packagene® Lambda DNA Packaging System. For the first library, recombinant phages were plated and screened with a partial cowpea ferritin cDNA fragment (215 bp) specific for CP2 that had been labelled with [α-^{32}P]dATP. For the second library, recombinant phages were plated and screened with a partial cowpea ferritin cDNA fragment (215 bp) specific for CP5 that had been labelled with [α-^{32}P]dATP. Hybridizations were performed in 6 $\times$ SSC [1 $\times$ SSC being 0.15 M NaCl/16.5 mM sodium citrate (pH 7.0)]/0.5 % SDS/0.5 % polyvinylpyrrolidone/0.5 % Ficoll containing 20 μg of herring sperm DNA/ml at 65 °C. Filters were washed twice in 2 $\times$ SSC/0.1 % SDS at 60 °C. Inserts from positive recombinant phages were subcloned into M13; their sequences were determined by single-stranded sequencing in both directions with *Taq* polymerase and a series of primers to walk along the complete clones (Promega *Taq*Track® Sequencing System).

Quantitative RT–PCR

RT–PCR [17,18] was used for the quantification of mRNA in cowpea roots and leaves. The RT–PCR was performed as a two-step reaction, i.e. synthesis of cDNA followed by PCR.

RNA preparation

Total RNA was extracted from 1 g of frozen plant tissue by a guanidinium thiocyanate procedure [16]. Plant tissue samples were taken from leaves at lamina lengths specified in Table 3, or from root tips (2 cm). All total RNA preparations were treated with RNase-free DNase. RNA quality was checked by gel electrophoresis and then quantified by spectrophotometry.

cDNA synthesis

First-strand cDNA (forming a DNA/RNA duplex) was synthesized from 1 μg of total RNA by the method of Dallman and Porter [18] with MMLV reverse transcriptase and oligo(dT) in the presence of a ribonuclease inhibitor. Reaction conditions were set to give a maximum yield of cDNA. After annealing at 37 °C, 200 units of MMLV reverse transcriptase was added and the reaction was left for 60 min at 37 °C. A second 200 units of enzyme was added and the reaction was left for a further 60 min at 37 °C, followed by inactivation of the enzyme at 70 °C for 10 min. The final solution was made to 100 μl; 4 μl samples were used in each PCR reaction.

Competitive templates

Templates were fragments of genomic cowpea DNA that had been subcloned in pUC18. The genomic fragments for CP1 and CP2 contained small introns and were used unchanged (see Table 2). The genomic fragments for CP3 and CP5 contained large introns and were engineered to remove most of the intron (by using the Clontech Mutagenesis Kit and specific oligonucleotides to loop out a large segment of DNA) so that a suitable fragment size was obtained (see Table 2). The templates were used as linearized pUC18, and stock solutions from 8 to 8 $\times$ 10^7 ag/μl (expressed as concentration of unique template DNA) were prepared by 10-fold serial dilutions.

PCR

Components of the PCR reactions were tested for contaminants by performing 40-cycle reactions without added DNA template. Each reaction mixture contained, dNTPs (200 μM final concentration of each), 0.5 μM each primer, 1.5 mM $MgCl_2$, PCR buffer [67 mM Tris/HCl (pH 8.8)/16.6 mM $(NH_4)_2SO_4$/0.2 mg/ml gelatin/0.45 % (v/v) Triton X-100], 1 unit of *Taq* polymerase and water to a final volume of 20 μl. Samples were amplified, with the PTC-100 thermal cycler from MJ Research, by 30 cycles of the following protocol: 94 °C for 60 s, 57 °C for 60 s and 72 °C for 90 s. An aliquot of each reaction mixture was subjected to electrophoresis on 2 % (w/v) agarose (see Figure 3).

Isolation of soybean ferritin genomic DNA fragments

Total soybean DNA was prepared by using standard techniques [19]. Specific oligonucleotides for the amplification of CP1 or CP2 (see Table 2) were used in a touchdown PCR programme with total soybean DNA as template. The reaction mixture was the same as in the PCR section above. Samples were amplified for 30 cycles, starting with five cycles at 94 °C for 60 s, 65 °C for 60 s and 72 °C for 90 s. The annealing temperature was decreased to 60 °C for five cycles and then to 55 °C for the last 20 cycles. All PCR products were cloned into pUC18 vector and sequenced for identification.

RESULTS

Isolation and sequence of cowpea cDNA

We reported previously [14] the presence of at least three different genes for ferritin subunits in cowpeas, on the basis of PCR amplification of an internal protein-coding segment of mRNA

© 1999 Biochemical Society

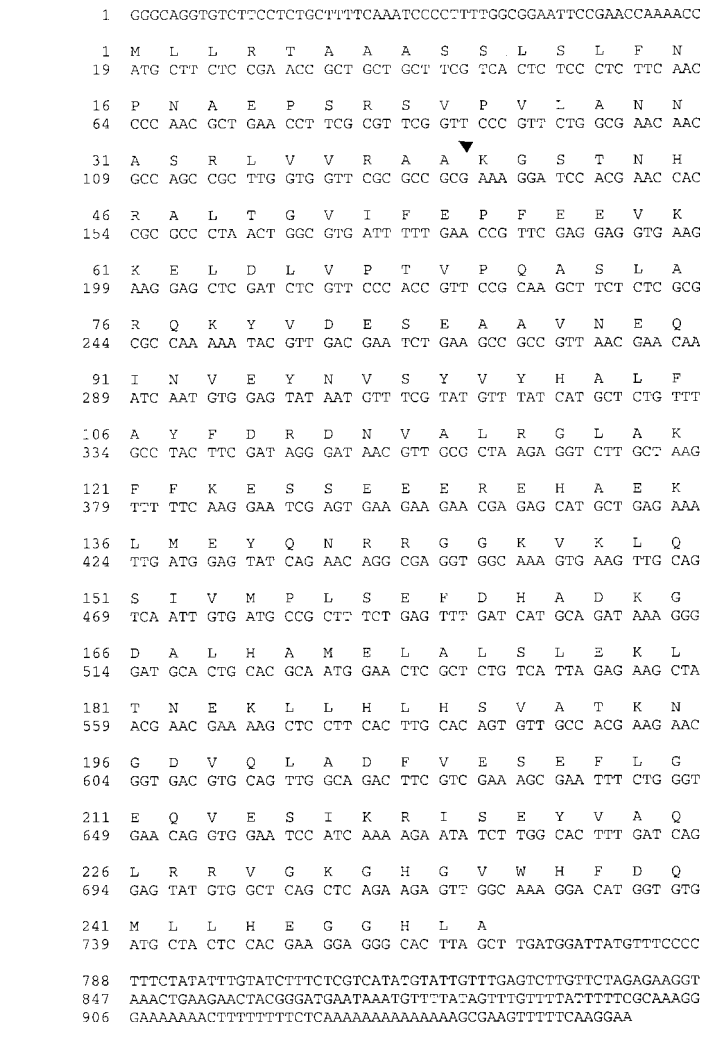

Figure 1 Nucleotide and deduced amino acid sequences of the CP2 cDNA from cowpeas

The amino acid translation is shown for the open reading frame. The putative transit peptide cleavage site is indicated by an arrowhead.

isolated from developing leaves. The unique clones were named, in order of isolation, cowpeas 1, 2 and 5. We have used these segments of DNA as probes to isolate two complete coding regions for expressed cowpea ferritin subunits. Libraries of cDNA were constructed from mRNA isolated from expanding cowpea leaves (5–7 cm in lamina length). The cDNA synthesized was cloned into λgt10.

To isolate CP2 (the full cDNA corresponding to cowpea 2), approx. 60000 clones were screened with a unique segment of DNA of 215 bp (prepared from the cowpea 2 isolate) cloned in M13. Only one positive clone was detected; this was sequenced completely. It contained one open reading frame (Figure 1). Translation of the DNA showed that it coded for a ferritin subunit plus a transit peptide and the cDNA was identical with the segment of sequence encoded by the probe. The transit peptide had a unique sequence (no similarity to the other published plant ferritin sequences) but contained the general characteristics of a signal peptide for transit into chloroplasts or mitochondria [20]. The sequence of the mature polypeptide was

Table 1 Comparison of cowpea and soybean ferritin subunits for mature proteins

Amino acid sequences for mature soybean ferritin [6], cowpea 2 and cowpea 3 are compared. Only 72 residues for cowpeas 1 and 5 have been identified; these were compared with the corresponding region in the soybean sequence.

Source	Amino acid identity (%)	Identity plus conservative substitutions (%)
Soybean	100	100
Cowpea 3	84	89
Cowpea 2	77	81
Cowpea 1	72	79
Cowpea 5	97	97

only 77% identical (Table 1) with the first published sequence for soybean used as the model example of a legume ferritin [6].

A second library from developing leaves was screened with a unique segment of DNA of 215 bp (prepared from cowpea 5 isolate) cloned in M13. From approx. 50000 clones, three positives were identified at a stringency that was known to show some annealing to related ferritin sequences. Sequence information for the three clones provided a surprise. One contained no ferritin code and the other two were identical ferritin codes. However, the new clones did not contain the probe sequence but a related ferritin homologue that was different from all three species found previously [14]. We named this new ferritin CP3 (see Figure 2). CP3 had 84% identity in the mature protein to the soybean sequence used as a point of comparison (see Table 1). In these studies we did not find a cowpea 5 clone. This was not pursued because the cowpea 5 fragment is the product of the cowpea gene very closely related (97% similarity; Table 1) to the previously published protein sequences from legumes. The translated sequence of CP3, like that of CP2, contained a transit peptide sequence at the N-terminus (Figure 2). This segment had some similarity to the transit peptide in the published soybean sequence, which has been shown to target the protein to the chloroplast [3]. Thus the chloroplast is the probable location for CP3.

Both lambda libraries described above were probed with the unique cowpea 1 fragment without success. Subsequent results (see Table 3) on the expression of this polypeptide in leaves clearly showed that cowpea 1 is not part of the ferritin in normal developing leaves.

Expression of ferritin genes in cowpea

Our work with cowpeas has shown that at least four distinct polypeptides could be produced to assemble the ferritin molecule of 24 polypeptides. It was therefore important to discover the physiological function of these members of a gene family. Cloning cDNA established only that three of the polypeptides were probably located in plastids (transit peptides). We decided that measurement of the levels of mRNA was the best chance of obtaining information about when and where these genes were active. Preliminary experiments quickly established that RNA blots were not sensitive enough for normal ferritin expression and also were not sufficiently specific to discriminate between genes. It had been established before [14] that oligonucleotide primers could be designed to amplify specifically only one of the genes or its mRNA product. Thus we turned to RT–PCR to provide a sensitive and specific method of discrimination.

© 1999 Biochemical Society

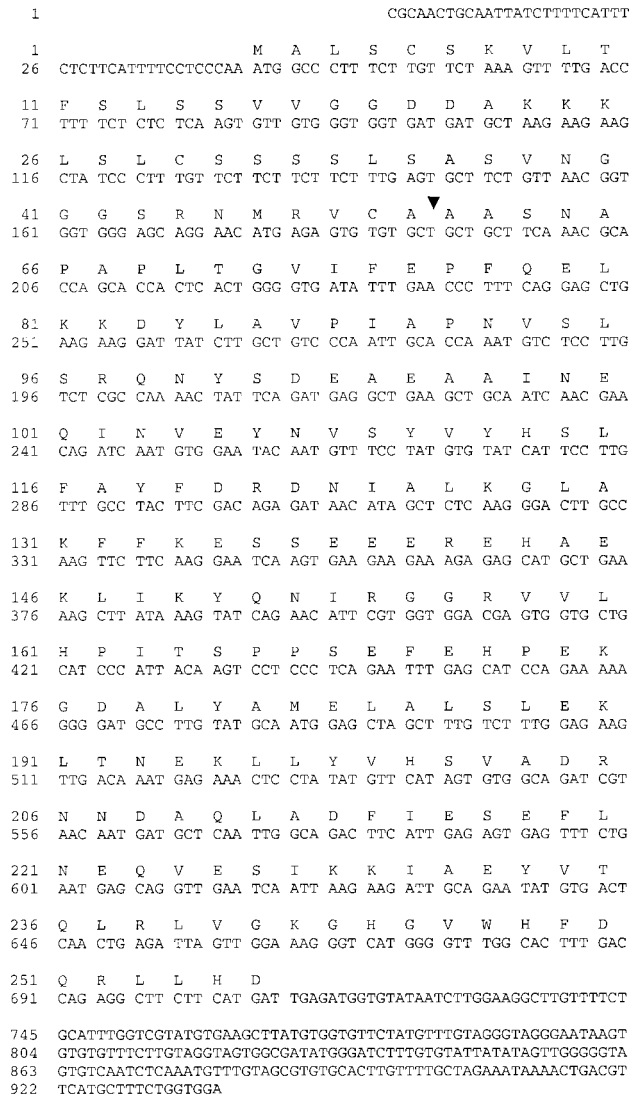

Figure 2 Nucleotide and deduced amino acid sequence of the CP3 cDNA from cowpeas

The amino acid translation is shown for the open reading frame. The putative transit peptide cleavage site is indicated by an arrowhead.

There are many problems in using PCR for quantitative or semi-quantitative measurement. The difficulties were very clearly expressed in quantitative terms by Raeymaekers [21]. A fundamental requirement is to have internal standards that will amplify to the same efficiency as the test material over the number of cycles used in the amplification. To achieve this, we developed a unique standard for each of the four ferritin gene transcripts. Each standard was identical in sequence with the target sequence because each standard was a fragment of the corresponding gene. However, the gene fragments had a small intron in the middle that made the standards slightly larger than the targets (Table 2). The size of the fragments was also manipulated so that there could not be any mistake in identification of the DNA amplified on an agarose gel. Each target sequence and corresponding standard was amplified by a unique pair of oligonucleotides (see Table 2). Under the conditions used for amplification (see the Materials and methods section), only the target sequences were formed without any trace of other target sequences. To test amplification efficiencies, pure samples of DNA containing standard and test sequences in mixtures of different concentrations were subjected to PCR. Only equimolar amounts of standard and test gave equal amounts of amplified product bands. If efficiencies were not the same, then equal amounts of product bands would be formed when the amounts of sample and test were not equimolar [21]. These PCR trials were run under the same conditions as used with plant extracts.

Quantification was based on the formation of a total first-strand cDNA synthesis from total RNA from a tissue. Trials were run to test the total yield of cDNA with time. The best indication was the subsequent amount of cDNA detected by PCR. A uniform set of conditions for all samples was then set for reverse transcriptase reactions.

We confined the analysis of RNA preparations to differences in concentrations of an order of magnitude. It was expected (and shown below, in Table 3), that concentrations of specific mRNA species would differ by very large amounts, depending on tissue and stage of development. We wished to achieve a guide to the patterns of expression. It did not seem sensible to try to refine measured concentrations of mRNA to a greater precision: it would have been an enormous amount of work and would not add to the information, partly because errors occur in RT–PCR that are impossible to control. These errors are: variable recovery of RNA from tissue samples, variation in yields of cDNA in the reverse transcriptase reaction, and the possible bias in amplification owing to the comparison of PCR of DNA (the internal standard) with a DNA/RNA hybrid (the product of reverse transcriptase). We have confidence in the precision of the results reported in Table 3 because each value quoted was determined three times on separate tissue samples. In all cases, the same result was obtained from separate tissue samples, provided that results were based on total RNA extracted from tissues. However, the absolute concentrations of mRNA quoted in Table 3 could be in error owing to the RT–PCR technique.

To illustrate the results obtained in test RT–PCR experiments, two quantification series are shown in Figure 3. It can be seen that the unknowns were equated directly with a standard or could be positioned between two standards. This type of information produced the final results in Table 3. Several important results were obtained from seedlings provided with all growth requirements, including plentiful iron. First of all, mRNA for ferritin subunits 2, 3 and 5 (numbers correspond to the original numbering of isolated clones, and are used in Table 3) was produced during leaf development in similar amounts from each gene. Subunits 2 and 5 were produced earlier than subunit 3 in leaf expansion. However, subunit 3 remained in developed leaves, even 1 week after complete expansion (see biofoliate leaf values) when the amounts of subunits 2 and 5 were very low. The second notable feature was that the mRNA for subunit 1 was never present in any sample in amounts comparable with the maximum concentrations of the other mRNA species. The third point to emphasize is the similarity of the situation in developing leaves to that in developing root tips. However, the mRNA for subunit 2 was produced in roots at higher levels than that for other ferritin species.

Homologues of cowpea ferritin genes in soybeans

Since the publication of the first cDNA sequences for plant ferritin in 1990 and 1991 [3,6], only one gene product has been identified in soybeans (*G. max*), which have been used as a model for ferritin synthesis and function in plants [3,22–24]. The

© 1999 Biochemical Society

Table 2 Specific oligonucleotides and standards used for RT–PCR

The RT–PCR method used required the inclusion of competitive standards that amplified with the same efficiency as the cDNA being measured [18]. For cowpeas 1 and 2, cloned fragments of genomic DNA (in pUC18) that span the first intron were used as standards because they were amplified with the same efficiency as the corresponding cDNA test material. For both cowpeas 3 and 5, the first intron was approx. 900 bp in length and genomic fragments spanning intron 1 were much larger than, and showed large differences in amplification efficiency compared with, their respective cDNA fragments. Mutagenesis was used to remove most of these large introns to create effective internal standards for cowpeas 3 and 5 (see the Materials and methods section). The forward (f) and reverse (r) primers used to amplify unique products for each target cDNA and standard sequence are shown, together with the expected product sizes.

Target ferritin	Oligonucleotides	Internal standard	Size of amplified fragment (bp)	
			Test	Standard
CP1	f, 5′-GAAGTGAAAAAAGAGCTTCTT-3′ r, 5′-GCCAATCCTTTCAGAGCAAC-3′	pUC18/genomic DNA fragment (complete intron)	185	285
CP2	f, 5′-GAGGTGAAGAAGGAGCTCGAT-3′ r, 5′-GCAAGACCTCTTAGCGCAAC-3′	pUC18/genomic DNA fragment (complete intron)	185	265
CP3	f, 5′-AAATGTCTCCTTGTCTCGCC-3′ r, 5′-CTTTTCAGCATGCTCTCT-3′	pUC18/genomic DNA fragment (partial intron)	195	290
CP5	f, 5′-GAGGTTAAGAAGGAAGAGCTT-3′ r, 5′-GCAAATCCCTTGAGAGCTAC-3′	pUC18/genomic DNA fragment (partial intron)	185	246

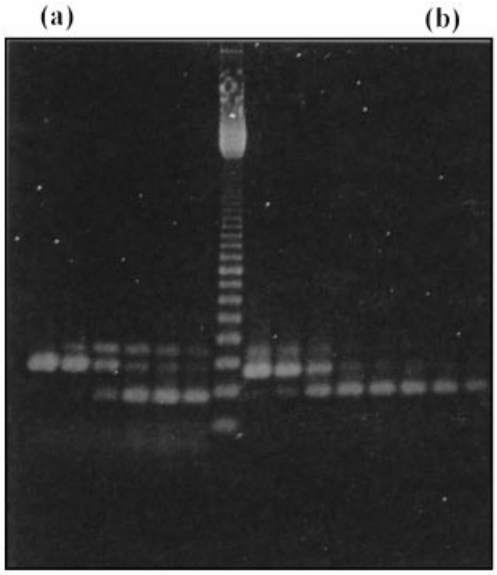

Figure 3 Examples of quantification of cDNA by using competitive PCR

Total first-strand cDNA prepared from total RNA from each plant sample was co-amplified by PCR with a separate standard dilution series of competitive templates for each target ferritin (see the Materials and methods section, the Results section and Table 3). The competitive templates were identical with the target cDNA sequences except for the inclusion in the middle of the sequence of a segment of the natural intron from each specific gene. Thus each standard was slightly larger than its corresponding target sequence. All amplified bands were between 185 and 290 bp in length. Products from PCR were separated on 2% (w/v) agarose gels. (**a**) To the left of the molecular mass ladder are amplification products of PCR with the use of oligonucleotides specific for CP1. Bands for amplified test cDNA (195 bp), standard (285 bp) and a third, apparently larger, product can be identified. This third product is a heteroduplex that results from annealing between heterologous strands of test and standard amplified fragments [17]. This was established separately by annealing pure samples of test and standard bands. The equivalence point falls between two dilutions of competitive standard in the series. In all cases in which this occurred, both standard quantities were used in calculations of mRNA levels and results were given as a range (see CP1, 5–6 cm, Table 3). (**b**) To the right of the molecular mass ladder are amplification products of PCR with the use of oligonucleotides specific for CP2. Bands can be identified as amplified test cDNA (185 bp), standard (265 bp) and heteroduplex. Amounts of amplified target and standard were clearly equivalent at a single dilution of standard; in all cases in which this occurred, a single figure was used when expressing final results (see CP2, 7–8 cm, Table 3).

Table 3 Estimation of cowpea ferritin mRNA levels by using RT–PCR

Total first-strand cDNA was prepared from total RNA from each plant sample (see the Materials and methods section). For each cowpea ferritin target, a master mix was made up containing designated amounts of standard PCR components (see the Materials and methods section) and 1 μl of first-strand cDNA in each 20 μl of final reaction volume. The mix was dispensed as 18 μl aliquots into tubes containing 1 μl of each dilution of competitive standard, followed by 1 unit of *Taq* polymerase. An aliquot of each reaction in a series was run on a 2% (w/v) agarose gel (see the example in Figure 3). Equivalence occurred when amplified products from standard and target were judged to be equal on the gel. Then the amount of input standard DNA equalled the amount of target present in the added cDNA. Results are expressed as pg of target messenger fragment (equivalent to pg of input standard) per μg of total RNA used to prepare cDNA. The analysis was repeated for each tisue with separate plant samples (mostly in triplicate).

Tissue	Size (cm)	Target mRNA ...	mRNA (equivalent to pg of standard/μg of total RNA)			
			CP1	CP2	CP3	CP5
Leaf						
Trifoliate	2–3		0.2	2	0.2–2	2–20
	5–6		0.2–2	2–20	2	2–20
	7–8		2	20	2–20	20
	10–12*		2	0.2–2	2–20	0.2–2
Bifoliate	8–10†		2	0.2	20	0.2
Root tips	2		0.2	20–200	2–20	20

* Fully expanded leaf, harvested immediately on reaching this size.
† Fully expanded leaf, harvested approx. 1 week after reaching this size.

published cDNA for soybean ferritin is highly homologous to the ferritin subunit 5 fragment that we have described in cowpeas [14]. Given the gene family for ferritin in cowpeas, it seemed likely that a family would also exist in other legumes, including soybeans. We used the unique oligonucleotides that we had developed to amplify specifically cowpea fragments 1 and 2 (which have sequences substantially different from the published soybean; see Table 1) and tried to amplify similar sequences in preparations of soybean DNA. The PCR conditions used for these experiments are described in the Materials and methods section. When the cowpea 1 oligonucleotides were used, three

© 1999 Biochemical Society

```
          10         20         30         40         50
soybean   FEPFEEVKKS ELAVPTAPQV SLARQNYADE CESAINEQIN VEYNASYVYH
soybean1             VIPSVPKA SLARQMYTDQ REAALNAQIN VEYNVSYVYH
soybean2             LVPTVPQA SLARQKYVDE SESAVNEQIN VEYNVSYVYH

          60         70
soybean   SLFAYFDRDN VALKGFAKFF
soybean1  AMYAYFDRDN
soybean2  AMFAYFDRDN
```

Figure 4 Comparison of amino acid sequences of soybean ferritin fragments 1 and 2 with published soybean ferritin sequence

The amino acid sequences of soybean ferritin fragments 1 and 2 are compared with the relevant portion of the published soybean ferritin amino acid sequence [3]. Numbering is from the first residue of the mature published subunit [6].

pieces of DNA were amplified; the largest and most prominent of these (1070 bp), when sequenced, showed that it was part of the code for a new soybean ferritin gene. We called this fragment soybean 1 because it was closest in DNA and protein sequence to our existing cowpea 1 (compare sequences in Figures 4 and 5). When the cowpea 2 oligonucleotides were used, two pieces of DNA were amplified. The more prominent product was 280 bp long; when it had been sequenced it was clearly part of another new soybean ferritin gene. This fragment was called soybean 2 because it was very closely homologous to cowpea 2 in DNA and protein sequence (compare sequences in Figures 4 and 5). The soybean 1 fragment contained one large, unique intron in the same position as the equivalent position in cowpea [14] and maize genes [9]. The soybean 2 fragment contained one small, unique intron in the same position as the others studied. Thus it is clear that the gene family in cowpeas has its counterparts in soybeans. It should be noted that we have sequenced more than one clone of each of the soybean 1 and 2 fragments. This exercise showed that there must even be multiple forms of the new genes described, because the intron sequences were not identical in different clones. In addition, we have made specific oligonucleotides for the soybean 1 gene fragment and have used these to amplify segments of soybean 1 protein code from a cDNA preparation made from RNA extracted from developing soybean leaves. A single band of exactly the predicted size for the protein code of soybean 1 was formed. Thus at least one of the new soybean genes discovered is transcribed.

DISCUSSION

Since 1989 a number of primary sequences for plant ferritin have been determined by the isolation of cDNA clones made from mRNA preparations [3,11–14,25]. Most attempts at cDNA library probing have resulted in a single isolate from each species. However, there have been exceptions to this outcome. We published a preliminary report based on the use of PCR technology that demonstrated at least three ferritin mRNA species in cowpeas [14]. The segments of protein sequence obtained showed substantial variation in the highly conserved ferritin protein subunit. Briat and co-workers have published several papers in recent years that examine their finding of highly homologous ferritin sequences in a single species: maize [9–11,13,25]. Their isolates have varied mostly in the control regions of the genes; they have found that the different subunits are expressed under different physiological conditions [9,10]. The protein sequences reported for the ferritin subunit from this monocotyledon are almost identical. In the present paper we have presented new results that expand on our original observation in 1993. It is clear there is much more still to be learned about the gene family for plant ferritin.

```
                                                                  -42
    soybean                                                  MALAPSKVS
    cowpea5
    cowpea3                                                  MALSCSKVL
    cowpea2                                                     ML
    cowpea1
arabidopsis                                                  MASNALSSF
maize (FM2)  MAPRPPRAYI HPSSYSYPAT YSTTYYSAPD AQNSRERTRP PPPSVPIQMM

                -32        -22        -12         -2          9
    soybean  TFSGFSPKPS VGGAQKNPTC .SVSLSFLNE KLGSRNLRVC A    STVPL
    cowpea5                                                  |
    cowpea3  TFSLSSVVGG DDAKKKLSLC SSSSLSASVN GGGSRNMRVC AAASNAPAPL
    cowpea2  LRTAAASSLS LFNPNAEPSR SVPV....LA NNASRLVVRA AKGSTNHRAL
    cowpea1                                                  |
arabidopsis  TAANPALSPK PLLPHGSASP SVSLGFSRKV GGGRAVVVAA ATVDTNNMPM
maize (FM2)  LRVSSSPAAA VANHLSGGAA ATTAPARVTA QRSGVSLSAA AAAGKGKEVL

                P helix                                A helix
             19         29         39         49         59
    soybean  TGVIFEPFEE VKKSELAVPT APQVSLARQN YADECESAIN EQINVEYNAS
    cowpea5      FEPFEE VKKEELAVPT APQVSLARQY YADECESAIN EQINVEYNAS
    cowpea3  TGVIFEPFQE LKKDYLAVPI APNVSLSRQN YSDEAEAAIN EQINVEYNVS
    cowpea2  TGVIFEPFEE VKKELDLVPT VPQASLARQK YVDESEAAVN EQINVEYNVS
    cowpea1      FEPFEE VKKELLVIPT ELHASLARQK YTDQSEAALN AQINVEYNVS
arabidopsis  TGVVFQPFEE VKKADLAIPI TSHASLARQR FADASEAVIN EQINVEYNVS
maize (FM2)  SGVVFQPFEE IKGELALVPQ SPDRSLARHK FVDDCEAAIN EQINVEYNAS

             A helix                B helix
          69         79         89         99        109
    soybean  YVYHSLFAYF DRDNVALKGF AKFFKESSEE EREHAEKLMK YQNTRGGRVV
    cowpea5  YVYHSLFAYF DRDNVALKGF AKFFKE
    cowpea3  YVYHSLFAYF DRDNIALKGL AKFFKESSEE EREHAEKLIK YQNIRGGRVV
    cowpea2  YVYHALFAYF DRDNVALRGL AKFFKESSEE EREHAEKLME YQNRRGGKVK
    cowpea1  YVYHAMYAYF DRDNVALKGL AKFFKE
arabidopsis  YVYHSMYAYF DRDNVAMKGL AKFFKESSEE ERGHAEKFME YQNQRGGRVK
maize (FM2)  YAYHSLFAYF DRDNVALKGF AKFFKESSDE EREHAEKLME YQNKRGGRVR

                                C helix
          119        129        139        149        159
    soybean  LHPIKNAPSE FEHVEKGDAL YAMELALSLE KLVNEKLLNV HSVADRNNDP
    cowpea5
    cowpea3  LHPITSPPSE FEHPEKGDAL YAMELALSLE KLTNEKLLYV HSVADRNNDA
    cowpea2  LQSIVMPLSE FDHADKGDAL HAMELALSLE KLTNEKLLHL HSVATKNGDV
    cowpea1
arabidopsis  LHPIVSPISE FEHAEKGDAL YAMELALSLE KLTNEKLLNV HKVASENNDP
maize (FM2)  LQSIVAPLTE FDHPEKGDAL YAMELTLALE KLVNEKLHSL HGVATRCNDP

             D helix                E helix
          169        179        189        199
    soybean  QMADFIESEF LSEQVESIKK ISEYVAQLRR VG.KGHGVWHF DQRLLD
    cowpea5
    cowpea3  QLADFIESEF LNEQVESIKK IAEYVTQLRL VG.KGHGVWHF DQRLLHD
    cowpea2  QLADFVESEF LGEQVESIKR ISEYVAQLRR VG.KGHGVWHF DQMLLHEGGHLA
    cowpea1
arabidopsis  QLADFVESEF LGEQIEAIKK ISDYITQLRM IG.KGHGVWHF DQMLLN
maize (FM2)  QLIDFIESEF LEEQGEAINK VSKYVAQLRR VGNKGHGVWHF DQMLLQEAA
```

Figure 5 Comparison of primary structures of some plant ferritins

Plant ferritin sequences are from soybean [6], cowpea leaves (cowpeas 2 and 3 are from the present study; cowpeas 1 and 5 are from [14]), *A. thaliana* seedlings [13] and maize [25]. Helices are identified by horizontal lines and include the plant-specific proposed P-helix [12]. The seven residues involved in the ferroxidase centre are shown in bold [2,23]. Points in a sequence represent gaps required for alignment. Numbering of the plant ferritin sequences starts at the cleavage site of the transit peptide (vertical line). The cleavage point has been determined experimentally for soybean [6]; the others are deduced.

We previously reported three fragments of cDNA for ferritin subunits in cowpea leaf extracts [14], and here we have expanded the list to four isolates (see Figure 2 for the complete sequence of the fourth isolate, CP3). The earlier isolates (called cowpea 1, 2 and 5) are shown in full in Figure 1 (CP2) and in part for isolates 1 and 5 in Figure 5. This paper reports the full sequence of the original cowpea 2. The isolates cowpea 1, 2 and 3 are most interesting because the subunit sequences diverge considerably from the other cowpea isolate (cowpea 5), which is almost identical with the published sequences from other legumes as represented by soybean in Figure 5 [6,8,11]. We have tried to isolate the full cDNA for cowpea 1 but without success so far. Expression results (Table 3) show that cowpea 1 mRNA is present in very low levels even in developing leaves. Thus it is not surprising that we failed to abstract a cDNA clone from our lambda phage libraries. We have also prepared a cDNA library from developing leaf extracts at the 7–8 cm stage (see Table 3), a sample that should contain a maximum amount of cowpea

© 1999 Biochemical Society

mRNA. This cDNA has been used for specific nested primer amplifications and the product has been probed with specific cowpea 1 fragment in Southern blots. We have not achieved a positive result and thus no clone has been isolated. By comparison, specific nested primer amplification worked well when cowpea 2 was used as a test of the system.

At this point it is instructive to examine the comparisons between published subunit sequences from other legumes and cowpea 2, a subunit sequence that must have separated from the other sequences a long time ago: Table 1 shows that there is 77% identity between mature (no transit peptide) cowpea 2 and the published soybean subunit. Examination of the aligned sequences in Figure 5 shows that these subunits are probably very similar in three-dimensional structure because most sequence variation is concentrated in the N-terminus, the A-helix and the long loop between helices B and C (the proposed structure of plant ferritin subunits can be found in [2,12]). On the basis of the structure of human H subunits, plant subunits have a core structure of four long helices (A to D), with a possible short helix at the C-terminus (see annotations in Figure 5). Plant subunits have a unique extra sequence at the N-terminus with a structural relationship to the rest of the subunit that is not really understood. The residues involved in iron oxidation and storage in the ferritin structure [2,23] are completely conserved in cowpea 2 (highlighted in Figure 5). Thus we have no evidence for functional differences between mature cowpea 2 and other plant subunits. The translated sequence of CP 2 contains a putative transit peptide that has no similarity to published sequences (see Figure 5). However, this transit peptide has the general characteristics required to target the nascent polypeptide to either plastids or mitochondria [20]. The same conclusions were also reached from the translated sequence for CP 3, with the exception that the transit peptide in this species has partial similarity to the transit peptide in the published soybean sequence (Figure 5).

Sequence comparisons alone cannot tell us why the gene for cowpea 2 (and cowpea 1) has diverged substantially from other members of the family. Large differences in ferritin sequence between plant species are not a unique observation: the recently published *Arabidopsis* sequence (Figure 5) [13] has diverged from anything found in legumes and in maize. The published maize sequences (Figure 5) [25] are different again; however, the amino acid substitutions are (as in cowpea 2) concentrated in the same structural elements. The four active genes that we have detected in cowpea form a unique group because of the protein sequence divergences demonstrated within one species. We have used our unique primers for cowpea 1 and 2 sequences to probe genomic DNA preparations from soybean. We have found proof of homologues of these species in soybean (see Figure 4). Briat and co-workers [13] suggest that the *A. thaliana* library of expressed sequence tags might contain fragments of quite divergent ferritin sequences. Taken together, this information implies strongly that there is a substantial family of ferritin genes in each higher-plant genome. Some genes will code for nearly identical protein subunits that are expressed in the plastids in response to different environmental signals (as discovered by Briat and co-workers in maize [9,10]). Some genes will code for protein subunits that have substantially different sequences but might have the same basic protein function and be expressed in the plastids (as we have found for cowpeas). Perhaps this second group of genes are expressed in different tissues (as discussed in the next section).

To try to begin to understand the function of the four different subunits in cowpeas, we decided to examine transcription by following the amounts of mRNA produced in different tissues at different stages of development. We had to make the assumption that each subunit sequence is the product of a unique gene. The work with maize [9] has shown that this might not be an accurate assumption because in maize, more than one gene can code for a single protein subunit sequence. There might be more than one gene coding for each of our designated cowpea sequences. The variations in the intron sequences of clones produced by PCR from soybean DNA (mentioned in the Results section) also supports the conclusion for maize. It is clear that we need more investigations into the total ferritin genes in one species. This should come partly out of total genome analyses.

The method of RT–PCR provided the best combination of specificity and sensitivity to start to answer the questions about expression. The outcome of many months of work is summarized in Table 3. The results obtained from RT–PCR are dependent on the design of unique primers for each ferritin subunit of interest and also on the design of an appropriate unique internal standard. All fragments had to be short and of similar lengths to prevent variations in the efficiency of amplification. To obtain an effective guide to changes in mRNA concentrations, we settled on 10-fold differences in amounts of internal standard between reactions (see Figure 3). These differences in standards explain why individual results have a low precision (Table 3). The results show very interesting changes. In developing leaves provided with a complete supply of nutrients, cowpea 2 and 5 subunits are likely to be major components of ferritin involved in leaf development. In contrast, cowpea 3 is initially transcribed at lower levels but remains present as a transcript in mature leaves. Cowpea 1 is not a significant transcript in leaves under the conditions studied. These observations show that at least three genes are employed to produce ferritin in normal leaf development. Roots are much more difficult to study and only a single type of root tissue sample has been used. Table 3 shows that cowpea 2 is the dominant transcription species, with smaller amounts of cowpeas 3 and 5. Again, cowpea 1 is detected but cannot be considered significant. These results do suggest (on the basis of mRNA concentrations) that ferritin is expressed in roots in amounts that might be similar to those in developing leaves. Is ferritin in roots acting as a staging point for iron before transport to the shoot, or is it present just for cellular development?

The results presented here are a substantial addition to our knowledge of plant ferritin and establish that this knowledge of the gene family is still rudimentary. The method of quantitative RT–PCR is potentially valuable for a more detailed examination of the selective expression of multiple ferritin genes in specific tissues under a variety of physiological conditions. This information needs to be complemented by some convenient refined method to analyse the subunit composition of ferritin isolated in micro quantities from tissue samples.

A.J.W. was supported by a Postdoctoral Fellowship awarded under the University of New England Women's Reentry Scheme.

REFERENCES

1 Theil, E. C. (1987) Annu. Rev. Biochem. **56**, 289–315
2 Andrews, S., Arosio, P., Bottke, W., Briat, J. F., von Darl, M., Harrison, P. M., Laulhère, J. P., Levi, S., Lobréaux, S. and Yewdall, S. (1992) J. Inorg. Biochem. **47**, 161–174
3 Ragland, M., Briat, J. F., Gagnon, J., Laulhère, J. P., Massenet, O. and Theil, E. C. (1990) J. Biol. Chem. **265**, 18339–18344
4 Seckbach, J. J. (1982) J. Plant Nutr. **5**, 369–394
5 Van der Mark, F., van den Briel, W. and Huisman, H. G. (1983) Biochem. J. **214**, 943–950
6 Lescure, A. M., Proudhon, D., Pesey, H., Ragland, M., Theil, E. C. and Briat, J. F. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 8222–8226
7 Klausner, R. D., Rouault, T. A. and Harford, J. B. (1993) Cell **72**, 19–28
8 Van Wuytswinkel, O., Savino, G. and Briat, J. F. (1995) Biochem. J. **305**, 253–261

© 1999 Biochemical Society

9 Fobis-Loisy, I., Loridon, K., Lobréaux, S., Lebrun, M. and Briat, J. F. (1995) Eur. J. Biochem. **231**, 609–619
10 Savino, G., Briat, J. F. and Lobréaux, S. (1997) J. Biol. Chem. **272**, 33319–33326
11 Spence, M. J., Henzl, M. and Lammers, P. J. (1991) Plant Mol. Biol. **17**, 499–504
12 Lobréaux, S., Yewdall, S., Briat, J. F. and Harrison, P. M. (1992) Biochem. J. **288**, 931–939
13 Gaymard, F., Boucherez, J. and Briat, J. F. (1996) Biochem. J. **318**, 67–73
14 Wicks, R. E. and Entsch, B. (1993) Biochem. Biophys. Res. Commun. **192**, 813–819
15 Brown, T. A. (1991) Molecular Biology Labfax, pp. 2–7, BIOS Scientific Publishers, London
16 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
17 Gilliland, G., Perrin, S., Blanchard, K. and Bunn, H. F. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 2725–2729
18 Dallman, M. J. and Porter, A. C. (1991) in PCR: A Practical Approach (McPherson, M. J. and Taylor, G. R., eds.), pp. 215–224, IRL Press, Oxford
19 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20 Von Heijne, G., Steppuhn, J. and Herrmann, R. G. (1989) Eur. J. Biochem. **180**, 535–545
21 Raeymaekers, L. (1993) Anal. Biochem. **214**, 582–585
22 Proudhon, D., Briat, J. F. and Lescure, A. M. (1989) Plant Physiol. **90**, 586–590
23 Lawson, D. M., Artymiuk, P. J., Yewdall, S. J., Smith, J. M. A., Livingstone, J. C., Treffry, A., Luzzago, A., Levi, S., Arosio, P., Cesareni, G. et al. (1991) Nature (London) **349**, 541–544
24 Ragland, M. and Theil, E. C. (1993) Plant Mol. Biol. **21**, 555–560
25 Lobréaux, S., Massenet, O. and Briat, J. F. (1992) Plant Mol. Biol. **19**, 563–575

Received 5 June 1998/12 October 1998; accepted 18 November 1998

© 1999 Biochemical Society