Mol Gen Genet (1993) 238:285–293

MGG

pas de 40

The *Nicotiana tabacum* genome encodes two cytoplasmic thioredoxin genes which are differently expressed

Christophe Brugidou*, Isabelle Marty, Yvette Chartier, Yves Meyer**

Laboratoire de Biologie Moléculaire et Physiologie Végétale, Unité associée au CNRS 565, Université de Perpignan, avenue de Villeneuve, F-66860 Perpignan – Cedex, France

Received: 1 July 1992 / Accepted: 9 October 1992

Abstract. A *Nicotiana tabacum* thioredoxin h gene (EMBL Accession No. Z11803) encoding a new thioredoxin (called h_2) was isolated using thioredoxin h_1 cDNA (X58527), and represents the first thioredoxin h gene isolated from a higher plant. It encodes a polypeptide of 118 amino acids with the conserved thioredoxin active site Trp – Cys – Gly – Pro – Cys. This gene comprises two introns which have lengths of 1071 and 147 bp respectively, and three exons which encode peptides of 29, 41 and 48 amino acids, respectively. This thioredoxin h shows 66% identity with the amino acid sequence of thioredoxin h_1 (X58527) and only around 35% with the choroplastic thioredoxins. The two thioredoxins, h_1 and h_2, do not have any signal peptides and are most probably cytoplasmic. Using the 3′ regions of the mRNAs, two probes specific for thioredoxins h_1 and h_2 have been prepared. Southern blot analysis shows that thioredoxin sequences are present in only two genomic *Eco*RI fragments: a 3.3 kb fragment encodes h_1 and a 4.5 kb fragment encodes h_2. Analysis of the ancestors of the allotetraploid *N. tabacum* shows that thioredoxin h_2 is present in *N. sylvestris* and *N. tomentosiformis* but that thioredoxin h_1 is absent from both putative ancestors. Thus, the thioredoxin h_1 gene has probably been recently introduced in to *N. tabacum* as a gene of agronomic importance, or linked to such genes. Northern blot analysis shows that both genes are expressed in *N. tabacum*, mostly in organs or tissues that contain growing cells. Thioredoxin h_1 is always expressed at a lower level than h_2 in tobacco plants. In contrast, the thioredoxin h_1 gene is abundantly expressed in freshly isolated protoplasts, while h_2 mRNAs are not detectable.

* *Present address*: ILTAB The Scripps Research Institute, Division of Plant Biology, MRC7 10666 North Torrey Pines Road, La Jolla, CA 92037, USA

** *Present address*: CSIC Dept. of Molecular Genetics, Jorge Girona 18, E-08034 Barcelona, Spain

Communicated by A. Kondorosi

Correspondence to: Y. Meyer

Key words: Thioredoxins – Cell growth – Protoplast – *Nicotina tabacum*

Introduction

Thioredoxins are small, ubiquitous oxidoreductase proteins (M_r, approximately 12 000 daltons) with a conserved active site having the structure Trp-Cys-Gly-Pro-Cys. The oxidized form contains a disulphide bridge which is reduced by a thioredoxin reductase; the reduced form is a powerful protein disulphide oxidoreductase (Holmgren 1989). Consequently, it plays a post-translational regulatory role in breaking disulphide bridges and modifying the enzymatic activity of target proteins. In mammals, reduced thioredoxin activates the glucorticoid and androgen receptors, ornithine decarboxylase and protein synthesis in rabbit reticulocyte lysates and deactivates a protein kinase that controls the activity of initiation factors (Tagaya et al. 1989; Muller 1991). In addition, reduced thioredoxin serves as a hydrogen donor to electron transfer chains leading to the reduction of small molecules such as ribonucleotides (Reichard 1988; Holmgren 1989), sulphate and methionine sulphoxide (Tsang and Schiff 1976; Gleason and Holmgren 1988). Finally, thioredoxin is an essential subunit of the bacteriophage T7 DNA polymerase (Modrich and Richardson 1976) and participates in bacteriophage f1 and M13 assembly (Russel and Model 1985). Nevertheless, the in vivo functions of thioredoxin have not always been clearly demonstrated (Gleason and Holmgren 1988; Muller 1991), partly because other disulphide proteins (for example glutaredoxin; Holmgren 1989) and protein disulphide isomerase (Freedman et al. 1988; Tasanen et al. 1988; Lundström and Homgren 1990) are able to replace thioredoxins.

The amino acid or nucleotide sequences of thioredoxins have been established in a wide variety of prokaryotes (Gleason and Holmgren 1988) and eukaryotes, including vertebrates (Johnson et al. 1988, Jones

and Luk 1988; Tonissen et al. 1989; Wollman et al. 1988; Tonissen and Wells 1991), algae (Gleason et al. 1985; Muller and Buchanan 1989; Decottignies et al. 1990, 1991) and fungi (Gan 1991; Muller 1991; Wetterauer et al. 1992). Apart from the active dithiol centre which is common to all thioredoxins and the small size (around 100 amino acids) the rest of the molecule presents considerable variability.

In higher plants and green algae, two thioredoxin systems have been characterized: the first one is localized in chloroplasts and has two nuclear-encoded thioredoxins (f and m types), which play a key role in photosynthesis (Jacquot 1984; Droux et al. 1987; Muller and Buchanan 1989; Decottignies et al. 1990, 1991; Van Langendonckt and Vanden Driessche 1992). They are both reduced by a ferredoxin-dependent thioredoxin reductase and regulate enzymes of CO_2 fixation (Buchanan 1980; Cséke and Buchanan 1986; Huppe et al. 1990). cDNA clones of both types have been isolated and sequenced (Kamo et al. 1989; Maeda et al. 1986; Wedel et al. 1992). The second thioredoxin system (h type) is present in heterotrophic cells and has been detected in wheat, carrot and spinach (Vogt and Follmann 1986; Johnson et al. 1987a, b). The thioredoxin h activities, which are activated by NADPH-dependent thioredoxin reductases, are most probably localized in the cytoplasm (Florencio et al. 1988). One enzyme has been purified from spinach leaves and partially sequenced (Marcus et al. 1991). Recently we have isolated from *N. tabacum* cells cultivated in vitro a cDNA that encodes an h type thioredoxin (EMBL accession number NTTRNA = X58527; Marty and Meyer 1991).

In this paper we report the isolation of the first genomic clone encoding a higher plant thioredoxin. The cloned gene was sequenced and found to have an open reading frame that encodes a protein presenting 66% identity with that encoded by X58527. The molecular organization, the presence of introns and the putative promoter sequences were analysed. In addition two specific probes have been isolated, which allowed us to demonstrate that both genes are expressed in tobacco plants but that they are differently regulated.

Materials and methods

Plant material. N. tabacum ev. Xanthi cell suspensions derived from pith by Jouanneau and Tandeau de Marsac (1973) were subcultured every week by × 10 dilution. Three-day-old cultures contain dividing cells while 11-day-old cultures consist of stationary phase non-dividing cells. Tobacco plants (*N. tabacum* ev. Maryland) and other *Nicotiana* species were grown in a greenhouse under controlled conditions on a mixture of peat and vermiculite (1:3). *N. sylvestris* and *N. tomentosiformis* were obtained from Dr. Delon at the Institut National du tabac SEITA, Bergerac, France. Fully expanded leaves, roots and flower buds were collected from 3-month-old plants. Developing leaves and roots were collected from 2-week-old plant. *N. tabacum* protoplasts were isolated from mature leaves of greenhouse-grown tobacco plants according to the method described previously (Meyer and Abel 1975).

Preparation of a partial library and isolation of a genomic clone. DNA was digested with *Eco*RI and electrophoresed on a preparative 0.8% agarose gel at 35 V. Fragments were collected according to their size in five fractions: 2.7–4.0 kb, 3.4–4.9 kb, 3.7–9.0 kb, 5.0–9.5 kb and 6.5–23 kb. A *N. tabacum* partial genomic library was constructed with the second fraction (3.4–4.9 kb) by cloning in predigested lambda ZAPII-*Eco*RI (Stratagene) and packaging with a Gigapack Gold packaging kit from Stratagene, according to the recommendations of the manufacturer. From the partial genomic phage library, 400 000 plaques of the recombinant bacteriophage were plated at a density of about 50 000 per plate, transferred onto Hybond N+ membrane (Amersham) and screened with the thioredoxin X58527 probe (Marty and Meyer 1991). The hybridization and washing conditions were identical to those described in the protocols provided by Stratagene and Amersham. One positive lambda clone was purified. The Bluescript phagemid containing the cloned DNA insert was excised in vivo by coinfection with the helper phage according to the Stratagene protocol. XL1 Blue was used as host strain; the selection and growth of the recombinant cells were as described by Stratagene.

DNA probes. The double-stranded thioredoxin probe (NTTRNA = X58527) was obtained by *Pst*I-*Bam*HI digestion of the corresponding plasmid (pTZ18R) and the 750 bp insert separated from the vector on an agarose gel. A 509 bp fragment corresponding to the 5′ coding region was recovered by *Pst*I and *Eco*RV digestion. A 240 bp fragment corresponding to the 3′ region was recovered by *Bam*HI and *Eco*RV digestion. Finally, a specific 468 bp fragment corresponding to the 3′ region of the isolated gene (Z11803) was obtained by *Pst*I and *Eco*RI digestion. Separation, recovery and elution of these different inserts were carried out using Schleicher and Schuell NA 45 paper according to the procedures of Sambrook et al. (1990). The probes were labelled by the Multiprime technique (Amersham kit and procedures), and are designated thioredoxin h_1 for the full-length thioredoxin cDNA (X58527, 750 bp), 5′ thioredoxin h_1 for the 509 bp fragment, 3′ thioredoxin h_1 for the 240 bp fragment and 3′ thioredoxin h_2 for the 468 bp fragment. The *Arabidopsis* histone H4 clone, A748, was provided by C. Gigot (Chaboute et al. 1987). A748 contains a 196 bp fragment in the H4 coding region, which was excised and labelled using the same Multiprime technique.

Southern blot and gene copy-number reconstruction analysis. DNA was extracted from different plant species as previously described and using a CsCl-ethidium bromide centrifugation protocol (Delcasso-Tremousaygue et al. 1988). After digestion with appropriate restriction endonucleases and analysis on a 1% agarose gel, DNA fragments were transferred onto Gene-screen or Hybond N^+ membranes according to the recommendations of the

manufacturer. The membrane was prehybridized for 2 h at 42° C with the following solution: 50% deionized formamide, 2× Denhardt's, 5× SSC (1× SSC is 0.15 M NaCl, 15 mM sodium citrate), 1% SDS and 100 µg/ml yeast tRNA. It was hybridized in the same solution with denatured probe at 42° C for 20 h, washed twice with 2× SSC at room temperature for 15 min, and twice at 65° C with 2× SSC, 1% SDS for 20 min. A stringent wash was carried out at 70° C with 1× SSC, 1% SDS for 10 min. Hybridization to heterologous DNA was carried out at 37° C and without the stringent wash. To estimate the copy number of the thioredoxin gene in the *N. tabacum* genome, amounts of genomic *Eco*RI-digested DNA equivalent to 0.5 (0.19 µg), 1 (0.38 µg), 2 (0.77 µg) and 5 (1.92 µg) times 10^5 haploid genomes were electrophoresed with known amounts of thioredoxin h_1 cDNA (0.2 pg, 0.4 pg, 0.8 pg and 2 pg). The DNA was transferred to Hybond N^+ membrane and hybridized with ^{32}P-labelled thioredoxin h_1 cDNA.

Northern blot analysis. Total RNA was extracted from different tissues by previously described methods (Grosset et al. 1990). The same quantity of RNA (10 µg) was denatured and fractionated on denaturing gels as described by Meinkoth and Wahl (1984). Hybridization and washing conditions were the same as for Southern blots except that hybridization solutions contained 100 µg/ml polyU.

Nucleotide sequences. Nucleotide sequences were determined by the dideoxynucleotide chain-termination method (Sanger et al. 1977) with double-stranded DNA templates according to the recommendations of the Amersham Multiwell system kit. For deletion subcloning of the genomic clone, the *Exo*III/mung bean nuclease kit (Stratagene) was used. Oligonucleotide primers were used to sequence the complementary strands. Supercoiled plasmid DNA was isolated and purified on CsCl ethidium bromide gradients (Sambrook et al. 1990). Denaturing acrylamide gels (6%) were performed using a Sequi-gen apparatus (Biorad) and run at 2000 V from 90 to 300 min. Sequence analyses were done using Apple Macintosh computer programs, DNA Strider (Marck 1988), L FastA (Pearson 1990) and ClustalV (Higgins et al. 1992).

Results

Identification of thioredoxin sequences in the tobacco genome

Using the complete thioredoxin cDNA h_1 as a probe, we assayed for the presence of homologous sequences in the *N. tabacum* genome. Southern blot analysis of genomic DNA digested with various restriction endonucleases indicates that thioredoxin is encoded by a small family of genes (Fig. 1). Two bands of similar intensities were obtained in *Eco*RI (4.5 and 3.3 kb), *Hin*dIII (2.8 and 2.5 kb) and *Sph*I (19.0 and 4.2 kb) digests. Two bands of similar intensity and one less intense band were obtained

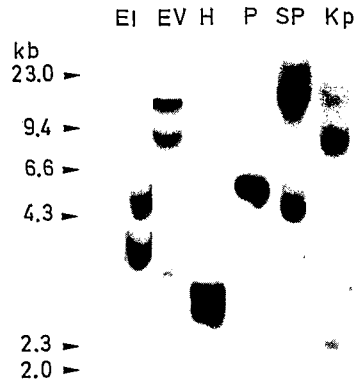

Fig. 1. Southern hybridization analysis of *Nicotiana tabacum* genomic DNA with the thioredoxin h_1 cDNA probe: 30 µg of genomic DNA was digested by *Eco*RI (EI), *Eco*RV (EV), *Hin*dIII (H), *Pst*I (P), *Sph*I (Sp) and *Kpn*I (Kp), separated on an agarose gel and transferred onto nylon membrane. Subsequent processing of the Southern blot is described in the Materials and methods

in *Eco*RV digests (15.3, 7.5 and 3.0), one in *Pst*I (4.8 kb) and one faint and one intense band were obtained in *Kpn*I (7.0 and 2.2 kb) digests. Additional bands of very low intensity (not visible in Fig. 1) were obtained with *Hin*dIII (one band), *Pst*I (three), *Sph*I (one band) and *Kpn*I (three bands). Copy number reconstruction with DNA digested with *Eco*RI indicates that there is approximately one gene copy per haploid genome. The band patterns and intensity of hybridization can be consistently interpreted to give a total per genome of two genes which are homologous to thioredoxin h_1.

Isolation and characterization of a genomic thioredoxin clone

On the basis of the Southern analysis and of the restriction map of thioredoxin h_1 cDNA (an *Eco*RV site is located 20 nucleotides downstream from the stop codon), the five size fractions (see the Materials and methods) were digested by *Eco*RV and tested by hybridization with the thioredoxin h_1 5′ probe (data not shown). The double digestion with *Eco*RI/*Eco*RV suggests that the coding sequence ends at 3.2 and 2.2 kb respectively on the 4.5 and 3.3 kb *Eco*RI fragments. We chose to clone the 4.5 kb fragment which presents the largest upstream sequence. One positive genomic clone was isolated and approximately 2.5 kb of the DNA sequence of the genomic clone was determined. The complete sequence of the gene and its immediate 3′ and 5′ flanking regions is given in Fig. 2 and has been submitted to the EMBL data bank under the accession number Z11803.

The sequence contains an open reading frame corresponding to a polypeptide of 118 amino acids with the conserved active site Trp-Cys-Gly-Pro-Cys. The coding region is interrupted by two introns which have lengths of 1071 and 147 bp respectively. The exons encode peptides of 29, 41 and 48 amino acids respectively. The intron sequences are rich in A/T (65%) compared to the coding sequence (52%), as described for other plant in-

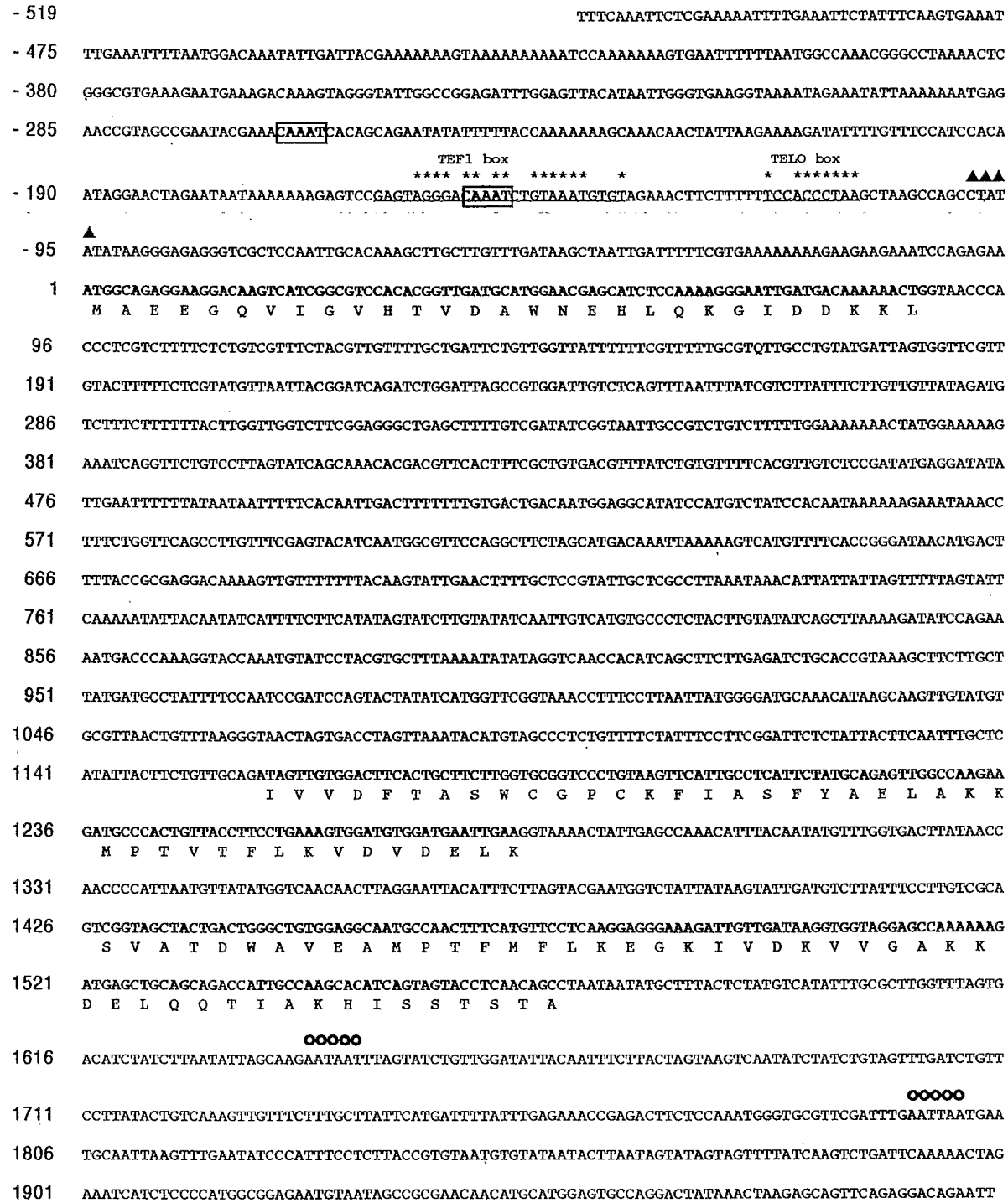

Fig. 2. Nucleotide and deduced amino acid sequences of thioredoxin h_2 gene. The putative TATA box is indicated above the sequence by *filled triangles,* the putative CAT boxes are *boxed,* the putative TEF1 and TELO boxes are indicated by *asterisks* and the putative polyadenylation signals are shown by *open circles*

trons (Sauer et al. 1990). Both introns start with the typical three nucleotides GTA and end with CAG according to Brown (1986) and Rogers (1990). Furthermore, we have found the characteristic branch point according to the consensus published by Brown (1986) and Sauer et al. (1990). The 3′ branch sites (CTCAT and CTTAT) begin 24 and 18 nucleotides upstream from the end of the respective intron.

The 5′ flanking sequences of the thioredoxin gene contain the sequence TATATATAA at position −98 from the ATG. The 5′ flanking regions also contain two conserved boxes containing the pentanucleotide CAAAT at positions −266 and −152 from the ATG. In addition, sequences homologous to the telo box (position −120) and to the tef1 box (position −161) are detectable and located as in the promoter of the translation elongation

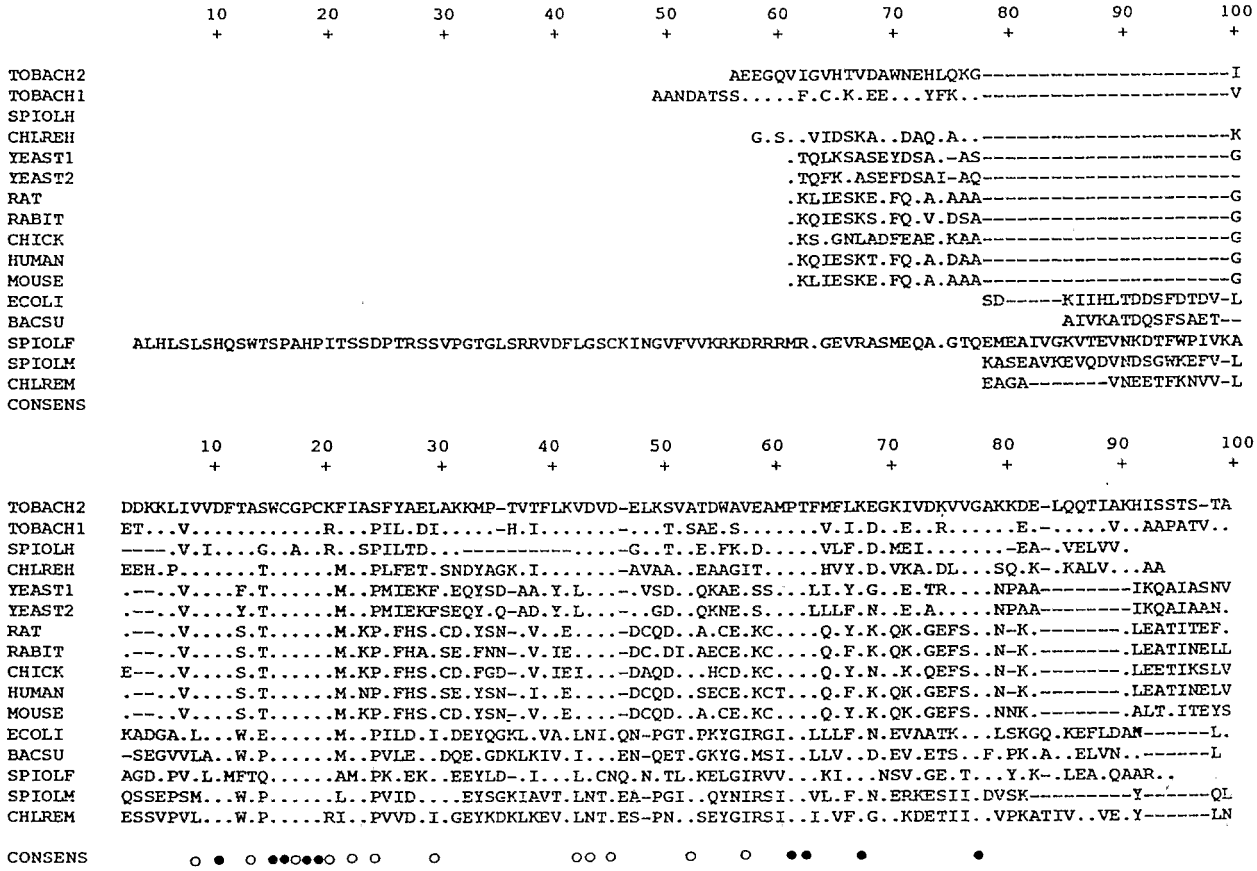

Fig. 3. Homologies between the deduced amino acid sequences of thioredoxin h_2 (TOBACH2) and the protein sequences of other thioredoxins: *N. tabacum* thioredoxin h_1 (TOBACH1; Marty and Meyer 1991), spinach (SPIOLH, Marcus et al. 1991), h-Chlamydomonas (CHLREH; Decottignies et al. 1990), yeast 1 and 2 (Gan 1991), rat (Tonissen et al. 1989), rabbit (Johnson et al. 1988), chicken (Jones and Luk 1988), human (Wollman et al. 1988), mouse (Tagaya et al. 1989), *Escherichia coli* (Holmgren 1985), *Bacillus subtilis* (BACSU; Chen et al. 1989), f-spinach (SPIOLF; Kamo et al. 1986), m-spinach1 (SPIOLM; Maeda et al. 1986) and m-*Chlamydomonas* (CHLREM; Decottignies et al. 1991). *Dots* correspond to identical and the *dashes* to deleted amino acid residues

factor EF-1α1 of *Arabidopsis thaliana* (Axelos et al. 1989; Curie et al. 1991, 1992).

In the 3′ flanking region two putative polyadenylation signals were identified and 248 nucleotides downstream from the termination codon, which conform to the published consensus sequence (AATAAA; Joshi 1987) with substitution of one base. Nevertheless, the second polyadenylation signal is more probable based on the length of the 3′ sequence of thioredoxin h_1 cDNA and the length of the mRNA (approximately 900 bp) deduced from Northern blot analysis of thioredoxin h_2 expression (see below).

Analysis of the deduced amino acid sequence of the open reading frame shows that we have isolated a gene which encodes a new tobacco thioredoxin (Fig. 3) having only 66% identity with thioredoxin h_1. Thioredoxin h_2 is shorter (118 amino acids) than h_1 (126 amino acids) and very different in the C-terminal region. Indeed, the N-terminal sequence ANDATSS is not present in thioredoxin h_2. Thioredoxin h_2 shows 66% identity with the thioredoxin h_1 amino acid sequence, 57% with the spinach thioredoxin h, 46% with the *Chlamydomonas* thioredoxin h, from 44 to 42% with the animal thioredoxins, about 38% with the prokaryotic thioredoxins and only 33% (type m) to 38% (type f) with the spinach chloroplastic thioredoxins.

Characterization of sequences homologous to thioredoxins h_1 and h_2 in higher plants

In order to perform a detailed analysis of the thioredoxin h genes in tobacco and other plants, two probes were prepared corresponding to the 3′ regions of the mRNAs encoding thioredoxins h_1 and h_2. These sequences show sufficient divergence to permit specific detection of each gene (Fig. 4). Southern blots performed with the complete thioredoxin h_1 cDNA or with the specific probes (3′ thioredoxin h_1 and 3′ thioredoxin h_2) on *Eco*RI digests of *N. tabacum* confirmed that thioredoxin h_2 is encoded in the 4.5 kb fragment and show that thioredoxin h_1 is encoded by a 3.3 kb fragment. *N. sylvestris* and *tomentosiformis* each show only one band, identical in size and homologous to thioredoxin h_2 but have no sequence homology to the 3′ region of thioredoxin h_1. Genomic DNA from various dicots, including rapeseed (*Brassica*

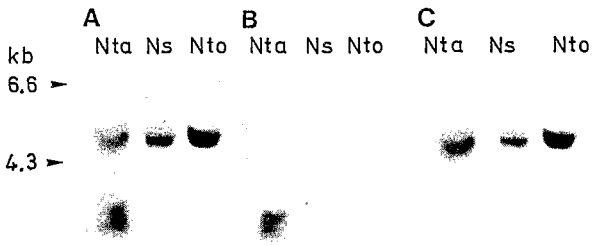

Fig. 4A–C. Southern hybridization analysis of DNA from different tobacco species with thioredoxin probes. Ten micrograms of genomic DNAs from *N. tabacum* (Nta), *N. sylvestris* (Ns) and *N. tomentosiformis* (Nto) were digested with *Eco*RI, separated on an agarose gel, transferred onto nylon membrane and hybridized with **A** thioredoxin h_1 cDNA **B** 3′ thioredoxin h_1 **C** and 3′ thioredoxin h_2 probes at 37° C and without the stringent wash

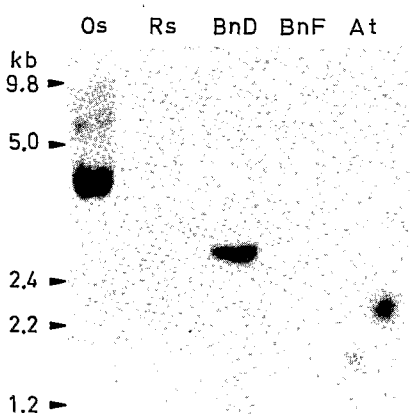

Fig. 5. Southern hybridization analysis of DNA from different plant species with thioredoxin h_1 cDNA as probe. Ten micrograms of genomic DNAs from *Oryza sativa* (Os), *Raphanus sativus* (Rs), *Brassica napus* var. Darmor (BnD), *B. napus* var. Furax (Bnf) and *Arabidopsis thaliana* (At) were digested with *Eco*RI, separated on an agarose gel, transferred onto nylon membrane, then hybridized to thioredoxin h_1 cDNA at 37° C and without the stringent wash

napus ev. Darmor and Furax), radish (*Raphanus sativus*) *A. thaliana* and from one monocot, rice (*Oryza sativa* ev. Columbia), were digested with *Eco*RI (Fig. 5). After electrophoresis and transfer to nylon membrane, the corresponding Southern blot was hybridized with the thioredoxin h_1 probe. One hybridization band was observed with *O. sativa* (4.0 kb) and *B. napus* ev. Darmor (3.0 kb) The same pattern was obtained with the 3′ thioredoxin h_2 probe but not with the 3′ thioredoxin h_1 probe; in contrast, no hybridization signals were observed for the other species. Similar results were obtained after *Hin*dIII, *Eco*RV and *Sph*I digestion (data not shown).

Expression of h-thioredoxins in tobacco plants

Total RNA from various organs of adult and young tobacco plants were separated on agarose gels and the Northern blot was successively hybridized with the thioredoxin h_1 probe and the specific probes (Fig. 6A, B–C). The choice of tissues was based on their develop-

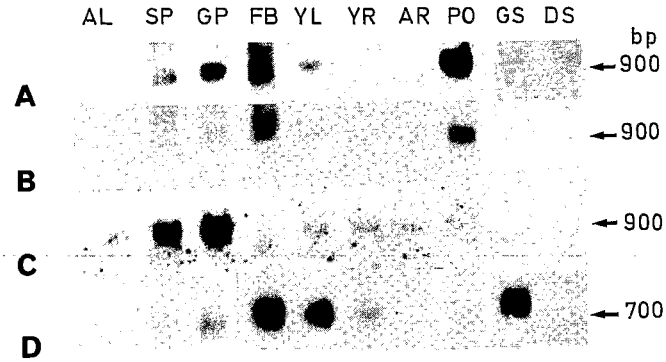

Fig. 6A–D. Expression of tobacco thioredoxin genes (h_1 and h_2) in different tissues. Ten micrograms of total RNAs extracted from different tissues of adult and young tobacco plants, cells and protoplasts were blotted onto nylon membrane and hybridized to different specific, ^{32}P-labelled probes. **A** thioredoxin h_1 probe, **B** 3′ thioredoxin h_1 probe, **C** 3′ thioredoxin h_2 probe and **D** *Arabidopsis* histone H4 probe. The ^{32}P-labelled tobacco probes and *Arabidopsis* histone H4 cDNA probe were successively hybridized to the Northern blot. AL, Adult leaves; SP, stationary phase (cells of 12-day-old culture); GP, growth phase (cells of 3-day-old culture); FB, flower buds; YL, young leaves; YR, young roots; AR, adult roots; PO, freshly isolated protoplast; GS, germinating seeds; DS, dry seeds

mental stage: on the one hand, tissues that contain dividing cells such as young leaves, flower buds, young roots, germinating seeds and, on the other hand tissues having no cell division, such as adult leaves, roots and dry seeds. The corresponding Northern blot was hybridized to *Arabidopsis* histone H4 as an internal standard indicating cell division (Fig. 6D). The mRNAs for thioredoxins h were both localized in the 0.9 kb region: however mRNAs encoding the chloroplastic thioredoxins are larger 1.0 to 1.1 kb in size (Kamo et al. 1989; Wedel et al. 1992) and do not hybridize with our probes. Thioredoxin h mRNAs are detectable in all young tissues that contain dividing cells (Fig. 6A), while in contrast the mRNAs are barely detectable in adult tissues. Nevertheless, in germinating seedlings the level of thioredoxin h mRNA is low (Fig. 6A). Analysis by Northern blots using the specific probes shows that most of mRNA is homologous to the thioredoxin h_2 gene (Fig. 6C) but that transcripts homologous to thioredoxin h_1 are always detectable in growing tissues of tobacco plants (Fig. 6B).

Expression of h-thioredoxins in tobacco cells and protoplasts

The thioredoxin h_1 cDNA clone was isolated from a cDNA library prepared from in vitro cultured cells, which was screened for genes expressed at a higher level in growing cells than in stationary phase cells. Accordingly, tobacco cells in culture contain 0.9 kb mRNAs which hybridize to the complete thioredoxin h_1 probe, accumulate abundantly in growing cells and are present at low levels in stationary phase cells (Fig. 6A). When the Northern blots were probed with the specific sequences, in vitro cultured cells were found to contain mRNAs of similar size homologous to both genes, the thioredoxin

h_2 transcripts (Fig. 6C) being more abundant than the thioredoxin h_1 transcript (Fig. 6B). In contrast, freshly isolated mesophyll protoplasts showed only mRNA corresponding to the thioredoxin h_1 sequence.

Discussion

Similarity of the tobacco h-thioredoxins with other thioredoxins

Two tobacco thioredoxin sequences have been isolated. The deduced protein sequences present 57 to 68% amino acid identity with each other and with the partial spinach thioredoxin h sequence (Marcus et al. 1991), indicating that the three enzymes are closely related. The spinach enzyme is localized in the cytoplasm, and is reduced by an NADPH-dependent thioredoxin reductase. Thioredoxins h_1 and h_2 encode proteins that contain the thioredoxin active centre Cys-Gly-Pro-Cys and are of small size (14 kDa). In addition, the thoredoxin h_1 has recently been produced in *Escherichia coli,* using the pET expression vector (Dubendorff and Studier 1991). The enzyme shows typical thioredoxin activity in vitro (P. Marinho et al. in preparation). The thioredoxins encoded by the two tobacco clones do not show any signal peptides and are most probably cytoplasmic. Thioredoxin activities with the same enzymatic characteristics and localization have also been detected in carrot cells cultured in vitro (Buchanan 1980; Cséke and Buchanan 1986; Johnson et al. 1987a) and in wheat seeds (Vogt and Follmann 1986; Johnson et al. 1987b; Kobrehel et al. 1991), indicating that h-thioredoxins are widely distributed in plants.

This is the first cloning of a gene encoding this protein in a higher plant, where studies are limited by the considerable variability of thioredoxin sequences which impairs screening with heterologous probes. Nevertheless, all thioredoxins have very similar enzymatic properties in vitro. For example, they are all able to activate the chloroplastic NADP malate dehydrogenase (Jacquot et al. 1981) or insulin (Holmgren 1989). In contrast, the in vivo function of thioredoxins is well established only for the chloroplastic thioredoxins, which have been shown to regulate CO_2 assimilation and to be essential for photosynthetic growth (Cséke and Buchanan 1986; Muller and Buchanan 1989). All other eukaryotic thioredoxins have been named heterotrophous thioredoxins, but it is not clear if they control similar pathways in different organisms. The clones we have isolated encode proteins that are closely related to the spinach thioredoxin h, and more loosely to the yeast, *Chlamydomonas* and to vertebrate thioredoxins. Other thioredoxins, including the eukaryotic *Dictyostelium* (Wetterauer et al. 1992) and the nuclear-encoded chloroplastic plant thioredoxins have only the active dithiol site in common and a few amino acids.

Further proof of the large variability of the thioredoxin sequences arises from the analysis of Southern blots. Although sequences homologous to tobacco thioredoxin h_1 were detected in a monocot (rice) and in *B. napus* (cv. Darmor), under identical conditions no hybridization was found in other *Brassicaceae* (radish and *Arabidopsis*). Nevertheless using this h_1 probe we recently isolated a thioredoxin h clone from an *Arabidopsis* cDNA library prepared in lambda ZAPII (Rivera-Madrid et al. 1992). It shows 67% nucleotide identity with thioredoxin h_1, i.e. sufficient homology to produce a signal on phage plates (enriched in the sequence) but not on Southern blots.

Analysis of the structure of thioredoxin h_2 gene

Using the sequence of the thioredoxin h_1 cDNA, we have localized the coding sequence and deduced the intron sites and sequences in the genomic clone of thioredoxin h_2. The proposed scheme is in part hypothetical and must be confirmed by the isolation and sequencing of an h_2 cDNA. This is the first genomic clone encoding a thioredoxin isolated from a higher plant and may be compared to thioredoxin genomic sequences. Genomic clones isolated from yeasts (Gan 1991) do not have introns, a situation common to nuclear genes in yeast. One genomic clone recently isolated from human cells (Tonissen and Wells 1991) contains five introns, which interrupt the coding sequence at positions different from the tobacco introns. In addition, a human gene encoding a human prolyl 4-hydroxylase beta-subunit (Tasanen et al. 1988) presents two thioredoxin domains. Each domain is interrupted by two introns at positions different from the human and tobacco thioredoxin introns. Thus it appears that thioredoxin genes contain introns but their positions in the genes are not conserved, suggesting that the various thioredoxin genes are not closely related.

Origin of thioredoxin h_1 and h_2 in N. tabacum *genome*

N. tabacum is an allotetraploid plant which originates from hybridization between *N. sylvestris* and *N. tomentosiformis*. Surprisingly the presumed ancestors contain only thioredoxin h_2 genes although, for example, each homologous gene coding for nitrate reductase present in *N. tabacum* originates from a different ancestor (Vaucheret et al. 1989). Due to the recent appearance of *N. tabacum* it is unlikely that thioredoxin h_1 derives from thioredoxin h_2. Since its domestication *N. tabacum* has been crossed with different *Nicotiana* species in order to introduce genes of agricultural importance, mostly associated with resistance to pathogens. Thus, it is likely that the thioredoxin h_1 gene originates from another *Nicotiana* and possibly confers on *N. tabacum* a phenotype that is not confered by thioredoxin h_2, due to differences between the coding sequences or due to the different regulation in gene expression. In this study we have used two varieties of *N. tabacum:* the cell suspension was derived from Xanthi, the greenhouse plants were Maryland. Both contain and express both genes, but this is not necessarily the case for all tobacco varieties. An analysis of the genetic origin of thioredoxin h_1 may shed light on the function of this gene in *N. tabacum*.

Differential expression of thioredoxin h_1 and h_2 genes

Multiple thioredoxin genes appears to be an exceptional situation: *E. coli* and vertebrates possess only one thioredoxin gene per haploid genome. Nevertheless, *Dictyostelium discoideum* has at least three genes which are co-ordinately expressed (Wetterauer et al. 1992). In *Saccharomyces cerevisiae* two similar genes are functionally exchangeable (Gan 1991; Muller 1991) while a third is very divergent (Bolle et al. 1992) but there is currently no information concerning its expression. In higher plants the presence of multiple thioredoxins h has been reported (Vog and Follmann 1986); however whether they are encoded by different genes or arise through post-translational modifications is not known. In this paper we show that *N. sylvestris* and *N. tomentosiformis* have only one thioredoxin h gene that is hybridizable with our probe. In contrast *N. tabacum* contains at least two genes encoding thioredoxins h and the levels of the corresponding mRNAs levels are not regulated in the same way. The characterization of the phylogenetic origin of thioredoxin h_1 should allow us to determine if thioredoxin h_1 is specialized for a particular protein target and is always associated with another h-thioredoxin or if this protein can catalyze alone the reduction of the disulphide bridges of all cellular proteins.

Acknowledgements. C. Brugidou was supported by an OSTROM molecular biology programme. The authors thank P. Gallois, M. Raynal, R. Cooke and F. Grellet (Perpignan) for their fruitful, stimulating, scientific and technical discussions. Information on the possible origin of the thioredoxin h_1 gene from R. Delon (Seita-Bergerac) was extremely useful. B. Lescure (Toulouse) helped us very efficiently in the analysis of the promoter.

References

Axelos M, Barbet C, Liboz T, le Van Thaï A, Curie C, Lescure B (1989) The gene family encoding the *Arabidopsis thaliana* translation elongation factor EF-1α: molecular cloning, characterization and expression. Mol Gen Genet 219:106–112

Bolle PA, Gilliquet V, Berben G, Dumont J, Hilger F (1992) Yeast sequencing reports: The complete sequence of K3B, a 7.9 kb fragment between PGK1 on chromosome III, reveals the presence of seven open reading frames. Yeast 8:205–213

Brown JWS (1986) A catalogue of splice junction and putative branch point sequences from plant introns. Nucleic Acids Res 14:9549–9559

Buchanan BB (1980) Role of light in chloroplast enzymes. Annu Rev Plant Physiol 31:341–374

Chaboute ME, Chaubet N, Philipps G, Gigot C (1987) Genomic organization and nucleotide sequences of two histone H3 and two histones H4 genes of *Arabidopsis thaliana*. Plant Mol Biol 8:179–191

Chen NY, Zhang JJ, Paulus H (1989) Nucleotide sequence of the overlapping genes for the subunits of Bacillus subtilis aspartokinase II and their control regions. J Biol Chem 262:8787–8798

Curie C, Liboz T, Barbet C, Gander E, Medale C, Axelos M, Lescure B (1991) Cis and trans-acting elements involved in the activation of *Arabidopsis thaliana* A1 gene encoding the translation elongation factor EF-1α. Nucleic Acids Res 19:1305–1310

Curie C, Liboz T, Montané MH, Rouan D, Axelos M, Lescure B (1992) The activation process of *Arabidopsis thaliana* A1 gene encoding the translation elongation factor EF-1α is conserved among angiosperms. Plant Mol Biol 18:1083–1090

Cséke C, Buchanan BB (1986) Regulation of the formation and utilization of photosynthate in leaves. Biochem Biophys Acta 853:43–63

Decottignies P, Schmitter JM, Jacquot JP, Dutka S, Picaud A, Gadal P (1990) Purification, characterization, and complete amino acid sequence of a thioredoxin from a green alga, *Chlamydomonas reinhardtii*. Arch Biochem Biophys 280:112–121

Decottignies P, Schmitter JM, Dutka S, Jacquot JP, Miginiac-Maslow M (1991) Characterization and primary structure of a second thioredoxin from the green alga, *Chlamydomonas reinhardtii*. Eur J Biochem 198:505–512

Delcasso-Tremousaygue F, Pannabieres F, Grellet F, Ananiev E, Delseny M (1988) Structural and transcriptional characterization of the external spacer of a ribosomal RNA nuclear gene from a higher plant. Eur J Biochem 172:767–776

Droux M, Jacquot JP, Miginiac-Maslow M, Gadal P, Huet JC, Crawford NA, Yee BC, Buchanan BB (1987) Ferredoxin-thioredoxin reductase: an iron-sulfur enzyme linking light to enzyme regulation in oxygenic photosynthesis. Purification and properties of the enzyme from C3, C4 and cyanobacterial species. Arch Biochem Biophys 252:426–439

Dubbendorff JW, Studier FW (1991) Controlling basal expression in an inducible T7 expression system by blocking the target T7 promotor with lac repressor. J Mol Biol 219:45–59

Florencio FJ, Yee BC, Johnson TC, Buchanan BB (1988) An NADP/thioredoxin system in leaves: purification and characterization of NADP-thioredoxin reductase and thioredoxin h from spinach. Arch Biochem Biophys 266:496–507

Freedman RB, Hawkins HC, Murant SJ, Reid L (1988) Protein disulphide-isomerase: a homologue of thioredoxin implicated in the biosynthesis of secretory proteins. Biochem Soc Trans 16:96–99

Gan ZR (1991) Yeast thioredoxin genes. J Biol Chem 266:1692–1696

Gleason FK, Holmgren A (1988) Thioredoxin and related proteins in procaryotes. FEMS Microbiol Rev 54:271–298

Gleason FK, Whittaker MM, Holmgren A, Jörnvall H (1985) The primary structure of thioredoxin from the filamentous cyanobacterium *Anabaena* sp. 7119. J Biol Chem 260:9567–9573

Grosset J, Marty I, Chartier Y, Meyer Y (1990) mRNAs newly synthesized by tobacco mesophyll protoplasts are wound inducible. Plant Mol Biol 15:485–496

Higgins DG, Bleasby AJ, Fuchs R (1992) ClustalV: Improved software for multiple sequence alignment. Comp Appl Biol Sci (in press)

Holmgren A (1989) Thioredoxin and glutaredoxin systems. J Biol Chem 264:13963–13966

Huppe HC, de Lamotte-Guéry F, Jacquot JP, Buchanan BB (1990) The ferredoxin-thioredoxin system of a green alga, *Chlamydomonas reinhardtii*. Planta 180:341–351

Jacquot JP (1984) Post translational modification of proteins in higher plant chloroplasts: enzyme regulation by thiol disulfite interchange. Physiol Veg 22:487–507

Jacquot JP, Buchanan BB, Martin F, Vidal J (1981) Enzyme regulation in C4 photosynthesis. Purification and properties of thioredoxin-linked NADP-malate dehydrogenase from corn leaves. Plant Physiol 68:300–304

Jouanneau JP, Tandeau de Marsac N (1973) Stepwise effects of cytokinin activity and DNA synthesis upon mitotic cycle events in partially synchronised tobacco cells. Exp Cell Res 77:167–174

Johnson RS, Mathews WR, Biemann K, Hopper S (1988) Amino acid sequence of thioredoxin isolated from rabbit bone marrow determined by tandem mass spectrometry. J Biol Chem 263:9589–9597

Johnson TC, Cao RQ, Kung JE, Buchanan BB (1987a) Thioredoxin and NADP-thioredoxin reductase from cultured carrot cells. Planta 171:321–331

Johnson TC, Wada K, Buchanan BB, Holmgren A (1987b) Reduction of purothionin by the wheat seed thioredoxin system. Plant Physiol 85:446–451

Jones SW, Luk KC (1988) Isolation of a chicken thioredoxin cDNA clone. J Biol Chem 263:9607–9611

Joshi CP (1987) Putative polyadenylation signals in nuclear genes of higher plants: a compilation analysis. Nucleic Acids Res 15:9627–9640

Kamo M, Tsugita A, Wiessner C, Wedel N, Bartling D, Herrmann RG, Aguilar F, Gardet-Salvi L, Schürmann P (1989) Primary structure of spinach chloroplast thioredoxin F: protein sequencing and analysis of complete cDNA clones for spinach chloroplast. Eur J Biochem 154:197–203

Kobrehel K, Yee BC, Buchanan BB (1991) Role of the NADP thioredoxin system in the reduction of alpha-amylase and trypsin-inhibitor proteins. J Biol Chem 266:6135–6140

Lundström J, Holmgren A (1990) Protein disulfide-isomerase is a substrate for thioredoxin reductase and has thioredoxin-like activity. J Biol Chem 265:9114–9120

Maeda K, Tsugita A, Dalzoppo D, Vilbois F, Schürmann (1986) Further characterization and amino acid sequence of m-type thioredoxins from spinach choroplasts. Eur J Biochem 154:197–203

Marck C (1988) Strider 1–1. Nucleic Acids Res 16:1829–1836

Marcus F, Chamberlain SH, Chu C, Masiarz FR, Shin S, Yee BC, Buchanan BB (1991) Plant thioredoxin-H-an animal-like thioredoxin occurring in multiple cell compartments. Arch Biochem Biophys 287:195–198

Marty I, Meyer Y (1991) Nucleotide sequence of a cDNA encoding a tobacco thioredoxin. Plant Mol Biol 17:143–147

Meinkoth J, Wahl G (1984) Hybridisation of nucleic acid immobilized on solid support. Anal Biochem 138:267–284

Meyer Y, Abel WO (1975) Importance of the wall for cell division and in the activity of the cytoplasm in cultured tobacco protoplasts. Planta 123:33–40

Modrich P, Richardson CC (1976) Bacteriophage T7 deoxyribonucleic acid replication *in vitro*. A protein of *Escherichia coli* required for bacteriophage T7 DNA polymerase activity. J Biol Chem 250:5508–5514

Muller EGD (1991) Thioredoxin deficiency in yeast prolongs S phase and shortens the G1 interval of the cell cycle. J Biol Chem 266:9194–9202

Muller EGD, Buchanan BB (1989) Thioredoxin is essential for photosynthetic growth. The thioredoxin gene of *Anacystis nidulans*. J Biol Chem 264:4008–4014

Pearson WR (1990) Rapid and sensitive comparison with FASTAP and FASTA. Methods Enzymol 183:63–93

Reichard P (1988) Interactions between deoxyribonucleotide and DNA synthesis. Annu Rev Biochem 57:349–374

Rivera-Madrid R, Marinho P, Brugidou C, Chartier Y, Meyer Y (1992) Nucleotide sequence of a cDNA clone encoding an *Arabidopsis thaliana* thioredoxin h. Plant Physiol, in press

Rogers JH (1990) The role of introns in evolution. FEBS Lett 268:339–343

Russel M, Model P (1985) Thioredoxin is required for filamentous phage assembly. Proc Natl Acad Sci USA 82:29–33

Sanger F, Nicklens S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Sambrook J, Fritsch EF, Maniatis T (1990) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Sauer J, Friedländer K, Gräm-Wicke U (1990) Primary structure, genomic organization and heterologous expression of a glucose transporter from *Arabidopsis thaliana*. EMBO J 9:3045–3050

Tagaya Y, Maeda Y, Mitsui A, Kondo N, Matsui H, Hamuro J, Brown N, Arai K, Yokota T, Wakasugi H, Yodoi J (1989) ALT-derived factor (Adf), an Il-2 receptor Tac inducer homologous to thioredoxin – possible involvement of dithiol-reduction in the Il-2 receptor induction. EMBO J 8:757–764

Tasanen K, Parkkonen T, Chow LT, Kivirikko KI, Pihlajaniemi T (1988) Characterization of the human gene for a polypeptide that acts both as the beta subunit of prolyl 4-hydroxylase and as protein disulfide isomerase. J Biol Chem 263:16218–16224

Tonissen KF, Wells JRE (1991) Isolation and characterization of human thioredoxin encoding genes. Gene 102:221–228

Tonissen KF, Robin AJ, Wells JRE (1989) Nucleotide sequence of a cDNA encoding rat thioredoxin. Nucleic Acids Res 17:3973

Tsang MLS, Schiff JA (1976) Sulfate-reducing pathway in *Escherichia coli* involving bound intermediates J Bacteriol 125:923–933

Van Langendonckt A, Vanden Driessche (1992) Isolation and characterization of different forms of thioredoxins from the green alga *Acetabularia mediterranea:* identification of an NADP/thioredoxin system in the extrachloroplastic fraction. Arch Biochem Biophys 292:156–164

Vaucheret H, Vincentz M, Kronenberger J, Caboche M, Rouzé P (1989) Molecular cloning and characterisation of the two homologous genes coding for nitrate reductase in tobacco. Mol Gen Genet 216:10–5

Vogt K, Follmann H (1986) Characterization of three different thioredoxins in wheat. Biochim Biophys Acta 873:415–418

Wedel N, Clausmeyer S, Herrmann RG, Gardet-Salvi L, Schürmann P (1992) Nucleotide sequence of cDNAs encoding the entire precursor polypeptide for thioredoxin m from spinach chloroplasts. Plant Mol Biol 18:527–533

Wetterauer B, Jacquot JB, Veron M (1992) Thioredoxin from *Dictyostelium discoideum* are a developmentally regulated multigene family. J Biol Chem, in press

Wollman EE, d'Auriol L, Rimsky L, Shaw A, Jacquot JP, Wingfield P, Graber P, Dessarps F, Robin P, Galibert F, Bertoglio J, Fradelizi D (1988) Cloning and expression of a cDNA for a human thioredoxin. J Biol Chem 263:15506–15512

Mol Gen Genet (1993) 238:294–303

MGG

Genetic and physical mapping of barley telomeres

Marion S. Röder[1], **Nora L.V. Lapitan**[2], **Mark E. Sorrells**[1], **Steven D. Tanksley**[1]

[1] Department of Plant Breeding and Biometry, Cornell University, 252 Emerson Hall, Ithaca, NY 14853, USA
[2] Department of Agronomy, Colorado State University, Fort Collins, CO 80523, USA

Received: 17 August 1992 / Accepted: 19 October 1992

Abstract. Barley (*Hordeum vulgare* L.) telomeres were investigated by means of pulsed field gel electrophoresis (PFGE) and in situ hybridization. In situ hybridization showed that a tandemly repeated satellite sequence has a subtelomeric location, and is present at thirteen of the fourteen chromosome ends. PFGE revealed that this satellite sequence is physically close to the telomeric repeat. Pulsed field gel electrophoresis was then used for segregation analysis and linkage mapping of several telomeric and satellite loci in a segregating doubled-haploid population. The telomeric repeat displayed a hypervariable segregation pattern with new alleles occurring in the progeny. Eight satellite and telomeric sites were mapped on an restriction fragment length polymorphism (RFLP)-map of barley, defining the ends of chromosome arms 1L, 2S, 3L, 4S, 4L, 5S and 6. One satellite locus mapped to an interstitial site on the long arm of chromosome 3. The pyhsical location of this locus was confirmed by in situ hybridization to wheat/barley addition line 3.

Key words: Telomere mapping – Subtelomeric DNA satellite – Pulsed field gel electrophoresis – In situ hybridization – Barley (*Hordeum vulgare*)

Introduction

Telomeres constitute the outermost ends of eukaryotic chromosomes and are important in maintaining their integrity. Their specialized structure enables the replication of the ends of linear chromosomes. Telomeric DNA consists of stretches of a simple, tandemly repeated oligonucleotide with the general consensus sequence $(T/A)_{1-4}G_{1-8}$ (Zakian 1989). Oligonucleotide units can be added by the enzyme telomerase, which is a specialized reverse transcriptase (Blackburn 1991). This enzyme activity results in eukaryotic telomeres having a dynamic structure with lengths varying during the life cycle and in successive generations (Shampay and Blackburn 1988).

In higher plants, Richards and Ausubel (1988) have cloned telomeric sequences from *Arabidopsis thaliana* and determined the consensus sequence of the telomeric oligonucleotide to be TTTAGGG. In comparison to *Arabidopsis thaliana,* the consensus sequence of the telomeric repeat from tomato was determined as TT(T/A)AGGG (Ganal et al. 1991), suggesting a high level of conservation among the telomeric repeats in higher plant species. The *Arabidopsis* telomere clone and a synthetic oligonucleotide with the corresponding sequence have been used for in situ hybridization studies in barley and rye (Schwarzacher and Heslop-Harrison 1991; Wang et al. 1991) and were shown to hybridize exclusively to the outermost ends of chromosomes in both species. In barley, the terminal location of the 7 bp repeat was confirmed by Bal31 exonuclease digestion experiments (S. Wang et al., submitted).

In contrast to the evolutionarily conserved telomeric sequences, subtelomeric repeated sequences are much more divergent and are often species-specific. Subtelomeric satellite DNA sequences, which are organized in long arrays of tandemly repeated units, have been described for rye (Bedbrook et al. 1980), onion (Barnes et al. 1985), *Chironomus* (Saiga and Edström 1985), *Aegilops squarrosa* (Rayburn and Gill 1987), tomato (Lapitan et al. 1989) and rice (Wu et al. 1991). In tomato, the subtelomeric satellite DNA was determined to be very close (i. e. <150 kb) to the telomeres, based on pulsed field gel electrophoresis (Ganal et al. 1991). In other species, the analysis of subtelomeric regions has revealed more complex structures often containing several repeated elements (Chan and Tye 1983; Young et al. 1983; Corcoran et al. 1988; Cross et al. 1990).

In this study, we describe the subtelomeric localization of a barley satellite sequence by means of in situ hybridization. Using pulsed field gel electrophoresis, we mapped arrays of this satellite sequence and of the telomeric repeat on an RFLP map of barley.

Communicated by H. Saedler

Correspondence to: M. Röder