Biochem. J. (1997) **328**, 795–800 (Printed in Great Britain)

Alternative mRNA splicing of 3′-terminal exons generates ascorbate peroxidase isoenzymes in spinach (*Spinacia oleracea*) chloroplasts

Takahiro ISHIKAWA[1], Kazuya YOSHIMURA, Masahiro TAMOI, Toru TAKEDA and Shigeru SHIGEOKA[2]

Department of Food and Nutrition, Faculty of Agriculture, Kinki University, Nakamachi, Nara 631, Japan

We have isolated two cDNA clones encoding spinach (*Spinacia oleracea*) stromal and thylakoid-bound ascorbate peroxidase isoenzymes [Ishikawa, Sakai, Yoshimura, Takeda and Shigeoka (1996) FEBS Lett. **384**, 289–293]. The gene (*ApxII*) encoding both chloroplastic ascorbate peroxidase isoenzymes was isolated and the organization of the gene was determined. Alignment between the cDNAs and the gene for chloroplastic ascorbate peroxidase isoenzymes indicates that both enzymes arise from a common pre-mRNA by alternative splicing of two 3′-terminal exons. Genomic Southern-blot analysis supported this finding. The gene spanned nearly 8.5 kbp and contained 13 exons split by 12 introns. The penultimate exon 12 (residues 7376–7530) for the stromal ascorbate peroxidase mRNA consisted of one codon for Asp^{365} before the TAA termination codon, and the entire 3′-untranslated region, including a potential polyadenylation signal (AATAAA). The final exon 13 (residues 7545–7756) for the thylakoid-bound ascorbate peroxidase mRNA consisted of the corresponding coding sequence of the hydrophobic C-terminal region, the TGA termination codon and the entire 3′-untranslated region, including a potential polyadenylation signal (AATATA). Both exons were interrupted by a 14 bp non-coding sequence. Northern-blot and reverse transcription-PCR analysis showed that the transcripts for stromal and thylakoid-bound ascorbate peroxidase are present in spinach leaves.

INTRODUCTION

In plant tissues the chloroplasts are potentially the most powerful source of oxidants and sites within the cell most at risk from photo-oxidative damage. The photoreduction of molecular oxygen in chloroplasts is unavoidable and leads to the production of a superoxide radical ($O_2^{-\bullet}$), which is followed by formation of H_2O_2 via the disproportionation of $O_2^{-\bullet}$ in enzymic and non-enzymic reactions. The production rate of $O_2^{-\bullet}$ and H_2O_2 is estimated to be about 160 and 80 $\mu M \cdot s^{-1}$ in chloroplasts under normal conditions [1]. Because 10 μM H_2O_2 is sufficient to inhibit the photosynthetic assimilation of CO_2 by 50 % [2], photosynthesis is significantly impaired within a fraction of a second if H_2O_2 is not promptly scavenged. To rid themselves of excess H_2O_2, therefore, chloroplasts of higher plants develop two ascorbate peroxidase (AsAP) isoenzymes which exist as stromal soluble (sAsAP) and thylakoid-bound (tAsAP) forms [3–5]. Moreover, other AsAP isoenzymes have been known as a cytosolic form (cAsAP) [4,6–8] and microbody-bound form (mAsAP) [9,10].

Miyake et al. [5,11] have reported that the enzymological properties of tAsAP are very similar to those of sAsAP with respect to high specificity for ascorbate (AsA), K_m values for H_2O_2 and AsA, inhibition by cyanide, thiol-modifying reagents, thiols and suicide inhibitors, such as hydroxyurea, and inactivation in AsA-depleted medium . The sole difference in terms of properties between tAsAP and sAsAP is the higher molecular mass of the membrane-bound enzyme compared with the soluble enzyme. Thus tAsAP seems to be bound to thylakoid membranes in such a form that the active site of the enzyme is exposed to the stroma for the access of the substrate. In the previous study, the first complete cloning and molecular characterizations of sAsAP and tAsAP from spinach have shown that the nucleotide sequence encoding the tAsAP isoenzyme is identical with that of sAsAP through the coding region up to amino acid position 364, where the remainder of the C-terminal coding region is substituted by a different sequence that encodes 50 amino acids which constitutes a hydrophobic thylakoid-membrane-binding domain [12]. These data and the presence of different polyadenylation tracts in cDNAs encoding the sAsAP and tAsAP variants suggest that both enzymes are generated by a common pre-mRNA from an identical gene by alternative splicing of the 3′-terminal exons. The cDNAs encoding cAsAP isoenzymes have been isolated and characterized from several plant sources, including spinach (*Spinacia oleracea*) [7,13,14]. The genes from pea (*Pisum sativum*) [15] and *Arabidopsis thaliana* (thale cress) [16] have been isolated and characterized. Accordingly, it seems likely that the AsAP of spinach is a multigene family and there are at least four AsAP genes. In order to elucidate the occurrence of an identical nuclear gene by alternative splicing of the 3′-terminal exons and to understand the gene regulation of chloroplastic AsAP isoenzymes in higher plants, we sought to clone and characterize the nuclear gene (*ApxII*) from spinach.

MATERIALS AND METHODS

Materials

Spinach seeds were germinated on moist gauze at 15 °C in the dark. The cotyledons from seedlings grown for 10–14 days in the dark were transferred to illumination (140 μE/s per m^2) for 24 h

---

Abbreviations used: $O_2^{-\bullet}$, superoxide radical; AsAP, ascorbate peroxidase; sAsAP, stromal ascorbate peroxidase; tAsAP, thylakoid-bound ascorbate peroxidase; cAsAP, cytosolic ascorbate peroxidase; mAsAP, microbody-bound ascorbate peroxidase; AsA, ascorbate; poly(A)$^+$, polyadenylated; SSC, standard saline/citrate.

[1] Present address: Department of Biochemistry, Wakayama Medical College, 27 Kyubancho, Wakayama 640, Japan.

[2] To whom correspondence should be addressed.

The nucleotide sequence data reported in this paper have been submitted to the DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number AB002467.

to obtain the greening cotyledons as previously described [12]. Restriction enzymes and modifying enzymes were purchased from Takara Shuzo Co., Ltd. (Kyoto, Japan). The cDNAs coding for spinach chloroplastic AsAP isoenzymes were originally cloned into plasmid pBluescript SK(+) [7,12]. The *Escherichia coli* strain Y1090r$^-$ and DH5α F′ were obtained from Amersham (Amersham, Bucks., U.K.). All other chemicals were of analytical grade and were used without further purification.

Genomic DNA extraction

Genomic DNA was isolated from mature green spinach leaves. The leaves (20 g) were frozen in liquid nitrogen and ground to a slurry using a pestle and mortar while adding 20 ml of extraction buffer [50 mM Tris/HCl (pH 7.8)/100 mM NaCl/10 mM EDTA/200 μg/ml proteinase K/2 % SDS]. The slurry was incubated at 37 °C for 1 h. The solution was phenol-extracted three times, and the DNA was collected by precipitation with 0.1 vol. of 3 M ammonium acetate and 2 vol. of 100 % ethanol. The DNA was further purified by ethidium bromide/CsCl density-gradient centrifugation at 100000 ***g*** for 16 h at 20 °C.

Genomic Southern hybridization

Total DNA (20 μg) from spinach leaves was digested to completion with various restriction enzymes, separated by agarose-gel electrophoresis (1 % gels) and transferred to a Hybond N+ filter (Amersham) using a model 785 vacuum blotter according to the manufacturer's (Bio-Rad, Richmond, CA, U.S.A.) instructions. After transfer of the DNA to a nylon membrane, the genomic DNA fragments for chloroplastic AsAP isoenzymes were detected by probing with radiolabelled full-length tAsAP cDNA using the a random-primed DNA labelling kit (Takara, Shuzo). Washing the membrane in 1 × standard saline citrate (SSC)/0.1 % SDS at 65 °C was referred to as 'low stringency' and 0.1 × SSC/0.1 % SDS at 65 °C as 'high stringency'.

Construction and screening of partial genomic library

Spinach genomic DNA (40 μg) was digested with restriction endonuclease *Eco*RI and fractionated by electrophoresis in 1 % agarose gel. The 3.0–6.0 kbp fragments, which were identified by Southern hybridization, were excised and purified using the GENECLEAN II kit (Bio 101). The size-fractionated *Eco*RI DNA fragments were ligated into the *Eco*RI site of vector λgt11 (Amersham), packaged in a λ-*in vitro* packaging module (Amersham) and plated using *E. coli* strain Y1090 r$^-$. Phage harbouring appropriate inserts were screened by hybridization to the non-radiolabelled full-length tAsAP cDNA probe using a DIG DNA labelling kit (Boehringer-Mannheim G.m.b.H., Mannheim, Germany). Three positive *Eco*RI fragments, designated PAsAP1 (4.6 kbp), 2 (3.5 kbp) and 3 (5.6 kbp) were subcloned into a pBluescript II SK(+) vector (Stratagene, Heidelberg, Germany) and sequenced using the dideoxy-chain-primer method modified for double-stranded plasmid DNA. The overlapping insert DNA fragments were obtained by subcloning after restrict digestion and partial deletion by exonuclease III.

Primer-extension analysis

An antisense oligonucleotide primer 5′-TATTGTGGGTATG-AACGGTC-3′ spanning bases +78 to +97 from the initiation codon was 5′-end-labelled with [α-^{32}P]ATP using T4 polynucleotide kinase and gel-purified. The labelled primer was hybridized with total RNA (150 μg) from spinach leaves at 95 °C for 15 min, followed by incubation at 25 °C overnight in 50 % formamide/400 mM NaCl/20 mM Pipes (pH 6.4)/1 mM EDTA. The primer was then extended with Superscript II RNase H$^-$ reverse transcriptase (Gibco–BRL, Rockville, MD, U.S.A.) at 42 °C for 60 min in a reaction mixture that included 40 mM KCl, 50 mM Tris, pH 8.3, 6 mM $MgCl_2$, 1 mM dithiothreitol, and 1 mM dNTPs. The samples were then precipitated with ethanol and the extended products analysed by denaturing 8 %-PAGE.

Northern-blot analysis

Total RNA was isolated from spinach mature green leaves (5.0 g wet weight) with the guanidine thiocyanate method [7]. Total RNA (30 μg each) was subjected to electrophoresis on 1.2 % agarose gel containing 2.2 M formaldehyde and transferred to a Hybond N membrane (Amersham). Prehybridization took place at 55 °C for 6 h in buffer containing 5 × SSC, 5 × Denhardt's solution, 1 % SDS and 100 μg/ml denatured salmon sperm DNA. The blot was hybridized at 55 °C for 18 h in the presence of the ^{32}P-random-primed spinach tAsAP cDNA. The blot was washed twice at room temperature in 2 × SSC/0.1 % SDS for 10 min each, and in 0.1 × SSC/0.1 % SDS at 60 °C for 60 min. The blot was then exposed to X-ray film.

Reverse transcription PCR

Polyadenylated [poly(A)$^+$] RNA was purified using the PolyATtract mRNA Isolation Systems (Promega, Madison, WI). First-strand cDNA synthesis was performed using spinach greening cotyledons, poly(A)$^+$RNA and SuperScript II RNase H$^-$ reverse transcriptase (Gibco–BRL) with a random hexanucleotide primer. The reaction mixture (25 μl) contained standard enzyme buffer, 2 μg of poly(A)$^+$ RNA, 7 μM hexanucleotide primer, 10 mM dithiothreitol and 1 mM each dNTP, with 200 units of reverse transcriptase, and was heated to 95 °C for 5 min. A common 5′ primer of both sAsAP and tAsAP (P-1; 5′-AAAACCACCCAATCTCACTCACT-3′) was combined with an sAsAP-specific 3′ primer P-2 (5′-TTTTTAAGCAACAC-AACACACTCG-3′) or a tAsAP-specific 3′ primer P-3 (5′-ACTGCCAAAACTCCAATCACAATC-3′) to detect mRNA encoding the sAsAP and tAsAP variants, respectively. The PCR cycle was 94 °C for 60 s, 60 °C for 60 s and 72 °C for 90 s and was repeated 30 times.

RESULTS

Genomic Southern-blot analysis

The finding that the cDNAs encoding chloroplastic AsAP isoenzymes either contained or lacked a sequence with a part of the intact reading frame strongly suggests the possibility that only one gene generates pre-mRNA for chloroplastic AsAP isoenzymes by alternative splicing [12]. The genomic organization of the gene for chloroplastic AsAP in spinach leaves was examined by Southern-blot analysis. When the genomic DNA was digested with several restriction endonucleases and the fragments were probed with tAsAP cDNA, one to three hybridization signals were detected, depending on the restriction enzymes (results not shown). When the membrane was washed under low-stringency conditions, the result showed the same hybrid pattern as that under high-stringency conditions (results not shown). These observations suggest that there was one copy of the gene for chloroplastic AsAP in the genome of spinach. To determine whether this hypothesis was compatible with the structure of the gene for chloroplastic AsAP, the genomic DNA fragments

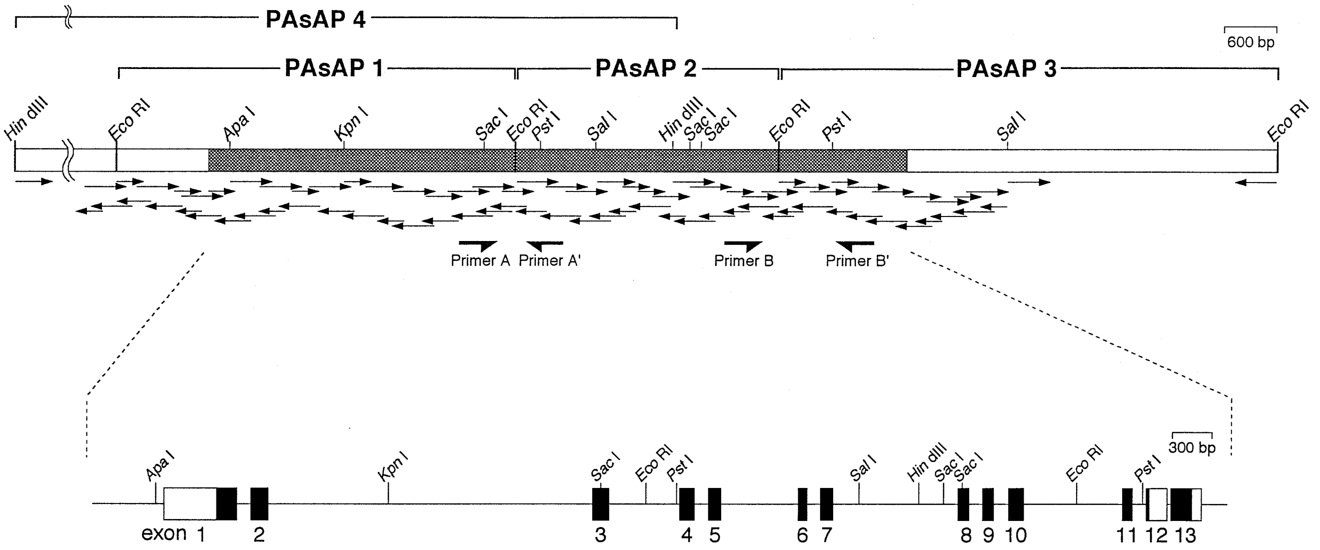

Figure 1 Organization of the spinach chloroplastic AsAP gene (*ApxII*)

Horizontal thin lines indicate sequentially overlapping clones derived from the spinach partial genomic library. The shaded box indicates the proposed region of spinach genomic DNA. Important restriction-enzyme sites used for mapping and subcloning are indicated. The locations of primers to detect the vicinity of connection of each clone are shown by arrows. Exons are represented by open (untranslated) and closed (translated) boxes and numbered at the bottom.

spanning the coding sequence for chloroplastic AsAP isoenzymes were cloned from a partial genomic library of spinach leaves. Three positive clones, PAsAP1 (4.6 kbp), 2 (3.1 kbp), 3 (5.6 kbp) hybridized with tAsAP cDNA and were found to contain the distinct portion of the entire chloroplastic AsAP gene (*ApxII*). Therefore the three fragments contained a full-length, 8.5 kbp *ApxII* gene in addition to 0.3 and 4.5 kbp respectively of the 5′-upstream and 3′-downstream regions. Furthermore, PAsAP4, one of the clones harbouring 10.2 kbp restriction-endonuclease-*Hin*dIII fragments, was found to overlap the whole region of PAsAP1 and the upstream region of PAsAP2 by restriction-enzyme mapping and partial sequencing of PAsAP4. Synthetic primers corresponding to the vicinity of connection of each *Eco*RI fragment shown in Figures 1 and 2 were used to identify the connection sites of the AsAP gene. The restriction map and cloning strategy of genomic clones encoding the chloroplastic AsAP isoenzymes are shown in Figure 1.

Organization of exon and intron structure

By alignment between the cDNAs for chloroplastic AsAP isoenzymes and the gene, the *ApxII* nuclear gene was found to contain 12 introns (Figures 1 and 2). The coding region of the chloroplastic AsAP gene consisted of 13 exons, referred to as 1 to 11, 12 (penultimate) and 13 (final) in order from 5′ to 3′ (Figures 1 and 2). The TATA box was found in positions −43 to −46, and two putative CCAAT boxes were found 42 and 67 bp upstream from the TATA box. Significant potential regulatory elements were not found upstream of the transcription start site. Primer-extension analysis identified the transcription start site within exon 1 (Figure 3); the 5′-terminus of exon 1 contained the long 5′-untranslated region and the ATG initiation codon. The cDNA clones encoding chloroplastic AsAP isoenzymes were more than 200 nucleotides shorter than expected from the primer extension analysis [12]. It is worth noting that three small upstream open reading frames (residues 16–156, 201–236 and 233–277), which may influence expression from the authentic one [17], are present in the 5′-leader sequence. Exons 1–11 encoded the common amino acid sequence for stromal and thylakoid-bound AsAP isoenzymes. The important point to note is the two 3′-terminal exons. Two splice acceptor sites occur in exons 12 and 13 separated by a 14 bp-encoding sequence (residues 7531–7544). The penultimate exon 12 (residues 7376–7530) consisted of one codon for Asp^{365} before the TAA termination codon, and the entire 3′-untranslated region including a potential polyadenylation signal (AATAAA) of the sAsAP mRNA. The final exon 13 (residues 7545–7756) consisted of the corresponding coding sequence of the hydrophobic C-terminal region, the TGA termination codon, and the entire 3′-untranslated region, including a potential polyadenylation signal (AATATA) of the tAsAP mRNA. These results show that the mRNAs for the chloroplastic AsAP isoenzymes arose from only one gene by alternative usage in the final two exons. For all of the coding exons, the intron–exon splice junctions were readily identifiable, which conformed to the consensus sequences GT at the donor site and AG at the acceptor site (Table 1), which was in agreement with those of general genomic genes [18]. The 5′-splice site at 7275 was used in combination with the 3′-splice site at 7544, generating a spliced transcript for tAsAP in which 270 nt of the intron 12 were removed.

Detection of chloroplastic AsAP isoenzyme mRNAs

To analyse the expression of mRNA for chloroplastic AsAP isoenzymes, the total RNAs from spinach leaves under normal conditions (non-stress conditions) were subjected to Northern blotting using labelled tAsAP cDNA as a hybridization probe. As a result, the mRNAs of chloroplastic AsAP isoenzymes were expressed under normal conditions with a length of 1.4 kb (results not shown). However, we could not distinguish the level of sAsAP mRNA from that of the tAsAP mRNA because of the similar size of the mRNAs for the chloroplastic AsAP isoenzymes. To provide the expression of both isoenzymes in detail, oligo(dT)-primed cDNAs were synthesized by reverse transcriptase from

```
gcttttagcgtaaaaccccaacgaaccaacacctctccactttaactggtctgagcaaaa   -491
aaatcctatttaaaattaaacccaaaattggactcagttttgtaagaagctcgccccgac   -431
ctgaccaatatatagtcacctcttggctcaattggtccaccaaaaatacaaagccacata   -371
ccgccaacacaaagatggttatgggtctggccgacccgaggcccaatgtaacatgacaca   -311
aagcccgacccgacccgacccaaacataagccaggctcgacacgagcacaaagtcgtgga   -251
ctgagccacatttttaagaaaaagcacgacgatgcactacccgatatgacaaaacacac    -191
taaatttactaaaattaatcctagacaagacggtttggccccactcagtatgagggtata   -131
tgtgggcttgggcccaatctttagtttttcagcccgacccaattaggtcagtttcaaaca    -71
aacttgacatgggccgagaccgtttatacccacaatccatgggctcagtctgaaatctga    -11
          .
cctggtctccACTCGGCATAAGGGTATGTGGGCTTGGGCCAACCCGACCCAACTTGATCA     50
                                   ext-1r
GTTACAAACAGACTTGACGTGCGCTGAGACCGTTCATACCCACAATACACGGGCTCAGTC    110
TGAAATCCGACCCGATCCAAGACCATCTTTACCACCAACCCACTAATCCACAATTGACTA    170
ACTACAGTCTACCGCTTCACAGGTCACAGCATGACCCAGGCCCTAAAGCTCTTATTGAAA    230
CAATGAAACCCCAAAAAAACGCGTGGTTGGCGCACTTCTTCCTGTAATCCTTCAAAACCA    290
GTAGTGTCAGGAGAGAGAAAAACCACCCAATCTCACTCACTTTCTCTCTCTATATTTTCA    350
AACCACCACCCCGCAACCAATGACTGATCGTCTAGCAATGGCATCCTTCACTACCACCAC    410
                                      M  A  S  F  T  T  T  T
CGCCGCCGCTGCATCTCGTCTGCTTCCTTCTTCTTCCTCCTCCATCTCTCGACTTTCTCT    470
 A  A  A  A  S  R  L  L  P  S  S  S  S  S  I  S  R  L  S  L
CTCCTCTTCTTCCTCCTCCTCCTCCTCACTCAAATGTCTCCGATCATCTCCACTCGTCTC    530
 S  S  S  S  S  S  S  S  S  L  K  C  L  R  S  S  P  L  V  S
TCACCTCTTCCTTCGACAGgttctccttccatttccgccgtttttgccttcaattttaag    590
 H  L  F  L  R  Q
ctaatgtgttctcaatttaatctgaagtgtttgttgttgctgtttttgcagAGAGGAGGT    650
                                                    R  G  G
TCAGCTTATGTGACGAAGACGAGGTTTAGCACGAAATGCTACGCTTCTGATCCTGCGCAG    710
 S  A  Y  V  T  K  T  R  F  S  T  K  C  Y  A  S  D  P  A  Q
CTGAAGAATGCTAGGGAAGATATTAAAGAGCTTCTTCAATCTAAGTTCTGTCATCCTATT    770
 L  K  N  A  R  E  D  I  K  E  L  L  Q  S  K  F  C  H  P  I
ATGgtgattttcctctgtacttgttgatttctgtgttattcgtttgatttcctgcctttt    830
 M
gtgattttgaggtttaggttttatgttgtctgaagaatttttttggattgtgtttgttta    890
ttattcgagttggtagtgtaatttaaactacggagtaatattttgatgactatatccgga    950
cggagggaatgataagatgattctacattatttctacataggtttgactgagttaggtgt   1010
attagcagctagtgaaattcgaccgtttttatgattgattgatgtgtgatattcgggttg   1070
agccttcctcggcgttggatgtttgcttagatgcaaaattaataccactatttgaacatt   1130
attgtgagctttactcgaggagtctatgtgcacttactattcttgtgttttcctttgaca   1190
atgaggtctttgtgttagtttatgtttatggtgatgtaccgtgcaaacgttgattgattt   1250
agaaacctcttgaactaagttaatatgtgaaaaactggggaaatgtgaagggaaatgg     1310
ctcaatcatataggcttatagcatgaaattgaatcattaatttgagacctatgcacctac   1370
catatttaggagagtttatggggcggaggggttacaatgatgggttcctcgttgtgattt   1430
accactgcaaagcctataccacactgacttcccaaggtaccaacacaagcatataacatt   1490
tggataccatcttcctataatcctattgaagaatgaagatagcagtgccatatcaattgc   1550
aaatatttcagggttggattcttccgttcttcctgccctcatggggaagaacaaaacatt   1610
tcccgaatgaggtggagagttagttataataaaatgctaagattgtgaatgtggtacaat   1670
tttccatatcgcacaactctttttactaaaatttggcaacatttgttccctgaccgcccc   1730
gatatctttaaaatctgtatggacctacaatataaaagtatcatgctctattcattacat   1790
aagtaattttataagagtatgttcttttaaatttcactgttaagaaaaataaaggtcgta   1850
cctagtgcacaaggctcccgctatttagcccatgggtttggggaaggtccaaatgtacgc   1910
agccttacccttatttaccgttatgtaataataataaaatatgattatgattttactcca   1970
aagattgcaaaaagtgtaacaggacatataatatggaaactatgaagttatttcaatatt   2030
gatctggtgttttgaaaaatatttcgatgcataggcgcacagagacaagcaccaagctg    2090
atggtgactaggagtttaggaccccaaaaggaaatggggagatgatgaaacccatccatt   2150
tgtgttgagttgttgctagcttcatctgtagtgaacttaaatgaaataattcacccgttt   2210
ttaaaggaggaatctgtagattactagatactacctgataacgaaggacgtcatactgtc   2270
agatgctataatattcccctacttattcagcttggatataagtctgacagtcttacttgg   2330
gaaatacttatggaattgcaggtgcctttctcgtcaggtgtcctgaaaataatctccttt   2390
gataagagagcttatagtttgctcactgtccacacacatgcaaagcagggtagttagtaa   2450
aactactgtagacctctcttctttccatcaggtagtctattttcttacttgagttacaaa   2510
attggtacattgggggtgtgtgtgtgtagccgtgtaggtgttgatgtcgtgggcgaggtg   2570
gggatcagggatgagacctttccatttcattacaagtctttgaagtacagttttctggta   2630
ctctatgccttttacgaatgtatattagtggctgagtctgaacaacttgaacaatgaagt   2690
gaaaagtttaatttaattacttggcgcgtctaaggctgaattggccctgatctttgacct   2750
aaagaaaagtgttattggttataatgaaatagttatctgttcctgtcaaaaattgtgcca   2810
cctccgtatggagaaaattttaatcgtttttatttaaaacatttcatcttctaggcaatt   2870
aacaaagtagtatgtctgaaattttagtatcgactattcttatttctatagacacgaaaa   2930
gggtcatggaaaaagatgcaaaatgcaaattgaggaaaatgaaaagaagttattgcattt   2990
atcttccaccgtctctcaagggctcaaggccataaacatgtgttcaattatgcagttttc   3050
cagctggcagctgttgacctcgactttaacatagttccaacttccaagttttacagtttt   3110
ccgtatgcctactttacaattgaaattaatctgcccactgtttaatacgtggcgcatgtc   3170
tttcttcctcacgattttctgggtgaacttatgaaatcactctctgcatgcatgtgaatg   3230
aagatgctatctttctttcaagactttgagcaggaattgaggaatttcaacactttattt   3290
gtagGTTCGCTTAGGTTGGCACGATGCCGGTACTTATAATAAGGACATTAAAGAATGGCC   3350
     V  R  L  G  W  H  D  A  G  T  Y  N  K  D  I  K  E  W  P
                                   Primer A
ACAAAGAGGTGGAGCCAATGGCAGTCTGAGCTTTGATGTTGAGCTCAGGCATGGAGCTAA   3410
 Q  R  G  G  A  N  G  S  L  S  F  D  V  E  L  R  H  G  A  N
TGCAGgtaagaaccagttccccaactccttcatatgctccctttgacggtggattttctt   3470
 A  G
ttcttttcccggctacttttgaaaatatatgatacaaatagactactccaagtgatacct   3530
ctatctactgttgaccttcttcaatgctaatgtttttgaaatataaaaaaggaaacagtg   3590
tattatctcaaatgatttttcttttaaaaaagagtgaaagataaatagctttttgttag    3650
aggtacatgaacaatgcttgtcaacttcagtatagctttcaaaattatattatcccgcat   3710
cctttacttctctatgatacttcattcctagatttcgggttaagtattatatttgaattc   3770
actcctaaatttatgtttgtactagggaaaccttgtcttcagtagggtctaatctataaa   3830
agaaaggaacttgtgagtgaaagaagaaaaactttatctcttgcgttgcactttcctctt   3890
ataattgcttgttatgtgaacgaaaacatgggttaattaattggtactgcagGTCTTGTT   3950
                                                     L  V
                 Primer A'
AATGCCCTGAAACTTCTACAGCCCATAAAAGACAAGTACTCTGGAGTTACATATGCAGAT   4010
 N  A  L  K  L  L  Q  P  I  K  D  K  Y  S  G  V  T  Y  A  D
CTATTCCAGCTGGCTAGTGCTACTGCAATAGAGgtcttctcatcttttctcatttatagt   4070
 L  F  Q  L  A  S  A  T  A  I  E
ctagtaatagttgacatagatggaagttgtagctttattgatcgaattattttctgaaca   4130
ttttagGAGGCTGGTGGTCCAACAATACCCATGAAGTATGGAAGAGTGGATGCCACAGGG   4190
       E  A  G  G  P  T  I  P  M  K  Y  G  R  V  D  A  T  G
CCGGAGCAGTGCCCAGAAGAAGGAAGGCTTCCTGgtaagagaactttccatagtttatat   4250
 P  E  Q  C  P  E  E  G  R  L  P  D
tgtgatggtttgaaaatataatccggatagaagtgccaacgttgccatggaactttcaac   4310
aggactgtggctggattatgtacgtgtacataattatgtattatcagggactttcaaatt   4370
tcaaacttatatacagattatacacttgtctaatttgtattaagtgaagggtttctccta   4430
aatattgttggtgtttatatcctggaggggaggtggcaataaccctggttgttgctgttg   4490
ttgaatctgagaaatccaagatggtttcctaaaaataatgcaactttcttggtgtctctc   4550
aatttcaatttttgttgggcttttctatggataaaaaattgtttacgcttggtcgttttg   4610
cgaagtttctagctttgattcgataatggtcggtactaatttggaagaattgtatcatgt   4670
tattgcttgtggtcctttgattgtctcttctgaatagacttcttggtctaagtcttgctg   4730
atattttatgctgcttgtgcttctatgtatctttgctgccttactgactcttagattctc   4790
cctgttcagATGCTGGACCTCCTTCACCTGCTCAACATCTACGTGATGTTTTCTACAGAA   4850
          A  G  P  P  S  P  A  Q  H  L  R  D  V  F  Y  R  M
TGGGTCTTGATGATAAGgtaaattaaactgttacctggattatcaactgatgcctcaatt   4910
  G  L  D  D  K
tggccatcgtgttcattatgaatttgtgagctgtttgtgttgatagGATATAGTAGCATT   4970
                                              D  I  V  A  L
ATCTGGAGCACATACGTTGGGAAGGTCTAGACCTGAACGCAGTGGTTGGGGCAAGCCAGA   5030
 S  G  A  H  T  L  G  R  S  R  P  E  R  S  G  W  G  K  P  E
GACTAAATACACGgtacttctattaatttttctgaatgggaagaaaatttaccaccaggt   5090
 T  K  Y  T
ccttctgtggtctctgaatagcagttccaagttggaaaattacattgtcgacacagaagt   5150
gtagattttataactaataatctagattgtttcaatgatttgcgcttagaagctcttata   5210
tacttttacttcctctatttatttttatagtttcaacgttctgcattttacactattcac   5270
cacacactactttgactatatttgaggatttatacttaataaaaaagagtcatgtgggat   5330
ctttcgtctcaatatatattttcaaaatatcaacttttttatattttttcaatttttggtaa   5390
ttaaatatattaatggttgaaattttgcactgtcaaagcgtgacaatagaagtgttgcat   5450
ctaataataaacggagggcatatatatttctcttggcaatataattctttgaaatagcac   5510
gatattgctgcataaaatgcacttgtcttaagtctgaagtatctccttatgcccctcctc   5570
tttgatggattcttggtattttctttattcactcagcaattagaataggtttaaattgtt   5630
aattaaatcgaaggtgtgaaagtgaaaatcggcagcattttggaggaaacaaagattacg   5690
aggaaaagcttctaaggtgtcctaaaaaaagacatttacacttgtcccaagtcccaattt   5750
taatcccccacttgatataggaaaaacccgaaaaagtgtggtgcctcagatatactggta   5810
tgtttttttagggttaccgtagtgacatgaagctattctcagagctccaactgtttccta   5870
ccattgccctccctctactatatttttttggtcatctttgtttgtatttttttttgcgct   5930
ttctggaatttgtttttgggattatacgcaatttgttaagctgatgaaatcatacttttc   5990
tccttcactagAAAGATGGACCTGGAGCTCCAGGAGGGCAGTCATGGACTGCGGAGTGGT   6050
          K  D  G  P  G  A  P  G  G  Q  S  W  T  A  E  W  L
TGAAGTTTGATAATTCCTATTTCAAGgttttcatttaaatttgttctagacatcaaaaag   6110
  K  F  D  N  S  Y  F  K
acacaaacaaaaaaaggtttgcatattcttatgctaacaatatattgcctaattccagGAC   6170
                                                          D
ATCAAAGAAAAGAGAGATGCAGATTTGCTTGTTTTGCCAACTGATGCTGCTCTTTTCGAA   6230
 I  K  E  K  R  D  A  D  L  L  V  L  P  T  D  A  A  L  F  E
GATCCGTCTTTCAAGgtaaagaaagcaatgcataatccatgagaaatgcaacaaaatcta   6290
 D  P  S  F  K
cacatggtattcgcttgtgtttgtgattcgtaaatattttcatgattgtagGTATATGCA   6350
                                                    V  Y  A
                   Primer B
GAGAAATATGCAGCTGACCAAGAAGCATTTTTCAAGGATTACGCTGAAGCCCATGCCAAA   6410
 E  K  Y  A  A  D  Q  E  A  F  F  K  D  Y  A  E  A  H  A  K
CTCAGCAACCAAGGAGCCAAATTTGACCCTGCTGAGgtctgtcttgttcctaattctcat   6470
 L  S  N  Q  G  A  K  F  D  P  A  E
gtatctgtgttccttaattttgcatcttaattttgtatctgaatccttctcctgtatg     6530
catttttttccttccttttggggtgggtgaggtggtgactctccatactaggctagctatg   6590
acagtaaaaaagacacaaaaggagggctaaggtgatttttttttggcaagcaaggtgac    6650
attttggtggcctccccacatatactccctccgtccctaaaagattgtctcaattgactt   6710
ttcacggtctccaagacactacttaaaacgtaattaattataatactacacgcgtaaaaa   6770
ttataaaaaattgatgttccgaaaatatatattaaaacgaatcttacaagaccccattct   6830
tataggataatgagaattcatggtcaaagttttcaaaatttgacccaagatattacgatg   6890
gggatcttttagggactgagggggtattagcttgagggaagatggtataatattgctgtt   6950
aattcttcccaaaactctgtatgtgttattgatataaacggtgtatgcaagtatctatta   7010
tctttcatcgtgtttagtgatgcaagatctttttttttcattcgaaatggtcaagaacttt   7070
cgagtttttttcctagacttctgaaatcttccccattaatctgttttttttttaatttcc   7130
ttcaaaaaaattcctatattgtttttacttgaccttcaaactagactaataggacttgct   7190
tatggacagGGTATCACTCTTAATGGAACCCCTGCCGGAGCAGCTCCAGAGAAGTTTGTA   7250
          G  I  T  L  N  G  T  P  A  G  A  A  P  E  K  F  V
GCAGCCAAGTACTCATCTAACAAGgtttgttattataatcattcttagtaaaaatacaga   7310
 A  A  K  Y  S  S  N  K
tttttttttccccatgacattgttagattattcatctttgttgctttttttcatggtttgc   7370
tgcagGATTAAAGAAAAAAAGAAAACTCAAAAGAATCATGTTTTCATGTGACTTTTGTCT   7430
    D  *
TTTTTTTAATAAGCAAGATGAAGCTATTTTCTTGTGTAGTCAGTCACATCTAAAACATGG   7490
TGGAGAAGAATAAAATTCAAACGAGTGTGTTGTGTTGCTTatttgaatttgcagAGATCA   7550

                                                        R  S
                                                     Primer B'
GAGCTTTCGGATTCTATGAAGGAAAAGATTCGCGCTGAATATGAAGGTTTTGGAGGTAGC   7610
 E  L  S  D  S  M  K  E  K  I  R  A  E  Y  E  G  F  G  G  S
CCTAATAAGCCTCTACCAACAAACTACTTCCTAAACATTATGATTGTGATTGGAGTTTTG   7670
 P  N  K  P  L  P  T  N  Y  F  L  N  I  M  I  V  I  G  V  L
GCAGTTCTATCATATCTTGCGGGAAATTGATTTGTGGTTTGATGAGTTTTTTTCCATTTA   7730
 A  V  L  S  Y  L  A  G  N  *
TAAATATAACGGCAGTTGATTATATGaaatttcttcttgaaaatcgtccattcacatgat   7790
attttgaataaatcagagttaaaattcttaaccttctgtttgtgctggaattgtaactaa   7850
aaagcataactttcttataaaaaaaataggaggatcagattggaatatactagccaagac   7910
tgatttgtgcaaataagggttagatccactagttctagagcggccgcaccgcggtggagc   7970
```

Figure 2 For legend see opposite page

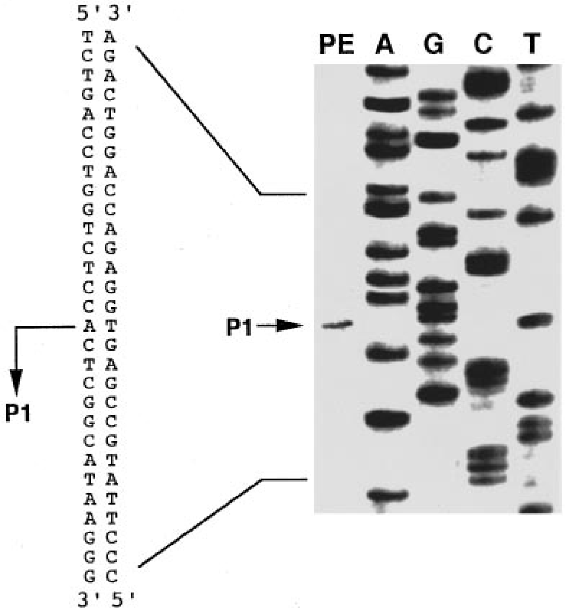

Figure 3 Determination of the transcription initiation site for the spinach chloroplastic AsAP gene by primer extension

The primer ext-1r was labelled with ^{32}P and annealed with 150 μg of total RNA from spinach leaves. The analysis was performed as described in the Materials and methods section. PE, sequencing ladder (A, G, C and T). The arrow indicates the major transcription initiation site.

Table 1 Splice sites (emboldened), length and percentage A+T of introns in the gene encoding chloroplastic AsAP isoenzymes

Intron	Length (bp)	5′-Splice site	3′-Splice site	A+T (%)
1	92	CAG **GT**TCT	TGC**AG** AGA	63.1
2	2521	ATG **GT**GAT	TGT**AG** GTT	63.2
3	527	CAG **GT**AAG	TGC**AG** GTC	67.4
4	93	GAG **GT**CTT	TTT**AG** GAG	71.0
5	575	CTG **GT**AAG	TTC**AG** ATG	64.0
6	89	AAG **GT**AAA	GAT**AG** GAT	64.1
7	958	ACG **GT**ACT	ACT**AG** AAA	67.1
8	91	AAG **GT**TTT	TCC**AG** GAC	72.5
9	96	AAG **GT**AAA	TGT**AG** GTA	67.7
10	753	GAG **GT**CTG	GAC**AG** GGT	65.3
11	101	AAG **GT**TTG	TGC**AG** GAT	72.3
12	270	AAG **GT**TTG	TGC**AG** ACA	71.1

purified poly(A)$^{+}$ mRNA and subjected to PCR. The specific primer pairs P-1/P-2 and P-1/P-3 for sAsAP and tAsAP revealed the 1333 bp and 1302 bp products respectively, which was in agreement with the previously obtained data [12]. These results indicated the presence of the sAsAP and tAsAP mRNAs in spinach leaves. When PCR cycles were changed to less than 30 cycles, equal amounts of reverse transcription PCR products were detected (results not shown), suggesting that each mRNA exists at almost the same level.

DISCUSSION

Here the first entire genomic structure (*ApxII*) encoding chloroplastic AsAP isoenzymes from spinach was identified. The *ApxII* gene consisted of essentially 13 exons interrupted by 12 introns. In previous studies, genomic structures encoding cAsAP isoenzymes (*ApxI*) have been characterized from pea [15] and *Arabidopsis* [16]. The structure of the *ApxII* gene for the chloroplastic AsAP isoenzymes from spinach was considerably different from those of the *ApxI* genes for cAsAP. The *ApxI* genes are interrupted by 9 introns, the first of which was located in the 5′-untranslated regions of the mature transcript. The intron in the 5′-noncoding region was not observed in the gene of *ApxII* (Figure 1).

From the results of a comparison with the *ApxII* genomic structure and cDNA sequences and genomic Southern blot, it is clear that the mRNAs of spinach chloroplastic AsAP isoenzymes are produced from a single gene by alternative exons 12 and 13 and share the same open reading frame, except for the hydrophobic thylakoid-membrane-binding domain. How the chloroplastic AsAP isoenzymes are produced by different selection of their 3′-ends is an interesting problem. One possibility is that primary transcripts with 3′-end heterogeneity arise from cleavage at different alternative polyadenylation sites. Intact exons 12 and 13 of the *ApxII* gene both contain their own termination codons and potential polyadenylation signals respectively. The production of heterogeneity in the amino acid sequences by regulation systems such as alternative splicing, 3′-end cleavage and polyadenylation of mRNA precursors is well documented in animal systems [19,20]. In higher plants, heterogenous C-terminal production by alternative splicing has been shown in spinach ribulosebisphosphate carboxylase/oxygenase (rubisco) activase [21]. In this case, the alternative usage of a part of the intron results in the synthesis of two rubisco activases that are different only in the presence or absence of their 37 additional amino acids at the C-terminal regions. Another example is revealed by pumpkin (*Cucurbita* sp. cv. Kurokawa Amakuri nankin) hydroxypyruvate reductase isoenzymes with or without a putative C-terminal signal for targeting to microbodies [22]. Although these cases are similar to spinach chloroplastic AsAP isoenzymes with respect to the production of C-terminal regions by alternative splicing, the generation of two different polypeptides from the resulting in-frame shift caused by insertion of part of their introns is different in the case of the *ApxII* gene. Two spinach chloroplastic AsAP isoenzymes generate the resulting alternative usage of the independent 3′-end nucleotide sequences which are located in the penultimate exon 12 and the final exon 13 of the gene (Figure 4). This is the first report describing the addition of a peptide sequence to a protein C-terminus by alternative 3′-terminal exons in higher plants. Therefore spinach chloroplastic AsAP isoenzymes may provide a good example of a new type of gene regulation mechanism in higher plants. Recently, the cDNA sequence encoding tAsAP from pumpkin has been reported [23]. The deduced amino acid sequence from pumpkin mature tAsAP showed a high degree of identity (83.7%) and a hydropathy profile similar to those of tAsAP from spinach. Interestingly, antibodies raised against the C-terminal 82-residue polypeptide of the pumpkin tAsAP cross-reacted with both tAsAP and

Figure 2 Nucleotide and deduced amino acid sequences of the gene (*ApxII*) encoding chloroplastic AsAP isoenzymes

The nucleotide sequences of exons and introns are shown in upper-case letters together with assigned amino acids and in lower-case letters respectively. The potential CCAAT box, the TATA box and the polyadenylation signal are double-underlined. Numbers on the right are positions from the major transcription initiation site (+1), which is marked by dot. Primers using primer extension and PCR analysis are underlined.

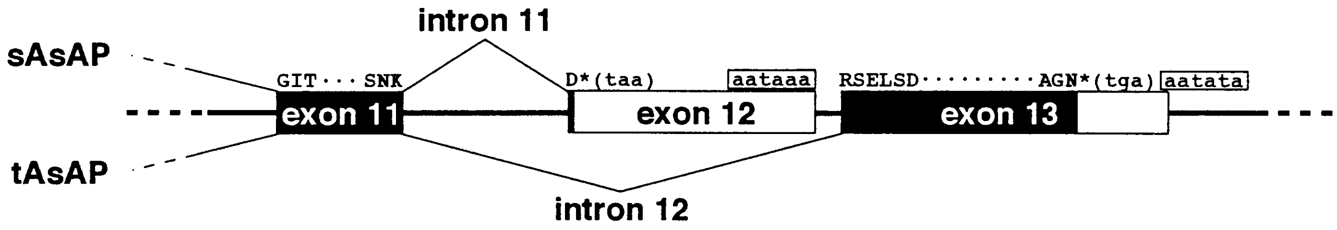

Figure 4 Partial organization of the use of the alternative two terminal exons of *ApxII* for the generation of the stromal (sAsAP) and thylakoid-bound AsAP (tAsAP) mRNAs

Boxes denote exons; open boxes represent 3′-non-coding transcribed regions and closed boxes indicate an open reading frame.

sAsAP from pumpkin. It seems likely that the alternative-splicing mechanism at the 3′-end for the chloroplastic AsAP isoenzymes occurs in higher plants, including pumpkin. Recently, cDNA sequences of sAsAP (accession No. X98925) and tAsAP (accession no. X98926) from *Arabidopsis* have been appeared in the GenBank database. A comparison of nucleotide sequences of cDNAs for both sAsAP and tAsAP from *Arabidopsis* showed 66.1 % identity, suggesting that both isoenzymes may be encoded by different genes. However, mature native sAsAP and tAsAP proteins in *Arabidopsis* chloroplasts are not identical.

The existence of only one gene encoding the chloroplastic AsAP isoenzymes and the one TATA box in the 5′-end flanking region of the *ApxII* gene suggests that these enzymes are transcribed from the same promoter that is active. Preliminary evidence for transcriptional levels of *ApxII* products using reverse transcription PCR showed that the levels of mRNA for both chloroplastic AsAPs are in almost equal quantities throughout the spinach leaves under normal conditions (results not shown). The activity of tAsAP was almost the same as that of sAsAP in spinach chloroplasts [5]. These results clearly support the hypothesis that the localization of the AsAP isoenzymes in both the thylakoid membrane and the stroma soluble fraction has important implications in the protection from photo-oxidative stress caused by active oxygen species, including H_2O_2, in the chloroplasts of higher plants.

This work was supported by a grant from NEDO/RITE's International Joint Research Program and by Kansai Research Foundation for Technology Promotion.

Received 9 June 1997/19 August 1997; accepted 28 August 1997

REFERENCES

1 Asada, K. and Takahashi, M. (1987) in Photoinhibition (Kyle, D. L., Osmond, C. B. and Arntzen, C. J., eds.), pp. 227–287, Elsevier Scientific Publishers, Amsterdam
2 Kaiser, W. (1976) Biochim. Biophys. Acta **440**, 476–482
3 Asada, K. (1997) in Oxidative Stress and the Molecular Biology of Antioxidant Defenses (Scandalios, J. G., ed.), pp. 715–735, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
4 Chen, G.-X. and Asada, K. (1989) Plant Cell Physiol. **30**, 987–998
5 Miyake, C. and Asada, K. (1992) Plant Cell Physiol. **33**, 541–553
6 Mittler, R. and Zilinskas, B. A. (1991) Plant Physiol. **97**, 962–968
7 Ishikawa, T., Sakai, K., Takeda, T. and Shigeoka, S. (1995) FEBS Lett. **367**, 28–32
8 Ishikawa, T., Takeda, T. and Shigeoka, S. (1996) Plant Sci. **120**, 11–18
9 Yamaguchi, K., Mori, H. and Nishimura, M. (1995) Plant Cell Physiol. **36**, 1157–1162
10 Bunkelmann, J. R. and Trelease, R. N. (1996) Plant Physiol. **110**, 589–598
11 Miyake, C., Cao, W.-H. and Asada, K. (1993) Plant Cell Physiol. **34**, 881–889
12 Ishikawa, T., Sakai, K., Yoshimura, K., Takeda, T. and Shigeoka, S. (1996) FEBS Lett. **384**, 289–293
13 Mittler, R. and Zilinskas, B. A. (1991) FEBS Lett. **289**, 257–259
14 Kubo, A., Saji, H., Tanaka, K., Tanaka, K. and Kondo, N. (1992) Plant Mol. Biol. **18**, 691–701
15 Mittler, R. and Zilinskas, B. A. (1992) J. Biol. Chem. **267**, 21802–21807
16 Kubo, A., Saji, H., Tanaka, K. and Kondo, N. (1993) FEBS Lett. **315**, 313–317
17 Gallie, D. R. (1996) Plant Mol. Biol. **32**, 145–168
18 Brown, J. W. S. (1986) Nucleic Acids Res. **14**, 9549–9559
19 Chabot, B. (1996) Trends Genet. **12**, 472–478
20 Wahle, E. (1995) Biochim. Biophys. Acta **1261**, 183–194
21 Werneke, J. M., Chatfield, J. M. and Ogren, W. L. (1989) Plant Cell **1**, 815–825
22 Hayashi, M., Tugeki, R., Kondo, M., Mori, H. and Nishimura, M. (1996) Plant Mol. Biol. **30**, 183–189
23 Yamaguchi, K., Hayashi, M. and Nishimura, M. (1996) Plant Cell Physiol. **37**, 405–409