Volume 315, number 3, 313–317 FEBS 11995 January 1993

Genomic DNA structure of a gene encoding cytosolic ascorbate peroxidase from *Arabidopsis thaliana*

Akihiro Kubo[a], Hikaru Saji[b], Kiyoshi Tanaka[a] and Noriaki Kondo[b]

[a]*Environmental Biology Division* and [b]*Regional Environment Division, National Institute for Environmental Studies, Onogawa, Tsukuba, Ibaraki 305, Japan*

Received 2 November 1992; revised version received 30 November 1992

A genomic DNA clone encoding cytosolic ascorbate peroxidase was isolated from a genomic library of *Arabidopsis thaliana*, using a cDNA for the enzyme as a probe. Nucleotide sequence and primer extension analyses of this gene (*APX1*) revealed nine exons split by eight introns, one of which is inserted in the 5′-untranslated region. The exon/intron organization of the *APX1* gene differs from that of the guaiacol peroxidase genes.

Active oxygen; Ascorbate peroxidase; Exon and intron; Genomic DNA sequence; Molecular cloning; *Arabidopsis thaliana*

1. INTRODUCTION

Among plant peroxidases (Prx), ascorbate Prx (APX, EC 1.11.1.11), which shows high specificity for ascorbate as an electron donor, is distinct from guaiacol Prx, represented by horseradish peroxidase (HRP), with regard to enzymatic and molecular properties [1]. The primary structure of APX is more similar to that of yeast cytochrome *c* Prx (CCP) than that of guaiacol Prx [2–4]. To date, no reports have demonstrated the genomic DNA structure of APX, although several guaiacol Prx genes have been sequenced.

Chloroplastic APX has been well characterized, and proved to be a component of the system for scavenging active oxygens produced through photosynthesis under normal and stress conditions [1,5,6]. Nonchloroplastic APX has been found in several plants, and purified from several species [5,7–9]. However, the physiological role of cytosolic APX in green tissue of higher plants is poorly understood.

In this paper, we report the isolation of a genomic DNA clone encoding the cytosolic APX of *Arabidopsis thaliana*. The genomic DNA structure of this gene is also discussed.

Correspondence address: A. Kubo, Environmental Biology Division, National Institute for Environmental Studies, 16–2 Onogawa, Tsukuba, Ibaraki 305, Japan. Fax: (81) (298) 51 4732.

Abbreviations: Prx, peroxidase(s); APX, ascorbate peroxidase; HRP, horseradish peroxidase; CCP, cytochrome *c* peroxidase.

2. MATERIALS AND METHODS

2.1. *Construction and screening of a genomic library*

Genomic DNA from *Arabidopsis thaliana* ecotype Columbia was a gift from Dr Y. Komeda (University of Tokyo). A portion of the DNA (4 μg) was digested with *Eco*RI to completion and electrophoresed in a 0.6% agarose gel. The 7.5- to 9.5-kb fragments, which were presumed to include an APX gene [3], were electroeluted from the gel and ligated with the *Eco*RI-digested plasmid vector pBluescript SK(−) (Stratagene Cloning Systems, USA). *Escherichia coli* strain DH1a was transformed with the recombinant plasmids. The subgenomic library was screened by colony hybridization using the digoxigenin-labeled cDNA encoding *Arabidopsis* APX [3] as a probe.

2.2. *DNA sequencing*

A 3.1-kb *Eco*O109I fragment from a genomic clone (pAPg2) was subcloned into the plasmid pBluescriptII KS(+) (Stratagene Cloning Systems) and several deletion derivatives were generated. The double-stranded DNA templates were sequenced by the cycle sequencing method with dye primers on an ABI Model 373A sequencer (Applied Biosystems, USA). Sequences were obtained from both strands of the genomic DNA.

2.3. *Primer extension*

A 30-mer oligonucleotide, at positions 315–344 in Fig. 2, complementary to the 5′ end region of the APX cDNA [3], was synthesized and labeled at the 5′ end with ^{32}P. Poly(A)$^+$ RNA (2 μg) prepared from *Arabidopsis* leaves [3] was hybridized with the oligonucleotide primer (5×10^4 cpm). The primer was extended with Superscript RNase H$^-$ reverse transcriptase (Life Technologies, USA) and analyzed on a sequencing gel.

2.4. *Cell fractionation and Western blotting*

Cell fractionation of leaf cells of *Arabidopsis thaliana* ecotype Columbia was achieved by differential centrifugation of the gently-ruptured protoplasts from the leaves according to [10], except that the $2{,}500 \times g$ (chloroplastic) pellet was lysed with 50 mM potassium phosphate (pH 7.8) containing 1 mM sodium ascorbate and centrifuged at $10{,}000 \times g$ for 5 min to separate the insoluble (chloroplastic membrane) fraction from the soluble (stroma) fraction. Western blotting of proteins in the subcellular fractions was performed as described using an anti-APX monoclonal antibody, AP3 [11].

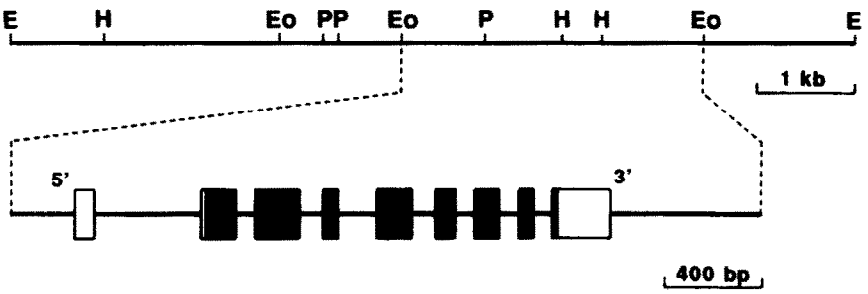


Fig. 1. Restriction map and structure of the *Arabidopsis* cytosolic APX gene (*APX1*). E, *Eco*RI; Eo, *Eco*O109I; H, *Hin*dIII; P, *Pst*I. Boxes represent exons: solid boxes, coding regions; clear boxes, non-coding regions.

3. RESULTS AND DISCUSSION

3.1. *Isolation of genomic DNA clones*

The *Arabidopsis* subgenomic library consisting of 1.7×10^3 independent recombinants was screened with a previously identified cDNA [3] encoding an APX of *Arabidopsis*. Two positive clones were isolated. One of them (pAPg2) contained an 8.7-kb *Eco*RI fragment and the other contained the same insert except a deletion spanning one of the vector-insert junctions. A restriction map of the genomic DNA inserted in pAPg2 is shown in Fig. 1.

3.2. *Structure of the APX gene*

A 3.1-kb *Eco*O109I fragment in pAPg2 (Fig. 1), hybridized with the APX cDNA, was subcloned and sequenced (Fig. 2), and a gene corresponding to the cDNA was found. The intron positions of the gene were unambiguously determined by comparison with the cDNA sequence and by the GT-AG rule [12]. The sequences of the exons match the cDNA exactly. The 5′ end of the gene mRNA was mapped within the 3.1-kb *Eco*O109I fragment by Northern analysis using various DNA fragments as probes. The exact position of the transcription start point was determined as shown in Fig. 2 by primer extension analysis (data not shown). The results revealed nine exons interrupted by eight introns. The predicted TATA box [13] is located 26 bp upstream from the transcription start site. The only possible start codon is located just before the N-terminal codon for the mature protein. This gene encodes no signal or transit peptide sequence, indicating the cytosolic location of the enzyme. One of the introns is inserted in the 5′-untranslated region. The exon/intron organization is depicted in Fig. 1.

We assume that this gene is a single-copy gene, since the genomic clone pAPg2 contains only one APX gene and the restriction map of the genomic DNA in pAPg2 is consistent with the results of genomic Southern analysis of the *Arabidopsis* genome [3], which showed an 8.7-kb *Eco*RI fragment and two *Pst*I fragments (19.5 and 1.5 kb), hybridizing with the APX cDNA. This gene was tentatively designated as *APX1*.

3.3. *Western blot analysis of proteins in the subcellular fractions*

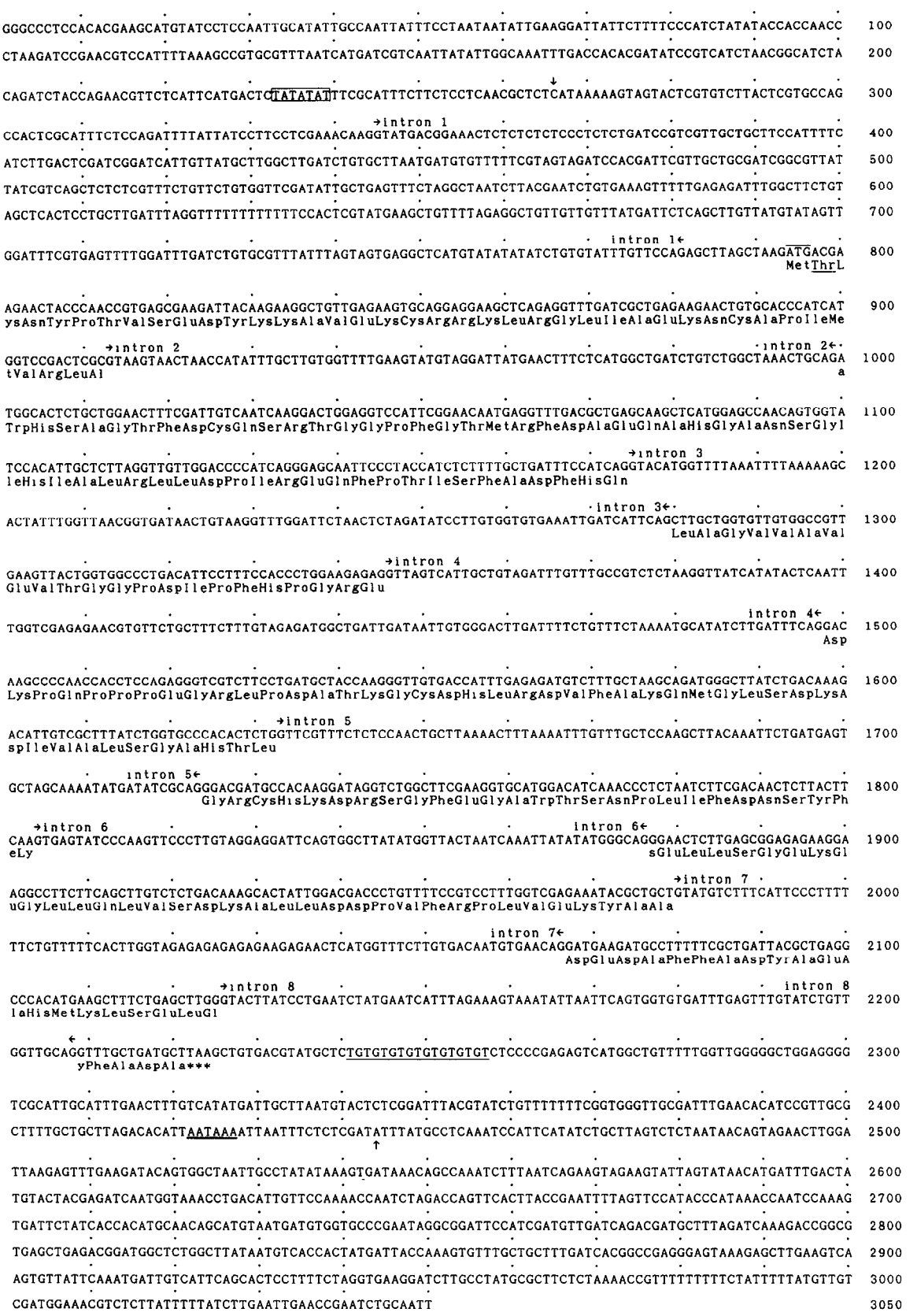
The subcellular localization of the *Arabidopsis* APX encoded by the *APX1* gene was examined by Western blotting of proteins in the subcellular fractions (Fig. 3). One of the anti-APX monoclonal antibodies, reactive with the *Arabidopsis* APX [3], reacted predominantly with a polypeptide corresponding to APX (28 kDa) in the $12{,}000 \times g$ supernatant fraction (lane 5). It is not clear whether faint bands with different molecular masses resulted from specific or nonspecific binding of the antibody. This provides additional support for the cytosolic location of the enzyme. Furthermore, the amino acid sequence of the *Arabidopsis* APX is more similar to that of the cytosolic APX of pea [2] (78% identity) than the partly determined sequence of the chloroplastic APX of tea [4] (38% identity between corresponding sequences). The subcellular localization of an isozyme of spinach APX, AP-II [14], against which the monoclonal antibodies had been raised, was examined by the same method. The results suggested that it is also a cytosolic enzyme (data not shown). APX activity in spinach leaves is elevated under stress conditions such as ozone exposure [15] and water deficiency [16]. However, the subcellular type of the APX induced remains to be determined.

3.4. *Comparison of Prx gene structure*

It has been found so far that the guaiacol Prx genes of higher plants have 1, 2, or 3 introns, located in the mature protein region [17–21; Theilade and Rasmussen,

→

Fig. 2. Nucleotide sequence of the *APX1* gene encoding the cytosolic APX of *A. thaliana*. The transcription start and the poly(A) addition sites are denoted by downward and upward arrows, respectively. The 5′ and the 3′ ends of each intron are marked with right- and left-oriented arrows, respectively. The deduced amino acid sequence of the APX is written under the nucleotide sequence. The start and the stop codons are indicated by overlining and asterisks, respectively. The N-terminal amino acid residue (Thr) of the mature protein [3] is underlined. The predicted TATA box is boxed. The predicted polyadenylation signal is bold underlined. The poly(dG-dT) sequence [3] is underlined.



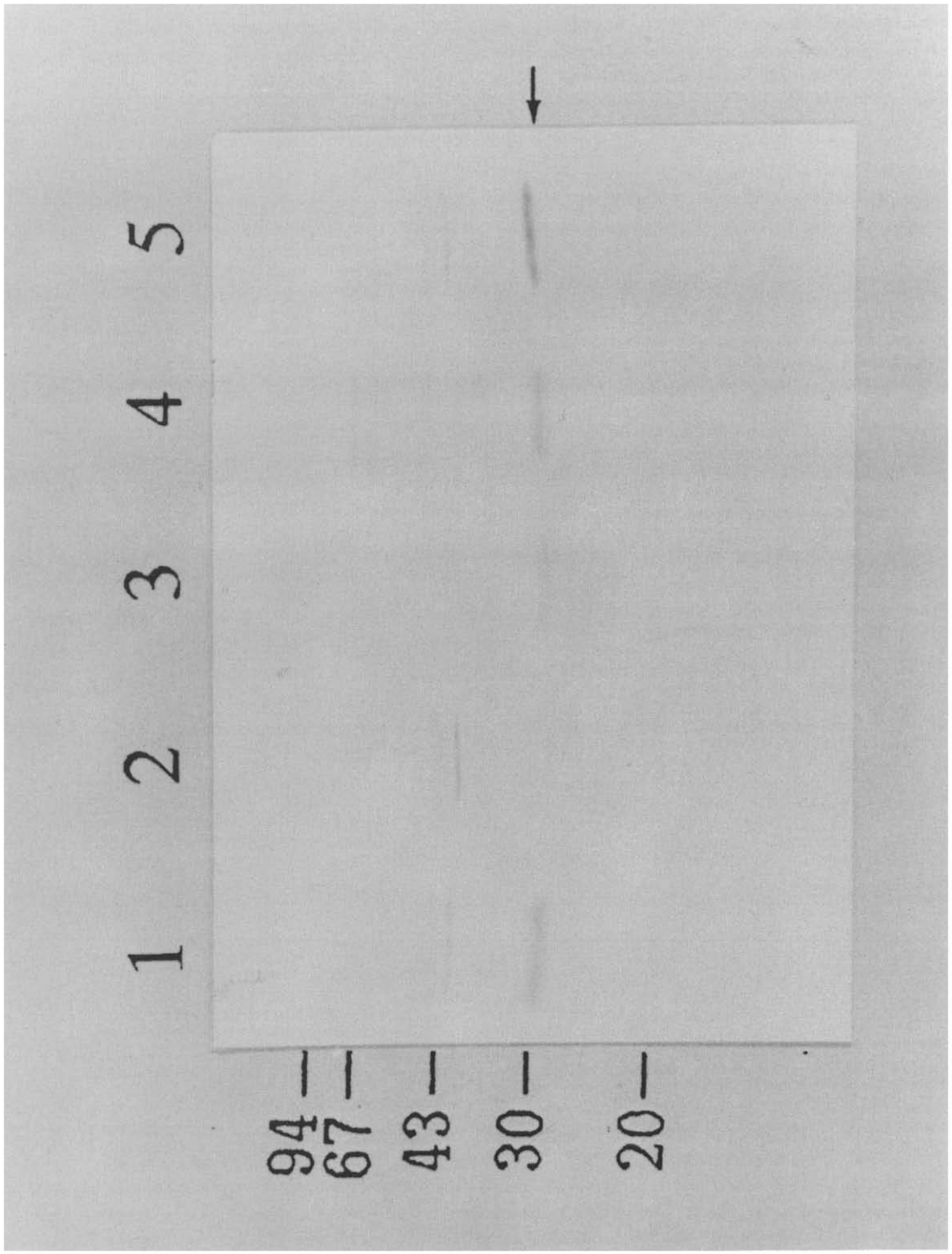


Fig. 3. Western blot analysis of proteins in the subcellular fractions from leaf cells of *Arabidopsis*. After cell fractionation, 9 μg protein was subjected to SDS-PAGE followed by western blotting using an anti-APX monoclonal antibody, AP3 [11]. Lane 1, protoplasts before cell fractionation; lane 2, soluble fraction from 2,500 × *g* pellet (stroma); lane 3, insoluble fraction from 2,500 × *g* pellet (chloroplastic membrane); lane 4, 12,000 × *g* pellet; lane 5, 12,000 × *g* supernatant. Arrow indicates the position of APX.

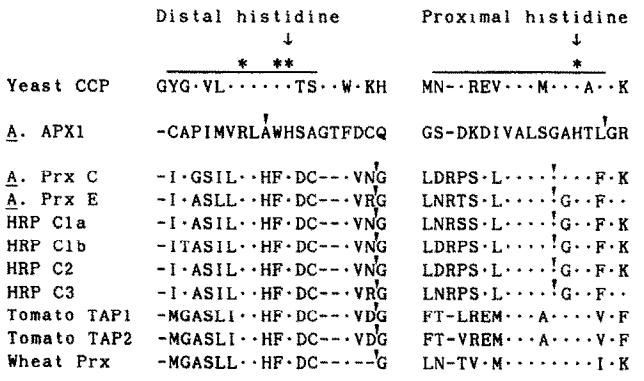


Fig. 4. Intron positions relative to the active site of Prx. Amino acid sequences near the proximal and distal histidine residues of the *Arabidopsis* cytosolic APX (APX1) are aligned with those of yeast CCP [24] and some guaiacol Prx [17–21]. Dots represent amino acid residues identical to those of APX1. Gaps in sequence, which have been introduced to improve the alignment, are indicated by dashes. Asterisks indicate the amino acid residues present at the heme pocket of CCP [23]. Amino acid residues which form the helix in CCP are overlined [23,24]. Positions in the peptide that correspond to intron placement in the coding sequence are denoted by arrowheads.

accession number M83671]. By contrast, the *APX1* gene has as many as eight introns, one of which is inserted in the 5′-untranslated region. The yeast CCP gene has no intron [22], although the primary structure of yeast CCP resembles that of APX [2–4].

The introns of the guaiacol Prx genes are inserted in the same positions, despite the fact that some of the genes lack some of the introns. Fig. 4 shows intron positions relative to the active site of Prx [23,24]. The intron splitting the active site domain seems to have been inserted after establishment of the domain in the course of evolution. The intron positions of the *APX1* gene are different from those of the guaiacol Prx genes.

The distinct difference in exon/intron organization between the *APX1* and the guaiacol Prx genes suggests a long distance between APX and guaiacol Prx in terms of molecular evolution.

Acknowledgements: We thank Dr. Y. Komeda for *Arabidopsis* DNA, Dr. S. Satoh for help in DNA sequencing, and Dr. Kunisuke Tanaka for discussions. This study was supported in part by Special Coordination Funds of the Science and Technology Agency of the Japanese Government.

REFERENCES

[1] Asada, K. (1992) Physiol. Plant. 85, 235–241.
[2] Mittler, R. and Zilinskas, B.A. (1991) FEBS Lett. 289, 257–259.
[3] Kubo, A., Saji, H., Tanaka, K., Tanaka, K. and Kondo, N. (1992) Plant Mol. Biol. 18, 691–701.
[4] Chen, G-X., Sano, S. and Asada, K. (1992) Plant Cell Physiol. 33, 109–116.
[5] Chen, G-X. and Asada, K. (1989) Plant Cell Physiol. 30, 987–998.
[6] Miyake, C. and Asada, K. (1992) Plant Cell Physiol. 33, 541–553.
[7] Dalton, D.A., Hanus, F.J., Russell, S.A. and Evans, H.J. (1987) Plant Physiol. 83, 789–794.
[8] Mittler, R. and Zilinskas, B.A. (1991) Plant Physiol. 97, 962–968.
[9] Elia, M.R., Borraccino, G. and Dipierro, S. (1992) Plant Sci. 85, 17–21.
[10] Edwards, E.A., Rawsthorne, S. and Mullineaux, P.M. (1990) Planta 180, 278–284.
[11] Saji, H., Tanaka, K. and Kondo, N. (1990) Plant Sci. 69, 1–9.
[12] Senapathy, P., Shapiro, M.B. and Harris, N.L. (1990) Methods Enzymol. 183, 252–278.
[13] Joshi, C.P. (1987) Nucleic Acids Res. 15, 6643–6653.
[14] Tanaka, K., Takeuchi, E., Kubo, A., Sakaki, T., Haraguchi, K. and Kawamura, Y. (1991) Arch. Biochem. Biophys. 286, 371–375.
[15] Tanaka, K., Suda, Y., Kondo, N. and Sugahara, K. (1985) Plant Cell Physiol. 26, 1425–1431.
[16] Tanaka, K., Masuda R., Sugimoto, T., Omasa, K. and Sakaki, T. (1990) Agric. Biol. Chem. 54, 2629–2634.
[17] Fujiyama, K., Takemura, H., Shibayama, S., Kobayashi, K., Choi, J-K., Shinmyo, A., Takano, M., Yamada, Y. and Okada, H. (1988) Eur. J. Biochem. 173, 681–687.
[18] Roberts, E. and Kolattukudy, P.E. (1989) Mol. Gen. Genet. 217, 223–232.
[19] Fujiyama, K., Takemura, H., Shinmyo, A., Okada, H. and Takano, M. (1990) Gene 89, 163–169.
[20] Hertig, C., Rebmann, G., Bull, J., Mauch, F. and Dudler, R. (1991) Plant Mol. Biol. 16, 171–174.
[21] Intapruk, C., Higashimura, N., Yamamoto, K., Okada, N., Shinmyo, A. and Takano, M. (1991) Gene 98, 237–241.
[22] Kaput, J., Goltz, S. and Blobel, G. (1982) J. Biol. Chem. 257, 15054–15058.
[23] Finzel, B.C., Poulos, T.L. and Kraut, J. (1984) J. Biol. Chem. 259, 13027–13036.
[24] Welinder, K.G. (1985) Eur. J. Biochem. 151, 497–450.