Volume 301, number 2, 185–189 FEBS 10940 April 1992

Genomic structure of the gene for copper/zinc-superoxide dismutase in rice

Atsushi Sakamoto, Takekazu Okumura, Hiroyuki Ohsuga and Kunisuke Tanaka

Department of Biochemistry, College of Agriculture, Kyoto Prefectural University, Shimogamo, Kyoto 606, Japan

Received 25 February 1992

We have isolated and determined the nucleotide sequence of the structural gene (*sodB*) coding for one of the copper/zinc-superoxide dismutase (Cu/Zn-SOD) isozymes from rice plants (*Oryza sativa* L.). The *sodB* gene is split into 8 exons spread about 2 kb of the chromosomal DNA, with the coding sequence beginning in the 2nd and ending in the 8th. Although none of *sodB* introns are inserted into similar positions in Cu/Zn-SOD genes from other sources, four of six introns in the protein-coding region are located at discriminating positions within structural domains of the protein. Genomic Southern analysis indicated that cytosolic Cu/Zn-SOD is coded for by 2–3 genes in the rice genome.

Active oxygen; Exon/intron; Genomic DNA sequence; SOD gene; Superoxide dismutase; Rice (*Oryza sativa* L.)

1. INTRODUCTION

Superoxide dismutase (SOD; superoxide:superoxide oxidoreductase, EC 1.15.1.1) catalyzes the first step in reactive oxygen-scavenging systems by conversion of superoxide anion radicals to hydrogen peroxide and molecular oxygen. This metalloenzyme is believed to be indispensable for protection of almost all living cells against oxidative stress generated under aerobic conditions [1]. According to prosthetic metals, SOD is classified into three distinct types (Mn-, Fe- and Cu/Zn-isozymes) [2]. In higher plants, the most prominent SODs are Cu/Zn-isozymes which occur as multiple isozymes, at least one of which is localized in the cytosol and the other in plastids [2]. It has been observed that the activity of plant SOD increases in response to a variety of environmental and chemical stimuli (i.e., exposure to air pollutants, high light intensities, low temperature under illumination, treatment with agents generating oxygen radicals, and so on) [3–6]. Recently, studies using specific cDNA probes have shown that the increased levels of SOD activities result from differential regulation of individual SOD genes at the transcriptional level [7,8]. Although several plant SOD cDNAs have been sequenced, thus far, no reports have demonstrated the genomic structure of plant SOD. To elucidate the regulation of this gene, a prime requisite is to understand the structure.

We have investigated the molecular basis of Cu/Zn-SOD isozymes from both spinach and rice [9,10]. Rice plants have at least two genes for cytosolic Cu/Zn-SOD as shown by cDNA analysis [10]. Furthermore, we isolated two corresponding genomic clones and characterized one of them. As the first step towards understanding gene regulation in response to environmental and oxidative stress, we report here the exon/intron structure of a cytosolic Cu/Zn-SOD gene from a rice plant, *Oryza sativa* L.

Abbreviations: SDS, sodium dodecyl sulfate; SOD, superoxide dismutase; SSC, 150 mM NaCl/15 mM trisodium citrate; 5′ exon, 5′ non-coding exon; 5′ intron, 5′ non-coding intron.

Correspondence address: K. Tanaka, Department of Biochemistry, College of Agriculture, Kyoto Prefectural University, Shimogamo, Kyoto 606, Japan. Fax: (81) (75) 781 0392.

2. MATERIALS AND METHODS

2.1. *Materials*

Two similar but distinct cDNA clones for cytosolic Cu/Zn-SOD (RSODA and RSODB) were previously isolated from a λgt11 library of developing rice seeds [10]. DNA was isolated from rice germs (*Oryza sativa* L. cv. Nipponbare).

2.2. *Construction and screening of a genomic library*

Nuclei were prepared from rice germs according to Spiker et al. [11]. High molecular weight genomic DNA was partially digested with *Sau*3AI and size-fractionated by sucrose density gradient centrifugation. DNA fragments ranging in size between 9 and 20 kb were pooled, dephosphorylated, and ligated with *Bam*HI-digested arms of the λ replacement vector EMBL3 (Stratagene, La Jolla, CA, USA). Concatenated λ DNAs were packaged using an in vitro packaging extract (Stratagene). Recombinant clones were screened for plaques which hybridized to the ^{32}P-labeled probe prepared from a rice Cu/Zn-SOD cDNA (RSODB).

2.3. *Nucleotide sequencing*

Appropriate restriction fragments from a genomic clone (gSOD7) were subcloned into the Bluescript SK(+) vector (Stratagene) and overlapping deletion mutants were generated by exonuclease III/mung-bean nuclease digestion of the asymmetrically cut plasmid DNA

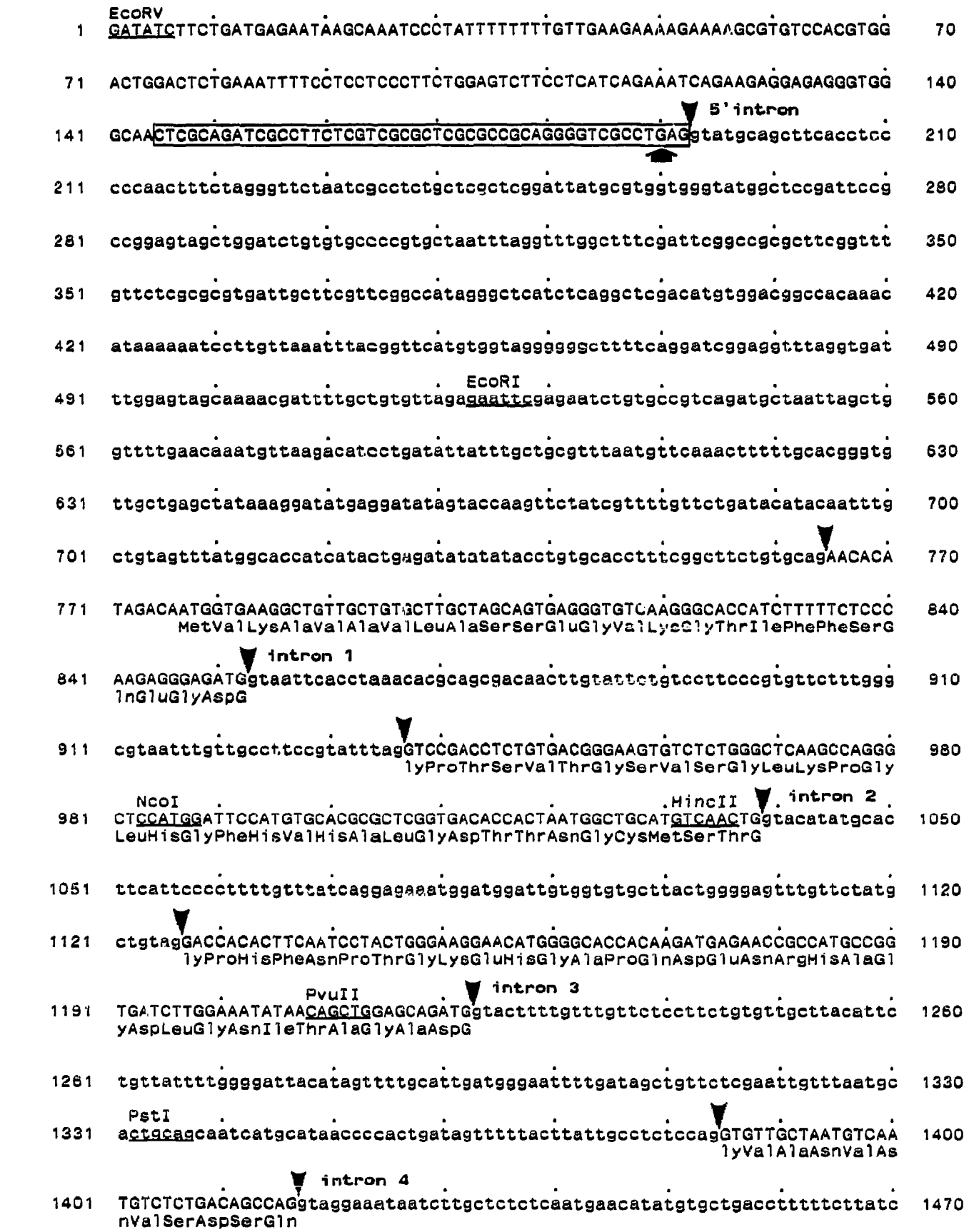

Fig. 1. For legend see next page

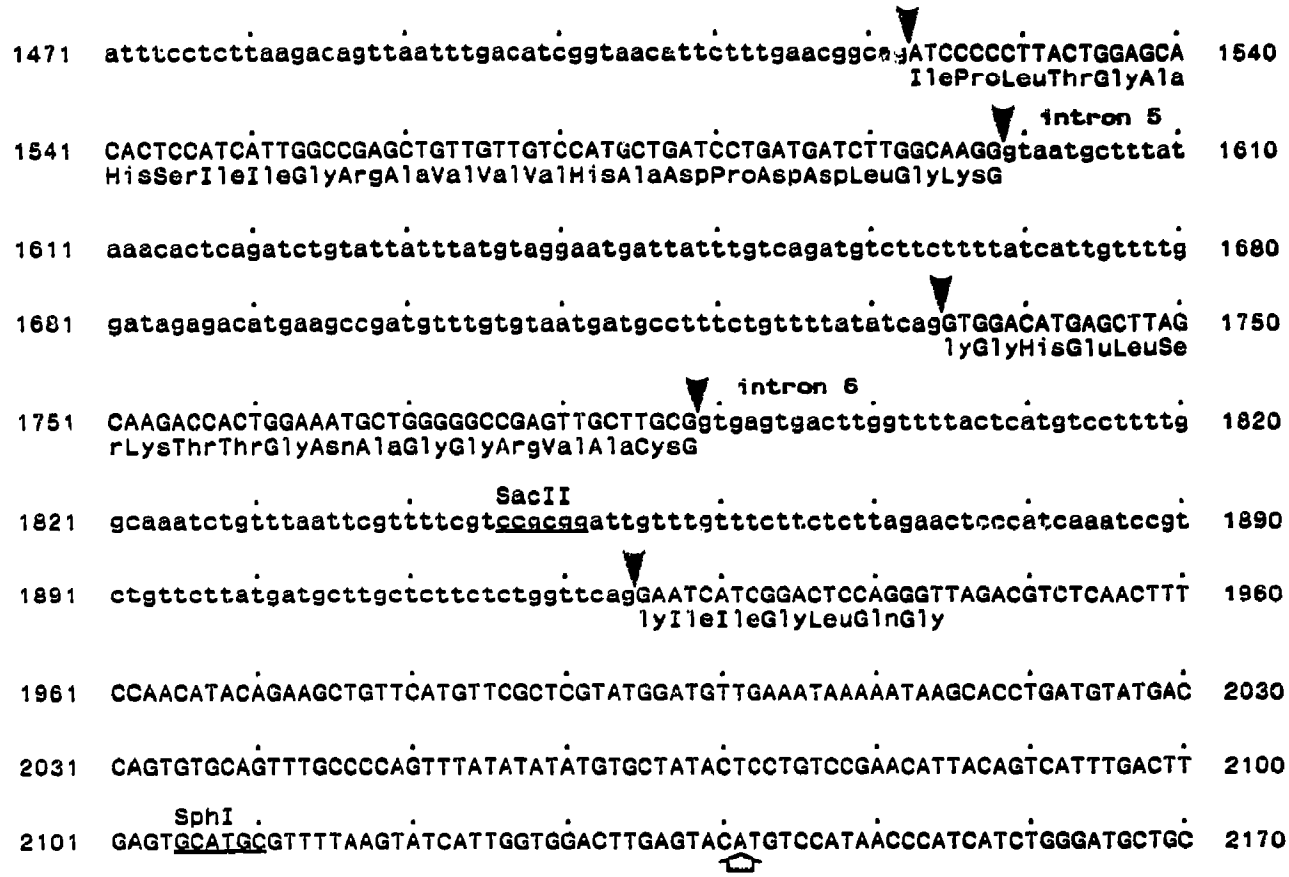

Fig. 1. Nucleotide sequence of the *sodB* gene in rice. The 5′ end of the corresponding cDNA (RSODB) is marked by a closed vertical arrow. The boxed sequence refers to a highly homologous region to the 5′ non-coding sequence of RSODA cDNA (73.5% homology) [10]. Introns are in lower case while exons and the flanking regions are in upper case letters. Several features are indicated as follows: the restriction sites are marked with underlined sequences with their names above; the polyadenylation site is marked by an open vertical arrow; exon/intron boundaries are indicated by arrowheads.

[12]. The nucleotide sequence was determined in both directions by means of automated fluorescence sequencing with dye primers using the Applied Biosystem Model 370A/373A DNA sequencing system (Applied Biosystems, Foster City, CA, USA).

2.4. *Genomic Southern analysis*

Total DNA extracted from rice germs was purified by CsCl equilibrium ultracentrifugation. Aliquots (each 10 μg) were digested with restriction endonucleases (*Eco*RI and *Eco*RV), separated on a 0.7% agarose gel and transferred to a nylon membrane [13]. Hybridization was performed using RSODA cDNA as the probe and then the membrane was washed twice at 50°C in 1 × SSC and 0.1% SDS.

3. RESULTS AND DISCUSSION

3.1. *Isolation of SOD genomic clones*

A genomic λ EMBL3 library (consisting of 1.1×10^6 independent recombinants) was screened using a previously identified rice Cu/Zn-SOD cDNA (RSODB) as the probe. Based on Southern hybridization and restriction mapping, twelve genomic clones were categorized into two representative clones, one (termed gSOD27) for RSODA cDNA and the other (termed gSOD7) for RSODB cDNA. We designated these transcriptionally active Cu/Zn-SOD genes as *sodA* (gSOD27) and *sodB* (gSOD7), respectively.

3.2. *Structure of the SOD gene*

The complete structure of the *sodB* gene was determined (Fig. 1). Structural alignment provided not only perfect agreement between the cDNA sequence and the putative exons in the *sodB* gene, but also revealed six introns in the coding region (introns 1, 84 bp; 2, 88 bp; 3, 162 bp; 4, 106 bp; 5, 136 bp; 6, 136 bp). Every intron except the fourth interrupts between the first two bases (dG-dG) of a glycine codon. Exon/intron junctions, in all cases, conform to the splice donor and acceptor consensus motifs [14]. The 5′ transcribed but untranslated region is interrupted by an additional intron (5′ intron) at 12 bp upstream from the translation start site. Insertion of the 5′ intron results in the creation of an upstream exon lacking protein-coding information (5′ exon). The occurrence of a 5′ intron in the 5′ upstream region could not be confirmed by direct sequence com-

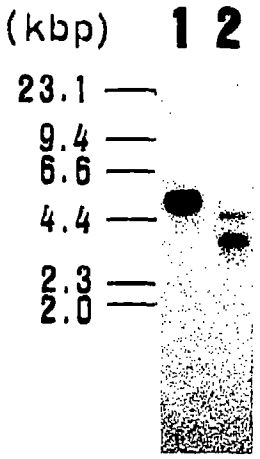

Fig. 2. Genomic Southern analysis of the rice gene encoding cytosolic Cu/Zn-SOD. Total DNA (10 μg) was digested with *Eco*RI (lane 1) or *Eco*RV (lane 2), separated on 0.7% agarose gel and transferred to a nylon membrane. The membrane was hybridized with the ^{32}P-labeled cDNA insert of RSODA. Molecular sizes were calibrated by reference to the migration of λ/*Hin*dIII DNA fragments (indicated in kilobase pairs).

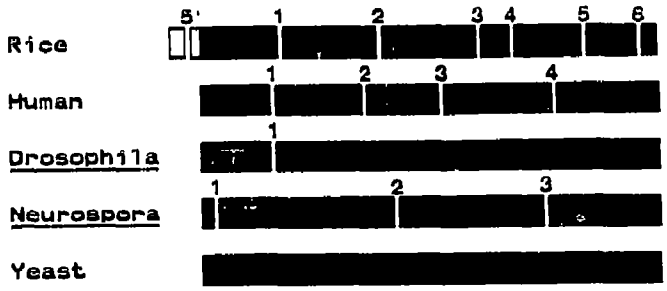

Fig. 3. Comparison of exon/intron organization of Cu/Zn-SOD genes. Closed boxes refer to protein-coding exons which are arranged for maximum matching. Introns are not shown but numbered above exon/exon junctions. Open boxes in the rice gene indicate the 5' non-coding sequences interrupted by insertion of the 5' intron (marked above as '5''). The yeast gene has no introns.

parison since the region of RSODB cDNA is quite short (15 bp). However, the possibility of the 5' intron is supported by the following observations: (i) the gene for RSODA cDNA (*sodA*) contains a 5' intron inserted at a similar position (unpublished results), and (ii) there is a highly homologous region (73.5%, boxed in Fig. 1) in the 5' upstream region of the *sodB* gene with the 5' non-coding sequence of RSODA cDNA [10], which corresponds to the 5' exon in the *sodA* gene. It is noted that the resultant 5' exon also has high sequence similarity to the 5' non-coding regions of cytosolic Cu/Zn-SOD cDNA sequences from other plant species [10]. The presumed 5' intron of the *sodB* gene is 572 bp long, corresponding to the largest one within the gene. To summarize, the *sodB* gene is composed of 8 exons interrupted by 7 introns (one in the 5' non-coding, and the rest in the structural regions, respectively), spanning about 2 kb of genomic DNA.

3.3. *Copy number of the SOD gene*

Characterization of cDNA clones has indicated that there are at least two genes coding for cytosolic Cu/Zn-SOD in rice [10]. In order to estimate the copy number of the Cu/Zn-SOD gene, rice DNA was cleaved with *Eco*RI or *Eco*RV (neither of which digests within the Cu/Zn-SOD cDNA sequences) and hybridized with RSODA cDNA as the probe (Fig. 2). A single intense signal was detected in *Eco*RI-digested genomic DNA, whereas two bands appeared in the DNA treated with *Eco*RV. In the latter, the hybridization signals must be at least composed of three bands, since the lower is denser than the upper. The restriction endonuclease *Eco*RV did not cut the structural region of the *sodB* gene (see Fig. 1), but split the *sodA* coding region into two fragments (unpublished results: the restriction site is located in an internal intron). Considering the restriction pattern of SOD genomic clones, it seems most likely that the rice genome contains 2–3 genes for cytosolic Cu/Zn-SOD.

3.4. *Comparison of SOD gene structure*

Having determined the genomic sequence for plant Cu/Zn-SOD, we compared SOD gene structures derived from a broad range of eukaryotes (Fig. 3). The SOD genes have been determined in humans [15], *Drosophila melanogaster* [16], *Neurospora crassa* [17] and yeast [18]. It appears that the exon/intron organization for plant Cu/Zn-SOD differs from those found from other sources, especially in terms of intron numbers and inserted positions (introns 1 and 2 are in close, but not identical positions to those found in human, *Drosophila* or *Neurospora* genes). Moreover, the presence of the 5' intron is unique in that it is observed only in the plant gene. Structural analysis of the crystal form of bovine Cu/Zn-SOD has indicated that modern Cu/Zn-SOD has arisen through duplication and fusion of short primitive gene(s) (exons) [19–21]. Chary and his colleagues have supported this hypothesis by comparing intron positions in *Neurospora*, human and *Drosophila* genes with the structural domains of bovine Cu/Zn-SOD [17]. In this comparison, although the inserted positions of plant introns are not conserved in other Cu/Zn-SOD genes (Fig. 3), four of the six introns in the protein-coding region are situated at or near structural junctions which distinguish between β-strands (introns 1 and 3), between loops (intron 2), or between β-strand and loop (intron 4) (Fig. 4). Therefore, this distribution of rice introns also suggests that plant Cu/Zn-SOD genes have been generated by the proposed evolutionary pathway. The existence of three other introns (introns 5, 6 and 5' intron) may be the result of more recent insertion events after the gene was established.

We are attempting to define *cis*- and *trans*-acting regulatory elements concerning the molecular mechanisms

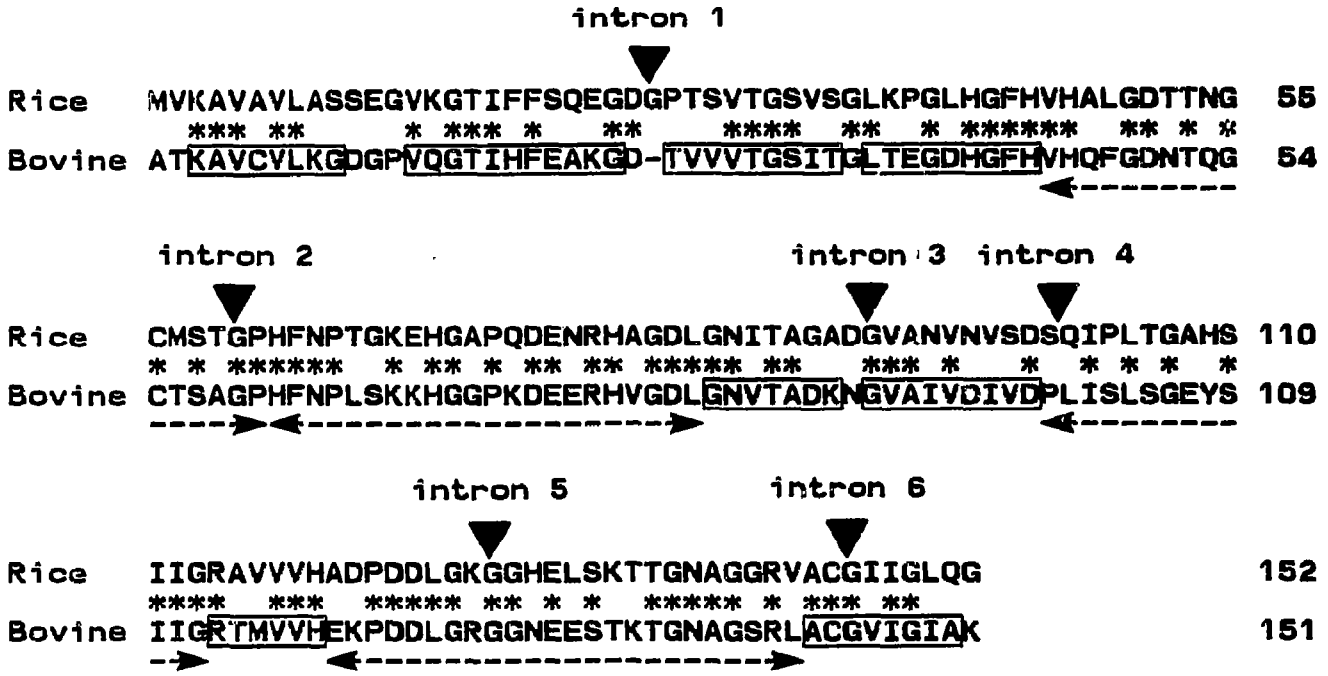

Fig. 4. Intron positions relative to the secondary structure of Cu/Zn-SOD. The deduced amino acid sequence for *sodB* gene is aligned with the primary sequence of bovine Cu/Zn-SOD [22]. *Shows identity between two sequences (57.2% identity within the rice sequence). Insertions of introns with respect to the rice sequence are marked above by triangles. Secondary structures assigned for the crystal form of bovine Cu/Zn-SOD are indicated by boxes (for β-strands) and dashed arrows (for loops) below the corresponding sequence, respectively [20].

responsible for transcriptional regulation and induction of the SOD gene in rice. In addition, we are also focusing up on how SOD participates and functions in plant defense systems. These studies are in progress using molecular approaches, including expression in transgenic tobacco and rice systems.

Acknowledgements: This work was funded in part by Grant 63860004 from the Ministry of Education, Science and Culture of Japan.

REFERENCES

[1] Fridovich, I. (1978) Science 201, 875–880.
[2] Asada, K., Kanematsu, S., Okada, S. and Hayakawa, T. (1980) in: Chemical and Biochemical Aspects of Superoxide and Superoxide Dismutase (Bannister, J.V. and Hill, H.A.O. eds.) Elsevier, Amsterdam, pp. 136–153.
[3] Tanaka, K. and Sugawara, K. (1980) Plant Cell Physiol. 21, 601–611.
[4] Lee, E.H. and Bennet, J.H. (1982) Plant Physiol. 69, 1444–1449.
[5] Furusawa, I., Tanaka, K., Thanutong, P., Mizuguchi, A., Yazaki, M. and Asada, K. (1984) Plant Cell Physiol. 25, 1247–1254.
[6] Rabinowitch, H.D. and Sklan, D. (1980) Planta 148, 162–167.
[7] Tsang, E.W.T., Bowler, C., Herouart, D., Van Camp, W., Villarroel, R., Genetello, C., Van Montagu, M. and Inze, D. (1991) Plant Cell 3, 783–792.
[8] Perl-Treves, R. and Galun, E. (1991) Plant Mol. Biol. 17, 745–760.
[9] Sakamoto, A., Ohsuga, H., Wakaura, M., Mitsukawa, N., Hibino, T., Masumura, T., Sasaki, Y. and Tanaka, K. (1990) Nucleic Acids Res. 18, 4923.
[10] Sakamoto, A., Ohsuga, H. and Tanaka, K. (1992) Plant Mol. Biol. (in press).
[11] Spiker, S., Murray, M.G. and Thompson, W.F. (1983) Proc. Natl. Acad. Sci. USA 80, 815–819.
[12] Henikoff, S. (1984) Gene 28, 351–359.
[13] Reed, K.C. and Mann, D.A. (1985) Nucleic Acids Res. 13, 7207–7221.
[14] Mount, S.M. (1982) Nucleic Acids Res. 10, 459–472.
[15] Levanon, D., Lieman-Hurwitz, J., Dafni, N., Wigderson, M., Sherman, L., Bernstein, Y., Laver-Rudich, Z., Danciger, E., Stein, O. and Groner, Y. (1985) EMBO J. 4, 77–84.
[16] Seto, N.O.L., Hayashi, S. and Tener, G.M. (1987) Nucleic Acids Res. 15, 10601.
[17] Chary, P., Hallewell, R.A. and Natvig, D.O. (1990) J. Biol. Chem. 265, 18961–18967.
[18] Bermingham-McDonogh, O., Gralla, E.B. and Valentine, J.S. (1988) Proc. Natl. Acad. Sci. USA 85, 4789–4793.
[19] MacLachlan, A.D. (1980) Nature 285, 267–268.
[20] Tainer, J.A., Getzoff, E.D., Beem, K.M., Richardson, J.S. and Richardson, D.C. (1982) J. Mol. Biol. 160, 181–217.
[21] Getzoff, E.D., Tainer, J.A., Stempien, M.M., Bell, G.I. and Hallewell, R.A. (1989) Proteins 5, 322–336.
[22] Steinman, H.M., Naik, V.R., Abernethy, J.L. and Hill, R.L. (1974) J. Biol. Chem. 249, 7326–7338.