

# Molecular Characterization of Tobacco Sulfite Reductase: Enzyme Purification, Gene Cloning, and Gene Expression Analysis<sup>1</sup>

Keiko Yonekura-Sakakibara,<sup>\*2</sup> Toshihiko Ashikari,<sup>\*</sup> Yoshikazu Tanaka,<sup>\*</sup> Taka-aki Kusumi,<sup>\*</sup> and Toshiharu Hase<sup>†</sup>

<sup>\*</sup>Institute for Fundamental Research, Suntory Ltd., 1-1-1 Wakayama-dai, Shimamoto-cho, Mishima, Osaka 618-8503; and <sup>†</sup>Division of Enzymology, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871

Received for publication, April 27, 1998

A cDNA clone, *NtSiR1*, that encodes the precursor of ferredoxin-dependent sulfite reductase (Fd-SiR) has been isolated from a cDNA library of tobacco (*Nicotiana tabacum* cv. SR1). The identity of the cDNA was established by comparison of the purified protein and the predicted structure with the nucleotide sequence. The amino terminus of the purified enzyme was Thr<sub>62</sub> of the precursor protein, and the mature region of *NtSiR1* consisted of 632 amino acids. Tobacco Fd-SiR is 82, 77, and 48% identical with Fd-SiRs from *Zea mays*, *Arabidopsis thaliana*, and a cyanobacterium, respectively. Significant similarity was also found with *Escherichia coli* NADPH-SiR in the region involved in ligation of siroheme and the [4Fe-4S] cluster. On Northern blot analysis, a transcript of *NtSiR1* was detected in leaves, stems, roots, and petals in similar amounts. We also isolated a genomic SiR clone named *gNtSiR1*. It consists of 8 exons and 7 introns. Genomic Southern blot analysis indicated that at least two SiR genes are present in the tobacco genome.

**Key words:** cDNA and gene cloning, gene expression, *Nicotiana tabacum*, sulfite reductase, sulfur-assimilatory gene.

Higher plants assimilate inorganic sulfur and synthesize various sulfur-containing compounds, such as amino acids, lipids, and coenzymes. Assimilation of sulfur is one of the most important metabolic processes. After uptake from the soil by sulfate transporter(s), sulfate ions are first activated with ATP by ATP sulfurylase to form adenosine 5'-phosphosulfate (APS). APS is converted to 3'-phosphoadenosine-5'-phosphosulfate (PAPS) by APS kinase. PAPS is reduced to sulfite by PAPS reductase and sulfite is further reduced to sulfide by sulfite reductase (SiR). The resulting sulfide is finally assimilated into cysteine by cysteine synthase.

cDNAs for sulfate transporters (1, 2), ATP sulfurylase (3, 4), and APS kinase (5) have been cloned from higher plants. Expression of genes encoding sulfur metabolizing enzymes has recently been performed. In *Stylosanthes hamata* the genes for high-affinity sulfate transporters, *shst1* and *shst2*, were markedly induced in roots when

external sulfate ions were limited (1). Expression of ATP sulfurylase occurred in leaves, stems, roots, and flower buds but not in tubers in potato (3). Expression of ATP sulfurylase was enhanced by sulfate deprivation in *Arabidopsis thaliana* (4). The gene for APS reductase was also positively regulated by sulfur starvation in *A. thaliana* (6, 7). Although putative cDNAs encoding ferredoxin-dependent sulfite reductase (Fd-SiR) have been reported in *Zea mays* (8) and *A. thaliana* (9), the regulation of gene expression and extrapolation of the cDNAs to Fd-SiR polypeptides remain to be investigated.

SiR is localized in chloroplasts or plastids and catalyzes the six electron reduction of sulfite to sulfide. SiR contains a siroheme and a [4Fe-4S] cluster, and electrons required for the reductive reaction are transferred *via* ferredoxin (Fd). *Escherichia coli* has another type of SiR, which is composed of two subunits, a flavoprotein (SiRFP) containing FAD and FMN, and a hemoprotein (SiRHP) with the same prosthetic groups as Fd-SiR (10). *E. coli* SiR utilizes NADPH as an electron donor. Recently, the X-ray crystal structure of SiRHP was determined at 1.6 Å resolution, and the three-domain fold of the polypeptide backbone that controls cofactor assembly and catalytic reactivity has been revealed (11). Four important cysteine residues have been identified in the three-dimensional structure of SiRHP; three of which are liganded to three Fe atoms of the [4Fe-4S] cluster and the other of which bridges the Fe atoms of the [4Fe-4S] cluster and siroheme. An extensive hydrogen-bonding network of positive side chains of basic amino acids, water molecules and siroheme carboxylates is proposed to be involved in substrate recognition and catalysis.

<sup>1</sup> The nucleotide sequence data in this paper have been submitted to the DDBJ, EMBL, and GenBank nucleotide sequence databases under accession numbers D83583 (*NtSiR1*) and AB010717 (*gNtSiR1*).

<sup>2</sup> To whom correspondence should be addressed. Tel: +81-75-962-8807, Fax: +81-75-962-8262, E-mail: Keiko.Sakakibara@suntory.co.jp

Abbreviations: APS, adenosine 5'-phosphosulfate; Cbf1, centromere binding factor 1; CDEI, centromere DNA element I; Fd-, ferredoxin-dependent; NADPH-, NADPH-dependent; NiR, nitrite reductase; PAPS, 3'-phosphoadenosine-5'-phosphosulfate; 5'-RACE, 5'-rapid amplification of cDNA ends; RT, reverse transcription; SiR, sulfite reductase.

In this study, we have cloned both the cDNA and the gene encoding the precursor of tobacco Fd-SiR, and report here its genetic organization including its promoter region and expression in various organs. We also discuss the structural characteristics of Fd-SiR found on comparison with other siroheme-containing enzymes.

#### MATERIALS AND METHODS

**Plant Materials and Growth Conditions**—Tobacco (*Nicotiana tabacum* cv. SR1) plants were grown in the field from April to August in Japan, and 2.4 kg of green leaves was used for the purification of Fd-SiR. To obtain total RNA from green leaves, stems, roots, and petals, plants were grown in a standard greenhouse for 3 months.

**Fd-SiR Activity Assay and Determination of Protein Concentrations**—The activity of Fd-SiR was measured by the method of Arb and Brunold (12) with Fd or methyl viologen as an electron donor. One unit of Fd-SiR activity was defined as the amount that formed  $1 \mu\text{mol S}^{2-}$  per minute. Proteins were quantitated by Bradford's method (13) with bovine serum albumin as a standard.

**Purification of Tobacco Fd-SiR**—Approximately 2.5 kg of frozen leaf tissue was homogenized in 8 liters of buffer A [50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM EDTA, 1 mM MgCl<sub>2</sub>, 14 mM 2-mercaptoethanol, 1% (w/v) sodium ascorbate, and 0.5 mM phenylmethylsulfonyl fluoride] with 250 g of Polyclar AT in a Waring blender CB-6. The homogenate was passed through four layers of cheesecloth and then centrifuged at  $10,000 \times g$  at 4°C for 10 min. The resulting supernatant was fractionated with ammonium sulfate at 40–70% saturation and the precipitate was dissolved in a small volume of buffer B [20 mM Tris-HCl (pH 7.5), 1 mM MgCl<sub>2</sub>, and 1 mM EDTA], the solution was then desalted on a column of Sephadex G-25 (10 cm i.d.  $\times$  40 cm) equilibrated with buffer B. The desalted crude enzyme was directly loaded onto a column of DE-52 (5.4 cm i.d.  $\times$  25 cm) and then eluted with a linear gradient of NaCl, from 0 to 400 mM, in buffer B. Fractions containing Fd-SiR activity were pooled, concentrated by 70% ammonium sulfate precipitation, and then desalted on a column of Superdex 200 with buffer C [50 mM Tris-HCl (pH 7.5), 1 mM MgCl<sub>2</sub>]. The crude SiR was adsorbed on a small column of Fd-Sepharose (14) and eluted with a linear gradient of 0 to 400 mM NaCl in buffer C. Fd-SiR was further purified by two successive column chromatographies on Phenyl Superose (FPLC system) and Superdex 200 (SMART system), with development with a linear gradient, 40 to 0% saturation, of ammonium sulfate in buffer C, and with isocratic elution with buffer D [50 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 1 mM MgCl<sub>2</sub>], respectively.

**Determination of the Amino Terminal Sequence**—The purified Fd-SiR was subjected to SDS-PAGE (15) to remove a trace amount of contaminating proteins, electroblotted onto a polyvinylidene difluoride membrane (Immobilon™-P<sup>50</sup> transfer membrane; Millipore), and then analyzed with a gas-phase amino acid sequencer (model 477A/120A; Applied Biosystems).

**Screening of a cDNA Library**—A  $\lambda$ ZAPII cDNA library prepared from tobacco (cv. SR1) leaves was purchased from Stratagene. Fd-SiR cDNA was screened by nucleic acid hybridization using a cDNA for maize Fd-SiR (8) as a probe. The probe was labeled with <sup>32</sup>P-dCTP by the random

priming method (16). General methods for screening were used, as described by Sambrook *et al.* (17).

**5'-Rapid Amplification of cDNA Ends (5'-RACE)**—5'-RACE was performed with a 5'-RACE kit (GIBCO BRL) essentially as described by the manufacturer. The nucleotide sequence of the primer for PCR was 5'-GTCGTATTG-CATATAGCTTCC-3'. The amplified cDNA fragment was directly sequenced.

**Isolation of Genomic Clones**—A tobacco genomic library (*N. tabacum* cv. Xhanti-nc, Clontech) in EMBL3 was screened using *NtSiR1* as a probe. The probe was labeled with <sup>32</sup>P-dCTP as described previously. Four genomic clones was isolated from  $5.5 \times 10^6$  plaques. An approximately 6.3 kb genomic fragment was subcloned into the pUC19 plasmid.

**DNA Sequencing and Sequence Analysis**—Sequencing of cDNAs was performed by the dideoxy chain-termination method (18) using a PRISM DNA sequencing kit (Applied Biosystems) with an automated DNA sequencer (Model 373A; Applied Biosystems). Geneworks (IntelliGenetics) and GENETYX (Software Development) were used for computer analysis of nucleotide and deduced amino acid sequences.

**Northern Blot Analysis**—Total RNA was prepared from leaves, stems, roots, and petals by the method of McGookin (19). The RNA was size-fractionated by electrophoresis in a formaldehyde-containing agarose gel, and then transferred to a nylon membrane (Hybond N+; Amersham) according to the protocol recommended by the manufacturer. The probe was labeled by the random-primed method in the presence of <sup>32</sup>P-dCTP.

#### RESULTS AND DISCUSSION

**Purification and Amino-Terminal Analysis of Fd-SiR from Tobacco Leaves**—To obtain information on the primary structure of tobacco Fd-SiR, the enzyme was purified from green leaves as described under "MATERIALS AND METHODS." The final preparation of the enzyme was purified about 500-fold, and gave a major single band corresponding to a molecular mass of 64 kDa on SDS-PAGE (Fig. 1). The specific activity was 1.9 units/mg protein. Tobacco Fd-SiR gave absorption peaks at 395 and 580 nm in the visible region, which were derived from siroheme. The  $A_{395}/A_{280}$  and  $A_{580}/A_{280}$  ratios were 0.38 and 0.08, respectively. These spectroscopic characteristics were comparable to those of spinach Fd-SiR (20). The amino-terminal sequence was determined up to the eleventh residue to be X-Pro-Ala-Lys-Pro-Ala-Ala-Val-Glu-Pro-Lys, and was compared with those of Fd-SiRs from leek leaves (21), turnip leaves and roots (22) (Fig. 2). These amino-terminal sequences show significant homology, although there is considerable variation in the amino-terminal position. The amino-terminus of tobacco Fd-SiR starts at the same position as in turnip leaf Fd-SiR.

In leek leaves, Fd-SiR activity was separated into three materials on anion-exchange column chromatography, although they had the same amino-terminal sequence (21). In tobacco leaves, however, Fd-SiR activity was eluted as a single fraction at every step of purification, except for the "side-activity" due to Fd-NiR (23).

**Isolation of a cDNA Clone Encoding Fd-SiR**—The  $\lambda$ ZAPII tobacco cDNA library was screened with a maize

Fd-SiR cDNA (8) as a probe. Twelve positive plaques were obtained from  $3 \times 10^6$  recombinant phages. The inserts of the isolated clones ranged from 1.5 to 2.3 kb, and all clones analyzed seemed to be siblings on the basis of their restriction maps (data not shown). The longest clone among them, designated as *pNtSiR1*, was chosen for determination of the nucleotide sequence (Fig. 3). The insert of *pNtSiR1* had a reading frame encoding 672 amino acids, although the initiation codon was missing. The amino-terminal sequence of the purified Fd-SiR was found on the basis of the deduced amino acid sequence (Fig. 3, underlined). The missing region was amplified by 5'-rapid amplification of cDNA ends (5'-RACE), and the resulting amplified cDNA was directly sequenced. The overlapping region of *pNtSiR1* and the resultant DNA fragment obtained on 5'-RACE matched completely, and 167 bases extended beyond the 5'-terminus of *pNtSiR1*. A putative initiation methionine codon was found in the extended region. The precursor of tobacco Fd-SiR, designated as *NtSiR1*, therefore consists of 693 amino acids, and is processed into a mature polypeptide of 632 amino acids, whose calculated molecular mass is 70.2 kDa.

**Structural Comparison of Siroheme-Containing Enzymes**—Fd-SiR in higher plants is a plastidic protein requiring Fd as an electron donor (24). SiR in enterobacteria, composed of distinct subunits referred to as the flavoprotein and hemoprotein, utilizes NADPH as an electron donor (25). Biochemical studies have revealed that Fd-SiR and the hemoprotein contain a [4Fe-4S] cluster and one siroheme per subunit as cofactors, respectively (23). The mature region of *NtSiR1* exhibited significant homology to SiRs from various species. Figure 4 shows alignment of the amino acid sequences of Fd-SiRs from *Z. mays* (8), *A. thaliana* (9), and *Synechococcus* PCC7942 (26) with the hemoprotein of *E. coli* (25). Nitrite reductase (NiR), which catalyzes the reduction of nitrite to ammonia using Fd as an electron donor, is known to be another siroheme-containing enzyme, and Fd-NiR from tobacco (27) was included in the sequence comparison. At the amino acid level, *NtSiR1* is 77% homologous to the Fd-SiR of *Z. mays*, 82% to that of *A. thaliana*, 48% to that of *Synechococcus* PCC7942, 32% to the hemoprotein of *E. coli* SiR, and 18% to Fd-NiR of *N. tabacum*. It is noteworthy that only the tobacco and cyanobacterium Fd-SiRs have an extended

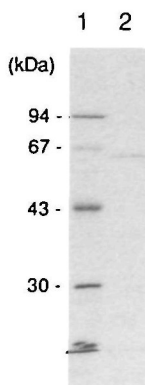


Fig. 1. SDS-PAGE of the purified tobacco Fd-SiR. About 0.5  $\mu$ g of purified Fd-SiR was subjected to SDS-PAGE and then stained with Coomassie Brilliant Blue. Molecular mass markers are shown at the left.

sequence in their carboxy-terminal region, and they are weakly homologous.

The three-dimensional structure of the hemoprotein of *E. coli* NADPH-SiR has been determined by X-ray analysis (11), which revealed that the siroheme and [4Fe-4S] cluster are retained in the active site of the enzyme through four cysteine ligands, Cys<sub>434</sub>, Cys<sub>440</sub>, Cys<sub>479</sub>, and Cys<sub>483</sub>. The four cysteines are strictly conserved in all plant Fd-SiRs (positions 502, 508, 548, and 552 in tobacco Fd-SiR), and their proximal regions show high similarity (Fig. 4), suggesting that the configurations around their redox centers are similar. The basic residues involved in the substrate coordination to siroheme in *E. coli* NADPH-SiR (11) are also highly conserved in cyanobacteria and higher plant Fd-SiR (Arg<sub>133</sub>, Arg<sub>166</sub>, Arg<sub>202</sub>, Lys<sub>284</sub>, and Lys<sub>286</sub> in tobacco Fd-SiR). These structural characteristics are also found in Fd-NiR, although the overall homology to Fd-SiR is not so high. This conservation suggests that the two enzymes diverged from a common ancestral gene. Fd-SiRs and Fd-NiR have common insertions from Ala<sub>250</sub> to Phe<sub>268</sub>, in comparison with the hemoprotein of *E. coli* NADPH-SiR (Fig. 4, Region A). The insertion seems to be located on the surface of the protein based on the superimposition of Fd-SiRs to the three-dimensional structure of the hemoprotein of *E. coli* NADPH-SiR. This region, unique to Fd-dependent enzymes, might be related to the protein-protein interaction with Fd.

**Expression of the Fd-SiR Gene in Various Tobacco Organs**—Molecular cloning of the cDNA enabled us to analyze the pattern of expression of tobacco Fd-SiR. Total RNAs were separately prepared from leaves, stems, roots, and petals of mature plants, and then the distribution of the Fd-SiR transcript was examined by Northern blotting. A transcript of 2.4 kb in length accumulated in all the organs examined to comparable levels (Fig. 5). The enzymatic activity of Fd-SiR has been reported to be present in leaves (20) and roots (28). Our results confirm the wide distribution of Fd-SiR in whole plants at the transcription level, and furthermore suggest that sulfite reduction occurs not only in leaves, stems, and roots, but also in petals.

It has been reported that a non-photosynthetic type of Fd is constitutively expressed in the petals and fruits of *Citrus* (29), suggesting the existence of an electron-donor system for SiR in these reproductive organs. It has also been reported that the cysteine synthase gene is expressed in petals of *A. thaliana* (30). This implies that, at least, sulfite is assimilated into cysteine in petals. Some sulfur-containing substances such as thionin are known to be accumulated in flowers (31). At present, the expression of other sulfur-assimilatory genes such as those of sulfate transporters and

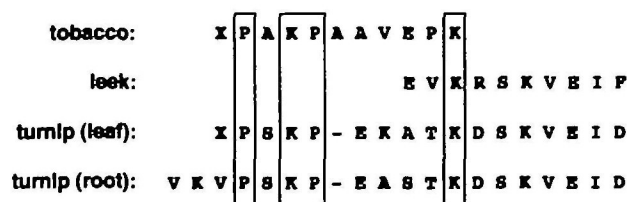


Fig. 2. Comparison of the amino-terminal sequences of Fd-SiRs. X indicates a residue not identified. Conserved amino acids in the sequences are boxed. Gaps are introduced to achieve maximum homology.

```

AAATATATCTGTAATACACTCTCTCCGCCATTTTCACAAATTTCTCTGTTCGTA AAAACCCCTAGCTTTATTTATCTATTTGTTATTCGCCT 90
AGGGGTAAAGGTGACGACGTCGTTTGGAGCAGCGATTAAACATCGCCGTCGCGATGACCCGAACCCCAAGCTCCAATTCACAACCTTCTC 180
      M T T S F G A A I N I A V A D D P N P K L Q I H N F S
CGGGTTAAAAGCACCTTCAATTCGGCTGTGCTTAGCAGGCGTCTTACAGTITTTTCAGTCCCTTTCCCCGTCGAACTAGTCTATTGT 270
G L K S T S N S L L L S R R L H V F Q S F S P S N P S S I V
CCGCCGCTATCTACGCCAGCAAAGCCAGCTGCAGTGGAGCCCAAGCGTAGTAAGGTTGAAATATTCAAAGAAGCAGAGTAACCTCAATAG 360
R A V S T P A K P A A V E P K R S K V E I F K E Q S N F I R
GTATCCTCTTAATGAGGAGATTCTAAATGATGCCCCCAACATCAATGAGGCTGCAACACAATTGATCAAGTTCATGGAAGCTATATGCA 450
Y P L N E E I L N D A P N I N E A A T Q L I K F H G S Y M Q
ATAGCACAGGATGAGCGTGGGGGAAGATCATCTCATCTCGGACAAAGAACCTGGTGGGGAGTACCAACACAGACTCTACTT 540
Y D R D E R G G R S Y S F M L R T K N P G G E V P N R L Y L
GGTCATGGATGATCTTGTGACCAATTTGGGATTGGGACACTTCGTTTGACAACAAGCAGAACCTTTCAGCTGCATGGGGTCTTGAAAA 630
V M D D L A D Q F G I G T L R L T T R Q T F Q L H G V L K K
AAATCTCAAGCAGTAATGAGTACAATCATCAAAAACATGGTTCAACTCTTGGTGCATGTGGTGACCTCAATAGGAACGTTCTTGCTCC 720
N L K I V M S T I I K N M G S T L G A C G D L N R N V L A P
AGTCCGCCCATTTGCTAAAAAGATTATATGTTTGCTAAACAAACAGCTGATAACATTGCAGCACTTTTAACTCCCCAGCTCGGATTTTA 810
A A P F A K K D Y M F A K Q T A D N I A A L L T P Q S G F Y
CTATGACGTTTGGGTGGATGGGGAGAAAGTTATGACAGCAGAACCTCCGAAAGTTGTGAAAGCTCGAAAATGATAACTCCCATGGAACAAA 900
Y D V W V D G E K V M T A E P P E V V K A R N D N S H G T N
CTTCCCTGACTCACTGAACCCATTTATGGAACCTCAGTCTTCCCAAGGAAGTTCAAAATGCAAGTACCGTGCCACTGCAATCACTCCGT 990
F P D S P E P I Y G T Q F L P R K F K I A V T V P T D N S V
GGACATTTTCACAAATGATATAGGTGTGTGTGTGTATCTAATGAGGATGGAGAGCCCTCAGGGATTCAACATATATGTTGGTGGTGGTAT 1080
D I F T N D I G V V V V S N E D G E P Q G F N I Y V G G G M
GGGGGCAACTCATAGGATGGAAACCACTTTTCTCGATTGGCAGAGCCATTAGGTTATGTGCCTAAAGAGGATATACTCTATGCTGTTAA 1170
G R T H R M E T T F P R L A E L L G Y V P K E D I L Y A V K
AGCCATTTGTTTACTCAAAGAGAAAACCGCAGAAAGATGATCGCAGATACAGCAGATTGAAATATTTACTCAGCTCAAGGAAATCGCA 1260
A I V V T Q R E N G R R D D R R Y S R L K Y L L S S W G I E
GAAGTTTCGATCTGCTCACTGAACAGTATTATGGAAAGAAATTTCAACCTTGCCGTGAATTCGCTGAGTGGGAAATCAAGAGTATTATGGG 1350
K F R S V T E Q Y Y G K K F Q P C R E L P E W E F K S Y L G
ATGGCAGCAGCGGGAGATGGTAGCTTGTGTTGTGGTCTCATGTTGACAATGGTTCGTA AAAAGGAGCGATGAAGAACGCTCAGGAA 1440
W H E A G D G S L F C G L H V D N G R V K G A M K K A L R E
AGTTATGAGAAGTATAATCTGAATGTCGCTCTCACACCAAAACAGAAATATTATTTGCAATATTCGACAAGCGTGGAGCGCCCAT 1530
V I E K Y N L N V R L T P N Q N I I L C N I R Q A W K R P I
CACCACCTTCTTGACAGGGTGGTTTGTGCAACCTAGGTATGTGATGCCACTCAATCTAACAGCAATGGCCCTGCCCGCTTTTCTCT 1620
T T V L A Q G G L L Q P R Y V D P L N L T A M A C P A F P L
TTGTCTCTTGAATAACTGAAGCTGAGCGTGGAAATACCTGCATCTCAAGCGTGTTCGAGCTATTTTGAAGGGTGGTCTGAAGTA 1710
C P L A I T E A E R G I P D I L K R V R A I F E R V G L K Y
CAGTGAATCTGTTGTCATAAGGATAACAGGATGTCCTAATGGGTGTGCTCGACCATACATGGCTGAACCTGGCTTGGTTGGAGATGGTCC 1800
S E S V V I R I T G C P N G C A R P Y M A E L G L V G D G P
AAACAGCTATCAGATCTGGCTCGGTGGAACCTCCCAATCAAACCTTCAAGTAAAGCTTAAGTTAAGTTCAAGTATCAGGATCTTGA 1890
N S Y Q I W L G G T P N Q T S L A K T F K D K L K V Q D L E
AAAAGTCTGGAGCTTTATTTTCCATTTGAGAAGAAAGCAGCAATCTAAAGAATCATTGGGACTTCCAAAACCCATGGGATTGA 1980
K V L E P L F F H W R R K R Q S K E S F G D F T N R M G F E
GAAACTTGGGGAGTTTGTGAAAAATGGGAAGGCATCCCTGAGTCATCATCTCGATATAACTTGAAGCTATTTGCTGATAGAGAGACTTA 2070
K L G E F V E K W E G I P E S S S R Y N L K L F A D R E T Y
CGAAGCCATGGATGCACTTGAAGCATCAAGAATAAAAATGCCATCAATTAGCAATTGAAGTGGTGGCAATATGTTGCTTPOCCAGCA 2160
E A M D A L A S I Q D K N A H Q L A I E V V R N Y V A S Q Q
AAATGGGAAAAGTATGGACTGACCATGCTCTTCTGTGCCAAATAGCAAGATTGAGCAAGGAGGTATAACTGGTCCCATGCTTTTAGC 2250
N G K S M D *
CTGTTATCTATGGTAACAGTTGAAACAAGTGGAGAATTTGACGTATATTTGCAAGCTTTGTAGTACTTCACTTGTAGACCAACAGTA 2340
CCCTGCAGTTTCTTTTCTGTTCAATTTTCTGTTCTATTTGAAGAGTAAACAAGATCTTACAATGGTCTGTTTTTCTTGTATTTTT 2430
TAGAGGATTGCTTTTAAACCTCAAAAAAAA

```

Fig. 3. The nucleotide and deduced amino acid sequences of tobacco Fd-SiR. The 5'-RACE product and the cDNA insert are combined to give one cDNA sequence for tobacco Fd-SiR. The amino acid sequence in the amino-terminal region of the mature form of Fd-

SiR is underlined. The direction and extent of the primer for 5'-RACE are indicated by an arrow. The dotted base indicates the 5'-end of the cDNA insert, *pNtSiR1*.

the sulfate-activating pathway in petals has not been reported. Further study of sulfur assimilation in petals is required.

**Isolation and Characterization of a Fd-SiR Gene**—A tobacco genomic library constructed with the EMBL3 vector was screened by hybridization with the tobacco cDNA for Fd-SiR, *NtSiR1*. Four hybridizing clones were obtained and mapped by digestion with a variety of restriction enzymes. They were divided into two groups, one group containing the entire region corresponding to *NtSiR1*. A clone in this group, designated as *gNtSiR1*, was

chosen for further analysis. The complete nucleotide sequence of *gNtSiR1* (5,190 bp) was determined and has been deposited in the DDBJ database under accession number AB010717. By aligning the genomic sequence with the cDNA sequence, the tobacco Fd-SiR gene was found to be composed of 8 exons interrupted by 7 introns, as shown in Fig. 6. All of the predicted introns had the GT-AG consensus border sequence (32). A single nucleotide difference between *gNtSiR1* and *NtSiR1* was found in the 3' untranslated region. The other group of positive clones, designated as *gNtSiR2*, was analyzed by partial sequencing, and was

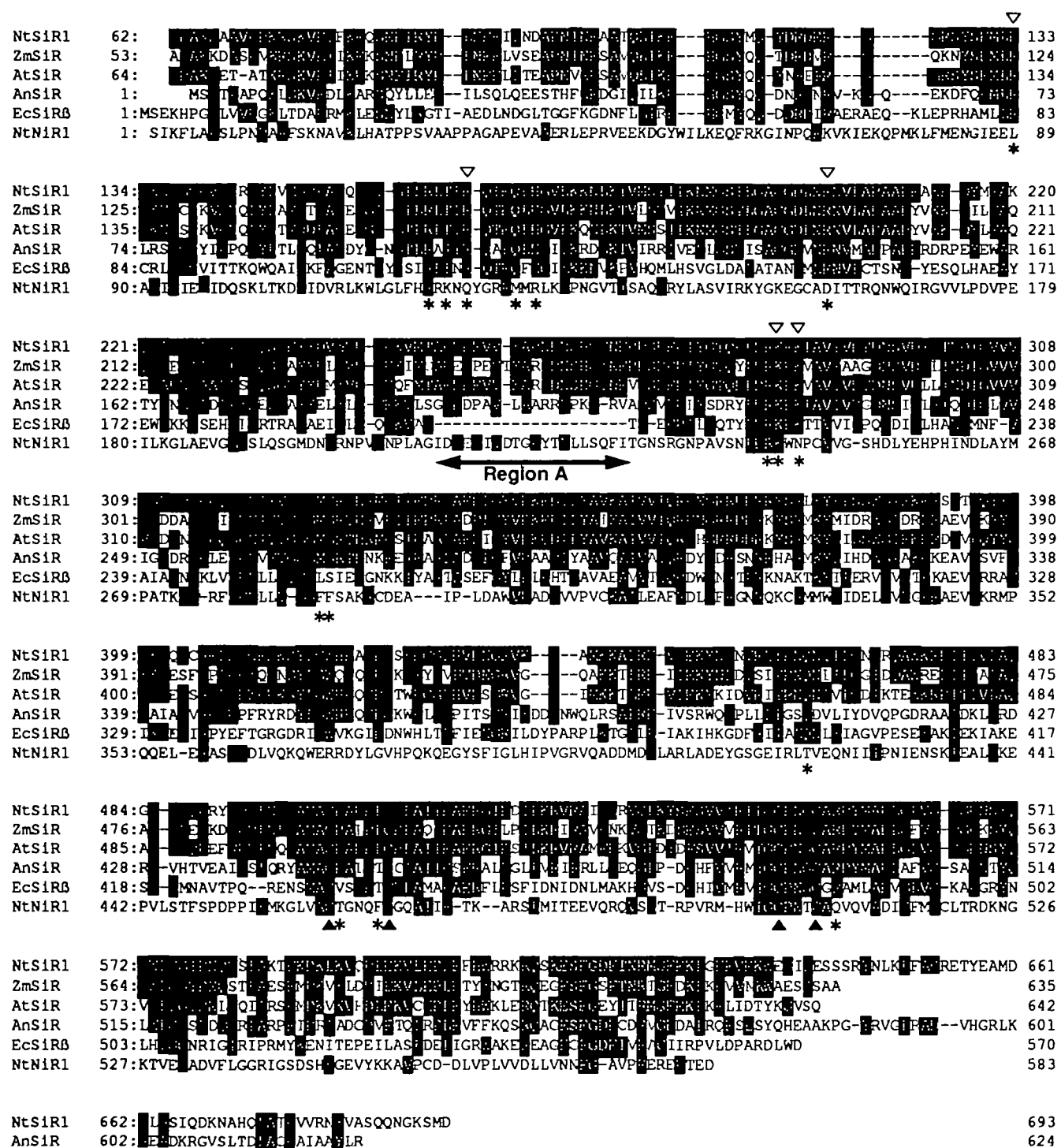


Fig. 4. Comparison of the amino acid sequences of SiRs. The mature region of *NtSiR1* is compared with those of *Z. mays* Fd-SiR (*ZmSiR*) (8), *A. thaliana* Fd-SiR (*AtSiR*) (9), *Synechococcus* PCC7942 Fd-SiR (*AnSiR*) (24), the hemoprotein of *E. coli* NADPH-SiR (*EcSiRβ*) (23), and *N. tabacum* Fd-NiR (*NtNiR1*) (25). Identical residues to *NtSiR1* are indicated by white letters on a black back-

ground. The amino acid residues that bind to the substrate (11) are shown by white triangles. The amino acid residues predicted to be ligated to the [4Fe-4S] cluster and the siroheme are shown by black triangles. Asterisks show the amino acid residues surrounding the siroheme. The region indicated by the arrow (region A) is common in Fd-linked enzymes (see the text).

found to be quite similar to *gNtSiR1* in the coding region. Some changes were found mainly in the intron region, suggesting that at least two copies of the SiR gene are present in the genome of tobacco.

To determine the organization and copy number of the tobacco Fd-SiR gene more directly, Southern blot analysis of the total genomic DNA was carried out using a DNA fragment located within the second exon of *gNtSiR1* as a

probe. The genomic DNA was cut with four different restriction enzymes, and all digests gave two or three hybridizable fragments, as shown in Fig. 7. The fragments with sizes of 2.8 and 3.4 kb detected in the digests of *EcoRI* and *HindIII*, respectively, were in good agreement with those postulated from the structure of *gNtSiR1*. This indicates that there are two copies of the SiR gene in the total genome of tobacco. The second gene detected on

Southern blotting seems to correspond to *gNtSiR2*. *N. tabacum* is an amphidiploid, and this implies a single gene for Fd-SiR per genome.

It has been proposed that the genes for Fd-SiR and Fd-NiR had a common ancestral gene. The two enzymes have the same prosthetic groups and comparable activities as to the reduction of sulfite and nitrite, although their substrate specificities are considerably different. The entire structures of the genes of *A. thaliana* and spinach are known, and they consist of 4 exons and 3 introns, although the Fd-NiR gene has not been cloned from tobacco. The sites of the intron insertions are quite different from those in Fd-SiR (data not shown), suggesting that the two genes diverged at an early stage of molecular evolution, probably before the appearance of eukaryotic photosynthetic organisms.

Potential *cis*-acting elements involved in the regulation of sulfur assimilatory genes were found in the sequence of 572 nucleotides upstream of the translation starting site of *gNtSiR1* on research with the PLACE database at the National Institute of Agrobiological Resources, Tsukuba. In yeast, most sulfur assimilatory genes require centromere binding factor 1 (Cbf1) for normal transcription. DNA binding sites for Cbf1 called the CDEI element (5'-TCACG-TGA-3') are found in the promoters of all sulfur assimilatory structural genes, *i.e.* MET4, MET5, MET10, MET14, MET16, and MET25. MET5 is the gene for the hemoprotein of yeast SiR. Cbf1 is not capable of gene activation by itself, but enhances the binding of MET4 to the promoters of the MET genes, and is a transcriptional activator for MET gene expression in response to sulfur availability (33). The MET4 binding sequence is known to be 5'-TGGC-

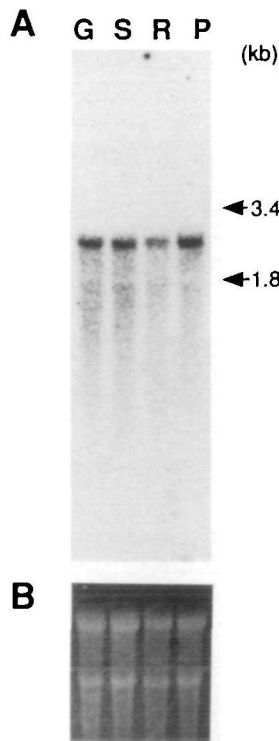


Fig. 5. (A) Northern blot analysis of the expression of *NtSiR1* in various organs of tobacco plants. Total RNAs (10  $\mu$ g), prepared from green leaves (G), stems (S), roots (R), and petals (P), were electrophoresed on a 1.2% agarose gel, transferred to a nylon membrane, and then hybridized with  $^{32}$ P-labeled cDNA for *NtSiR1*. (B) Ethidium bromide staining of total RNA on the agarose gel.

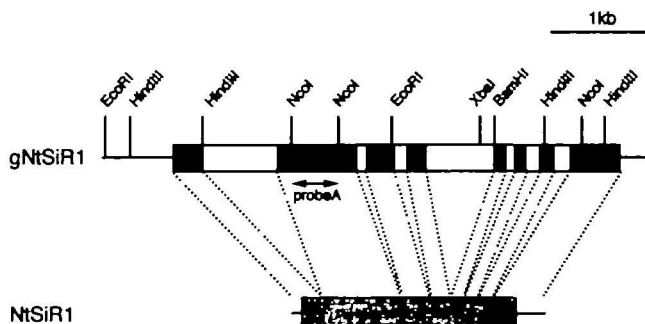


Fig. 6. Schematic representation of the genomic clone, *gNtSiR1*, and the cDNA clone, *NtSiR1*, coding for Fd-SiR. Exons and introns are indicated by black and white boxes, respectively. The gray and hatched boxes indicate the mature region and the transit peptide region of Fd-SiR, respectively. The position of the *NcoI* fragment, probe A, used for the Southern blot analysis in Fig. 7 is indicated by a double-headed arrow.

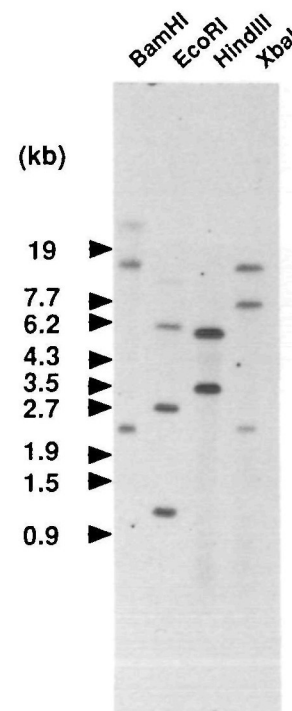


Fig. 7. Southern analysis of tobacco genomic DNA digested with the restriction enzymes, *BamHI*, *EcoRI*, *HindIII*, and *XbaI*, and hybridized with a  $^{32}$ P-labeled DNA probe prepared from *NtSiR1*.

AAATG-3'. These two *cis*-acting elements were found in the 5'-flanking region of *gNtSiR1*. The core motif of CDEI, CACGTG, and a sequence similar to the MET4 binding sequence, AGGCAAAAG, are found at position -183 and position -210, respectively. In the genomic sequences of the *A. thaliana* Fd-SiR gene submitted to the nucleotide sequence database (GenBank accession number Y10157) and the *A. thaliana* ATP sulfurylase gene (GenBank accession numbers U59738 and U59737), core motif and MET4 binding sequences also exist in the promoter regions. Characterization of the *cis*-elements in more detail and analysis of other sulfur assimilatory genes will be necessary to understand the mechanism of underlying regulation of sulfur metabolism in higher plants.

We wish to thank Dr. S.F. Chandler, Florigene, for critical reading of the manuscript, and Dr. H. Sakakibara, Nagoya University, for his support in preparing the manuscript.

## REFERENCES

- Smith, F.W., Ealing, P.M., Hawkesford, M.J., and Clarkson, D.T. (1995) Plant members of a family of sulfate transporters reveal functional subtypes. *Proc. Natl. Acad. Sci. USA* **92**, 9373-9377
- Takahashi, H., Sasakura, N., Noji, M., and Saito, K. (1996) Isolation and characterization of a cDNA encoding a sulfate transporter from *Arabidopsis thaliana*. *FEBS Lett.* **392**, 95-99
- Klonus, D., Höfgen, R., Willmitzer, L., and Riesmeier, J.W. (1994) Isolation and characterization of two cDNA clones encoding ATP-sulfurylases from potato by complementation of a yeast mutant. *Plant J.* **6**, 105-112
- Logan, H.M., Cathala, N., Grignon, C., and Davidian, J.-C. (1996) Cloning of a cDNA encoded by a member of the *Arabidopsis thaliana* ATP sulfurylase multigene family. *J. Biol. Chem.* **271**, 12227-12233
- Arz, H.E., Gisselmann, G., Schiffmann, S., and Schwenn, J.D. (1994) A cDNA for adenylyl sulphate (APS)-kinase from *Arabidopsis thaliana*. *Biochim. Biophys. Acta* **1218**, 447-452
- Gutierrez-Marcos, J.F., Roberts, M.A., Campbell, E.I., and Wray, J.L. (1996) Three members of a novel small gene-family from *Arabidopsis thaliana* able to complement functionally an *Escherichia coli* mutant defective in PAPS reductase activity encode proteins with a thioredoxin-like domain and "APS-reductase" activity. *Proc. Natl. Acad. Sci. USA* **93**, 13377-13382
- Setya, A., Murillo, M., and Leustek, T. (1996) Sulfate reduction in higher plants: molecular evidence for a novel 5'-adenylyl-sulfate reductase. *Proc. Natl. Acad. Sci. USA* **93**, 13383-13388
- Ideguchi, T., Akashi, T., Onda, Y., and Hase, T. (1995) cDNA cloning and functional expression of ferredoxin-dependent sulfite reductase from maize in *E. coli* cells in *Photosynthesis: from Light to Biosphere* (Mathis, P., ed.) Vol. II, pp. 713-716, Kluwer Academic Publishers, Dordrecht
- Brühl, A., Haverkamp, T., Gisselmann, G., and Schwenn, J.D. (1996) A cDNA clone from *Arabidopsis thaliana* encoding plastidic ferredoxin:sulfite reductase. *Biochim. Biophys. Acta* **1295**, 119-124
- Wu, J.Y., Siegel, L.M., and Kredich, N.M. (1991) High-level expression of *Escherichia coli* NADPH-sulfite reductase: requirement for a coned *cysG* plasmid to overcome limiting siroheme cofactor. *J. Bacteriol.* **173**, 325-333
- Crane, B.R., Siegel, L.M., and Getzoff, E.D. (1995) Sulfite reductase structure at 1.6 Å: evolution and catalysis for reduction of inorganic anions. *Science* **270**, 59-67
- Arb, C. and Brunold, C. (1983) Measurement of ferredoxin-dependent sulfite reductase activity in crude extracts from leaves using *O*-acetyl-L-serine sulphydrylase in a coupled assay system to measure the sulfide formed. *Anal. Biochem.* **131**, 198-204
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254
- Hirasawa, M. and Tamura, G. (1984) Flavin and iron-sulfur containing ferredoxin-linked glutamate synthase from spinach leaves. *J. Biochem.* **95**, 983-994
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
- Feinberg, A.P. and Vogelstein, B. (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **139**, 266-267
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual Ed. 2*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- Sanger, F., Nicklen, S., and Coulson, A. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
- McGookin, R. (1984) RNA extraction by the guanidine thiocyanate procedure in *Methods in Molecular Biology* (Walker, J.M., ed.) Vol. 2, pp. 113-116, Humana Press Inc., Clifton, New Jersey
- Aketagawa, J. and Tamura, G. (1980) Ferredoxin-sulfite reductase from spinach. *Agric. Biol. Chem.* **44**, 2371-2378
- Takahashi, S., Yoshida, Y., and Tamura, G. (1996) Purification and characterization of ferredoxin-sulfite reductases from leek (*Allium tuberosum*) leaves. *J. Plant Res.* **109**, 45-52
- Takahashi, S., Yip, W.C., and Tamura, G. (1997) Purification and characterization of ferredoxin-sulfite reductase from turnip (*Brassica rapa*) leaves and comparison of properties with ferredoxin-sulfite reductase from turnip roots. *Biosci. Biotech. Biochem.* **61**, 1486-1490
- Krueger, R.J. and Siegel, L.M. (1982) Spinach siroheme enzymes: isolation and characterization of ferredoxin-sulfite reductase and comparison of properties with ferredoxin-nitrite reductase. *Biochemistry* **21**, 2892-2904
- Asada, K., Tamura, G., and Bandurski, R.S. (1969) Methyl viologen-linked sulfite reductase from spinach leaves. *J. Biol. Chem.* **244**, 4904-4915
- Ostrowski, J., Wu, J.-Y., Rueger, D.C., Miller, B.E., Siegel, L.M., and Kredich, N.M. (1989) Characterization of the *cysJIIH* regions of *Salmonella typhimurium* and *Escherichia coli* B. *J. Biol. Chem.* **264**, 15726-15737
- Gisselmann, G., Klausmeier, P., and Schwenn, J.D. (1993) The ferredoxin:sulphite reductase gene from *Synechococcus* PCC-7942. *Biochim. Biophys. Acta* **1144**, 102-106
- Kronenberger, J., Lepingle, A., Caboche, M., and Vaucheret, H. (1993) Cloning and expression of distinct nitrite reductases in tobacco leaves and roots. *Mol. Gen. Genet.* **236**, 203-208
- Brunold, C. and Suter, M. (1989) Localization of enzymes of assimilatory sulfate reduction in pea roots. *Planta* **179**, 228-234
- Alonso, J.M., Chamorro, J., and Granell, A. (1995) A non-photosynthetic ferredoxin gene is induced by ethylene in *Citrus* organs. *Plant Mol. Biol.* **29**, 1211-1221
- Gotor, C., Cejudo, F.J., Barroso, C., and Vega, J. (1997) Tissue-specific expression of *ATCYS-3A*, a gene encoding the cytosolic isoform of *O*-acetylserine(thiol)ylase in *Arabidopsis*. *Plant J.* **11**, 347-352
- Gu, Q., Kawata, E.E., Morse, M.J., Wu, H.M., and Cheung, A.Y. (1992) A flower-specific cDNA encoding a novel thionin in tobacco. *Mol. Gen. Genet.* **234**, 89-96
- Brown, J.W. (1986) A catalogue of splice junction and putative branch point sequences from plant introns. *Nucleic Acids Res.* **14**, 9549-9559
- Kuras, L. and Thomas, D. (1995) Functional analysis of Met4, a yeast transcriptional activator responsive to S-adenosylmethionine. *Mol. Cell. Biol.* **15**, 208-216