

Short Communication

The 33 kDa Protein of the Oxygen-Evolving Complex: a Multi-Gene Family in Tomato

Jörn Görlach, Jürg Schmid¹ and Nikolaus Amrhein

*Institute of Plant Sciences, Swiss Federal Institute of Technology, Sonneggstrasse 5,
CH-8092 Zürich, Switzerland*

A cDNA was isolated by chance from tomato which had a high similarity to a cDNA clone from potato known to code for the 33 kDa protein of the oxygen-evolving complex [van Spanje et al. (1991) *Plant Mol. Biol.* 17: 157]. The sequence of a previously described partial cDNA clone from tomato [Ko et al. (1990) *Plant Mol. Biol.* 14: 217] which has also a high similarity but is not identical to the sequence described here indicates that tomato contains at least two genes coding for 33 kDa proteins per haploid genome. This conclusion is supported by Southern blot analysis. The tissue specific expression of the corresponding genes is described.

Key words: Multi-gene family — Oxygen-evolving complex — Photosynthesis — Tomato (*Lycopersicon esculentum* L.).

In plants, oxygen is evolved at the inner face of the thylakoid membrane in the so-called OEC, which is part of PSII. Three extrinsic luminal membrane proteins, the so called 33 kDa, 23 kDa and 16 kDa proteins, are involved in the manganese-dependent water splitting reaction (Anderson et al. 1984). The detailed structure of the OEC and the molecular mechanism of the reaction are not yet known (Babcock et al. 1989, Rutherford 1989). The 23 kDa and 16 kDa proteins function in the stabilization and regulation of the complex. The 33 kDa protein seems to be a constituent of the Mn-cluster and is directly involved in O₂-generation. The 33 kDa proteins from plants show high structural and functional homologies to the 33 kDa manganese-stabilizing polypeptide of the PSII from cyanobacteria (Koike and Inoue 1985). This 33 kDa protein is one of four well characterized proteins of the thylakoid lumen. Like plastocyanin and the other extrinsic components of the OEC, the 33 kDa protein is a nuclear-encoded protein (Westhoff et al. 1985, Jansen et al. 1987). The higher molecular weight precursors are synthesized in the cytosol and processed to the mature form by proteolytic cleavage while

traversing three membranes. This processing is a two step mechanism, which is mediated by N-terminal two-domain transit peptides (Hageman et al. 1986, Smeekens et al. 1986, James et al. 1989). A thylakoidal processing peptidase has been characterized (Kirwin et al. 1988, Shackleton and Robinson 1991). Here we present the nucleotide sequence of a tomato cDNA which codes for a precursor of the 33 kDa protein (Fig. 1A).

The basic molecular techniques were carried out according to Sambrook et al. (1989) or Ausubel et al. (1987). A tomato cDNA library in the expression vector Uni-ZAP™ XR (Schmid et al. 1992) was screened. Positive clones were plaque purified and the cDNA inserts subcloned into the vektor pSK(+). The sequences of both strands were determined. Recombinant pBluescript cDNA phagemids were excised in vivo from the λZAP vector by infecting *E. coli* XL1-Blue with 200 μl of the amplified library (10⁶ pfu) and helper phage R408 according to the procedure provided by Stratagene. After incubating *E. coli* XL1-Blue cells with the rescued phagemids, ampicillin resistant bacterial colonies were selected on LB-plates. The colonies (3 × 10⁵) were washed off the plates and plasmid DNA was isolated. This pool was used as template for anchored PCR. The primers used for the PCR were the commercially available T3 primer (17mer, Stratagene) and a cDNA specific oligonucleotide (16mer, corresponding to position 230–245 in Fig. 1A). 150 ng plasmid DNA were

Abbreviations: bp, base pair; OEC, oxygen-evolving complex; PCR, polymerase chain reaction; pfu, plaque forming units; RT, room temperature.

¹ Author to whom all correspondence should be addressed.

² EMBL Accession NO: Z11999.

for denaturation, 64°C/60 s for annealing as well as for primer extension. The resulting products were separated on a 1% low melting agarose gel and DNA fragments of the expected size were isolated, subcloned and sequenced. Poly(A)⁺-RNA (2 µg) from different tomato tissues were separated on 1.2% formaldehyde agarose gel (Sambrook et al. 1989) and transferred to a nylon membrane (Hybond N, Amersham) which was processed according to the manufacturer's instructions. Prehybridization and hybridization were in 50% formamide, 5 × SSC, 20 mM PIPES, pH 6.4, 200 µg ml⁻¹ denatured carrier DNA, 2 × Denhardt's solution, 0.5% SDS (1 × SSC: 150 mM NaCl, 15 mM sodium citrate, pH 7.0; 1 × Denhardt's: 0.02% of each Ficoll 400, bovine serum albumin and polyvinylpyrrolidone) at 42°C. The 33 kDa protein cDNA was used as radiolabeled probe (Feinberg and Vogelstein 1983). The filter was washed twice for a few minutes in 0.3 × SSC, 0.5% SDS at RT, three times for 30 minutes in the same buffer at 60°C and

once for a few minutes in 0.3 × SSC at RT and then exposed to Hyperfilm-MP (Amersham) at -80°C with intensifying screen. Genomic tomato DNA was isolated from cotyledons as described (Murray and Thompson 1980), digested with restriction enzymes and subjected to electrophoresis in a 0.8% agarose gel. Processing of the gel and blotting of the DNA to a Zeta-Probe membrane (BioRad) were done according to the manufacturer's instructions. Prehybridization and hybridization conditions were 0.5 M sodium phosphate, pH 7.2, 7% SDS and 1 mM EDTA at 55°C. The 33 kDa protein cDNA was used as radiolabeled probe (Feinberg and Vogelstein 1983). The membrane was washed for 45 minutes in 40 mM sodium phosphate, 5% SDS, 1 mM EDTA at 57°C and again at 62°C, twice for a short time in 2 × SSC at RT. The autoradiography was performed as outlined above. The sequence comparisons were done with the program PILEUP and LINEUP (GCG package, University of Wisconsin). Similarity was calculated for a threshold = 0.6.

Screening a tomato cDNA library for chorismate synthase clones, we isolated one clone which contained two cDNA inserts, one specific for chorismate synthase (J.G. and J.S., unpublished results). Based on sequence similarities with already known sequences the other cDNA was identified as one coding for a 33 kDa protein of the oxygen-evolving complex. This cDNA was incomplete at its 5'-end and ended at position 37 in Fig. 1A. To extend the cDNA sequence anchored PCR was performed and the resulting PCR fragments were subcloned and sequenced. The sequence of one PCR fragment which was identical as

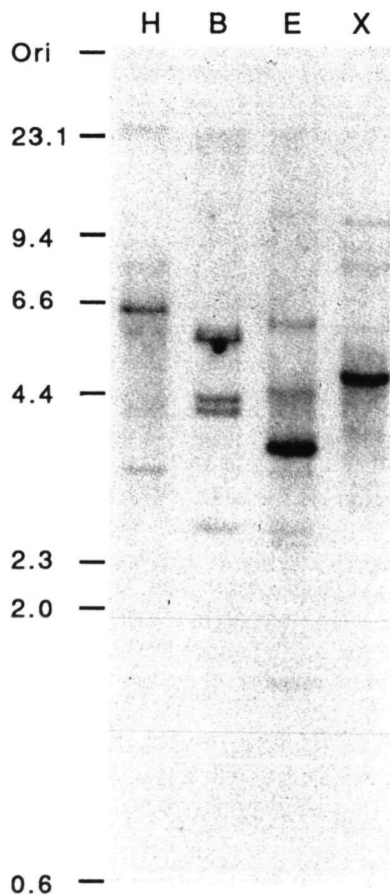


Fig. 2 Southern blot analysis. High molecular weight tomato DNA was digested with the restriction enzymes *Hind* III (lane H), *Bam*H I (lane B), *Eco*R I (lane E) and *Xba* I (lane X) and subjected to Southern blot analysis using the 33 kDa protein cDNA as radiolabeled probe. Fragments of *Hind* III digested λ -DNA were used as size marker (kb).

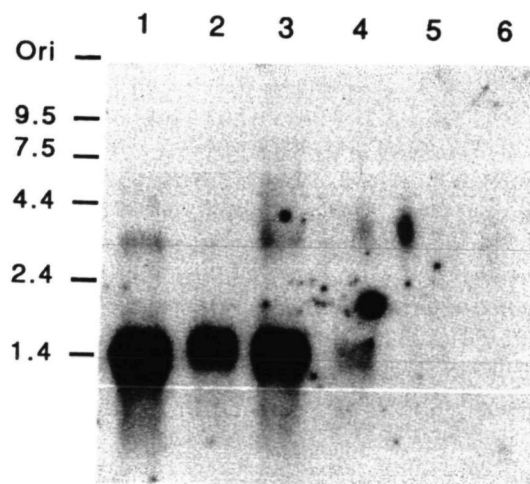


Fig. 3 Northern blot analysis. 2 µg poly(A)⁺-RNA from tomato leaves (lane 1), total flowers without sepals (lane 2), cotyledons (lane 3), stems (lane 4), roots (lane 5) and tomato cells in suspension culture (lane 6) were subjected to Northern blot analysis using the 33 kDa protein cDNA as radiolabeled probe. RNAs of different length (BRL) were used as size marker (nucleotides × 10³).

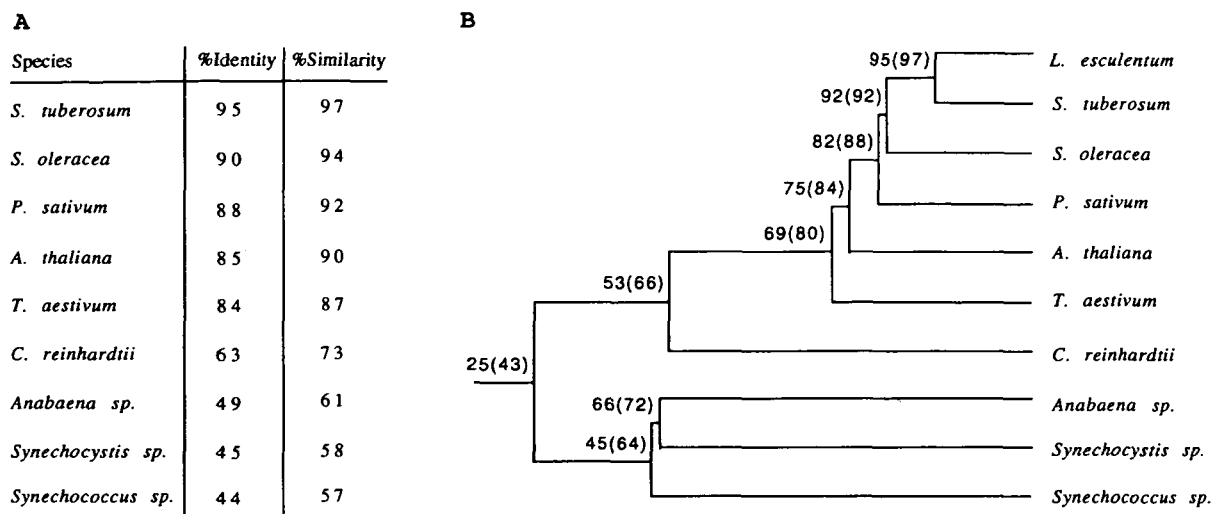


Fig. 4 Comparison of the mature 33 kDa protein from different plants and the mature Mn-stabilizing peptides from cyanobacteria. **A.** Amino acid identity and similarity of tomato 33 kDa protein with other 33 kDa protein sequences and 33 kDa PSII extrinsic proteins from prokaryotes. Sequences compared were published by: *S. tuberosum* (van Spanje et al. 1991), *S. oleracea* (Tyagi et al. 1987), *P. sativum* (Wales et al. 1989), *A. thaliana* (Ko et al. 1990), *T. aestivum* (Meadows et al. 1991), *C. reinhardtii* (Mayfield et al. 1989), *Anabaena sp.* (Borthakur and Haselkorn 1989), *Synechocystis sp.* (Philbrick and Zilinskas 1988), *Synechococcus sp.* (Kuwabara et al. 1987). **B.** Relationship of the compared amino acid sequences. Distances are drawn to scale. Numbers give percentage of overall identity (similarity) for the organisms listed. Calculations were done as in A.

far as it overlapped with the already known sequence was used to complete the cDNA sequence² shown in Fig. 1A. The first ATG of the sequence (Fig. 1A, bp position 25–27) resides within the open reading frame. The deduced amino acid sequence is almost identical to the primary structure of the precursor of the 33 kDa protein from potato (van Spanje et al. 1991). The N-terminal portion of the deduced amino acid sequence (Fig. 1A, amino acid position 1–82) resembles a transit peptide. It is known that for correct processing of the wheat precursor of the 33 kDa protein to the mature form, alanine residues are required at position –3 and –1 with respect to the terminal cleavage site (Shackleton and Robinson 1991). The amino acid sequence GANA|EG (Fig. 1A, amino acid position 79–84) fulfills this requirement and is almost identical to the cleavage site of the wheat precursor (GATA|EG). In the 3'-untranslated region of the 33 kDa protein cDNA, three potential polyadenylation signals (Joshi 1987) can be identified, suggesting differential polyadenylation of the primary transcription product. Similar heterogeneities have been observed for other plant genes (Schaller et al. 1991). So far isoforms of the 33 kDa protein have only been characterized in pea (Wales et al. 1989). A comparison of the tomato 33 kDa cDNA with a cDNA fragment isolated by Ko et al. (1990) from tomato indicates the existence of isoforms of the 33 kDa protein in tomato (Fig. 1B, 1C). The coding sequences of the two cDNA clones are similar, whereas the 3'-untranslated regions are unrelated to each other, reminis-

cent of many genes belonging to multi-gene families. This finding is further supported by a Southern blot analysis with chromosomal tomato DNA (Fig. 2). The rather complex pattern is a strong indication of a multi-gene family. The tissue specific expression of this gene family has been analyzed by Northern blots (Fig. 3). As expected for an enzyme which is involved in photosynthesis and in agreement with results of other groups (Ko et al. 1990, van Spanje et al. 1991), the 33 kDa protein gene family is only expressed in photosynthetically active tissues like leaves, total flowers without sepals, cotyledons and stems, but not in roots and tomato cells grown in suspension. The comparison of the mature tomato 33 kDa protein to corresponding sequences of other plant species and to the 33 kDa manganese-stabilizing polypeptide from cyanobacteria (Fig. 4) indicates the phylogenetic relationship of these organisms. The overall identity of all compared proteins is 25% (43% similarity). The 33 kDa proteins are well conserved between cyanobacteria and photosynthetic eukaryotes, even though cyanobacterial PSII does not contain peptides homologous to 23 kDa and 16 kDa proteins of higher plants.

References

- Andersson, B., Larsson, C., Jansson, C., Ljungberg, U. and Åkerlund, H.-E. (1984) Immunological studies on the organization of proteins in photosynthetic oxygen evolution. *Bio-*

- chim. Biophys. Acta* 766: 21–28.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience, New York.
- Babcock, G.T., Barry, B.A., Debus, R.J., Hoganson, C.W., Atamian, M., McIntosh, L., Sithole, I. and Yocum, C.F. (1989) Water oxidation in photosystem II: from radical chemistry to multielectron chemistry. *Biochemistry* 28: 9557–9565.
- Borthakur, D. and Haselkorn, R. (1989) Nucleotide sequence of the gene encoding the 33 kDa water oxidizing polypeptide in *Anabaena* sp. strain PCC 7120 and its expression in *Escherichia coli*. *Plant Mol. Biol.* 13: 427–439.
- Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132: 6–13.
- Hageman, J., Robinson, C., Smeeckens, S. and Weisbeek, P. (1986) A thylakoid processing protease is required for complete maturation of the lumen protein plastocyanin. *Nature* 324: 567–569.
- James, H.E., Bartling, D., Musgrove, J.E., Kirwin, P.M., Herrmann, R.G. and Robinson, C. (1989) Transport of proteins into chloroplasts. *J. Biol. Chem.* 264: 19573–19576.
- Jansen, T., Rother, C., Steppuhn, J., Reinke, H., Beyreuther, K., Jansson, C., Andersson, B. and Herrmann, R.G. (1987) Nucleotide sequence of cDNA clones encoding the complete '23 kDa' and '16 kDa' precursor proteins associated with the photosynthetic oxygen-evolving complex from spinach. *FEBS Lett.* 216: 234–240.
- Joshi, C.P. (1987) Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. *Nucl. Acids Res.* 15: 9627–9640.
- Kirwin, P.M., Elderfield, P.D., Williams, R.S. and Robinson, C. (1988) Transport of proteins into chloroplasts. *J. Biol. Chem.* 263: 18128–18132.
- Ko, K., Granell, A., Bennett, J. and Cashmore, A.R. (1990) Isolation and characterization of cDNAs from *Lycopersicon esculentum* and *Arabidopsis thaliana* encoding the 33 kDa protein of the photosystem II-associated oxygen-evolving complex. *Plant Mol. Biol.* 14: 217–227.
- Koike, H. and Inoue, Y. (1985) Properties of a peripheral 34 kDa protein in *Synechococcus vulcanus* photosystem II particles. Its exchangeability with spinach 33 kDa protein in reconstitution of O₂ evolution. *Biochim. Biophys. Acta* 807: 64–73.
- Kuwabara, T., Reddy, K.J. and Sherman, L.A. (1987) Nucleotide sequence of the gene from the cyanobacterium *Anacystis nidulans* R2 encoding the Mn-stabilizing protein involved in photosystem II water oxidation. *Proc. Natl. Acad. Sci. USA* 84: 8230–8234.
- Mayfield, S.P., Schirmer-Rahire, M., Frank, G., Zuber, H. and Rochaix, J.-D. (1989) Analysis of the genes of the OEE1 and OEE3 proteins of the photosystem II complex from *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* 12: 683–693.
- Meadows, J.W., Hulford, A., Raines, C.A. and Robinson, C. (1991) Nucleotide sequence of a cDNA clone encoding the precursor of the 33 kDa protein of the oxygen-evolving complex from wheat. *Plant Mol. Biol.* 16: 1085–1087.
- Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucl. Acids Res.* 8: 4321–4325.
- Philbrick, J.B. and Zilinskas, B.A. (1988) Cloning, nucleotide sequence and mutational analysis of the gene encoding the Photosystem II manganese-stabilizing polypeptide of *Synechocystis* 6803. *Mol. Gen. Genet.* 212: 418–425.
- Rutherford, A.W. (1989) Photosystem II, the water-splitting enzyme. *TIBS* 14: 227–232.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
- Schaller, A., Schmid, J., Leibinger, U. and Amrhein, N. (1991) Molecular cloning and analysis of a cDNA coding for choris-mate synthase from the higher plant *Corydalis sempervirens* Pers. *J. Biol. Chem.* 266: 21434–21438.
- Schmid, J., Schaller, A., Leibinger, U., Boll, W. and Amrhein, N. (1992) The in-vitro synthesized tomato shikimate kinase precursor is enzymatically active and is imported and processed to the mature enzyme by chloroplasts. *Plant J.* 2: 375–383.
- Shackleton, J.B. and Robinson, C. (1991) Transport of proteins into chloroplasts. *J. Biol. Chem.* 266: 12152–12156.
- Smeeckens, S., Bauerle, C., Hageman, J., Keegstra, K. and Weisbeek, P. (1986) The role of the transit peptide in the routing of precursors toward different chloroplast compartments. *Cell* 46: 365–375.
- Tyagi, A., Hermans, J., Steppuhn, J., Jansson, C., Vater, F. and Herrmann, R.G. (1987) Nucleotide sequence of cDNA clones encoding the complete "33 kDa" precursor protein associated with the photosynthetic oxygen-evolving complex from spinach. *Mol. Gen. Genet.* 207: 288–293.
- van Spanje, M., Dirkse, W.G., Nap, J.-P. and Stiekema, W.J. (1991) Isolation and analysis of cDNA encoding the 33 kDa precursor protein of the oxygen-evolving complex of potato. *Plant Mol. Biol.* 17: 157–160.
- Wales, R., Newman, B.J., Pappin, D. and Gray, J.C. (1989) The extrinsic 33 kDa polypeptide of the oxygen-evolving complex of photosystem II is a putative calcium-binding protein and is encoded by a multi-gene family in pea. *Plant Mol. Biol.* 12: 439–451.
- Westhoff, P., Jansson, C., Klein-Hitpaß, L., Berzborn, R., Larsson, C. and Bartlett, S.G. (1985) Intracellular coding sites of polypeptides associated with photosynthetic oxygen evolution of photosystem II. *Plant Mol. Biol.* 4: 137–146.

(Received September 28, 1992; Accepted February 15, 1993)

