

## Isolation of two cDNA clones from tomato containing two different superoxide dismutase sequences

Rafael Perl-Treves,<sup>1\*</sup> Benedetta Nacmias,<sup>1</sup> Dvora Aviv,<sup>1</sup> Elisha P. Zeelon<sup>2</sup> and Esra Galun<sup>1\*</sup>

<sup>1</sup>*Department of Plant Genetics, The Weizmann Institute of Science, Rehovot 76100, Israel;* <sup>2</sup>*Biotechnology General, Ness Ziona, Israel (\*authors for correspondence)*

Received 8 March 1988; accepted in revised form 16 August 1988

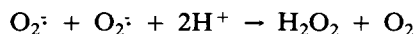
**Key words:** Cu,Zn superoxide dismutase, oligonucleotide probes, nucleotide sequences, isozymes, *Lycopersicon esculentum*

### Abstract

A cDNA library was derived from the poly(A)<sup>+</sup> RNA of young tomato leaves. The library was cloned in a  $\lambda$ gt11 system and screened by synthetic oligonucleotide probes having sequences that match the codes of conserved regions of amino acid sequences of Cu,Zn superoxide dismutase (SOD) proteins from a wide range of eukaryotic organisms. Two cDNAs were isolated, cloned and sequenced. One of the cDNAs, P31, had a full-size open reading frame of 456 bp with a deduced amino acid sequence having an 80% homology with the deduced amino acid sequence of the cytosolic SOD-2 cDNA of maize. The other cDNA, T10 (extended by T1), had a 651 bp open reading frame that revealed, upon computer translation, 90% homology to the amino acid sequence of mature spinach chloroplast SOD. The 5' end of the reading frame seems to code for a putative transit peptide. This work thus suggests for the first time an amino acid sequence for the transit peptide of chloroplast SOD. Northern hybridizations indicated that each of the P31 and T10 clones hybridized to a blotted poly(A)<sup>+</sup> RNA species. These two species are differentially expressed in the plant organs: e.g., the species having the T10 sequence was detected in the leaves but not in roots, while the one with the P31 sequence was expressed in both leaves and roots. The cDNA clones P31 and T10 were also hybridized to Southern blots of endonuclease fragmented tomato DNA. The clones hybridized to specific fragments and no cross hybridization between the two clones was revealed under stringent hybridization conditions; the hybridization pattern indicated that, most probably, only one locus is coding for each of the two mRNA species.

### Introduction

Superoxide dismutase (SOD; EC 1.15.1.1) is a house-keeping metalloenzyme of aerobic organisms. It catalyzes the disproportionation of superoxide radicals (O<sub>2</sub><sup>•−</sup>) in the following general equation:



(see [11] for review).

Plants contain three types of SOD: Cu,Zn-, Mn- and Fe-containing enzymes [1, 25]. The three types of SOD catalyze the same reaction with similar rates, but the amino acid sequence of Cu,Zn SOD differs substantially from the respective sequences of Mn SOD and Fe SOD. The latter are considered to differ from the former in phylogeny. There are at least two Cu,Zn SOD isozymes, one being imported into the chloroplasts and another remaining in the cytosol. In the chloroplast stroma, SOD constitutes the first

link in a cascade of enzymes and antioxidants which protect the photosynthetic apparatus from noxious oxygen derivatives [25]. In maize, two cytosolic (Cu,Zn SOD), one chloroplastic (Cu,Zn SOD) and one mitochondrial (Mn SOD) enzymes have been described [4, 5, 6], all of which are coded by nuclear genes. Antibodies against purified Cu,Zn SOD isozymes of maize were raised and a cDNA clone coding for the cytosolic SOD-2 was isolated and sequenced from this plant [7]. SOD enzymes from several other plants have been purified, such as spinach [2] and pea [10]. Spinach chloroplast Cu,Zn SOD has been isolated and its amino acid sequence reported [16].

We used the amino acid sequence information available for mammalian, fish, insect, yeast and spinach Cu,Zn SOD proteins to construct oligonucleotide probes covering the most conserved regions of these enzymes, to isolate SOD sequences from a tomato (*Lycopersicon esculentum*) cDNA library. Two different cDNA clones were identified, which most probably contain the coding sequences for the cytosolic and the chloroplast forms of Cu,Zn SOD.

## Materials and methods

### Plant material

Tomato seedlings of a cherry-type cultivar (E.C., an inbred maintained in our laboratory) were grown in the greenhouse. For the cDNA library preparation, RNA was extracted from two-week-old seedlings.

### Plant RNA extraction and fractionation

We used mainly two protocols. The first was a modification of method 2 in [15]. Leaves (95 g) were blended and homogenized for 10 min in 100 ml of 6% sodium-4-amino salicylate and 100 ml phenol-cresol mix. In total, four extractions with the organic mix and (after ethanol precipitation) three extractions with 3 M sodium acetate were performed, to yield 90 mg of undegraded RNA. The second method, used both for large quantities and minipreps of

total RNA, is from Logemann *et al.* [18]. After sodium acetate extraction, a LiCl extraction was added [23]. RNA fractionation on oligo-dT cellulose (Pharmacia) was according to Aviv and Leder [3].

### Construction of a cDNA library

Materials for cDNA synthesis were from the Amersham cDNA synthesis kit, other enzymes from New England Biolabs. Phosphatase  $\lambda$ gt11 arms and packaging extracts were from Promega.

First-strand synthesis was according to the "Amersham 5  $\mu$ g protocol" with 10  $\mu$ g of poly(A)<sup>+</sup>-enriched RNA as starting material, and the addition of 2  $\mu$ g actinomycin D. Phenol-chloroform extractions, followed by ammonium acetate and ethanol precipitation, were performed between enzymatic steps [12]; the total Cerenkov counts in the pellet were then measured.

Second-strand synthesis was according to the Okayama and Berg [22] and Gubler and Hoffman [12], followed by T4 DNA polymerase treatment (6 units in 50 mM Tris pH 8, 6 mM MgCl<sub>2</sub>, 25 mM KCl, 0.1  $\mu$ g/ml bovine serum albumin (BSA), 0.1 mM each dNTP in 20  $\mu$ l, 30 min at 37 °C). *Eco* RI methylase was applied (10 units in 80  $\mu$ M S-adenosyl methionine, 0.4 mg/ml BSA, 100 mM Tris pH 8, 1 mM EDTA in 10  $\mu$ l, 20 min at 37 °C). Phosphatase *Eco* RI linkers (Pharmacia) were ligated to cDNA and cut 3 times with excess *Eco* RI. Cut linkers were removed by passage through an S-300 Sephacryl column. The cDNA (200 ng) was ligated to 5  $\mu$ g  $\lambda$ gt11 arms and packaged with two portions of packaging extract, resulting in about 10<sup>6</sup> plaque-forming units, of which 65% were recombinants.

### Origins and use of probes

A human SOD-1 cDNA clone (h-SOD) in pBR322 was a gift of Prof. Y. Groner (Dept. of Virology, Weizmann Institute of Science). Nick translation reactions were according to Maniatis *et al.* [20].

Oligonucleotides were synthesized using an Applied Biosystems machine, model 380B and labelled with T4 kinase (NEN) [20]. Free ATP was removed

by a DEAE 52 column; the eluate was frozen and used for 3–4 weeks. Radioactivity of 100000 cpm (Cerenkov) per ml of hybridization reaction were found to be sufficient.

#### *Screening the cDNA library by plaque hybridization*

The cDNA library in the vector  $\lambda$ gt11 (460000 recombinants, non amplified) was grown on *Escherichia coli* strain Y1090 in LB medium with ampicillin. Duplicate nitrocellulose filters were prepared [20]. Prehybridization was done for at least 3 h in  $6\times$  SSC,  $5\times$  Denhardt's solution, 0.5% SDS and 50–100  $\mu$ g of denatured, sonicated salmon DNA. The probe was then added without buffer change and hybridization was performed overnight with agitation. The temperatures of prehybridization and hybridization were 37°C (and down to 25°C) for h-SOD and oligonucleotides B and D, and 45°C for oligonucleotides P and T. After hybridization, a few washes in  $2\times$  SSC, 0.25% SDS at 37–42°C were applied. The same nitrocellulose filters were used for screens with different probes up to 4 times, the old probe being washed off by 40 min in  $0.1\times$  SSC, 0.25% SDS at 70–80°C. After hybridization, positive plaques were picked, diluted, replated and rehybridized; 3–4 such cycles were done to purify library-isolates.

#### *Sequence analysis of cDNA clones*

*Eco* RI (and *Eco* RI-*Xba* I) fragments from  $\lambda$ gt11 were subcloned into the Bluescript KS-minus vector (Stratagene). VCS-M13 Helper Phage (Stratagene) was added to logarithmic-phase cultures and, after overnight co-culture, single-strand DNA for Sanger sequencing was prepared [26]. An IBI Sequel kit was used for sequencing. Sequence data were analyzed using UWGCG computer programs [9]. Nested deletion clones were prepared by exonuclease-III digestion using the "Erase a Base" kit of Promega.

#### *RNA gels and Northern blots*

RNA samples (and control DNA markers) were glyoxylated [8] and run on 1.2% agarose gels in circulating 12.5 mM buffer phosphate pH 6.5. Blotting to GeneScreen Plus (Du Pont) and hybridization were according to the supplier (50% formamide method), with the application of UV irradiation after blotting [14].

#### *Tomato DNA extraction*

Total DNA extraction from young tomato leaves was performed according to Pazkowski and Saul [24].

#### *Southern blots and hybridization*

DNA separation was performed on 1% agarose gels in TAE buffer [20]. Restriction endonucleases were from New England Biolabs or Pharmacia. Gels were blotted to GeneScreen Plus membrane (Du Pont, blotting and hybridization protocols of supplier). Following blotting, the DNA was fixed by UV irradiation [14].

### **Results**

#### *Strategy for the isolation of Cu,Zn SOD coding sequences from tomato*

We based our search for Cu,Zn SOD genes in tomato on the information that has become available in recent years on the amino acid sequences of Cu,Zn SODs in a wide range of eukaryotes. This information includes the full sequences of bovine, human, equine, fish, fly, yeast and spinach enzymes [16], as well as the amino acid sequence deduced from maize cytosolic SOD-2 cDNA [7]. Four single-stranded DNA probes of anti-mRNA orientation measuring 17 to 23 bases were thus constructed to match the codes for highly conserved amino acid sequences. Where specific amino acids varied among these conserved sequences, we prepared the oligonucleotides according to the spinach amino acid sequence be-

cause of its phylogenetic relatedness to tomato. For codons that had several possibilities at the 3rd position we either inserted deoxiinosine or split the synthesis into 2 or 4 parts, so that a compound mixture of oligonucleotides was generated. These probes, as well as a human SOD-1 cDNA probe, were then used to screen the tomato cDNA library.

#### *The cDNA library and its screening*

RNA from the foliage of two-week-old tomato seedlings was fractionated on an oligo-dT column, and 10 µg of the poly(A)<sup>+</sup> fraction were used, yielding 600 ng cDNA; 200 ng were cloned in λgt11. The

original library containing 460000 transformants was screened by hybridization with the human SOD-1 clone (applied to the whole library at 37°C, 6× SSC), as well as with each of the four synthetic oligonucleotide probes (T, P, B, D), each applied to half of the library.

All 4 oligonucleotides were tested and shown to hybridize with the human SOD (h-SOD) plasmid, but only two of them, P and T, succeeded to select clones from the cDNA library. These two are shown in Table 1, together with the corresponding sequence in maize, spinach and human SODs, as well as in the two clones they selected from the tomato library.

The thermal conditions for the oligonucleotide hybridization were based on the expected CG and

Table 1. Oligonucleotide probes used in this study and the corresponding amino acid sequences in various SODs.

| Probe P                 |    |                         |     |                      |                 |                         |                 |                         |    |
|-------------------------|----|-------------------------|-----|----------------------|-----------------|-------------------------|-----------------|-------------------------|----|
| Amino acid sequence     | 3' | T                       | G   | P                    | H               | F                       | N               | P                       | 5' |
| Synthesized sequence    |    | TGI                     | CCI | GGI                  | GT <sub>2</sub> | AA <sub>2</sub>         | TT <sub>2</sub> | GG                      |    |
| Human SOD-1             |    | A<br>(CGI)              | *   | *                    | *               | *                       | *               | *                       |    |
| Spinach chloroplast SOD |    | *                       | *   | *                    | *               | *                       | *               | *                       |    |
| Maize cytosolic SOD-2   |    | *                       | *   | *                    | *               | *                       | *               | *                       |    |
| Clone T10               |    | *                       | *   | A<br>(CGI)           | *               | *                       | *               | *                       |    |
| Clone P31               |    | *                       | *   | *                    | *               | Y<br>(AT <sub>2</sub> ) | *               | *                       |    |
| Probe T                 |    |                         |     |                      |                 |                         |                 |                         |    |
| Amino acid sequence     | 3' | T                       | T   | G                    | N               | A                       | G               | G                       | R  |
| Synthesized sequence    |    | TGI                     | TGI | CC <sub>2</sub><br>A | TT <sub>2</sub> | CG <sub>2</sub><br>A    | CCI             | CCI                     | GC |
| Human SOD-1             |    | K<br>(TT <sub>2</sub> ) | *   | *                    | *               | *                       | *               | S<br>(TC <sub>2</sub> ) | *  |
| Spinach chloroplast SOD |    | *                       | *   | *                    | *               | *                       | *               | *                       | *  |
| Maize cytosolic SOD-2   |    | S<br>(TC <sub>2</sub> ) | *   | *                    | *               | *                       | *               | *                       | *  |
| Clone T10               |    | *                       | *   | *                    | *               | *                       | *               | *                       | *  |
| Clone P31               |    | S<br>(TC <sub>2</sub> ) | *   | *                    | *               | *                       | *               | *                       | *  |

The nucleotide sequences of the synthetic probes P and T are presented. The synthesized strand was anti-mRNA; the corresponding amino acids are shown above it. The corresponding amino acid sequence of spinach SOD [16], and the deduced sequences from human SOD-1 [27] and maize SOD-2 [7] and the T10 and P31 tomato cDNA clones are shown below. When the reaction was split into 2 or 4 parts, the added nucleotides are written vertically.

\*Amino acids of the different sequences that conform to the corresponding sequences in the probes; only triplets that differ from the probe were written explicitly. I = deoxiinosine.

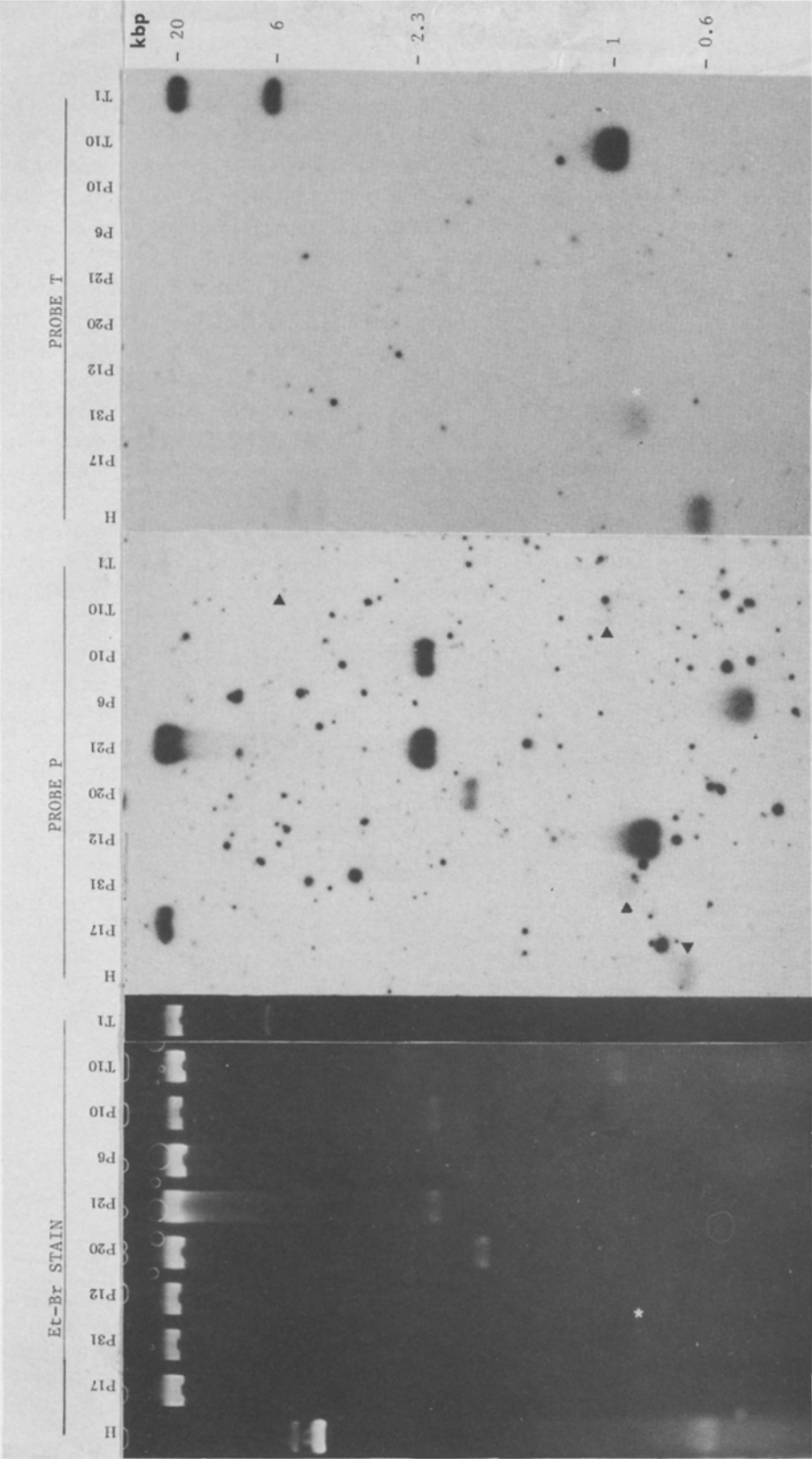


Fig. 1. Hybridization of cDNA library-isolates with oligonucleotide probes. DNA from several tomato cDNA library-isolates was digested with *Eco* RI to release cDNA inserts (P17 and T1 did not release *Eco* RI inserts; T1 was cut with *Eco* RI and *Xba* I) and run on an agarose gel alongside with *Pst* I-digested human SOD-cDNA (lane H). Gels were blotted and hybridized to end-labelled oligonucleotides P or T at 45 °C. Washes with 2× SSC, 42 °C were followed by a 2 min wash at 60 °C for probe T and 50 °C for probe P in 6× SSC. Clones P31, P17, P6, P10, P21, P12 and P20 were isolated by P probe screens, while T1 and T10 by probe T. The clones released inserts of 2.3 – 0.5 kbp, which showed different degrees of hybridization to the probes. Only clones T1, T10 and P31 did show cross hybridization to the second probe. White asterisks and black triangles mark weaker bands.

AT pairing between the probes and the respective cDNA sequence to which they should hybridize [21]. Thus the  $T_m$  estimates for the hybrids with probes P and T were 50°C and 60°C respectively (calculated as 4°C for every GC pair and 2°C for AT pairs, assuming no mismatches). These probes were thus hybridized at 44–46°C in 6× SSC. The expected  $T_m$  of probes B and D were 48°C and 50°C, but no positive hybridization was detected with either of these probes (even at low stringencies of 37°C to 25°C), nor with the h-SOD probe.

Two clones strongly hybridized to probe T and 5 clones hybridized, at varying strengths, to probe P (Fig. 1). Five of these clones were further analyzed: clones T1 and T10 strongly hybridized with probe T and weakly with probe P; clone P31 showed medium and weak hybridizations with probes P and T, respectively; clones P12 and P20 hybridized strongly with probe P only. Under the above conditions, almost no hybridization of these clones with the h-SOD probe could be detected (not shown).

### Sequencing of cDNA clones

The inserts from the  $\lambda$ gt11 clones P12, P20, P31, T1 and T10 were subcloned into Bluescript plasmids. Single-stranded DNA was prepared for sequencing. Subclones containing the two opposite orientations of the inserts were identified for most clones. The suffix designations A and B (e.g. T10A and T10B) refer, respectively, to clones yielding mRNA-like reading, starting at the 5' end of the cDNA, and to clones which yield the complementary reading sequences.

The sequences read were computer-translated in all possible frames and the derived amino-acid sequences were compared to those of spinach chloroplast Cu,Zn SOD and of maize cytosolic Cu,Zn SOD-2. Clones T10 and P31 readily showed strong resemblance to the respective SODs. Such a resemblance was not detected in clones P10 and P12, upon partial sequencing.

The sequencing strategy for reading the base sequences of clones T1 and T10 is summarized in Fig. 2; a similar strategy was used for sequencing

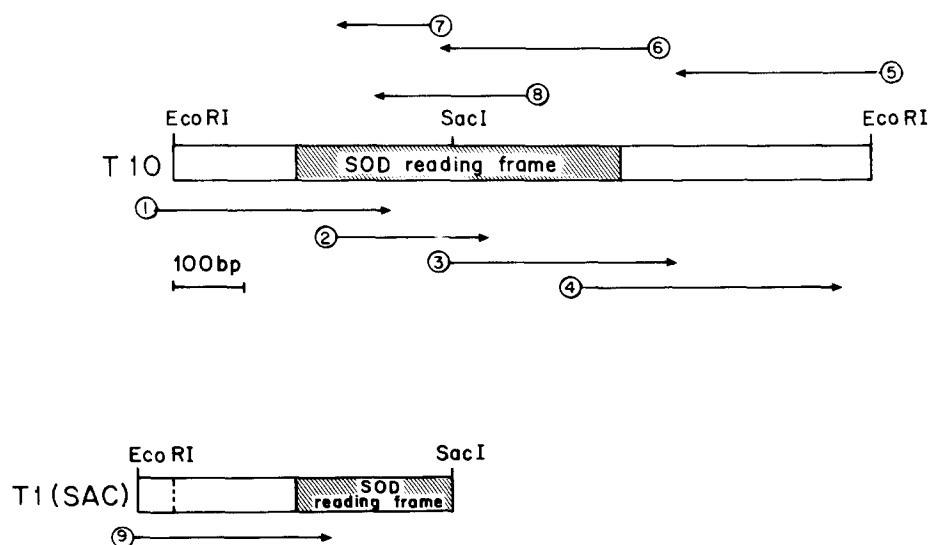


Fig. 2. Sequencing schemes of the T1 and T10 clones. Bluescript plasmids containing *Eco* RI inserts of T10, in both orientations (denoted as A and B), were used. T1 had an *Eco* RI-*Xba* I insert and was subcloned by *Sac* I cleavage to produce the T1 (SAC) subclone having a sequence which was identical to T10A except for a 44 bp extension at the former's 5' end. The shaded box is the coding region for the putative mature SOD protein. This box is preceded by a putative leader sequence. Arrows define the direction and length of the reading performed. Numbers in circles on the rear of arrows describe the clones and the primers used for sequence reading. 1, clone T10A with primer T3 (primer of the Bluescript vector); 2, clone T10A with primer X; 3, a T10A deletion-clone with primer T3; 4, a T10A deletion-clone with primer T3; 5, clone T10B with primer T3; 6, a T10B deletion-clone with primer T3; 7, clone T10B with primer Y; 8, clone T10B with primer T (probe T); 9, clone T1 and its sub-clone T1 (SAC) with primer T3.

clone P31. We used exonuclease-III digestions to read further into the cDNA inserts, by generating subclones with various deleted sizes. Furthermore, to bridge remaining gaps we synthesized 17 base oligonucleotides, (denoted in Figs. 2, 3, 4 as X, Y and W) to prime the borders of the gaps so that the readings could be completed and confirmed.

The inserts of clones P31 and T10 measured 860 and 1020 bp, respectively. The base sequences of T10 and T1 appeared to be identical but the latter's insert extended 44 bp, at the 5' end, beyond the T10 sequence. Upon *Sac* I cleavage of T1, a sub-clone (T1-SAC) was obtained, which included the 5' extension and a precise overlap with the T10 clone. Figures 3 and 4 show the complete sequences of the two cDNAs.

#### *Northern and Southern blot hybridizations*

The inserts of the plasmids containing the cDNA clones P31 and T10 were labelled and used as probes in blot hybridizations.

Total RNA, poly(A)<sup>+</sup> RNA and poly(A)<sup>-</sup> RNA from young tomato leaves were glyoxylated and applied to agarose gels for Northern blot hybridization. *Eco* RI digests of P31 and T10 (100 pg of insert) were also applied to the same gel as internal control. Stringent hybridization conditions (50% formamide, 42 °C) were applied (Fig. 5). The internal controls indicated that the two probes did not cross-hybridize. Each probe recognized an abundant mRNA species from the poly(A)<sup>+</sup> fraction. The sizes of these RNAs were similar to those of the clones inserts, i.e. about 1 kb. It was difficult to differentiate between these poly(A)<sup>+</sup> RNAs on the basis of their migration on such gels.

Total-RNA samples (20 µg) from different plant organs were similarly blotted and hybridized with each of the two cDNA probes (Fig. 6). Both RNA species were revealed in appreciable amounts in young leaves, as well as in young and full-size fruits, and in flowers (not shown). In roots and stems the T10 type RNA was undetected while the P31 type was abundant. This, and additional blots (not shown), indicated a quick diminution of SOD mRNAs with the expansion of the leaves.

For Southern blot hybridization about 10 µg total tomato DNA was digested by endonucleases and run on agarose gels. Internal control lanes contained *Eco* RI digests of P31 and T10 as in the Northern blots. The blots were hybridized with either of the two labelled probes P31 and T10 (Fig. 7). The two probes did not recognize each other but did hybridize to discrete tomato DNA fragments. Under stringent conditions, the hybridization pattern with many enzymes was of two bands. We believe these do not represent two loci because upon hybridization with a shorter probe (half of the cDNA insert) some of the fragments disappeared (not shown). Restriction sites that are not present in the cDNA indicate the existence of introns in the genes. Signal strength, as compared to dilutions of the plasmid controls, did not suggest more than one copy on the genomic fragments.

#### **Discussion**

##### *Use of oligonucleotides as heterologous probes*

Clone P31 had one nucleotide mismatch with respect to the synthetic probe P, while clones T1 and T10 had a perfect match with probe T. In Table 1 and Fig. 1 actual hybridization signals can be compared to the sequence homology between probes and target sequences. It is apparent that choice of proper stringency conditions was critical; the data also indicated that it is important to use a set of such probes in parallel.

##### *Identification and comparison of the two tomato cDNA clones*

Figure 8 shows the amino acid sequence alignment of four SOD enzymes: maize SOD-2 cDNA, the two tomato clones (this work), and spinach chloroplast SOD. The latter [16] was derived by sequencing the chloroplast SOD protein (K. Asada, personal communication) and represents therefore the mature protein, while the one from maize represents a cytosolic Cu,Zn SOD [7].

By counting the amino-acid substitutions between

```

      10              30              50
1  CAAAAATGGTGAAGGCCGTGCGCGTCCTTAACAGCAGTGAAGGTGTTAGTGGCACCTACC 60
   M V K A V A V L N S S E G V S G T Y L

      70              90              110
61 TCTTCACTCAAGTAGGAGTAGCTCCAACCACAGTTAATGGAAATATTTCTGGCCTAAAAC 120
   F T Q V G V A P T T V N G N I S G L K P

      130              150              170
121 CTGGACTTTCATGGCTTCCATGTCCATGCCCTTGGTGATACCACAAATGGCTGTATGTCAA 180
   G L H G F H V H A L G D T T N G C M S T

      P 190              210              230
181 CAGGACCACATTACAATCCTGCTGGTAAGGAGCATGGTGCTCCTGAAGATGAGGTGCGTC 240
   G P H Y N P A G K E H G A P E D E V R H

      250              270              290
241 ATGCTGGTGATCTTGGTAACATCACAGTTGGAGAAGATGGTACTGCATCTTTTACTATTA 300
   A G D L G N I T V G E D G T A S F T I T

      310              330              350
301 CCGACAAGCAGATTCCCTCTCACTGGTCCACAGTCCATCATTGGAAGAGCTGTTGTTGTTTC 360
   D K Q I P L T G P Q S I I G R A V V V H

      370              390              410
361 ATGCTGATCCTGATGATCTTGGAAAGGGAGGACATGAGCTCAGTAAAAGCACCAGGAAATG 420
   A D P D D L G K G G H E L S K S T G N A

      430              450              470
421 CTGGCGGAAGGATTGCTTGTGGTATTATTGGCCTCCAGGGTTAACTGCTCCGGTGAGGTG 480
   G G R I A C G I I G L Q G *

      490              510              530
481 AAGATTCAGCAACTAGCGGTGGTGTCTGTCTTAGAATAAAGTTATATTGGAGCTGCTGAG 540

      550              570              590
541 CCTTGTGTTTCTATGTATTTATCATTTGACGCCCTTTTGAATTGGGGTTGAACCATTCATAA 600

      610              630              650
601 TACATGACTGTAGTGCTGCTTTTCAGTGGATGTTGTACTAGTTTTTGCCGATCATTAAAC 660

      670              690              710
661 ATAAAAAACAGGGCTGTGTTAGCTTTGCTGGTTTCTTAAATTTGAGTGGATGGTTTGA 720
   W

      730              750              770
721 AACTAGACAAAATTTAGATGAATTTGCAGATTACAGTGCCTGCAGTTTTTGTTCATA86 777

```

Fig. 3. The complete sequence of the cDNA clone P31. The nucleotide sequence and the deduced amino acid sequence of P31 are shown. The open reading frame consists of 152 amino acids and the total length of the cloned insert is 863 bp. The sites corresponding to the oligonucleotide probes T and P as well as to the primer W are shown as lines above the nucleotide sequence.

|     |                                                                  |                     |       |     |
|-----|------------------------------------------------------------------|---------------------|-------|-----|
|     | 10                                                               | 30                  | 50    |     |
| 1   | CATT TTTATCAATT CATCAT TGTGGCAGCAGCATGGCCGCCCACTCAATCTTCACCACAAC |                     |       | 60  |
|     |                                                                  | M A A H S I F T T T |       |     |
|     | 70                                                               | 90                  | 110   |     |
| 61  | CAGCACTACCAATTCTTTCTTATACCCAATTTCTTCTTCTTCTTCTTCCCTAACATTAA      |                     |       | 120 |
|     | S T T N S F L Y P I S S S S S S P N I N                          |                     |       |     |
|     | 130                                                              | 150                 | 170   |     |
| 121 | CTCTTCTTTCTAGGTGTTTCACTTAATGTTAATGCAAAATTTGGCCAATCTTTGACCCT      |                     |       | 180 |
|     | S S F L G V S L N V N A K F G Q S L T L                          |                     |       |     |
|     | 190                                                              | 210                 | 230   |     |
| 181 | TTATGCTGTCACTACCCCTAAACCCCTTACTGTTTTTGCTGCTACTAAGAAAGCTGTTGC     |                     |       | 240 |
|     | Y A V T T P K P L T V F A A T K K A V A                          |                     |       |     |
|     | 250                                                              | 270                 | 290   |     |
| 241 | TGTCCTTAAGGGCAACTCTAATGTTGAGGGGTTGTCACTCTATCTCAAGATGATGATGG      |                     |       | 300 |
|     | V L K G N S N V E G V V T L S Q D D D G                          |                     |       |     |
|     | 310                                                              | 330                 | 350   |     |
| 301 | GCCAACTACGGTTAATGTTGCTATAACTGGACTTGCTCCTGGACTTCACGGGTTTCATTT     |                     |       | 360 |
|     | P T T V N V R I T G L A P G L H G F H L                          |                     |       |     |
|     | 370                                                              | 390                 | P 410 |     |
| 361 | ACATGAATATGGTGACACTACAAACGGATGTATGTCTACAGGAGCCCATTTTAATCCGAA     |                     |       | 420 |
|     | H E Y G D T T N G C M S T G A H F N P N                          |                     |       |     |
|     | 430                                                              | 450                 | Y 470 |     |
| 421 | TAAATTGACACATGGAGCTCCTGGAGATGAAATCCGTCATGCGGGTGACCTGGGAAACAT     |                     |       | 480 |
|     | K L T H G A P G D E I R H A G D L G N I                          |                     |       |     |
|     | 490                                                              | 510                 | 530   |     |
| 481 | AGTGGCCAATGCCGATGGCGTGGCTGAAGTGACACTTGTAGATAATCAGATACCATTGAC     |                     |       | 540 |
|     | V A N A D G V A E V T L V D N Q I P L T                          |                     |       |     |
|     | 550                                                              | 570                 | 590   |     |
| 541 | CGGTCCAAACTCAGTTGTTGGACGAGCACTTGTGGTTCATGAGCTTGAGGATGACCTCGG     |                     |       | 600 |
|     | G P N S V V G R A L V V H E L E D D L G                          |                     |       |     |
|     | 610                                                              | 630                 | I 650 |     |
| 601 | AAAGGGTGGCCATGAACTCAGTCTTACCCTGGCAATGCTGGTGGACGATTGGCATGCGG      |                     |       | 660 |
|     | K G G H E L S L T T G N A G G R L A C G                          |                     |       |     |
|     | 670                                                              | 690                 | 710   |     |
| 661 | TGTGGTTGGTTTGA CTCCAATATGAAGTAAATGCAGCAGCAACAACAGCTCTTTTGTGG     |                     |       | 720 |
|     | V V G L T P I *                                                  |                     |       |     |

```

              730              750              770
721  AATTATAATGAGTGAACGTATTTACCCTCCTCTCCGGGTGCCAATAAACCTCTACATGA  780

              790              810              830
781  ATTTGTAAGTGAAGAAATTCCTTTGATGAAGTAAATTTGAAGTTGTATGTACACACTTGCCAA  840

              850              870              890
841  CACTTTTTTGTCTTCAAACTGTAATTGGAGTTGGCTATTTACCTTCCTTTGTGTTTGGC  900

              910              930              950
901  TATCTAAAGTTACACTCAGTTAACTAAGTGTTGTCTTGGATTTTGAAGTCTGTGCATGTA  960

              970              990              1010
961  TAGTTTCCTCGAGACGAGGTCTATCAGAAACAACAATCTGTCATTGATTAAAGATCAGATT  1020

              1030
1021  AGTTTATCACTTCA19 1034

```

Fig. 4. The complete sequence of the cDNA clone T10 (extended by clone T1). The nucleotide sequence and the deduced amino acid sequence are based on T10 as well as on T1. As shown in Fig. 2, T1 extends towards the 5' end, by 44 nucleotides, beyond T10; from nucleotide 45 the sequences of T10 and T1 are identical. Lines above the sequence show the corresponding sites of probes P and T and primers X and Y.

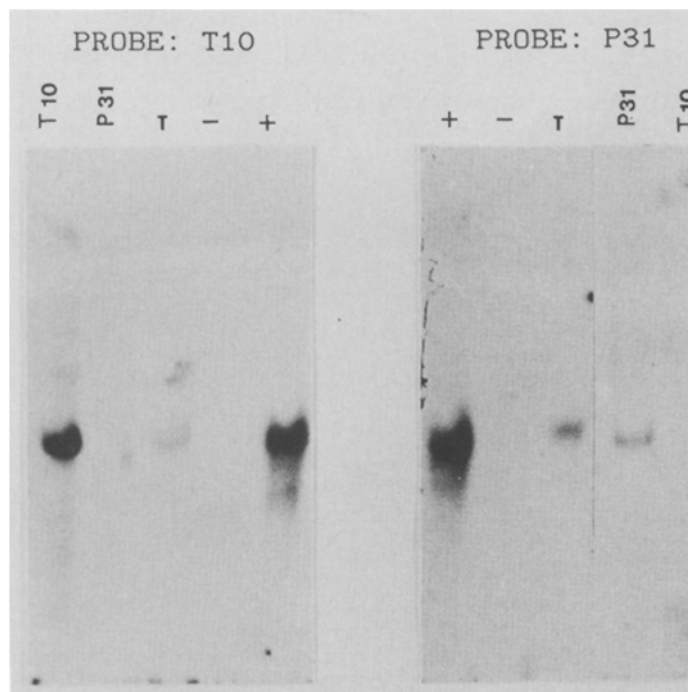


Fig. 5. Northern blot hybridization of tomato RNA to cDNA clones. Samples of tomato leaf RNA, before and after oligo-dT fractionation, were glyoxylated and run on agarose gels. Glyoxylated plasmid DNA was also applied to the gels as internal controls. Lane designations: +, 12.5  $\mu$ g poly(A)<sup>+</sup> RNA; -, 20  $\mu$ g poly(A)<sup>-</sup> RNA; T, 20  $\mu$ g unfractionated leaf RNA; P31, 0.5 ng of P31 plasmid digested with *Eco* RI and mixed with 10  $\mu$ g sonicated salmon sperm DNA; T10, 0.5 ng of T10 plasmid digested with *Eco* RI and mixed with 10  $\mu$ g sonicated salmon sperm DNA. The blots were hybridized with nick-translated probes (either P31 or T10 inserts).

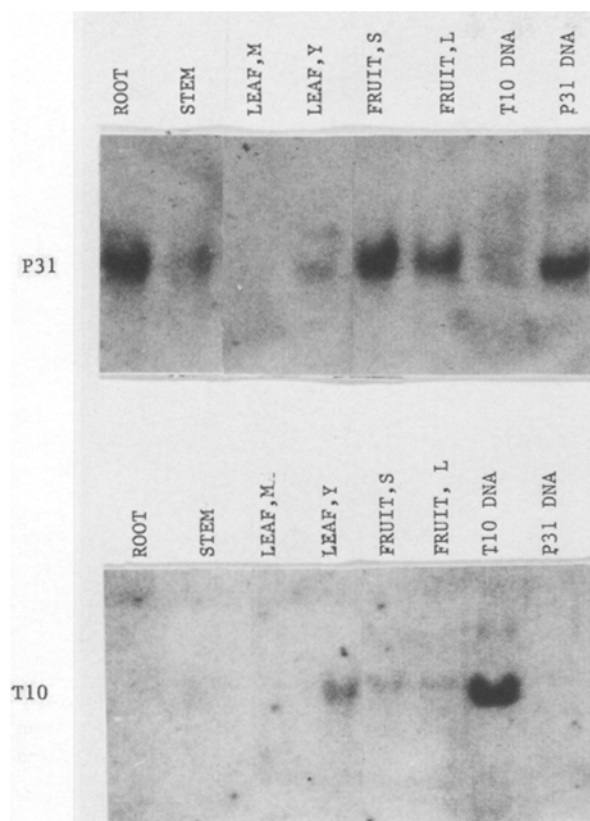


**Fig. 6.** Northern blot hybridization of RNA from different plant organs to cDNA clones. Total RNA samples (20  $\mu$ g) from various organs were glyoxylated and electrophoresed. Northern blots were prepared and hybridized with either P31 or T10 inserts. Control DNA lanes contain 0.5 ng of *Eco* RI-digested P31 and T10 plasmids, mixed with 20  $\mu$ g of poly(A)<sup>-</sup> tomato RNA and glyoxylated. Fruit, S – small (1 cm diameter) tomato fruits; Fruit, L – full-size, green fruits; Leaf, Y – top two youngest leaves of a mature plant; Leaf, M – fully expanded leaf of mature plant; root and stem samples were also from mature plants grown in pots.

the four sequences we concluded that T10 is related to spinach SOD (16 different amino acids out of 154, or 90% homology). P31 is related to maize SOD-2 (30 different amino acids out of 152, or 80% homology). All other pairwise combinations indicate 52 to 57 (34–37%) differences. If we design an evolutionary dendrogram with these substitutions, more than 30 substitutions will group the amino acid sequences of P31 and maize SOD together with respect to the other two sequences, whereas only 6 substitutions would divide them otherwise. Furthermore, most of

the substitutions within a pair are conservative (judged either from polarity/charge considerations or by the genetic code, i.e. whether one or more point mutations are required to change the codon). The differences between the two pairs are, however, bigger because most are nonconservative changes. This only contributes to increase the estimated divergence between the two pairs.

We conclude therefore that T10 must be the chloroplast SOD and P31 the cytoplasmic SOD homologous to maize SOD-2. The two tomato sequences are not more interrelated than spinach SOD and maize SOD-2 (because they come from different, analogous genes). Northern hybridization data indicating that both P31 and T10 are expressed in leaves, but only P31 in roots (Fig. 6) seem to support our identification.

The different distances (i.e., amino acid substitutions) between maize SOD and P31 and between T10 and spinach SOD (20% and 10% respectively) can be attributed to the larger phylogenetic distance in the first pair. The distance between all four plant SODs and human SOD is still larger (61–72 differences, 40–47%).

Another interesting feature that results from the comparison of all four sequences shown in Fig. 8 is that the first 24 amino acids (on the scale of Fig. 8) are the most variable ones; these are followed by a long conserved region. Between residues 88 and 103 there is another rather variable region; the carboxy end of the amino acid sequence is again highly conserved.

#### *Putative transit peptide in clone T10*

The P31 SOD-like sequence starts as expected with a methionine as translation start signal (as in maize SOD-2). The short sequence preceding the AUG in P31 is AAAA, which is in good agreement with the common translation initiation sequences of plants in this position [19]. The T10 mature SOD-like sequence (see Fig. 8) does not start with a methionine; a methionine is however found upstream, followed by 63 amino acids that precede the mature SOD sequence. This could be the transit polypeptide which would be necessary for the import of the enzyme

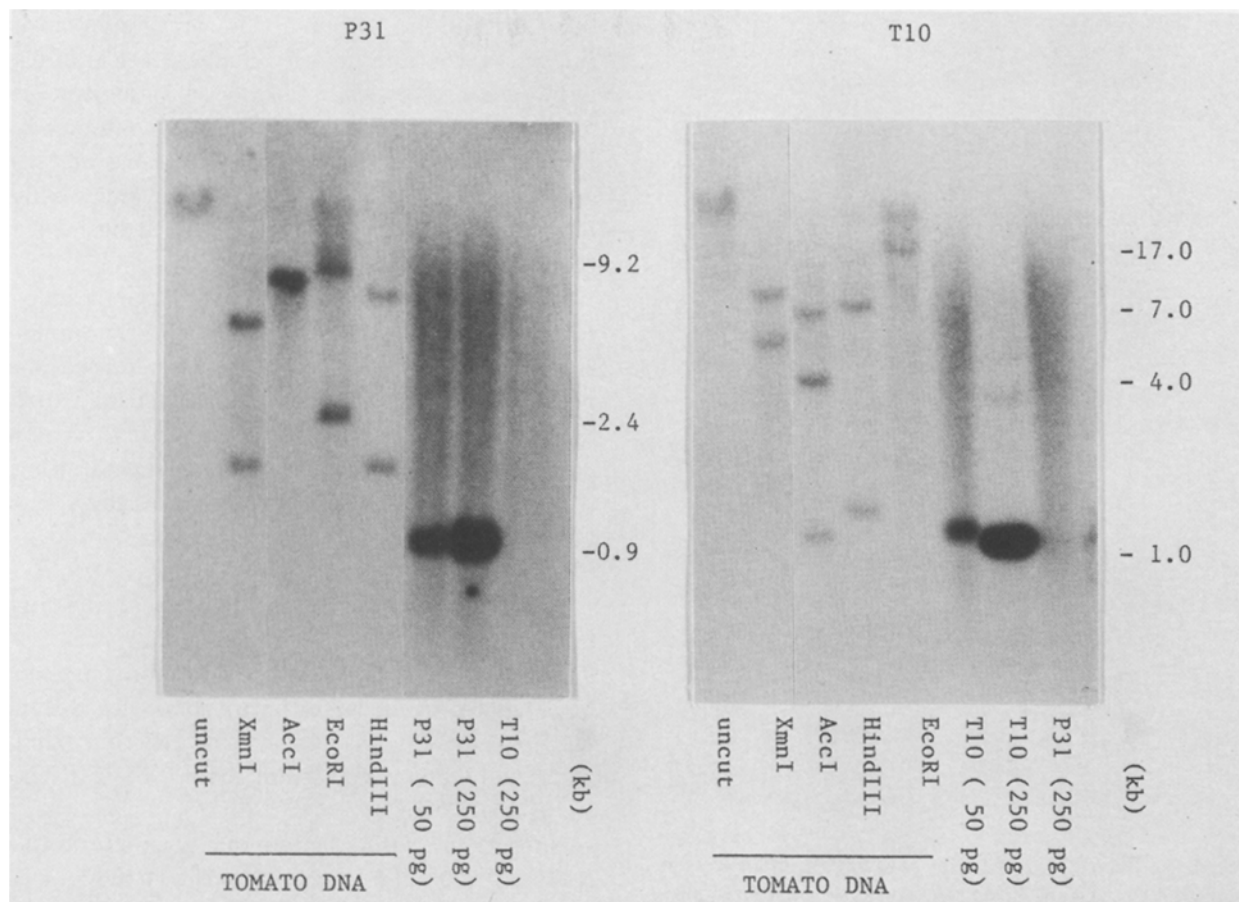


Fig. 7. Southern blot hybridization of tomato leaf DNA to cDNA clones. Total DNA (10  $\mu$ g) from tomato leaves was digested with endonucleases and run on agarose gels. *Eco* RI digests of either P31 or T10 plasmids (250 and 50 pg, which is, respectively, equivalent to 5 and 1 gene doses in the genomic DNA lanes) were mixed with 10  $\mu$ g salmon sperm DNA and run on the same gels as internal controls. The blots were hybridized with nick-translated probes (either P31 or T10 inserts) at 60 °C, 6 $\times$  SSC, followed by washes at 55 °C, 2 $\times$  SSC.

through the chloroplast double membrane into the stroma. The sequence immediately preceding this AUG is CAGC, which – except for the rarely occurring G – agrees with bases commonly found at this position [19]. This, together with our results which showed that on Northern blots the mRNA species is of about the same size as the cDNA, suggests that the 63 amino acids probably represent the full-size leader peptide. It could be, however, that another methionine, not included in our clone, exists further upstream.

Transit peptide sequences of 6 chloroplast proteins have been published ([28, 17], and reviews in [13, 29]). These sequences are quite variable, even when comparing the same protein among species.

Nevertheless, both Karlin-Neumann and Tobin [13] and Weisbeck *et al.* [29] describe some blocks of loose homology. The putative SOD transit peptide shows some of the described features: it starts with the amino acid sequence MAA like the light-harvesting chlorophyll a/b protein [13] and acetolactate synthase [17], and is processed between 2 alanines like plastocyanin. Its amino terminal half is unusually rich in serine and threonine. Negatively charged amino acids are absent; the only charged ones are 2 lysines. The AVTTPKPLT sequence near the carboxy-end resembles the VATPRMS sequence in *Silene* which was claimed to be a functional analog of the “GRV block” [29].

|                      |                                                                                                                                           |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| CLONE T10            | 1<br>M A A H S I F T T T S T T N S F L Y P I S S S S S S P                                                                                |
| CLONE T10            | N I N S S F L G V S L N V N A K F G Q S L T L Y A V T                                                                                     |
| CLONE T10<br>SPINACH | 63 1 10 20<br>T P K P L T V F A A T K K A V A V L K G N S N V E G V V T<br>A T K K A V A V L K G T S N V E G V V T                        |
| CLONE P31<br>MAIZE   | M V K A V A V L N S S E G V S G T Y L<br>M V K A V A V L A G T D V K G T I F                                                              |
| CLONE T10<br>SPINACH | 30 40 50<br>L S Q D D D G P T T V N V R I T G L A P G L H G F H L H E<br>L T Q L E D D G P T T V N V R I S G L A P G K H G F H L H E      |
| CLONE P31<br>MAIZE   | F T Q V G V A P T T V N G N I S G L K P G L H G F H V H A<br>F S Q E G D G P T T V T E S I S G L A P G L H E G H V H A                    |
| CLONE T10<br>SPINACH | 60 70 80<br>Y G D T T N G C M S T G A H F N P N K L T H G A P G D E I R<br>F G D T T N G C M S T G P H F N P D K K T H G A P E D E V R    |
| CLONE P31<br>MAIZE   | L G D T T N G C M S T G P H Y N P A G K E H G A P E D E V R<br>L G D T T N G C M S T G P H F N P V G K E H G A P E D E D R                |
| CLONE T10<br>SPINACH | 90 100 110<br>H A G D L G N I V A N A D G V A E V T L V D N Q I P L T G P<br>H A G D L G N I V A N T D G V A E A T I V D N Q I P L T G P  |
| CLONE P31<br>MAIZE   | H A G D L G N I T V G E D G T A S F T I T D K Q I P L T G P<br>H A G D L G N V T A G E D G V V N V N I T D S Q I P L A G P                |
| CLONE T10<br>SPINACH | 120 130 140<br>N S V V G R A L V V H E L E D D L G K G G H E L S L T T G N<br>N S V V G R A L V V H E L E D D L G K G G H E L S P T T G N |
| CLONE P31<br>MAIZE   | Q S I I G R A V V V H A D P D D L G K G G H E L S K S T G N<br>H S I I G R A V V V H A D P D D L G K G G H E L S L S T G N                |
| CLONE T10<br>SPINACH | 150<br>A G G R L A C G V V G L T P I<br>A G G R L A C G V V G L T P V                                                                     |
| CLONE P31<br>MAIZE   | A G G R I A C G I I G L Q G<br>A G G R V A C G I I G L Q G                                                                                |

*Fig. 8.* Comparison of P31 and T10 to plant SOD sequences. The amino acid sequence of spinach chloroplast SOD [16] and the deduced amino acid sequences of maize cytosolic SOD-2 [7] as well as of clones T10 and P31, were aligned ("Clone T10" stands for the complete sequence derived from both T1 and T10 clones). The T10 sequence extends 63 amino acids upstream of the corresponding three other SOD sequences, starting with methionine; we consider this 63 amino acids sequence as a putative leader polypeptide.

### Expression of the two Cu,Zn SOD genes

Our expression data refer to relative transcript abundance in total RNA samples. We found that transcript abundance of P31 and T10 decreases upon leaf maturation. Only P31 type RNA was detected in roots (and stems). The RNA that hybridized with P31 – which we identify as the cytoplasmic SOD – is apparently more abundant in the samples checked by us. In maize, the chloroplast isozyme was seen also in root samples [5]. This discrepancy may be explained by the use of maize seedlings which were germinated on filter paper where the roots were exposed to light, while our root samples were from under the soil surface.

### Acknowledgements

This work was supported by a grant (No 2730) from the European Community and Israel's National Council for Research and Development. We thank Dr O. Goldberg (Chemical Services, Weizmann Institute of Science) for the synthesis of oligonucleotides and Professor Y. Groner (Virology, Weizmann Institute of Science) for the h-SOD-1 clone. We are grateful to Ms M. Reichenbach, Ms D. Miron, Ms I. Leifer and Ms R. Stillmann for technical assistance and to Dr R. Fluhr, of our Department, for advice during several phases of this work.

### References

- Asada K, Kanematsu S, Okada S, Hayakawa T: Phylogenetic distribution of three types of superoxide dismutase in organisms and in cell organelles. In: Bennister JV, Hill HAO (eds) Chemical and Biochemical Aspects of Superoxide and Superoxide Dismutase, pp. 136–153. Elsevier/North Holland, New York (1980).
- Asada K, Urano M, Takahashi M: Subcellular localization of superoxide dismutase in spinach leaves and preparation and properties of crystalline spinach superoxide dismutase. *Eur J Biochem* 36: 257–266 (1973).
- Aviv H, Leder P: Purification of biologically active globin mRNA by chromatography on oligothymidilic acid-cellulose. *Proc Natl Acad Sci USA* 69: 1408–1412 (1972).
- Baum JA, Chandless JM, Scandalios JG: Purification and partial characterization of a genetically defined superoxide dismutase (SOD-1) associated with maize chloroplasts. *Plant Physiol* 73: 31–35 (1983).
- Baum JA, Scandalios JG: Developmental expression and intracellular localization of superoxide dismutase in maize. *Differentiation* 13: 133–140 (1979).
- Baum JA, Scandalios JG: Multiple genes controlling superoxide dismutase expression in maize. *J Heredity* 73: 95–100 (1982).
- Cannon RE, White JA, Scandalios JG: Cloning of cDNA for maize superoxide dismutase 2 (SOD-2). *Proc Natl Acad Sci USA* 84: 179–183 (1987).
- Carmichael GG, McMaster GK: The analysis of nucleic acids in gels using glyoxal and acridine orange. In: Grossman L, Moldave K (eds), *Methods in Enzymology*, Vol. 65, pp. 380–391. Academic Press, New York (1980).
- Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12: 387–395 (1984).
- Duke MV, Salin ML: Isozymes of Cuprozinic superoxide dismutase from *Pisum sativum*. *Phytochemistry* 22: 2369–2373 (1983).
- Fridovich I: Superoxide dismutases. *Ann Rev Biochem* 44: 147–159 (1975).
- Gubler U, Hoffman BJ: A simple and very efficient method for generating cDNA libraries. *Gene* 25: 263–269 (1983).
- Karlin-Neumann GA, Tobin EM: Transit peptides of nuclear encoded chloroplast proteins share a common amino acid framework. *EMBO J* 5: 9–13 (1986).
- Khandjian EW: Optimized hybridization of DNA blotted and fixed to nitrocellulose and nylon membranes. *Biotechnology* 5: 165–167 (1987).
- Kirby KS: Isolation of nucleic acids with phenolic solvents. In: Grossman L, Moldave K (eds) *Methods in Enzymology*, Vol. 12B, pp. 87–99. Academic Press, New York (1968).
- Kitagawa Y, Tsunasawa S, Tanaka N, Katsube Y, Sakiyama F, Asada K: Amino acid sequence of copper, zinc superoxide dismutase from spinach leaves. *J Biochem* 99: 1289–1298 (1986).
- Lee KY, Townsend J, Tepperman J, Black M, Chui CF, Mazur B, Dunsmuir P, Bedbrook J: The molecular basis of sulfonylurea herbicide resistance in tobacco. *EMBO J* 7: 1241–1248 (1988).
- Logemann J, Schell J, Willmitzer L: Improved method for the isolation of RNA from plant tissues. *Anal Biochem* 163: 16–20 (1987).
- Lutcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA: Selection of AUG initiation codons differs in plant and animals. *EMBO J* 6: 43–48 (1987).
- Maniatis F, Fritsch EF, Sambrook J: *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
- Mason PJ, Williams JG: Hybridization in the analysis of recombinant DNA. In: Hames BD, Higgins SJ (eds) *Nucleic Acid Hybridisation*, pp. 113–137. IRL Press, Oxford/Washington DC (1985).
- Okayama H, Berg P: High-efficiency cloning of full-length

- cDNA. *Mol Cell Biol* 2: 161–170 (1982).
23. Okita TW, Greene FC: Wheat storage proteins: isolation and characterization of the gliadin messenger RNAs. *Plant Physiol* 69: 834–839 (1982).
  24. Pazkowski JC, Saul MW: Direct gene transfer to plants. In: Weissbach A, Weissbach H (eds) *Methods in Enzymology*, Vol. 118, pp. 668–684. Academic Press, Orlando (1986).
  25. Rabinowitch HD, Fridovich I: Superoxide radicals, superoxide dismutase and oxygen toxicity in plants. *Rev Photochem Photobiol* 37: 679–690 (1983).
  26. Sanger F, Coulson AR, Barrel BG, Smith AJH, Roe BA: Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J Mol Biol* 143: 161–178 (1980).
  27. Sherman L, Dafni N, Lieman-Hurwitz J, Groner Y: Nucleotide sequence and expression of human chromosome 21-encoded superoxide dismutase mRNA. *Proc Natl Acad Sci USA* 80: 5465–5469 (1983).
  28. Stayton MM, Broslo P, Dunsmuir P: Characterization of a full-length petunia cDNA encoding a polypeptide of the light-harvesting complex associated with photosystem I. *Plant Mol Biol* 10: 127–137 (1987).
  29. Weisbeek P, Hageman J, Smeekeens S, de Boer D, Cremers F: Chloroplast-specific import and routing of proteins. In: Von Wettstein D, Chua N-H (eds) *Plant Molecular Biology*, pp. 77–91. Plenum Press, New York/London (1987).