

Alternative mRNA Splicing Generates the Two Ribulosebisphosphate Carboxylase/Oxygenase Activase Polypeptides in Spinach and *Arabidopsis*

Jeffrey M. Werneke, J. Mark Chatfield,¹ and William L. Ogren²

Agricultural Research Service, United States Department of Agriculture, 1102 South Goodwin Avenue, Urbana, Illinois 61801

Sequence analysis of ribulosebisphosphate carboxylase/oxygenase (rubisco) activase cDNA and genomic clones isolated from spinach and *Arabidopsis thaliana* indicates that the two polypeptides of rubisco activase arise from alternative splicing of a common pre-mRNA. In spinach, two 5' splice sites are used in processing a single 137-nucleotide intron near the 3' end of the primary transcript. This intron was either removed completely or, alternatively, the first 22 nucleotides of the intervening sequence were retained in the mature rubisco activase mRNA. The 22-nucleotide auxiliary exon contains an in-frame ochre termination codon and leads to the synthesis of a 41-kilodalton polypeptide. Removal of the entire 137-nucleotide intervening sequence results in the synthesis of a larger 45-kilodalton polypeptide. Thus, alternative splicing of the spinach rubisco activase mRNA results in the synthesis of two polypeptides that are identical except for 37 additional amino acids at the C terminus of the 45-kilodalton polypeptide. This conclusion was confirmed by Cleveland peptide mapping and by N-terminal and C-terminal amino acid sequence analyses of both purified polypeptides. This method of producing the two rubisco activase polypeptides may be an evolutionarily conserved feature in higher plants because a nearly identical process occurs in the production of the two rubisco activase polypeptides in *Arabidopsis*. In *Arabidopsis*, an alternatively spliced intron resides at precisely the same position as the alternatively spliced intron in spinach and results in the synthesis of 44-kilodalton and 47-kilodalton rubisco activase polypeptides. In contrast to spinach, however, the retained portion of the intervening sequence does not contain an in-frame termination codon. Rather, a shift in reading frame leads to termination of translation of the smaller polypeptide within the coding region of the larger polypeptide.

INTRODUCTION

Ribulosebisphosphate carboxylase/oxygenase (rubisco) initiates the pathways of photosynthetic carbon reduction and photorespiratory carbon oxidation by catalyzing either the carboxylation or oxygenation of ribulosebisphosphate. Rubisco must be activated before it exhibits catalytic activity, and in situ this activation requires the ATP-dependent enzyme rubisco activase (Salvucci, Portis, and Ogren, 1985, Streusand and Portis, 1987). Activation by rubisco activase involves the formation of a carbamate (Werneke, Chatfield, and Ogren, 1988), presumably on lysine 201 of the rubisco large subunit (Lorimer and Miziorko, 1980).

All plant species examined contain rubisco activase (Salvucci et al., 1987), a soluble, nuclear-encoded chloroplast enzyme. The protein is abundant, accounting for about 2% of the total soluble leaf protein in spinach (Ro-

binson et al., 1989). Rubisco activase comprises two polypeptides of approximately 41 kD and 45 kD in most plant species, including spinach and *Arabidopsis thaliana*. In maize leaves, however, immunoblot analyses have indicated the presence of only the smaller polypeptide (Salvucci et al., 1987). Both rubisco activase polypeptides were absent in an *Arabidopsis* mutant (*rca*), which required an elevated CO₂ concentration for growth (Salvucci, Portis, and Ogren, 1985). The *rca* mutant was unable to activate rubisco at standard atmospheric concentrations of CO₂ (Somerville, Portis, and Ogren, 1982) and led to the discovery of rubisco activase activity (Salvucci, Portis, and Ogren, 1985).

Spinach rubisco activase cDNAs have been cloned, sequenced, and expressed in *Escherichia coli* (Werneke, Zielinski, and Ogren, 1988). The deduced amino acid sequence of the cDNA indicated the presence of two conserved nucleotide binding regions (Werneke, Zielinski, and Ogren, 1988), which is consistent with the observation that ATP is required for activity. The recombinant protein

¹ Current address: Agricultural Research Service, United States Department of Agriculture, 3420 NW Orchard Avenue, Corvallis, OR 97330.

² To whom correspondence should be addressed.

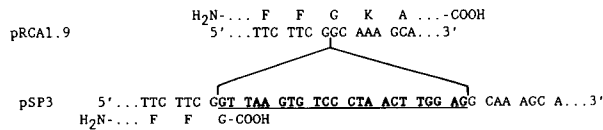


Figure 1. Alignment of Spinach Rubisco Activase cDNAs pRCA1.9 and pSP3 and Partial Nucleotide and Amino Acid Sequences.

The sequences shown begin at a position corresponding to nucleotide 1540 of pRCA1.9 (Werneke, Zielinski, and Ogren, 1988). pSP3 represents a subclass of rubisco activase mRNA containing a 22-nucleotide insertion (underlined) possessing an ochre codon that leads to early termination of translation relative to pRCA1.9.

expressed in *E. coli* consisted of a single polypeptide that catalyzed the activation of rubisco in an ATP-dependent reaction in vitro (Werneke, Chatfield, and Ogren, 1988), providing further evidence that the cDNA encoded rubisco activase. As noted above, however, rubisco activase comprises two polypeptides in vivo. Antibodies bound to and subsequently eluted from either the 41-kD or 45-kD spinach polypeptide reacted strongly with both polypeptides on immunoblots (Werneke, Zielinski, and Ogren, 1988). This cross-reactivity indicated that the two polypeptides were very similar, perhaps the products of two separate, but related, genes. However, genomic DNA gel blot analysis indicated that rubisco activase was encoded by a single-copy gene in both spinach and *Arabidopsis* (Werneke, Zielinski, and Ogren, 1988). The observations reported here have resolved this conundrum by demonstrating that the two polypeptides arise from alternative splicing of rubisco activase transcripts.

RESULTS

Isolation of Spinach Rubisco Activase cDNA Clones

A 1.6-kb spinach rubisco activase cDNA (pRCA1.6) was used to screen a spinach leaf cDNA library cloned in λ gt11 (Werneke, Zielinski, and Ogren, 1988). A full-length clone was recovered and designated pSP3. This rubisco activase cDNA possessed a 3'-untranslated region identical to the 1.9-kb cDNA (pRCA1.9) previously described (Werneke, Zielinski, and Ogren, 1988), indicating that both cDNAs represented mRNAs transcribed from a common gene. However, pSP3 differed from pRCA1.9 within the protein coding region by the presence of a 22-bp insertion at a position corresponding to nucleotide 1547 in pRCA1.9, as represented in Figure 1. The cDNA sequence of pSP3, beginning at a position corresponding with nucleotide 1548 in pRCA1.9, resumed complete nucleotide homology im-

mediately following the 22-bp insertion. The insertion in pSP3 contains an ochre translation termination signal in the first complete codon within the 22-bp sequence, and thus encodes a rubisco activase polypeptide shorter than that encoded by pRCA1.9. No other differences between these cDNA isolates were found to occur within the coding region. However, although the nucleotide sequences of pSP3 and pRCA1.9 were identical for the first 86 nucleotides upstream of a common ATG translation start codon, the 5'-untranslated regions of two clones showed no further homology. Examination of spinach genomic clones in this region indicates that the sequences 5' from this point are artifactual and that the 5'-untranslated region of the transcript is either 100 or 101 nucleotides in length, as determined by primer extension analysis.

Isolation of *Arabidopsis* Rubisco Activase cDNA Clones

Full-length cDNA clones encoding rubisco activase were recovered from an *Arabidopsis* leaf cDNA library constructed in λ gt10 (Werneke and Ogren, 1989) on the basis of hybridization to a 530-bp *Arabidopsis* rubisco activase cDNA clone (pARAB5) (Werneke, Zielinski, and Ogren, 1988). DNA sequence analysis of several clones indicated that there were also two classes of *Arabidopsis* rubisco activase cDNA. These cDNAs differed near the 3' end of the coding region by the presence or absence of the 11-bp sequence shown in Figure 2. Alignment of the spinach and *Arabidopsis* cDNA sequences indicated that the 11-bp insertion in the *Arabidopsis* clone occurred in the analogous codon, between the same nucleotides as the 22-bp insertion in the spinach clone. In contrast to the spinach clone, this 11-bp sequence caused a shift in the reading frame, leading to early termination of translation. It is unlikely that these differing cDNAs were derived from mRNAs transcribed from separate, but related, genes due to the absolute homology of their 3'-untranslated regions up to the sites of polyadenylation (data not shown). How-

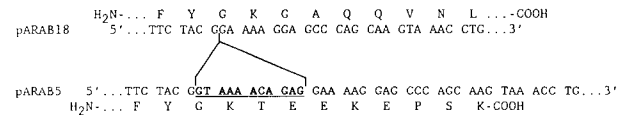


Figure 2. Alignment of *Arabidopsis* Rubisco Activase cDNAs pARAB18 and pARAB5 and Partial Nucleotide and Amino Acid Sequences.

The sequences shown begin at a position corresponding to nucleotide 1369 of pARAB18 (Werneke and Ogren, 1989). pARAB5 represents a subclass of rubisco activase mRNA containing an 11-nucleotide insertion (underlined) that results in a frameshift leading to early termination of translation, relative to pARAB18.

pARAB18 5'...TTTAAATTCGCTTGTGATATAAAAAAAAAAAAAAAAAA(ggaattcc)

pARAB5 5'...TTTAAATTCGCTTGTGATATAACTACACTTCTATAAAA(ggaattcc)

Figure 3. Alternative Polyadenylation Sites at the 3' End of Two *Arabidopsis* Rubisco Activase cDNAs.

Comparison of the 3' ends of several rubisco activase cDNAs indicated that at least two alternative polyadenylation sites are utilized. The positions of the EcoRI linkers used to construct the cDNA library are in lower case letters.

ever, at least two alternative polyadenylation sites were utilized in *Arabidopsis* rubisco activase mRNAs, as illustrated in Figure 3. Of five cDNAs sequenced through this region, use of the first or second polyadenylation site coincided with the absence or presence of the 11-bp sequence, respectively (not shown).

Isolation of Spinach Rubisco Activase Genomic Clones

Spinach genomic DNA was digested to completion with EcoRI and size-fractionated by electrophoresis, and 1.6-kb to 2.2-kb DNA fragments were cloned into λ gt11. Phage containing the spinach rubisco activase gene were then isolated by hybridization to pRCA1.9. The entire protein coding region of the spinach rubisco activase gene, shown in Figure 4, was found to reside on two EcoRI fragments of approximately 1.8 kb and 2.1 kb in length.

The genomic clones were completely sequenced and, by comparison with cDNA sequences, were found to contain six introns (Figure 4). Consensus G box and L box regions upstream of rubisco small subunit genes, implicated in the regulation of rubisco small subunit expression (Giuliano et al., 1988), were also identified upstream of the rubisco activase gene (Figure 4). These results are consistent with previous experiments demonstrating the expression of rubisco activase mRNAs and polypeptides in response to light (Zielinski, Werneke, and Jenkins, 1989).

One of the introns (IVS 6) occurred at the same position as the 22-bp insertion in pSP3 (Figure 1). A 22-bp sequence at the 5' end of the intron was identical to the 22-nucleotide insertion in pSP3. Thus, the two forms of rubisco activase mRNA can be generated from the single gene by splicing IVS 6 from alternative 5' sites to a common 3' site, as depicted in Figure 5. When the first 5' splice site is used, the entire intervening sequence is removed, and mRNA, represented by cDNA clones such as pRCA1.9, is produced. If the second 5' splice site is utilized, the 22-bp sequence containing the stop codon is

retained, generating an mRNA population represented by cDNA clones such as pSP3.

The 22-bp auxiliary exon (auxon) present in pSP3 contained an ochre translation termination signal in the first full in-frame codon. After import into the chloroplast, the derived molecular mass of the protein encoded by this cDNA is about 41 kD, the same apparent mass as the smaller rubisco activase polypeptide (Salvucci et al., 1987; Werneke, Zielinski, and Ogren, 1988). In clones without the auxon, such as pRCA1.9, translation proceeds for an additional 37 amino acids to produce a polypeptide with a deduced molecular mass of about 45 kD, similar to the observed mass of the larger of the two rubisco activase polypeptides.



Figure 4. Nucleotide Sequence of the Spinach Rubisco Activase Gene.

The bold upper case nucleotides represent the protein coding regions, and lower case nucleotides represent noncoding regions. Introns are lower case and underlined. Transcription begins at the C residue designated +1. The gene contains 7 exons, 6 introns, and the auxiliary exon (auxon) at +2436 to +2457 (between asterisks). Putative TATA and CAAT boxes (-30 to -34 and -64 to -68, respectively), consensus G box (-121 to -103 and -142 to -128) and L box (-201 to -185) regions upstream of the coding region are also underlined. The derived amino acid sequence of spinach rubisco activase is given in Werneke, Zielinski, and Ogren (1988).

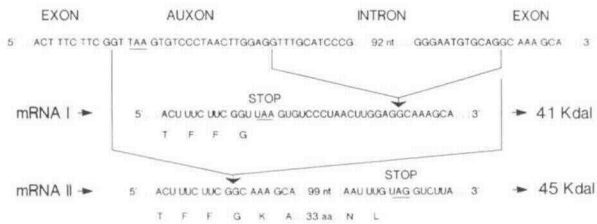


Figure 5. Alternative Splicing of IVS 6 of Spinach Rubisco Activase mRNA.

The top portion shows the sequence of IVS 6 and flanking regions. Shown below are the two types of mRNA produced by alternative splicing. When the second 5' splice site is utilized, the 22-nucleotide auxon is retained to produce mRNA I (center). The auxon contains an ochre termination signal (UAA) in the first codon within the sequence (underlined). Thus, when the auxon is retained, an mRNA is produced that encodes the 41-kD polypeptide. If the first 5' splice site is utilized, the auxon is absent and mRNA II is produced (bottom), so translation continues for an additional 37 amino acids, producing a 45-kD polypeptide.

Identification of the Spinach Rubisco Activase Transcription Start Site

The nucleotide sequences of cDNA clones pRCA1.9 and SP3 were identical for the first 86 nucleotides upstream of the translation start codon, but sequences 5' to this point diverged considerably. The authentic 5' end of the spinach rubisco activase transcript was identified by primer extension analysis using a ^{32}P -labeled synthetic oligonucleotide primer complementary to a nucleotide sequence within the coding region of the transit peptide. After hybridization of the primer to rubisco activase mRNA, first-strand cDNA was synthesized with reverse transcriptase. The transcription initiation site was mapped by subsequent electrophoresis through an acrylamide gel in a lane adjacent to dideoxy DNA sequencing reactions performed using the same primer and the genomic clone as a template, as shown in Figure 6. The 5' ends of the mRNAs were mapped to the C and A residues 30 and 31 nucleotides, respectively, downstream of the putative TATA box (Figure 4). The C residue has been designated +1 in this transcript (Figure 4). A small number of transcripts with 5' ends mapping to a point 28 nucleotides downstream of the TATA box were also observed (Figure 6). The detailed structure of the spinach rubisco activase genomic clone is shown in Figure 7.

Isolation of *Arabidopsis* Rubisco Activase Genomic Clones

A rubisco activase genomic clone was isolated from an *Arabidopsis* EMBL4 genomic library by hybridization to

pARAB5. This clone was partially sequenced and, by comparison with *Arabidopsis* cDNA clones, found to contain several introns. One of these introns occurred at a position corresponding to the 11-bp insertion in the *Arabidopsis* rubisco activase cDNAs (Figure 2). An identical 11-bp sequence was found to be present at the 5' end of the intron. Examination of the DNA sequence of this intron and flanking regions suggested that either of two 5' splice sites (separated by the 11-bp sequence) is utilized during splicing of the rubisco activase transcript. Thus, as illustrated in Figure 8, when the intervening sequence is completely removed, an mRNA encoding a 47-kD polypeptide

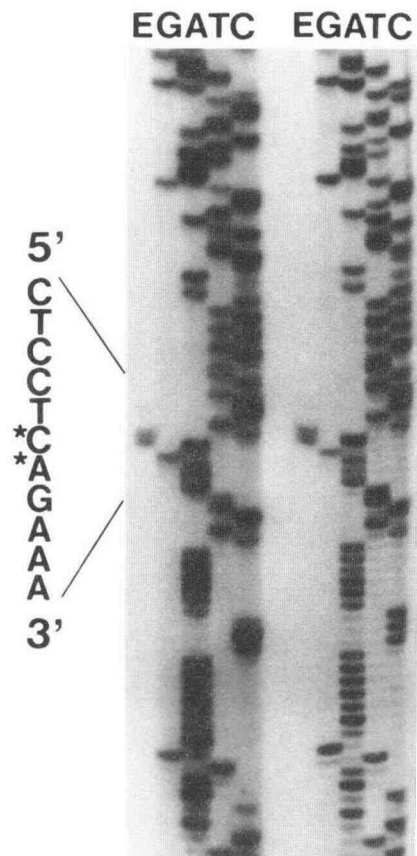


Figure 6. Primer Extension Analysis of Rubisco Activase mRNA.

The 5' end of spinach rubisco activase mRNA was mapped by primer extension using a ^{32}P end-labeled 33-mer. The products of the primer extension reaction (E) were electrophoresed adjacent to dideoxy DNA sequencing reactions (GATC) using either the ^{32}P -labeled 33-mer and *Taq* polymerase (left), or unlabeled 33-mer and ^{35}S -dATP (right). The 5' ends of most of the transcripts mapped to the C and A residues 30 and 31 nucleotides downstream of the TATA box, respectively (*). A small number of transcripts mapped to the C residue 28 nucleotides downstream of the TATA box.

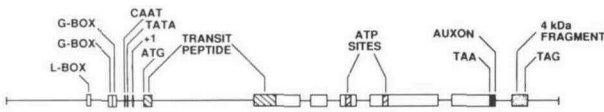


Figure 7. Schematic Map of the Spinach Rubisco Activase Genomic Clone.

For distances in kilobases, see Figure 4.

is produced. However, if splicing occurs at the alternative site, a population of rubisco activase mRNA that retains an 11-bp portion of the intervening sequence is generated. This 11-bp sequence changes the reading frame, and, as a consequence of the frame shift, termination of translation occurs prematurely to produce a shorter polypeptide of about 44 kD.

Cleveland Analysis

The alternative splicing model for rubisco activase suggested that the 41-kD and 45-kD polypeptides from spinach were identical from the N terminus to glycine 372 (the C terminus of the 41-kD polypeptide), and that the 45-kD species possessed an additional 37 amino acids at the C-terminal end. This hypothesis was supported by the observation that the two polypeptides were immunologically related (Werneke, Zielinski, and Ogren, 1988). To confirm the near identity of the two isoforms, each polypeptide was partially proteolyzed with chymotrypsin and then electrophoresed in a linear 10% to 18% acrylamide gel, according to the method described by Cleveland et al. (1977). Partial digestion and electrophoresis of purified 41-kD or 45-kD rubisco activase proteins resulted in nearly identical banding patterns, shown in Figure 9, a finding that is consistent with the conclusion that the two polypeptides arise from alternative splicing.

Amino-Terminal Sequence Analysis

To evaluate further the postulated close relationship between the two polypeptides, the first 20 amino acid residues of each were determined by Edman degradation. As shown in Table 1, both the 41-kD and the 45-kD species possessed identical N termini. These results are also consistent with the hypothesis that alternative mRNA splicing is responsible for the generation of the two forms of rubisco activase.

Carboxy-Terminal Sequence Analysis

The alternative splicing model predicted that the two forms of rubisco activase differed only at their C-terminal ends.

Purified 41-kD and 45-kD polypeptides were incubated with carboxypeptidase Y to release amino acids sequentially from the C termini. The amino acid sequence derived from cDNA clone pRCA1.9 indicated that the first C-terminal amino acid released from the 45-kD polypeptide should be a leucine residue (Figure 5) and, as shown in Figure 10, leucine was the first amino acid released by carboxypeptidase treatment of 45-kD spinach rubisco activase. Similar analysis of the purified 41-kD protein, shown in Figure 10, yielded glycine and then phenylalanine, in agreement with the C-terminal amino acids encoded by pSP3 (Figure 5).

Immunoblots of Spinach Leaf Sections

Spinach leaves were frozen and paradermal sections were removed with a freezing microtome. The thickness of the spinach leaves was estimated to be about 600 μm, with a palisade to spongy mesophyll transition occurring about halfway through the leaf (Terashima and Inoue, 1985). Sections 16 μm thick were made from both leaf surfaces to a depth of about 160 μm to ensure that only palisade or spongy mesophyll cells were collected. As shown in Figure 11, all sections contained equal ratios of the two rubisco activase polypeptides. These results suggest that both rubisco activase polypeptides are present in both

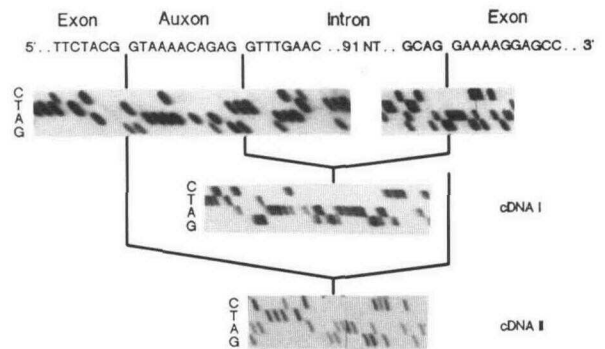


Figure 8. Alternative Splicing of *Arabidopsis* Rubisco Activase mRNA.

Alternative splicing of *Arabidopsis* rubisco activase mRNA proceeds in a manner very similar to that of spinach. Regions of the genomic sequence (top) are given as deduced from autoradiograms of sequencing reactions (loaded CTAG from top to bottom). Shown below the genomic data are autoradiograms of sequencing reactions (loaded CTAG from top to bottom) of the corresponding region of the two types of cDNA. When the second 5' splice site is utilized, the 11-bp auxon is retained (center). The presence of this auxon alters the reading frame and leads to early termination of translation and thus the 44-kD polypeptide. When the auxon is completely removed (bottom), translation continues for an additional 28 amino acids to produce the 47-kD polypeptide.

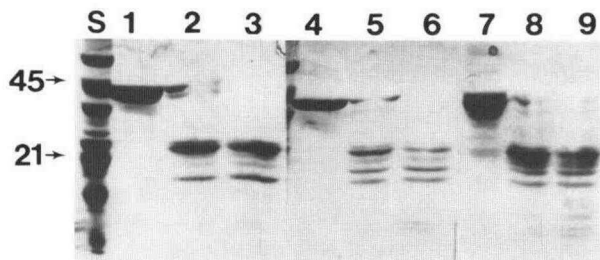


Figure 9. Peptide Mapping of Rubisco Activase Polypeptides by Cleveland Analysis.

The two spinach rubisco activase polypeptides were purified as described in Methods and subjected to partial proteolysis with chymotrypsin. Lane S, protein standards and molecular mass (kD); lane 1, purified 45-kD polypeptide; lanes 2 and 3, digested 45-kD polypeptide; lane 4, purified 41-kD polypeptide; lanes 5 and 6, digested 41-kD polypeptide; lane 7, mixed 45-kD and 41-kD polypeptides; lanes 8 and 9, mixed 45-kD and 41-kD polypeptides after partial proteolysis.

major cell types in the spinach leaf, and indicate that alternative splicing of rubisco activase mRNA may be a constitutive process in the leaf.

DISCUSSION

The nucleotide sequence of a spinach rubisco activase cDNA clone differed from the sequence of a spinach cDNA previously reported (Werneke, Zielinski, and Ogren, 1988) by the presence of a 22-bp insertion near the 3' end of the coding region. The nucleotide sequence of a spinach rubisco activase genomic clone indicated that this insertion occurred at an intron splice junction, and that the presence or absence of this sequence in the rubisco activase cDNAs was determined by alternative mRNA processing. The 22-bp auxon is retained as part of the mature rubisco activase message in some mRNAs and is removed as part of a larger intron in other rubisco activase mRNAs. The auxon contains an ochre codon that leads to premature termination of translation with respect to rubisco activase mRNAs from which the auxon has been removed with the adjacent intron. By this mechanism, both rubisco activase polypeptides are produced from the same single-copy gene.

All C_3 and C_4 plants examined produced polypeptides that cross-reacted with anti-spinach rubisco activase antibodies (Salvucci et al., 1987). To investigate the possibility that alternative mRNA splicing was responsible for the presence of two rubisco activase polypeptides in species other than spinach, we cloned and sequenced several *Arabidopsis* cDNAs encoding rubisco activase. As in spinach, two classes of cDNA were recovered, one containing

an insertion leading to early termination of translation. In contrast to spinach, the insertion in *Arabidopsis* was only 11 nucleotides in length and did not contain an in-frame termination codon. Rather, the insertion created a reading-frame shift that caused early termination of translation with respect to the cDNAs that did not contain this inserted sequence.

Upon sequencing genomic rubisco activase clones from *Arabidopsis*, we determined that, as in spinach, the insertion was an auxon that could either be removed or retained depending upon 5' splice site selection in an intron toward the 3' end of the coding region of the rubisco activase gene. Additionally, at least two alternative polyadenylation sites were identified and, of the five cDNAs sequenced, these alternative polyadenylation sites appeared to coincide with the presence or absence of the auxon (data not shown). The results indicate that the site of poly(A) addition may influence or result from alternative splice site selection in the rubisco activase transcript. The site of poly(A) addition has been shown to influence splice site choice in the alternative splicing of IgM heavy-chain RNAs (Galli et al., 1988).

The consequence of alternative splicing of the primary rubisco activase transcript is that two polypeptides differing at the C-terminal end of the enzyme are synthesized in the leaf. To demonstrate that alternative splicing was the source of the two rubisco activase polypeptides in

Table 1. Amino-Terminal Sequence Analysis of the Two Spinach Rubisco Activase Polypeptides as Determined by 20 Cycles of Edman Degradation

Cycle	41-kD Polypeptide		45-kD Polypeptide	
	Amino Acid	Amount	Amino Acid	Amount
		pmol		pmol
1	A	72	A	113
2	E	43	E	20
3	N	52	N	18
4	E	39	E	19
5	E	44	E	14
6	K	30	K	20
7	N	26	N	18
8	T	8	T	15
9	D	27	D	12
10	K	16	K	17
11	W	12	?	—
12	A	19	A	27
13	H	2	?	—
14	L	19	L	20
15	A	17	A	25
16	K	10	K	12
17	D	17	D	10
18	F	12	F	12
19	S	4	?	—
20	?	—	V	12

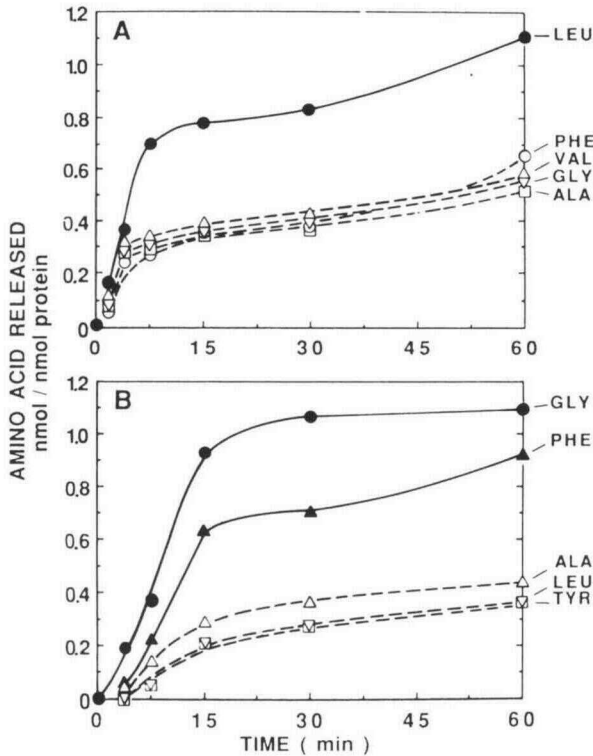


Figure 10. Carboxy-Terminal Amino Acid Analysis of Rubisco Activase Polypeptides.

(A) Time course of carboxypeptidase Y-dependent release of amino acids from the 45-kD polypeptide.

(B) Time course of carboxypeptidase Y-dependent release of amino acids from the 41-kD polypeptide.

vivo, we isolated the 41-kD and 45-kD polypeptides and incubated them separately with carboxypeptidase Y. The amino acids released from the C terminus were found to differ between the two polypeptides. The first amino acid residue released from the C-terminal of the 41-kD polypeptide was phenylalanine, and the C-terminal amino acid from the 45-kD polypeptide was leucine. These C-terminal amino acids are the same as those predicted from the alternative splicing model. Edman degradation of the two polypeptides indicated that the two polypeptides were identical at the N terminus. Thus, the two polypeptides share the same transit peptide and are most likely identical except for the additional 37 amino acids at the C terminus of the longer polypeptide.

Purified spinach rubisco activase appears to be a hexamer, with a molecular mass of about 250 kD (J.M. Chatfield and W.L. Ogren, unpublished results), as determined by Ferguson-plot analysis (Rodbard and Chrambach, 1971). It is not yet known whether the holoenzyme consists of one subunit type or a combination of the two

subunits in vivo. It is evident that activity does not require the presence of both subunit types because the two polypeptides have been separated by chromatography under native conditions (Werneke, Chatfield, and Ogren, 1988), and both purified polypeptides catalyzed the activation of rubisco in vitro independently (J.M. Chatfield, J.M. Werneke, and W.L. Ogren, unpublished results). Additionally, pRCA1.6, a clone encoding the larger polypeptide with several extra amino acids at the N-terminal end of the protein, yielded a polypeptide that catalyzed rubisco activation when expressed in *E. coli* (Werneke, Chatfield, and Ogren, 1988). Furthermore, the relative quantities of the two polypeptides in whole leaf extracts vary between species (Salvucci et al., 1987), indicating that the ratio of the two polypeptides in any presumptive holoenzyme is not a constant.

Alternative mRNA splicing amplifies genetic diversity by producing two or more forms of a protein from a single gene. The majority of examples of alternative mRNA splicing in animals exhibit developmental or tissue-specific regulation. These altered proteins generally fulfill identical functions and occur in response to unidentified requirements of development or tissue function (Breitbart, Andreadis, and Nadal-Ginard, 1987). Following the precedent

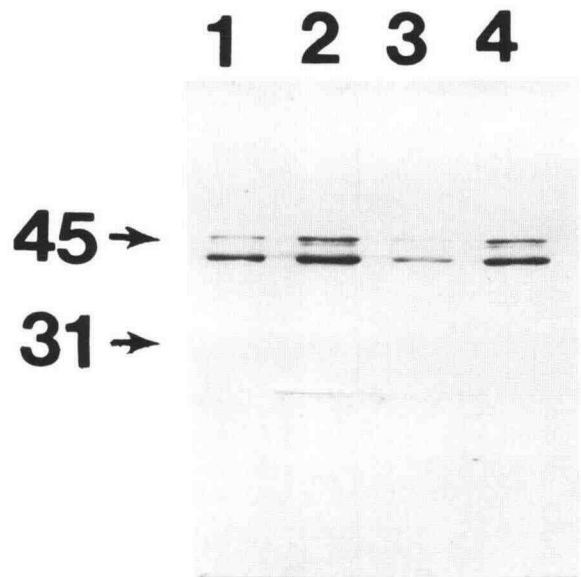


Figure 11. Immunoblot of Palisade and Spongy Mesophyll Polypeptides.

Protein isolated from either palisade or spongy mesophyll cells were fractionated by electrophoresis, transferred to nitrocellulose, and probed with anti-spinach rubisco activase antisera. Lane 1, palisade layer from 0 μm to 50 μm ; 2, palisade layer from 50 μm to 100 μm ; 3, spongy mesophyll layer from 0 μm to 50 μm ; 4, spongy mesophyll layer from 50 μm to 100 μm . Both the 45-kD and 41-kD polypeptides were present in equal ratios in all samples.

provided by examples of alternative splicing in animal systems, it seemed possible that factors such as cell age or cell type might influence the splice site selection in the rubisco activase transcripts, and thus alter relative amounts of the two polypeptides. To investigate the possibility of a spatial separation of expression of the polypeptides within the leaf, we obtained leaf samples greatly enriched in either palisade or mesophyll cells by slicing thin sections from the upper or lower surface of spinach leaves with a freezing microtome. Immunological analysis of extracts of these cells suggested that both rubisco activase polypeptides were synthesized throughout the leaf in equal quantities. The high degree of similarity between the two mRNAs and proteins may make more detailed evaluations of tissue-specific expression difficult in spinach or *Arabidopsis*.

The expression of rubisco activase mRNAs and polypeptides during greening and leaf development has recently been investigated in barley (Zielinski, Werneke, and Jenkins, 1989). Antisera directed against the spinach enzyme were used to probe immunoblots of soluble polypeptides isolated from the developmental gradient of cells along the axis of the barley leaf and from etiolated plants exposed to white light. In this system, the relative amounts of the two rubisco activase polypeptides did not appear to be regulated in a developmental or light-dependent fashion. It should be noted, however, that alternative splicing of rubisco activase transcripts has not yet been demonstrated in barley. Although these results suggest that both forms of the enzyme were constitutively expressed in all photosynthetic cells and were subject to the same light and developmental regulation, they do not preclude the possibility of a higher order of tissue, cell, or developmentally specific expression of the two mRNAs and polypeptides.

The data presented here demonstrate the existence of alternative splicing in plant systems, but the physiological significance of synthesizing two forms of rubisco activase remains unclear. However, this process may have important implications in photosynthesis. If these polypeptides were functionally equivalent enzymes in the chloroplast, there would be no need for the production of both polypeptides, and alternative splicing of the rubisco activase mRNA would likely become a dispensable process. Furthermore, spinach and *Arabidopsis*, the only two species for which both cDNA and genomic clones are available, generate the two polypeptides by essentially identical alternative splicing mechanisms. Also, the amino acid sequences of the 4-kD C-terminal fragments of these two species have been highly conserved, (Werneke, Zielinski, and Ogren, 1988; Werneke and Ogren, 1989). In examining several species of plants, only maize appeared to produce a single rubisco activase polypeptide (Salvucci et al., 1987). It is reasonable to assume that the two rubisco activase polypeptides in other species are produced by alternative splicing as well. Taken together, these findings

suggest that the alternative splicing of this particular intron has been conserved and that dual forms of rubisco activase may impart an advantage or be essential to photosynthesis in higher plants.

Most theories on the role of introns in the evolution of protein structure suggest that introns provide the means for exon shuffling (Gilbert, 1985), and that the exon unit encodes either structural or functional protein domains. In this evolutionary scheme, introns are advantageous in that they facilitate the exchange of these exons to create new domains on existing proteins. There is some evidence that exon shuffling has occurred during the evolution of the rubisco activase gene in that the two conserved regions that compose the ATP-binding site are on separate exons (Figure 4). It is possible that the additional amino acids at the C-terminal end of the larger rubisco activase polypeptide constitute an additional structural or functional domain essential for the regulation of photosynthesis under as-yet-undetermined conditions.

METHODS

Plant Material

Spinach (*Spinacea oleracea* L., American Hybrid No. 424) was grown at 22°C under a normal atmosphere in hydroponic nutrient culture (Heldt and Sauer, 1971) with an 11-hr photoperiod. The plants were illuminated at 300 $\mu\text{mol photons/m}^2 \text{ s}$ with fluorescent and incandescent light sources. *Arabidopsis thaliana* L. Heynh. "Columbia" wild-type was grown in a soil-peat mixture in a growth chamber supplemented with 1% CO_2 (Salvucci, Portis, and Ogren, 1985).

Purification of Rubisco Activase

Rubisco activase was purified from spinach chloroplasts (Werneke, Chatfield, and Ogren, 1988) except that fractions eluting at 170 mM to 200 mM KCl were pooled to include all of the larger rubisco activase polypeptide. Enzyme purity was determined from SDS-PAGE after Coomassie Blue staining and immunoblotting (Salvucci et al., 1987). All rubisco activase preparations used in subsequent analyses were determined to be greater than 90% pure by scanning densitometry.

Amino-Terminal Sequence Analysis

After fractionation of proteins from spinach stroma, 100- μg aliquots were electrophoresed in 12.5% resolving gels (Laemmli, 1970). Polypeptides were recovered from acrylamide gels by electroblotting onto QAE-activated glass membranes and detected with UV light after staining with 3,3'-dipentylloxycarbocyanine iodide (Aebersold et al., 1986). Bands containing about 20 μg of protein were cut from the glass fiber sheets, and the amino-terminal sequence was determined by automated Edman degradation at the University of Illinois Biotechnology Center.

Carboxy-Terminal Sequence Analysis

Polypeptides used for carboxy-terminal sequence analysis were purified, by re-chromatography of rubisco activase obtained as described above, on a QAE-Sepharose column (1 × 10 cm) equilibrated in 20 mM bis-Tris/propane/HCl (pH 7), 4 mM 2-mercaptoethanol, and 8 M urea. Rubisco activase protein was eluted from the column with a linear NaCl gradient of 0 M to 0.5 M in 100 mL of equilibration buffer. Peak fractions eluting at about 80 mM NaCl (41-kD polypeptide) and 120 mM NaCl (45-kD polypeptide) were collected and concentrated to about 1 mg/mL by ultrafiltration. The samples were then dialyzed overnight against 4 L of 10 mM Mes-NaOH (pH 6.3) containing 0.2 mM EDTA.

Carboxypeptidase Y (Sigma) was prepared by dissolving in 10 mM Mes-NaOH (pH 6.3) and 0.2 mM EDTA, and dialyzing against 4 L of the same buffer for 4 hr. The protein concentration was adjusted to 0.1 mM and stored at -20°C. The rubisco activase polypeptides (1.4 mg, 35 nmol) were pre-incubated at 25°C in 2 mL of dissolution buffer containing 0.5% (w/v) SDS. Prior to the addition of carboxypeptidase Y, an aliquot was removed to serve as a blank. To initiate the digestion, 0.3 nmol of carboxypeptidase Y was added and the release of amino acid residues followed by removing aliquots from the reaction mixture at the indicated times. The digestions were stopped by the addition of 2 volumes of methanol and immersion into a boiling water bath for 5 min. The samples were then centrifuged to remove precipitated protein, concentrated to dryness, and analyzed by HPLC as PTH-amino acid derivatives using a Waters Picotag system at the University of Illinois Biotechnology Center.

Cleveland Gel Analysis

Rubisco activase polypeptides (10 µg), purified by the QAE-Sepharose/urea method described above, were subjected to partial proteolysis with 0.5 µg to 5.0 µg of chymotrypsin (Sigma) and fractionated on a linear 10% to 18% acrylamide gel according to Cleveland et al. (1977). The partially digested peptides were visualized with Coomassie Blue.

cDNA Synthesis

Double-stranded cDNA was synthesized from spinach and *Arabidopsis* poly(A)⁺ mRNA using a modification of a published procedure (Gubler and Hoffman, 1983), as described previously (Werneke, Zielinski, and Ogren, 1988). The cDNAs were cloned into the EcoRI site of λgt11 (Huynh, Young, and Davis, 1985), and packaged in vitro according to manufacturer's directions (Promega Biotec).

Identification of Rubisco Activase cDNAs

Spinach and *Arabidopsis* rubisco activase clones were recovered from the cDNA libraries by hybridization screening (Maniatis, Fritsch, and Sambrook, 1982). The spinach library was probed with radiolabeled pSP26 (Werneke, Zielinski, and Ogren, 1988), and the *Arabidopsis* library with radiolabeled pARAB18 (Werneke and Ogren, 1989). Hybridization reactions were carried out over-

night using 10⁵ cpm of probe DNA/mL of hybridization solution at 65°C.

Genomic DNA Extraction

DNA was isolated from either leaves or roots by grinding the tissue to a fine powder in liquid nitrogen and adding the material 1:2 (w/v) to an extraction buffer containing 25 mM Tris-HCl (pH 9.0), 12.5 mM EDTA, and 200 mM NaCl. An equal volume of phenol was added, and the mixture was stirred for 20 min. The aqueous phase was recovered by centrifugation and extracted with an equal volume of chloroform. Nucleic acids were precipitated from the aqueous phase by the addition of an equal volume of isopropanol, pelleted, and resuspended in water. NaCl was added to 200 mM, and the nucleic acids were precipitated a second time by the addition of 2 volumes of 100% ethanol. The nucleic acids were recovered by centrifugation and resuspended in 10 mM Tris-HCl (pH 8.0), and DNase-free RNase was added to a final concentration of 50 µg/mL. The DNA was concentrated by ethanol precipitation, resuspended in sterile water, and stored at -20°C.

Construction and Screening of Subgenomic Libraries

Spinach genomic DNA was digested to completion with EcoRI and fractionated by electrophoresis on a 5% acrylamide gel using bisacrylylcystamine (Bio-Rad) as a cross-linking agent (Hansen, 1986). After electrophoresis, the gel was stained briefly with 0.5% methylene blue. The EcoRI fragments from 1.5 kb to 1.7 kb and 2.0 kb to 2.2 kb in length were excised, and gel slices were placed in a vial containing about 10 volumes of 10 mM Tris-HCl (pH 8.0), 2 mM EDTA. The gel was dissolved by addition of β-mercaptoethanol to a final concentration of 100 mM, and DNA fragments were concentrated by passage over a NACS Prepac cartridge (Bethesda Research Laboratories), according to manufacturer's directions.

The size-fractionated genomic DNAs were cloned into the EcoRI site of λgt11 (Huynh, Young, and Davis, 1985), packaged as described above, and plated using *Escherichia coli* strain Y1088. Phage harboring the appropriate inserts were identified by hybridization (Maniatis, Fritsch, and Sambrook, 1982) to the full-length rubisco activase cDNA clone pSP26. Approximately 1 out of every 10,000 plaques screened contained the EcoRI fragment of interest.

The spinach rubisco activase gene contained an EcoRI site within the second intron. If there were a second EcoRI site within this intron, the resulting genomic DNA would not have been isolated because a cDNA clone was used to identify the gene fragments. However, we consider a second EcoRI site to be unlikely because analysis of *Arabidopsis* genomic DNA (B.L. Marcotte, J.M. Werneke, and W.L. Ogren, unpublished data) has shown that introns 2 to 6 are located within the same codons as in the spinach gene and are similar in size.

Identification of *Arabidopsis* Rubisco Activase Genomic Clones

An *Arabidopsis* EMBL4 genomic library, generously provided by Dr. E. Meyerowitz (California Institute of Technology), was

screened by hybridization (Maniatis, Fritsch, and Sambrook, 1982) using radiolabeled pARAB18 (Werneke and Ogren, 1989) as a probe. Inserts from phage-producing signals were subcloned into M13 and identified as rubisco activase clones by nucleotide sequence analysis.

Nucleotide Sequencing

The spinach and *Arabidopsis* rubisco activase cDNA and genomic clones were sequenced using the dideoxynucleotide chain-termination method of Sanger, Nicklen, and Coulson (1977) using either single-stranded M13 templates or double-stranded plasmid DNA with Sequenase (United States Biochemical Corp.), or *Taq* polymerase (Promega Biotec), according to manufacturer's directions.

Primer Extension Analysis

The synthetic oligonucleotide 5'-GTGGCAGCTCCAACGGTTCGAGACAGCAGTAGCC-3' was labeled at the 5' end with [γ -³²P] ATP and T4 polynucleotide kinase (Maxam and Gilbert, 1980). Approximately 0.1 pmol (3×10^6 cpm) of the 33-mer was hybridized with 50 μ g of total spinach leaf RNA for 18 hr at 48°C in 50 μ L of 400 mM NaCl, 40 mM Pipes-NaOH (pH 6.8), 1 mM EDTA, and 50% formamide. The nucleic acids were precipitated by the addition of 2.5 volumes of ethanol, collected by centrifugation, and dried under vacuum. The pellets were resuspended in 25 μ L of a buffer consisting of 50 mM Tris-HCl; 70 mM KCl; 6 mM MgCl₂; 4 mM DTT; 500 μ M each dATP, dCTP, dTTP, and dGTP; and 250 μ g/mL BSA, containing 30 units of RNA-guard (Promega Biotec) and 5 units of avian myeloblastosis virus reverse transcriptase (Bethesda Research Laboratories). The reaction mixture was incubated at 37°C for 1 hr, extracted with phenol, precipitated twice from ethanol, and analyzed on 8% polyacrylamide-8.3 M urea DNA sequencing gels (Maxam and Gilbert, 1980).

Cryosectioning and Immunoblotting of Spinach Leaf Segments

Spinach leaves were cut into 4-mm squares with a razor blade and embedded in methylcellulose (O.T.C. Compound, Miles Laboratories) by freezing. Paradermal sections 16 μ m thick were removed from either the upper or lower surface of the leaves with a freezing microtome. Two or three sections were collected at a time, combined in a chilled microcentrifuge tube, and kept frozen. After all of the tissue was collected, 100 μ L of 20 mM Tris-HCl (pH 8.0), 2 mM MgCl₂, and 2 mM DTT was added and mixed with the sample on ice. SDS-PAGE sample buffer was added, and the samples were heated to 99°C for 3 min prior to electrophoresis (Laemmli, 1970). After electrophoresis, proteins were transferred to nitrocellulose and probed with anti-spinach rubisco activase antisera (Salvucci et al., 1987).

ACKNOWLEDGMENTS

We thank E.M. Orozco, Jr. for providing guidance in the primer extension analysis, R.F.E. Crang for helpful suggestions in obtain-

ing leaf sections, and E. Meyerowitz for providing an *Arabidopsis* genomic library.

Mention of a trademark, proprietary product, or vendor does not constitute a guarantee or warranty of the product by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products or vendors that may also be suitable.

Received May 3, 1989; revised June 20, 1989.

REFERENCES

- Aebersold, R.H., Teplow, D.B., Hood, L.E., and Kent, B.H. (1986). Electroblooming onto activated glass. High efficiency preparation of proteins from analytical sodium dodecyl sulfate polyacrylamide gels for direct sequence analysis. *J. Biol. Chem.* **261**, 4229-4238.
- Breitbart, R.E., Andreadis, A., and Nadal-Ginard, B. (1987). Alternative splicing: A ubiquitous mechanism for the generation of multiple protein isoforms from single genes. *Annu. Rev. Biochem.* **56**, 467-495.
- Cleveland, D.W., Fischer, S.G., Kirschner, M.W., and Laemmli, U.K. (1977). Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *J. Biol. Chem.* **252**, 1102-1106.
- Galli, G., Guise, J., Tucker, P., and Nevins, J. (1988). Poly(A) site choice rather than splice site choice governs the regulated production of IgM heavy-chain mRNAs. *Proc. Natl. Acad. Sci. USA* **85**, 2439-2443.
- Gilbert, W. (1985). Genes-in-pieces revisited. *Science* **228**, 823-824.
- Giuliano, G., Pichersky, E., Malik, V.S., Timko, M.P., Scolnik, P.A., and Cashmore, A.R. (1988). An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. *Proc. Natl. Acad. Sci. USA* **85**, 7089-7093.
- Gubler, U., and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263-269.
- Hansen, J.N. (1986). Use of solubilizable acrylamide disulfide gels for isolation of DNA fragments suitable for sequence analysis. *Anal. Biochem.* **116**, 146-151.
- Heldt, H.W., and Sauer, F. (1971). The inner membrane of the chloroplast envelope as the site of specific metabolite transport. *Biochim. Biophys. Acta* **234**, 83-91.
- Huynh, T.V., Young, R.A., and Davis, R.W. (1985). Constructing and screening cDNA libraries in λ gt10 and λ gt11. In *DNA Cloning*, Vol I, D.M. Glover, ed (Oxford: IRL Press), pp. 39-78.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lorimer, G.H., and Miziorko, H.M. (1980). Carbamate formation on the ϵ -amino group of lysyl residue as the basis for the activation of ribulosebisphosphate carboxylase by CO₂ and Mg⁺⁺. *Biochemistry* **19**, 5321-5326.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

- Maxam, A.M., and Gilbert, W.** (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–559.
- Robinson, S.R., Streusand, V.J., Chatfield, J.M., and Portis, A.R., Jr.** (1989). Purification of rubisco activase from leaves. *Plant Physiol.* **88**, 1008–1014.
- Rodbard, D., and Chrambach, A.** (1971). Estimation of molecular radius, free mobility, and valence using polyacrylamide gel electrophoresis. *Anal. Biochem.* **40**, 95–134.
- Sanger, F., Nicklen, S., and Coulson, A.R.** (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Salvucci, M.E., Portis, A.R., Jr., and Ogren, W.L.** (1985). A soluble chloroplast protein catalyzes ribulosebisphosphate carboxylase/oxygenase activation in vivo. *Photosynth. Res.* **7**, 191–203.
- Salvucci, M.E., Werneke, J.M., Ogren, W.L., and Portis, A.R., Jr.** (1987). Purification and species distribution of rubisco activase. *Plant Physiol.* **84**, 381–387.
- Somerville, C.R., Portis, A.R., Jr., and Ogren, W.L.** (1982). A mutant of *Arabidopsis thaliana* which lacks activation of RuBP carboxylase in vivo. *Plant Physiol.* **70**, 381–387.
- Streusand, V.J., and Portis, A.R., Jr.** (1987). Rubisco activase mediates ATP-dependent activation of ribulosebisphosphate carboxylase. *Plant Physiol.* **85**, 152–154.
- Terashima, I., and Inoue, Y.** (1985). Vertical gradient in photosynthetic properties of spinach chloroplasts dependent on intra-leaf light environment. *Plant Cell. Physiol.* **26**, 781–785.
- Werneke, J.M., and Ogren, W.L.** (1989). Structure of an *Arabidopsis thaliana* cDNA encoding rubisco activase. *Nucleic Acids Res.* **17**, 2871.
- Werneke, J.M., Chatfield, J.M., and Ogren, W.L.** (1988). Catalysis of ribulosebisphosphate carboxylase/oxygenase activation by the product of a rubisco activase cDNA clone expressed in *Escherichia coli*. *Plant Physiol.* **87**, 917–920.
- Werneke, J.M., Zielinski, R.E., and Ogren, W.L.** (1988). Structure and expression of a spinach leaf cDNA encoding rubisco activase. *Proc. Natl. Acad. Sci. USA* **85**, 787–791.
- Zielinski, R.E., Werneke, J.M., and Jenkins, M.E.** (1989). Coordinate expression of rubisco activase and rubisco during barley leaf cell development. *Plant Physiol.* **90**, 516–521.