Characterisation of three cDNA clones encoding different mRNAs for the precursor to the small subunit of wheat ribulosebisphosphate carboxylase

Steven M.Smith[1]*, John Bedbrook[2]* and Jim Speirs[3]+

*Division of Plant Industry, CSIRO, P.O. Box 1600, Canberra City, ACT 2601, and +Plant Physiology Unit, Division of Food Research, CSIRO, Macquarie University, North Ryde, NSW 2113, Australia

ABSTRACT
      We have isolated and sequenced three cDNA clones for the nuclear-
encoded precursor to the small subunit of the chloroplast enzyme, ribulose-
1,5-bisphosphate carboxylase of wheat. The nucleotide sequences of these
clones are different, indicating that they are probably derived from three
different mRNAs. This finding is consistent with the proposal that this
polypeptide is encoded by a multigene family in wheat, in support of
similar data reported by Broglie et al. (Bio/Technology 1:55–61, 1983).
We deduce that the mature small subunit polypeptide is comprised of 128
amino acids and that its precursor contains an N-terminal transit peptide
sequence. The sequences of both the mature small subunit and its transit
peptide differ at several positions from those determined by Broglie et al.
(1983) from a different wheat cultivar. Different wheat cultivars might
therefore contain different small subunit polypeptides. A comparison of
nucleotide and amino acid sequences of the small subunit from wheat, pea,
soybean and spinach shows that these sequences are not highly conserved,
particularly between monocotyledon and dicotyledon species.

INTRODUCTION

      The chloroplast enzyme, ribulose-1,5-bisphosphate carboxylase

(RuBPCase; E.C.4.1.1.39) is the most abundant protein of green plants (1).

This enzyme catalyses both the carboxylation and the oxygenation of ribu-

losebisphosphate as the first steps in photosynthetic carbon reduction and

photorespiration, respectively (2). The holoenzyme is an oligomer of

eight large subunit polypeptides (molecular weight 53,000) and eight small

subunit polypeptides (molecular weight 14,000) (3). The large subunit

contains the catalytic site for both carboxylase and oxygenase activities

(2), and its amino acid sequence is well conserved between different plant

species (4,5,6). In contrast, the function of the small subunit is

unknown, and its primary structure is less well conserved (7,8,9).

      The large subunit is encoded by a single gene in the chloroplast

genome (10) and synthesised in the chloroplast (11). The small subunit is

encoded in the nuclear genome (12) and synthesised in the cytosol as a

precursor of approximately 20,000 molecular weight (13,14,15,16). The

small subunit precursor is processed and imported into chloroplasts post-translationally (13,14,17,18). The amino acid extension removed during processing of this precursor is referred to as the transit peptide, because of its assumed role in polypeptide transport across the chloroplast envelope (17).

The synthesis of RuBPCase requires the coordinated, and develop-mentally-regulated, expression of chloroplast and nuclear genes. For example, in maize leaves, this enzyme is synthesised in bundle sheath cells but not in mesophyll cells (19,20). In some plant species the expression of nuclear and chloroplast genes is coordinately regulated by light (21,22). In the second leaf of wheat, the abundance of translatable mRNAs for both subunits declines in a coordinated manner as the leaf ages (16). While the gene encoding the large subunit polypeptide of wheat has been isolated (23), no cloned probe encoding the small subunit was avail-able to study, in more detail, the control of expression of RuBPCase genes in wheat. Cloned DNA sequences encoding this polypeptide had been obtained from pea (7,24) and soybean (8), but the cDNA clone for the pea polypeptide was found to hybridise very poorly with the mRNA of mono-cotyledon species (see RESULTS section). We have therefore isolated cDNA clones encoding the wheat (var. Falcon) small subunit polypeptide.

Recently, Broglie et al. (9) have reported the isolation and characterisation of a cDNA clone and a genomic clone encoding two different small subunit precursors of wheat (var. ERA). We have char-acterised three wheat small subunit cDNA clones and find that they differ from each other and from the sequences obtained by Broglie et al. (9). These results support the view that the wheat small subunit polypeptide is encoded by a family of expressed nuclear genes, but also might indicate that different cultivars of wheat contain different small subunit polypep-tides.

MATERIALS AND METHODS
Growth of plants
        Wheat (Triticum aestivum L. var. Falcon), pea (Pisum sativum), barley (Hordeum vulgare), petunia (Petunia hybrida) and bean (Phaseolus vulgaris) were grown as described previously for wheat plants (25).
Isolation of leaf RNA
        For wheat leaf RNA, the second leaves of wheat seedlings were harvested ten days after germination and total leaf RNA was isolated and

purified as described by Speirs and Brady (16). RNA was isolated by the
same method from young expanding leaf tissue of the other plants.

Isolation of wheat leaf polyadenylated RNA

Polyadenylated RNA (poly(A) RNA) was isolated by a modification of the
technique of Aviv and Leder (26). Total RNA (5 mg) dissolved in application
buffer (50 mM Tris-Cl, pH 7.5, 0.5 M NaCl, 0.1% (w/v) sodium dodecyl sulphate
(SDS), 1 mM EDTA) was applied to a column containing 0.5 g oligo(dT) cellulo-
se (Collaborative Research, type II) and recycled through the column several
times. The column was washed with 50 mM Tris-Cl, pH 7.5, 0.1 M NaCl, 0.1%
(w/v) SDS, 1 mM EDTA, and the poly(A) RNA eluted with distilled water. The
poly(A) RNA solution was adjusted to the concentration of application
buffer, reapplied to the column, washed, and eluted as before. This proc-
edure was repeated a further time before the poly(A) RNA was collected by
ethanol precipitation.

In vitro translation of RNA

RNA samples were translated in vitro in a wheat germ extract as des-
cribed previously (16). 6 µg of total RNA or 0.5–1 µg poly(A) RNA gave
optimum incorporations in 25 µl assays. The labelled isotope was L-[^{35}S]
methionine (Amersham). Hybrid-release-translations were carried out as
described previously (7,27).

Synthesis and cloning of cDNA

Poly(A) RNA from wheat leaves was used to direct the synthesis of
double–stranded cDNA as described previously (7). Following S_1 nuclease
treatment (7), DNA fragments of less than approximately 300 bp were removed
by chromatography on a column (30 x 0.6 cm) of Bio–Gel A150m (Bio–Rad) in
10 mM Tris-Cl, pH 7.6, 1 mM EDTA, 0.02% Na azide. The remaining cDNA was
treated with DNA polymerase and ligated (7) to phosphorylated BamHI linker
molecules (Collaborative Research). The cDNA was then digested with BamHI
and linker fragments removed by chromatography on a column (20 x 0.6 cm) of
Sepharose 4B in 10 mM Tris-Cl, pH 7.5, 1 mM EDTA, 200 mM Na acetate. The
cDNA was ligated into BamHI-linearised pBR322 (28), and Escherichia coli
K-12 strain 294 (Pro$^-$, EndA$^-$, r_k^- m_k^+, suIII$^+$, Thi$^-$) was transformed (29)
to ampicillin and fusaric acid resistance (30). Approximately 250 clones
were obtained from 1 µg poly(A) RNA.

Colony hybridisations

Clones were screened by colony hybridisation (31) initially using as
a probe, high specific activity ^{32}P-labelled DNA complementary to a sample
of the poly(A) RNA used to construct the clones. The reaction for synthesis-

ing this cDNA contained 10 μg poly(A) RNA, 50 mM Tris-Cl, pH 8, 8 mM $MgCl_2$,
0.4 mM dithiothreitol, 40 mM KCl, 12.5 μg ml^{-1} oligo(dT)$_{12-18}$, 0.2 mM each
of dATP, dGTP and dTTP, 50 μCi α[^{32}P]dCTP (2000 Ci $mmol^{-1}$, Amersham) and
5 μl reverse transcriptase (J. Beard) in a final volume of 50 μl. Incubation
was for 60 min at 37°C. The RNA was then hydrolysed by adding NaOH to 0.4 M,
SDS to 0.5% (w/v) and EDTA to 25 mM, in a final volume of 100 μl. After 120
min at 37°C, 2 μl glacial acetic acid was added and the cDNA purified by
phenol extraction and Sephadex G50 chromatography in the same buffer used
for Sepharose 4B chromatography (above). This protocol results in the
incorporation of approximately 50% of the isotope.

Bacterial colonies were grown on nitrocellulose filters (Millipore) on
1% (w/v) agar containing Luria broth and 50 μg ml^{-1} ampicillin. Plasmid
amplification was induced by incubating the filters on the same medium
supplemented with 100 μg ml^{-1} chloramphenicol, for 16 hours at 37°C.
Colonies were then denatured and baked as described previously (31). Filters
were prehybridised for 4 hours at 65°C, then hybridised for 16 hours at 65°C
in 4 x SET (0.6 M NaCl, 120 mM Tris-Cl, pH 8, 4 mM EDTA), 10 x Denhardt's
solution (32), 0.1% (w/v) SDS, 0.1% (w/v) Na pyrophosphate, 50 μg ml^{-1}
salmon sperm DNA, 100 μg ml^{-1} polyadenylic acid. The hybridisation reaction
contained 2 x 10^6 cpm (Cerenkov) of cDNA. After hybridisation, the filters
were washed extensively with 2 x SET, 0.1% (w/v) SDS, 0.1% (w/v) Na
pyrophosphate, at 60°C. After drying, the filters were exposed to X-ray
film together with an intensifying screen at –80°C for two days. When the
hybridisation probe employed was the nick-translated cDNA insert from
pTS234, HaeIII-digested pBR322 was included in the hybridisation reaction
at 1 μg ml^{-1}.

Purification of plasmid DNAs

Plasmid DNA was isolated from chloramphenicol-amplified cultures of
E. coli and purified by CsCl/ethidium bromide centrifugation (33,34).

DNA sequencing

DNA was sequenced according to the chemical modification procedure (35).
Restriction endonuclease fragments were labelled at their 3' ends using
α[^{32}P]dATP or α[^{32}P]dCTP and the Klenow fragment of DNA polymerase (36).
Labelled ends were separated either by a second restriction endonuclease
cleavage and fractionation on a 6% polyacrylamide gel (37) or by strand
separation (35). In addition to the four chemical modification reactions
recommended (35), two further reactions were usually employed, namely
alkali treatment (35) and hydroxylamine hydrochloride (38).

Preparation of RuBPCase DNA probes

cDNA sequences encoding the small subunit of wheat RuBPCase were
excised from pTS234 and pTS512 by BamHI restriction endonuclease cleavage.
The probe for the pea small subunit was isolated from pSSU1 (7). A BglII
fragment, encompassing the structural sequence of the maize large subunit
gene was excised from pZmc37 (10). The excised fragments were purified by
electrophoresis in 1% w/v agarose gels, recovered by absorption to hydro-
xylapatite and eluted with 0.6 M phosphate buffer according to Tabak and
Flavell (39). The purified fragments were labelled by nick-translation (40)
with $\alpha[^{32}P]dCTP$ or $\alpha[^{32}P]dATP$.

Hybridisation to electrophoretically-fractionated RNA

Samples of total leaf RNA (5 µg) were denatured in 5 mM methyl mercury
hydroxide and fractionated by electrophoresis in 1% (w/v) agarose gels
essentially by the method of Bailey and Davidson (41) except that the
electrophoresis buffer was 50 mM boric acid, 5 mM Na tetraborate, 10 mM Na
sulphate and 0.1 mM EDTA, pH 8.2. The gels were prepared for RNA transfer
to DBM-paper and photographed as described by Alwine et al. (42). RNA was
transferred to DBM-paper by blotting in 0.2 M Na acetate, pH 4. DBM-paper
filters were pre-hybridised at 42°C for 4 hours in 50% (v/v) formamide,
5 x Denhardt's solution (32), 0.1% (w/v) SDS, 100 µg ml^{-1} sonicated salmon
sperm DNA, 1% (w/v) glycine, 5 x SSPE (0.9 M NaCl, 5 mM EDTA, 50 mM NaH_2PO_4-
KOH, pH 7.7). Hybridisation reactions were for 12 hours at 42°C in the
same solution, except that glycine was omitted, Denhardt's solution reduced
to 2 x, and dextran sulphate included at 10% (w/v) (43). Approximately 2-3
x 10^6 dpm of ^{32}P-labelled DNA probe was included in the hybridisations.

The hybridised filters were washed twice for 15 min each in 2 x SSPE
buffer, 0.1% (w/v) SDS at room temperature, followed by two washes of 15 min
each in 0.1 x SSPE buffer, 0.1% (w/v) SDS at 50°C. The dried filters were
exposed to pre-flashed X-ray film at –80°C with an intensifying screen (44).

RESULTS

Isolation of cDNA clones for the small subunit of wheat RuBPCase

The products of in vitro translation of wheat leaf poly(A) RNA are
shown in Fig. 1(a). Prominent among these products is the precursor to the
small subunit (pSSU) of RuBPCase (16). Such an RNA preparation was used to
construct a set of cDNA clones which were subsequently screened by colony
hybridisation with ^{32}P-labelled cDNA, synthesised from a sample of the same
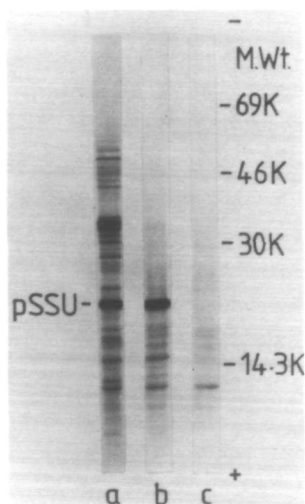RNA preparation. Clones giving the most intense hybridisation signals were

FIG. 1. Products of in vitro translation of wheat leaf RNA in a wheat germ system. [^{35}S]methionine-labelled products were separated by SDS gel electrophoresis in a 15–20% (w/v) polyacrylamide gradient and visualised by fluorography (16). (a) Products of total poly(A) wheat leaf RNA. (b) Products of RNA selected by hybridisation to pTS234. (c) Endogenous activity of the wheat germ system. pSSU is the precursor to the small subunit of RuBPCase (16). Molecular weight markers are ^{35}S-labelled methylated proteins (Amersham): Bovine serum albumin, 69K; Ovalbumin, 46K; Carbonic anhydrase, 30K; Lysozyme, 14.3K.

further analysed by hybrid-release-translation (27). Plasmid DNA from one of these clones (designated pTS234) hybridises with a mRNA whose translation product comigrates with pSSU (Fig. 1(b)). Confirmation of the identity of the cDNA insert of pTS234 as encoding pSSU was obtained by nucleotide sequencing (see below).

The purified cDNA from pTS234 was used as a probe to screen for further small subunit clones by colony hybridisation. Two further clones were obtained in this way, their plasmids being designated pTS406 and pTS512. The three cDNA clones obtained were subsequently characterised by nucleotide sequencing. Restriction endonuclease maps of the three small subunit cDNA clones are shown in Fig. 2. The three cDNAs are aligned relative to common HinfI and AvaI restriction endonuclease sites and orientated in the 5'-3' direction, relative to the mRNA, as determined by nucleotide sequencing.

3' untranslated sequences

The DNA sequences corresponding to the 3' untranslated region of the mRNA are shown in Fig. 3. Each cDNA encodes a complete 3' untranslated sequence separating a TAA termination codon of the coding region from a poly(A) tail. Comparison of these sequences shows that the three cDNA clones are derived from two, or possibly three, different mRNAs. The homology between the 3' ends of pTS234 and pTS406 is almost complete, there being only four nucleotide differences between the sequences (double underlined in Fig. 3). One of the differences, the presence of a thymidine in pTS406 at the 5' end of the poly(A) tail, results in the 3' non-coding
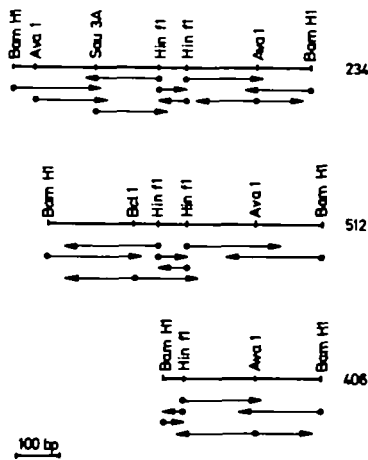
FIG. 2. Restriction endonuclease maps and sequencing strategies for cloned small subunit cDNAs. Sequences are aligned in the 5'-3' direction relative to the mRNA. Open circles and arrows represent the start points and directions of sequencing (35) from labelled 3' ends. 234, 512 and 406 are the cDNA clone numbers.

region of pTS406 consisting of 255 nucleotides whereas that of pTS234 is 250 nucleotides. The 3' non-coding region of pTS512 has many differences relative to the other two clones (single underlined in Fig. 3). Within the first 231 nucleotides there are 16 differences, two of which represent nucleotide additions or deletions. The following 38 nucleotides, which comprise the remainder of the 3' non-coding sequence, are apparently unrelated to the sequences of pTS234 and pTS406. The 3' untranslated sequences of two clones for the wheat small subunit polypeptide characterised by Broglie et al. (9) are different to each other, and to those shown here.

Coding sequences for the mature small subunit polypeptide

The complete nucleotide sequences of the three small subunit cDNA clones revealed that the cDNA of pTS234 is organised differently to that of pTS512 and to the sequences described by Broglie et al. (9). The organisa-tion of the cDNA of pTS234 is shown relative to the mRNA in Fig. 4. The coding information for amino acids 60-84 is not present in pTS234. Further-more the sequence encoding the 5' end of the mRNA (up to amino acid 59) is present in an inverted orientation relative to the remainder of its sequence, and relative to pTS512. The means by which the organisation of pTS234 has arisen is unknown. While it might reflect the presence of an equivalent mRNA in the wheat leaf, we believe it more probable that it was generated as

```
        coding      3' non-coding
        region         region

    ◄━━━━━━━━┃━━━━━━━━━━━►

512 --- TAA ACAGCTCACT CACCACGGGC CACATATAAA GTGCCATTGC AGTTTTGTCA ACTCTGACAT
234 --- TAA ACAGCTCACT CACGACAGGC CATATATAAA GTGCCATTGC AGTTTTGTCA ACTCTGACAT
406 --- TAA ACAGCTCACT CACGACAGGC CATATATAAA GTGCCATTGC AGTTTTGTCA ACTCTGACAT


512     TGCTTTGGGT TTTCCTTCTC CATTTGTCTT TCTTACTTGT TCCTAAGAAT ATGTGTATGT
234     TGCTTTGGGT TTTCCTTCTC CATTTATCTT TCTTATTTGT TCCTAAGAAT ATGTGTATGT
406     TGCTTTGGGT TTTCCTTCTC CATTTATCTT TCTTATTTGT TCCTAAGAAT ATGTGTATGT


512     CCATGTTCAT GTAACAACAT GGCTCGAGAA AGCGTGCTCG TATGTGAATG CAATCGGTGG
234     CCATGTTCAT GTACCAACAT GGCTCGAGAA AGCATGCTCG TATGAGAATG CTATCGGTGG
406     CCATGTTCAT GTACGAACAT GGCTCGAGAA AGCATGCTCG TATGTGAATG CTATCGGTGG


512     CTATGTATAT GCACATGCAC ATAATATATT ATACAATCTG TCGACTATGT ATTCGGGTAA
234     CTATGTATAT GAACGGACGC ATAATA  TT ATACAATATG TCGACTATAT ACGTTTTAAT
406     CTATGTATAT GCACGGACGC ATAATA  TT ATACAATATG TCGACTATAT ACGTTTTAAT


512     TTAATAATGA GAAACATCTT TGCCACCGC (A)
234     GGTAACCAAT GG (A)
406     GGTAACCAAT GGAAAAT (A)
```

FIG. 3. Nucleotide sequences of the 3' untranslated regions of cDNAs 512, 234 and 406. Shown here is the non-coding (mRNA sense) strand. Single underlinings show the positions of differences between the sequence of cDNA 512 and those of cDNAs 234 and 406. Double underlinings show the positions of differences between cDNAs 234 and 406. (A) denotes poly(A) tails, the lengths of which are: 512 – 30 nucleotides, 234 – 19 nucleotides and 406 – 44 nucleotides.

an artefact of the cloning process. A similar inversion has been found in an unrelated cDNA also cloned in this laboratory. Despite the organisation of pTS234, the sequence information which it contains is of value for elucidating the complete amino acid sequence of the mature small subunit polypeptide. This clone also encodes 35 amino acids of the transit peptide sequence (see below).

The nucleotide sequences of the three cDNA clones, together with the encoded amino acids, are shown in Fig. 5. The region of pTS234 which is inverted in the cDNA clone, is here presented in the same orientation as pTS512. The deduced amino acid sequence of the mature small subunit is consistent with similar data for pea (7,24), spinach (45), soybean (8) and wheat (9), although some differences exist between all of these sequences (see below). Clone pTS512 encodes 113 amino acids of the mature small subunit polypeptide, but does not include the information for amino acids 1-15 at the N-terminal end of the sequence. pTS234 does encode the N-terminal amino acids of the mature small subunit polypeptide but, as explained above,
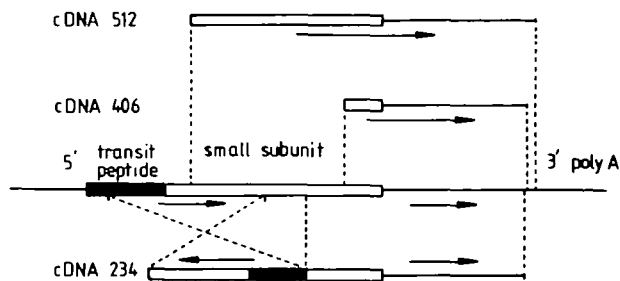
FIG. 4. Relationships between cloned cDNA sequences and the mRNA for the
small subunit precursor. Regions represented by open boxes correspond to
coding sequences for the mature small subunit, and closed boxes to the
transit peptide sequence. Arrows are orientated in the 5'-3' direction,
showing the inversion in cDNA 234 (see text).

lacks the coding information for amino acids 60-84. Taken together, these
two clones account for the whole small subunit polypeptide, which we deduce
to be composed of 128 amino acids, in agreement with Broglie et al. (9).

     In the overlapping coding regions of pTS234 and pTS406, all 64 nucleo-
tides are common to both clones. However, within the overlapping coding
sequences of pTS234 and pTS512, twelve nucleotide differences occur (under-
lined in Fig. 5). Ten of these differences occur at the third position of
the codons and do not alter the amino acid assignments. Interestingly, four
of these occur in adjacent codons (amino acids 55-58). The other two
differences do affect the amino acid assignments. The codon for amino acid
16 specifies serine in pTS234 but alanine in pTS512, while that for amino
acid 59 codes for valine in pTS234 and glycine in pTS512. There are two
reasons for doubting the assignment of alanine at position 16 in pTS512.
Firstly, serine occurs at this position in both sequences for the wheat
small subunit as determined by Broglie et al. (9), and in at least six other
plant species (7,8,45,46,47). Secondly, the alanine codon in pTS512 differs
from the serine codon in only the first nucleotide which, in pTS512 occurs
next to the BamHI linker molecule ligated to the cDNA during the construction
of the clone. The presence of guanidine at this position in pTS512 might be
a cloning artefact. Similarly we cannot feel confident about the assignment
of valine at position 59 in pTS234. This codon also occurs adjacent to a
BamHI linker molecule, and the amino acid assignment of Broglie et al. (9)
is glycine at this position.

     While we have no strong evidence for differences in the amino acid
sequences for the small subunit polypeptides encoded by these three cDNA

```
                                                                          -35
                                                                          Val
        512
        234                                              N-terminal ....C CTC


                                -30                 -25                 -20
                        Ala Pro Phe Gln Gly Leu Lys Ser Thr Ala Gly Leu Pro Val Ser Arg Arg
        512
        234             GCT CCT TTC CAG GGG CTC AAG TCC ACC GCC GGC CTC CCC CTC AGC CGC CGC


                                -15                 -10                 -5
                        Ser Asn Gly Ala Ser Leu Gly Ser Val Ser Asn Gly Gly Arg Ile Arg Cys
        512
        234             TCC AAC GGC GCT AGC CTC GGC AGC GTC ACC AAC GGT GGA AGG ATC AGG TGC


                                  5                  10              15 (Ala)
                        Met Gln Val Trp Pro Ile Glu Gly Ile Lys Lys Phe Glu Thr Leu Ser Tyr
        512                                                     Bam linker - GCT TAC
        234             ATG CAG GTC TGG CCC ATC GAG GGC ATC AAG AAG TTC CAG ACC CTG TCT TAC


                      20             Ser Thr 25              30
                        Leu Pro Pro Leu Ser Thr Glu Ala Leu Leu Lys Gln Val Asp Tyr Leu Ile
        512             CTG CCA CCG CTC AGC ACG GAG GCC CTC CTC AAG CAG CTG CAC TAC CTG ATC
        234             CTG CCA CCG CTC AGC ACG GAG GCC CTC CTC AAG CAG CTC CAC TAC CTG ATC


                      35             40              45              50
                        Arg Ser Lys Trp Val Pro Cys Leu Glu Phe Ser Lys Val Gly Phe Ile Phe
        512             CGC TCC AAG TGG GTG CCT TGC CTT GAG TTC AGC AAG GTC GGG TTC ATC TTC
        234             CGC TCC AAG TGG GTG CCT TGC CTC GAG TTC AGC AAG GTT GGG TTC ATC TTC


                                  55          (Val)60              65
                        Arg Glu His Asn Ala Ser Pro Gly Tyr Tyr Asp Gly Arg Tyr Trp Thr Met
        512             CGT CAG CAC AAT GCG TCT CCC GGG TAC TAC GAC GGC CGC TAC TGG ACA ATG
        234             CGT CAG CAC AAC GCA TCA CCT GTC - Bam-linker


                      70             75              80              85
                        Trp Lys Leu Pro Met Phe Gly Cys Thr Asp Ala Thr Gln Val Ile Asn Glu
        512             TGG AAG CTG CCT ATG TTC GGG TGC ACC GAC GCC ACA CAG CTG ATC AAC CAC
        234                                                                            CAC


                      90             95              100
                        Val Glu Glu Val Lys Lys Glu Tyr Pro Asp Ala Tyr Val Arg Ile Ile Gly
        512             GTC GAG CAG CTC AAG AAG GAG TAC CCT GAC GCG TAC GTC CGC ATC ATC GGA
        234             GTC GAG CAG CTC AAG AAG GAG TAC CCT GAC GCG TAT CTC CGC ATC ATC GGA


                      105            110             115
                        Phe Asp Asn Met Arg Gln Val Gln Cys Val Ser Phe Ile Ala Phe Lys Pro
        512             TTC GAC AAC ATC CGC CAG GTC CAG TGC GTC AGC TTC ATT GCC TTC AAG CCA
        234             TTC GAC AAC ATC CGT CAG GTC CAG TGC GTC AGC TTC ATC GCC TTC AAG CCA
        406                     Bam linker - T CAG GTC CAG TGC GTC AGC TTC ATC GCC TTC AAG CCA


                      120            125
                        Pro Gly Cys Glu Glu Ser Gly Lys Ala
        512             CCC GGC TGC CAG GAG TCC GGC AAG GCC TAA ........ 3' untranslated
        234             CCC GGC TGC CAG GAG TCC GGC AAG GCC TAA ........
        406             CCG GGC TGC CAG GAG TCC GGC AAG GCC TAA ........
```
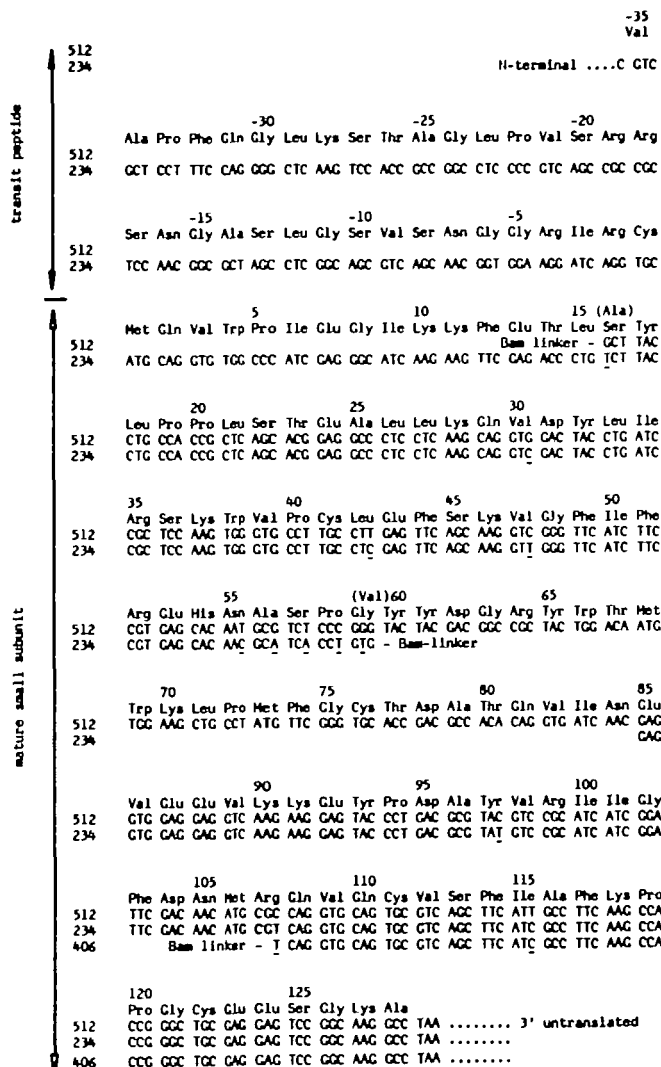
FIG. 5. Amino acid coding sequences of cloned cDNAs 512, 234 and 406.
Deduced amino acids of the mature small subunit have positive numbers, those
of the transit peptide, negative numbers. The assignments of amino acids
shown in brackets (positions 16 and 59) are tentative (see text). Differen-
ces in nucleotides between cDNA sequences 512 and 234 are underlined. The
position of the Bam-linker molecule in cDNA 234 is explained in the text,
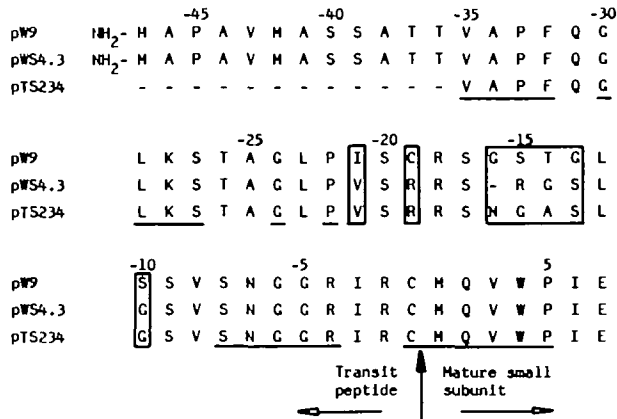and in Fig. 4.

```
              -45           -40           -35           -30
pW9    NH₂- H  A  P  A  V  H  A  S  S  A  T  T  V  A  P  F  Q  G
pWS4.3 NH₂- H  A  P  A  V  H  A  S  S  A  T  T  V  A  P  F  Q  G
pTS234    -  -  -  -  -  -  -  -  -  -  -  -  -  V  A  P  F  Q  G

                    -25           -20           -15
pW9    L  K  S  T  A  G  L  P [I] S [G] R  S [G  S  T  G] L
pWS4.3 L  K  S  T  A  G  L  P [V] S [R] R  S [-  R  G  S] L
pTS234 L  K  S  T  A  G  L  P [V] S [R] R  S [N  G  A  S] L

       -10           -5                           5
pW9   [S] S  V  S  N  G  G  R  I  R  C  H  Q  V  W  P  I  E
pWS4.3[G] S  V  S  N  G  G  R  I  R  C  H  Q  V  W  P  I  E
pTS234[G] S  V  S  N  G  G  R  I  R  C  H  Q  V  W  P  I  E

                         Transit  ↑  Mature small
                         peptide     subunit
```

FIG. 6. Comparison of transit peptide sequences deduced from nucleotide sequences of pWS4.3 and pW9 with that deduced from pTS234. Boxes show positions where these sequences differ. Regions which are absolutely conserved between pea (24), soybean (8) and wheat, are underlined. An arrow denotes the position at which the transit peptide is cleaved from the precursor polypeptide.

clones, the sequence determined here does differ from those determined by Broglie et al. (9). This sequence differs from pW9 in five amino acids (positions 50, 56, 83, 100 and 118) and from pWS4.3 in six (position 106 also). These differences all represent conservative amino acid changes, and none of them occurs in conserved regions of the small subunit polypeptide (see ref. 9). At the nucleotide sequence level, pTS234 differs from pTS512 by less than 4%, but both differ from the sequences characterised by Broglie et al. (9) by approximately 8%. The nucleotide sequences encoding the mature small subunit from clones pW9 and pWS4.3 (9) similarly are only about 4% different, even though the 3' and 5' flanking sequences of these two clones are completely different.

Transit peptide sequence

cDNA clone pTS234 encodes 35 amino acids of an N-terminal transit peptide sequence (Fig. 5). This sequence shows several differences relative to the equivalent sequences of the two small subunit precursors of wheat identified by Broglie et al. (9). These three sequences are shown for comparison in Fig. 6. All of the amino acid differences between these sequences occur between positions -11 and -21 (boxed in Fig. 6).

Hybridisation of cDNA to small subunit mRNA

One purpose for isolating cDNA clones encoding the small subunit of RuBPCase of wheat is to assay the synthesis of its mRNA during leaf develop-
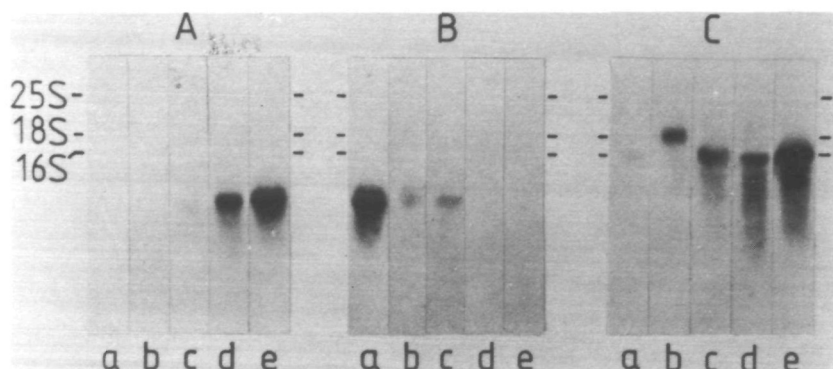
FIG. 7. Hybridisation of cloned probes for RuBPCase subunits to electro-
phoretically fractionated leaf RNAs. Total leaf RNA (5 µg) from (a) pea,
(b) Phaseolus, (c) petunia, (d) barley and (e) wheat was fractionated by
agarose gel electrophoresis in the presence of methyl mercury hydroxide and
transferred to DBM-paper (42). DBM-papers were hybridised with ^{32}P-labelled
probes for (A) wheat small subunit (cDNA 512), (B) pea small subunit (cDNA
SSU1, ref. 7) and (C) maize large subunit (from pZmc37 (ref. 10) – see
Materials and Methods). Shown here are the autoradiographs of the hybrid-
isation experiments. The electrophoretic mobilities of leaf rRNAs are
indicated.

ment (16). The use of pTS512 cDNA to detect this mRNA is demonstrated in
the experiment shown in Fig. 7. Total wheat leaf RNA was fractionated by
agarose gel electrophoresis under denaturing conditions, and transferred to
diazotised paper (42). Hybridisation with ^{32}P-labelled pTS512 cDNA
identifies the small subunit mRNA which has an electrophoretic mobility
equivalent to about 900 nucleotides (Fig. 7(A)). Included in this experi-
ment are RNA preparations from pea, petunia, Phaseolus and barley. The
wheat small subunit cDNA hybridises well to barley RNA, weakly to petunia
and almost undetectably to the two legumes (Fig. 7(A)). When cDNA for the
small subunit of pea (7) is used as a probe, strong hybridisation to pea RNA
is observed, and weak hybridisation to Phaseolus and petunia RNAs. However,
no hybridisation is evident to RNA of the two monocotyledon species (Fig.
7(B)). These observations confirm that the small subunit sequence is poorly
conserved, particularly between monocotyledon and dicotyledon species, and
demonstrates the importance of employing homologous probes for this mRNA.
In contrast, a DNA probe for the large subunit of RuBPCase from maize (10)
hybridises with RNA from each of the five species tested here (Fig. 7(C)),
demonstrating the relative conservation of this sequence between different

species. Interestingly, the mRNA for the large subunit of Phaseolus is
apparently much larger than that of the other four species tested here.

DISCUSSION

     The three cDNA clones for the small subunit of RuBPCase described here,
have different nucleotide sequences and are probably derived from three
different small subunit mRNAs. pTS512 differs markedly from pTS234 and
pTS406 both in the coding sequence and in the 3' untranslated sequence.
However, it is possible that the four nucleotide differences between pTS234
and pTS406 could have resulted from errors in either cDNA synthesis or in
plasmid replication. Thus, we have evidence for the expression of two, and
possibly three, different small subunit genes in the wheat leaf. Broglie
et al. (9) have demonstrated the expression of two further small subunit
genes in wheat, and shown that wheat probably contains more than ten small
subunit genes.

     It might be expected that wheat should contain several such different
genes because of its hexaploid nature, having been derived from three ancest-
ral diploid species (48). However, multiple small subunit genes have also
been detected in the diploid species, pea (24,49) and petunia (49) and in
the tetraploid soybean (8). In soybean, only one gene is known to be
expressed (8) but at least two different genes are expressed in pea (7,24)
and at least four different genes in petunia (49).

     Significant differences in the deduced sequences of the small subunit
polypeptides are observed when different cultivars of wheat and pea are
compared. Two sequences for the small subunit of wheat var. ERA differ in
only one amino acid (9), and in the present study employing wheat var.
Falcon, no convincing evidence for amino acid sequence heterogeneity is
found, and none is revealed by isoelectric focussing (50). However, the
sequence for the small subunit of Falcon differs from those of ERA by five
and six amino acids. Similarly, the amino acid sequence for the small
subunit of RuBPCase from Pisum sativum var. Feltham First (7,51) differs
from a sequence from P. sativum var. Progress No. 9 by eight amino acids (24).
Further studies are required to establish the full extent of amino acid
sequence heterogeneity within and between different cultivars of the same
species.

     The cDNA sequences corresponding to the 3' untranslated regions of the
small subunit mRNAs described here, are very rich in nucleotides A+T (60-62%)
relative to the coding regions (38-40%). However, none of these clones

encodes the sequence 5' AAUAAA 3' which is thought to be important for the polyadenylation of many eukaryotic mRNAs (52,53). This sequence is also absent from the 3' untranslated regions of the wheat small subunit mRNAs described by Broglie et al. (9) and from those of pea (7,24) and soybean (8). Neither is this sequence to be found in the 3' untranslated region of a mRNA for the major chlorophyll a/b binding protein (another nuclear-encoded chloroplast protein) of pea (24).

The accumulation of amino acid sequence data for the precursor to the small subunit of RuBPCase from different plant species, allows us to identify regions of this polypeptide which are conserved. These conserved regions are presumably important to its functions, for example in protein transport into the chloroplast or in the activity of the holoenzyme. Conserved features of the transit peptide sequence, particularly the positions of proline, lysine and arginine residues, have been described by Broglie et al. (9). The transit peptide sequence data for wheat described in the present study differs from two similar sequences obtained by Broglie et al. (9) in three and seven amino acids respectively (see Fig. 6). These differences all occur in a region of eleven amino acids which is also seen to be poorly conserved when pea and soybean are considered (9). Thus, the primary structure of this region may not be important to the function of the transit peptide. Comparison of the sequence data for the mature small subunit of pea (7,24), soybean (8), spinach (45), petunia (49) and wheat (9) reveals several regions which are completely conserved (see ref. 9). Most striking is a conserved hexadecapeptide which, in wheat, occurs between amino acid residues 60 and 75 (see Fig. 5). The conserved hexadecapeptide is rich in hydrophobic amino acid residues, relative to the rest of the small subunit polypeptide, suggesting that this part of the molecule may not be exposed at the surface of the RuBPCase holoenzyme. This sequence might therefore be involved in the binding of the small subunits to the central octomeric core of large subunits (2). The three charged amino acids in the conserved hexa-decapeptide could participate in ionic interactions with charged residues in the large subunit polypeptides. It is expected that the region of the small subunit which binds to the large subunit is highly conserved because of the conserved nature of the large subunit, and because the small subunits of one plant species will assemble with the large subunits of another (54).

The conserved hexadecapeptide is encoded by nucleotides which, between wheat and pea, show 81–85% sequence homology, whereas the overall nucleotide sequence homology for the complete small subunits of wheat and pea is approx-

imately 60% (7,9,24 and Fig. 5). This level of homology is insufficient to detect hybridisation between these sequences under the conditions employed in the experiment shown in Fig. 7. However, sufficient homology exists between sequences of wheat and barley small subunits to detect heterologous hybrid-isation, and similarly between pea, petunia and Phaseolus (Fig. 7). These observations demonstrate the relative sequence divergence of the small sub-unit between monocotyledon and dicotyledon plants.

[1]Present address: Botany Department, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JH, UK

[2]Present address: Advanced Genetic Sciences Inc., P.O. Box 3226, Berkeley, CA 94703, USA

[3]To whom correspondence should be addressed

REFERENCES
 1. Ellis, R.J. (1979) Trends in Biochem. Sci. 4, 241–244.
 2. Jensen, R.G. and Bahr, J.T. (1977) Ann. Rev. Plant Physiol. 28, 379–400.
 3. Baker, T.S., Eisenberg, D., Eiserling, F.A. and Weissman, L. (1975)
     J. Mol. Biol. 91, 391–399.
 4. McIntosh, L., Poulsen, C. and Bogorad, L. (1980) Nature 288, 556–560.
 5. Zurawski, G., Perrot, B., Bottomley, W. and Whitfield, P. (1981)
     Nucl. Acids Res. 14, 3251–3270.
 6. Shinozaki, K. and Sugiura, M. (1982) Gene 20, 91–102.
 7. Bedbrook, J.R., Smith, S.M. and Ellis, R.J. (1980) Nature 287, 692–697.
 8. Berry-Lowe, S.L., McKnight, T.D., Shah, D.M. and Meagher, R.B. (1982)
     J. Mol. Appl. Genet. 1, 483–498.
 9. Broglie, R., Coruzzi, G., Lamppa, G., Keith, B. and Chua, N.-H. (1983)
     Biotechnology 1, 55–61.
10. Bedbrook, J.R., Coen, D.M., Beaton, A.R., Bogorad, L. and Rich, A.
     (1979) J. Biol. Chem. 254, 905–910.
11. Blair, G.E. and Ellis, R.J. (1973) Biochim. Biophys. Acta 319, 223–234.
12. Kawashima, N. and Wildman, S.G. (1972) Biochim. Biophys. Acta 262,
     42–49.
13. Highfield, P.E. and Ellis, R.J. (1978) Nature 271, 420–424.
14. Chua, N.-H. and Schmidt, G.W. (1978) Proc. Nat. Acad. Sci. USA 75,
     6110–6114.
15. Cashmore, A.R., Broadhurst, M.K. and Gray, R.E. (1978) Proc. Nat. Acad.
     Sci. USA 75, 655–659.
16. Speirs, J. and Brady, C.J. (1981) Aust. J. Plant Physiol. 8, 603–618.
17. Chua, N.-H. and Schmidt, G.W. (1979) J. Cell Biol. 81, 461–483.
18. Smith, S.M. and Ellis, R.J. (1979) Nature 278, 662–664.
19. Huber, S.C., Hall, T.C. and Edwards, G.E. (1976) Plant Physiol. 57,
     730–733.
20. Link, G., Coen, D.M. and Bogorad, L. (1978) Cell 15, 725–731.

21. Smith, S.M. and Ellis, R.J. (1981) J. Mol. Appl. Genet. 1, 127-137.
22. Shinzaki, K., Sasaki, Y., Sakibama, T. and Kamikubo, T. (1982) FEBS Lett. 144, 73-76.
23. Bowman, C.M., Koller, B., Delius, H. and Dyer, T.A. (1981) Mol. Gen. Genet. 183, 93–101.
24. Coruzzi, G., Broglie, R., Cashmore, A. and Chua, N.-H. (1983) J. Biol. Chem. 258, 1399-1402.
25. Brady, C.J. and Scott, N.S. (1977) Aust. J. Plant Physiol. 4, 327–335.
26. Aviv, H. and Leder, P. (1972) Proc. Nat. Acad. Sci. USA 69, 1408-1412.
27. Smith, D.F., Searle, P.F. and Williams, J.G. (1979) Nucl. Acids Res. 6, 487-506.
28. Bolivar, F., Rodriguez, R.L., Greene, P.L., Betlach, M.C., Heyneker, H.K. and Boyer, H.W. (1977) Gene 2, 95–113.
29. Daghert, M. and Ehrlich, S.D. (1979) Gene 6, 23-28.
30. Maloy, S.R. and Nunn, W.D. (1981) J. Bact. 145, 1110-1112.
31. Grunstein, M. and Hogness, D.S. (1975) Proc. Nat. Acad. Sci. USA 72, 3961-3965.
32. Denhardt, D. (1966) Biochem. Biophys. Res. Comm. 23, 641-652.
33. Clewell, D.B. (1972) J. Bact. 110, 667-676.
34. Clewell, D.B. and Helinski, D.R. (1969) Proc. Nat. Acad. Sci. USA 62, 1159-1166.
35. Maxam, A. and Gilbert, W. (1980) in Methods in Enzymology, Grossman, L. and Moldave, K. Eds., 18, pp. 499-560, Academic Press, New York.
36. Wu, R. (1970) J. Mol. Biol. 51, 501-521.
37. Dingman, C.W. and Peacock, A.C. (1968) Biochemistry 7, 659-668.
38. Rubin, C.M. and Schmid, C.W. (1980) Nucl. Acids Res. 8, 4613–4619.
39. Tabak, H.F. and Flavell, R.A. (1978) Nucl. Acids Res. 5, 2321-2332.
40. Rigby, P.W.J., Dieckman, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
41. Bailey, J.M. and Davidson, N. (1976) Anal. Biochem. 70, 75–85.
42. Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) Proc. Nat. Acad. Sci. USA 74, 4835-4838.
43. Wahl, G.M., Stern, M. and Stark, G.R. (1979) Proc. Nat. Acad. Sci. USA 76, 3683-3687.
44. Laskey, R.A. and Mills, A.D. (1977) FEBS Lett. 82, 314-316.
45. Martin, P.G. (1979) Aust. J. Plant Physiol. 6, 401–408.
46. Haslett, B.G., Yarwood, A., Evans, I.M. and Boulter, D. (1976) Biochim. Biophys. Acta 420, 122-132.
47. Poulsen, C., Strobaek, S. and Haslett, B.G. (1976) in Genetics and Biogenesis of Chloroplasts and Mitochondria, Bucher, T.L. Ed., pp. 17-24, Elsevier, Amsterdam.
48. Feldman, M. and Sears, E.R. (1981) Sci. Amer. 244 (1), 102–112.
49. Dunsmuir, P., Smith, S. and Bedbrook, J. (1983) Nucl. Acids Res. 11, 4177-4183.
50. O'Connell, P.B.H. and Brady, C.J. (1981) Biochim. Biophys. Acta 670, 355–361.
51. Takruri, I.A.H., Boulter, D. and Ellis, R.J. (1981) Phytochem. 20, 413–415.
52. Cheng, C.C. (1976) J. Mol. Biol. 107, 527-547.
53. Proudfoot, N. (1982) Nature 298, 516-517.
54. Chua, N.-H. and Schmidt, G.W. (1978) in Photosynthetic Carbon Assimilation, Siegelman, H.W. and Hind, G. Eds., pp. 325–347, Plenum, New York and London.