

Identification and characterization of single-domain thiosulfate sulfurtransferases from *Arabidopsis thaliana*

Michael Bauer, Jutta Papenbrock*

Institute for Botany, University of Hannover, Herrenhäuserstr. 2, D-30419 Hannover, Germany

Received 10 October 2002; revised 8 November 2002; accepted 12 November 2002

First published online 22 November 2002

Edited by Marc Van Montague

Abstract Sulfurtransferases/rhodanases (ST) are a group of enzymes widely distributed in all three phyla that catalyze the transfer of sulfur from a donor to a thiophilic acceptor substrate. All ST contain distinct structural domains, and can exist as single-domain proteins, as tandemly repeated modules in which the C-terminal domain bears the active site, or as members of multi-domain proteins. We identified several ST in *Arabidopsis* resembling the C-terminus of the *Arabidopsis* two-domain ST1 and the single-domain GlpE protein from *Escherichia coli*. Two of them (accession numbers BAB10422 and BAB10409) were expressed in *E. coli* and purified. Both proteins showed thiosulfate-specific ST enzyme activity.

© 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: *Arabidopsis thaliana*; 3-Mercaptopyruvate; Protein domain; Rhodanese; Sulfurtransferase; Thiosulfate

1. Introduction

Sulfurtransferases/rhodanases (ST) catalyze the transfer of a sulfur atom from suitable sulfur donors to a nucleophilic acceptor. The most studied and best characterized ST is bovine rhodanese (thiosulfate:cyanide ST, EC 2.8.1.1) which catalyzes, in vitro, the transfer of a sulfane-sulfur atom from thiosulfate to cyanide, leading to the formation of sulfite and thiocyanate [1]. ST contain structural well-defined modules. They are found as tandem repeats, with the C-terminal domain hosting the active site cysteine residue, as single-domain proteins or as members of multi-domain proteins [2]. Although in vitro assays show ST activity associated with rhodanese or rhodanese-like domains, specific biological roles for most members of this homology superfamily have not been established [3]. Proposed roles include cyanide detoxification [4], sulfur metabolism [5], and mobilization of sulfur for iron-sulfur cluster biosynthesis or repair [6].

In the *Escherichia coli* genome eight open reading frames encoding proteins containing a rhodanese domain bearing the potentially catalytic cysteine have been identified. One of the genes encodes the 12 kDa protein GlpE; relative to the two-domain rhodanese enzymes of known three-dimensional struc-

ture, GlpE displays substantial shortening of loops connecting α -helices and β -sheets resulting in radical conformational changes surrounding the active site. Sequence searches through completed genomes indicate that GlpE can be considered to be the prototype structure for the ubiquitous single-domain rhodanese module [3].

In all eukaryotic and prokaryotic species investigated so far 3-mercaptopyruvate- and thiosulfate-specific ST were identified [1,7–9]. So far two *Arabidopsis* cDNA sequences encoding ST proteins were isolated and characterized [10–13]. Both protein sequences resemble the 3-mercaptopyruvate ST from mammals with respect to decisive amino acid residues. The search for thiosulfate ST in *Arabidopsis* showing homology to the mammalian thiosulfate ST by cDNA library screening had not been successful so far. In the meantime the complete sequence of the *Arabidopsis* genome has become available. A comprehensive database screening revealed the existence of at least 18 *Arabidopsis* proteins containing rhodanese domains. We were interested to identify thiosulfate ST among these proteins and to prove the occurrence of single-domain ST in *Arabidopsis*. We could demonstrate that at least two single-domain ST possess thiosulfate ST activity.

2. Materials and methods

2.1. DNA cloning techniques

The EST clones 149L5T7 and 211O17T7 were ordered from the *Arabidopsis* stock center at the Ohio State University and sequenced on both strands using the BIG-Dye Terminator with an ABI 310 (Applied Biosystems). The following primer pair was used to amplify a 411 bp coding sequence from EST 149L5T7: primer 104 (5'-CGG ATC CTC TCA ATC AAT CTC CTC C-3') extended by a *Bam*HI restriction site and primer 105 (5'-CAA GCT TAT TAG CAG ATG GCT CCT C-3') extended by a *Hind*III restriction site. For the amplification of the 423 bp coding sequence from EST 211O17T7 primer 108 (5'-GGA TCC GCG GAC GAG AGC AGA GT-3') extended by a *Bam*HI restriction site and primer 109 (5'-CCT GCA GTT GAA GAA GAA GGA GAC G-3') extended by a *Pst*I restriction site were used. A 549 bp cDNA for AAA80303 was amplified from an *Arabidopsis* cDNA library by primer 110 (5'-CGG ATC CGA AAC CAC TGC TTT TAA C-3') extended by a *Bam*HI restriction site and primer 111 (5'-CCT GCA GCT CTT CTA CCG GCA GCT-3') extended by a *Pst*I restriction site.

The PCR contained 0.2 mM dNTPs (Roth, Karlsruhe, Germany), 0.4 μ M of each primer (MWG, Ebersberg, Germany), 1 mM $MgCl_2$ (final concentration, respectively), 0.75 μ l RedTaq DNA-Polymerase (Sigma, Taufkirchen, Germany) and about 1 μ g template DNA in a final volume of 50 μ l. Before starting the first PCR cycle, the DNA was denatured for 180 s at 94°C followed by 28 PCR cycles of 45 s at 94°C, 45 s at 55°C, and 45 s at 72°C. The process was finished with an elongation phase of 420 s at 72°C. The amplified PCR fragments were ligated into the expression vector pQE-30 (Qiagen, Hilden, Germany) and introduced into *E. coli* strain XL1-blue.

*Corresponding author. Fax: (49)-511-762 3992.

E-mail address: jutta.papenbrock@botanik.uni-hannover.de (J. Papenbrock).

2.2. Expression and purification of the ST proteins in *E. coli*

The recombinant proteins were expressed according to the following protocol: after growth of the respective *E. coli* cultures at 37°C to an OD₆₀₀ of 0.6 in Luria–Bertani medium (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl) (Roth) containing ampicillin (100 µg/ml) (AppliChem, Darmstadt, Germany) induction was carried out for 3 h with 1 mM (final concentration) of isopropyl-β-D-galactoside (AppliChem). Cell lysis was obtained by adding lysozyme (final concentration 1 mg/ml) (Roth) and vigorous homogenizing using a glass homogenizer and a pestle. The recombinant proteins were purified under non-denaturing or denaturing conditions by affinity chromatography with the nickel affinity resin according to the manufacturer's instructions (Qiagen), dialyzed overnight and directly used for enzyme activity measurements. The purity of the protein preparations was checked by SDS-PAGE [14] and subsequent silver staining. The ST1 protein from *Arabidopsis* (accession number CAB64716) and the N-terminal domain of ST1 (ST1N-term) were expressed and purified as described [12,15].

2.3. Enzyme activity measurements

The enzyme assays with recombinant proteins were set up as described [12,15]. In general, 1 µg purified recombinant protein was used in each reaction. The determination of the K_m values was done in 1 ml assays under the following conditions: 0.1 M Tris–HCl pH 9.0, 60 mM KCN, and 5 mM 2-mercaptoethanol (Fluka, Taufkirchen, Germany) using 12 different substrate concentrations in the range of 50 µM–50 mM of 3-mercaptopyruvate (ICN, Eschwege, Germany) or thiosulfate (Fluka).

2.4. Miscellaneous

Protein contents were determined according to Bradford [16] using bovine serum albumin (Roth) as a protein standard. The DNA and amino acid sequence analyses and prediction of the molecular masses were performed with the programs MapDraw and Protean in DNASTAR (Madison, WI, USA). For the identification of protein domains several programs in <http://www.expasy.ch> were used. For the prediction of the protein localization different programs were applied (PSORT, Predator and TargetP, <http://www.expasy.ch/tools>). The multiple sequence alignment was done by ClustalW (<http://www.ebi.ac.uk/clustalw>). Statistical analysis was performed using the Student method (SigmaPlot for Windows version 7.0). The K_m values were

calculated from the non-linear Michaelis–Menten plot using an enzyme kinetic program (SigmaPlot 7.0).

3. Results

3.1. Identification of single-domain ST with a rhodanese signature in *Arabidopsis*

All members in the ST protein family in eubacteria, archaea, and eukaryotes are unified by characteristic well-defined sequence domains. In PROSITE two signatures were developed for this family: consensus pattern 1 [FY]-x(3)-H-[LIV]-P-G-A-x(2)-[LIVF] (PROSITE accession number PS00380) and consensus pattern 2 [AV]-x(2)-[FY]-[DEAP]-G-[GSA]-[WF]-x-E-[FYW] (Prosite access number PS00683). Pattern 1 is located in the N-terminal part of two-domain ST, pattern 2 at the C-terminal extremity of the enzyme. Only the C-terminal signature can be found in all ST proteins. We could identify 18 proteins in *Arabidopsis* containing a rhodanese signature. According to their sequence homology they can be classified into six groups (Table 1). In this work we analyzed only members of group VI, which comprises four closely related proteins and includes one less similar protein. The proteins in this group contain only the C-terminal pattern 2 and therefore possess similarity to the single-domain ST from bacteria (Fig. 1).

The four closely related proteins with the accession numbers AAA80303, BAB10409, AAD03573, and BAB10422 were characterized in more detail. The results of the global alignment using the Needleman–Wunsch algorithm underlines the high similarities (on average 50%) between the four proteins analyzed (Fig. 1). For comparison the C-terminal part of the ST1 (ST1C-term) protein without the linker sequence was included in the analysis; the average similarity to ST1C-term is about 30%. The active site cysteine, the rhodanese signature

Table 1
Summarizing table containing the members of the ST/rhodanese multi-protein family from *Arabidopsis*

	Gene identification	Description	No. ESTs	No. aa	Mol. mass (kDa)	Acc. no.
Group I						
1	At1g16460	thiosulfate sulfurtransferase (ST2)	8	318	34.7	CAB53639
2	At1g79230	thiosulfate sulfurtransferase (ST1)	14	379	41.9	CAB55306
Group II						
3	At5g23060	similar to unknown protein	24	387	41.3	BAB09823
4	At4g01050	hypothetical protein	20	457	49.9	CAB80914
Group III						
5	At5g03455	putative protein	–	132	14.8	CAB83305
6	At1g09280	At1g09280/T12M4_1	9	581	64.7	AAK82550
7	At2g40760	unknown protein	4	522	57.9	AAD32803
8	At1g17850	contains a rhodanese-like PF00581 domain	1	423	48.3	AAF97265
Group IV						
9	At2g42220	hypothetical protein	3	234	25.5	AAB88647
10	At3g08920	unknown protein	2	214	23.8	AAF07833
11	At4g24750	putative protein	4	260	28.7	CAA22988
Group V						
12	At5g19370	putative peptidyl-prolyl <i>cis</i> – <i>trans</i> isomerase	2	299	33.0	AAK43980
13	At5g55130	molybdopterin synthase sulfurlyase	5	464	50.6	AAD18050
Group VI						
14	At4g27700	hypothetical protein	2	237	26.3	CAB38282
15	At4g35770	senescence-associated protein	58	182	20.1	AAA80303
16	At5g66040	senescence-associated protein sen1-like protein; ketoconazole resistance protein-like	12	120	12.7	BAB10409
17	At2g17850	putative senescence-associated rhodanese protein	–	150	17.5	AAD03573
18	At5g66170	senescence-associated protein sen1-like protein	6	136	14.9	BAB10422

The proteins were grouped according to their amino acid sequence similarities using the ClustalW program. An update of the database mining was done on 8 March 2002. Proteins 1 and 2 in group I were analyzed previously [12,13,15]. Proteins 15–18 in group VI were analyzed in this study. No., number; aa, amino acids; Mol. mass, molecular mass; Acc. no., accession number in the EMBL database.

Table 2

Determination of the K_m values for the substrates 3-mercaptopyruvate and thiosulfate, the catalytic constants k_{cat} and the quotients k_{cat}/K_m using the recombinant ST proteins

Protein	Mol. mass (kDa)	3-Mercaptopyruvate			Thiosulfate		
		K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m
ST1	35.5	3.7	54.0	14.6	0.7	2.8	4.3
BAB10422	14.9	35.4	1.2	0.03	1.1	10.7	9.7
BAB10409	12.7	51.7	0.7	0.01	7.4	4.0	0.5

Enzyme assays were done as described in Section 2; each determination was done four times. The molecular masses (Mol. mass) were calculated for the recombinant proteins excluding the 6×His tag.

are true thiosulfate ST (Table 2). The K_m values of BAB10422 and BAB10409 for 3-mercaptopyruvate are unphysiologically high because the abundance of 3-mercaptopyruvate in the cell is very low, at least lower than the calculated K_m value as was shown and discussed by our group previously [13]. The velocity of product formation given by the catalytic constant k_{cat} differs for all three ST investigated and characterizes ST1 as a 3-mercaptopyruvate ST and BAB10422 as a thiosulfate ST. The catalytic efficiency k_{cat}/K_m includes both the enzyme activity and substrate affinity. The results indicate that although BAB10409 accepted almost exclusively thiosulfate as sulfur donor, the ST1 protein has the potential to function as a better thiosulfate ST than BAB10409. One has to search for a better substrate for BAB10409.

3.3. The influence of the ST1N-term on the activity of single-domain ST

As mentioned above both thiosulfate ST contain only the C-terminal rhodanese pattern 2; one could ask the question whether the addition of ST1N-term increases the activity of the single-domain ST. Increasing amounts of the purified ST1N-term [15] were added to the ST1C-term and the thiosulfate ST (1 µg ST1C-term or thiosulfate ST plus 1, 2, or 5 µg ST1N-term). As was shown previously the enzyme activity of the ST1C-term could be increased [15] whereas the activity of the single-domain thiosulfate ST remained unchanged (Fig. 3). Here the activity of the single-domain ST could not be increased by addition of a protein containing rhodanese pattern 1.

4. Discussion

The multi-protein family of ST/rhodanases in *Arabidopsis* comprises at least 18 members including two- and single-domain ST. In our previous experiments we characterized a couple of two-domain ST and demonstrated specificity for 3-mercaptopyruvate [12,13]. In this study the biochemical characterization of two single-domain ST from *Arabidopsis* allowed the identification of thiosulfate-specific ST. So far single-domain ST have been identified and characterized only in bacteria [3,9,17]. A prediction of substrate specificity based on the amino acid stretch around the active site cysteine (CRKGVTA for thiosulfate ST and CGS(T)GVTA for 3-mercaptopyruvate ST) as suggested by Colnaghi et al. [9] could not be applied for the proteins analyzed here. However, the four ST aligned are unified by another highly conserved stretch of amino acids around the active site cysteine residue: [LI]-[IL]-[LIV]-G-C-[KQE]-S-x-[RM]-S. All other *Arabidopsis* ST containing a rhodanese pattern differ in their amino acid sequence in this region (Papenbrock, unpublished). In contrast to the two-domain ST isolated from eukaryotes previ-

ously the thiosulfate-specific ST do not contain a second cysteine which was suggested to be involved in acceptor binding close to the active site cysteine [15]. Obviously the reaction mechanisms differ.

It was shown previously that the enzymatic activity of the *Arabidopsis* ST1 resides in the C-terminal domain but is boosted by the N-terminal domain and the linker peptide in the full-length enzyme [15]. Therefore it was tempting to analyze the effect of additional ST1N-term molecules on the enzyme activity of the single-domain proteins. So far the role of the N-terminal domain and the kind of interactions between both domains are not known; however, the activities of the single-domain proteins could not be enhanced. It was assumed that the two-domain ST evolved from a gene duplication event before the appearance of the *Arabidopsis* genus [10]. Obviously, both types could evolve independently in the *Arabidopsis* organism.

When comparing the kinetic parameters one has to keep in mind the putative different localizations of the proteins in the cell and different concentrations of the respective substrates. The ST1 protein was shown to be localized in mitochondria [12] whereas the predictions for the *Arabidopsis* single-domain ST are plastids, cytosol or nucleus. Currently, we are testing the in silico predictions in fusion experiments with the green fluorescent protein.

The more ST with different substrate specificities are identified in diverse organisms the stronger are the demands to clarify their function in the organism. The AAA80303 protein was identified previously as senescence-associated protein (Sen1) [18]. A homologous protein was isolated from radish (Din1) and was suggested to be the first gene for a chloroplast protein which is negatively controlled by light [19]. All members show high similarities to a variety of small stress-related proteins such as sulfide dehydrogenase from *Wolinella succinogenes* [20], the 11 kDa PspE protein encoded in the phage-shock protein operon [21], and the 12 kDa GlpE protein, a member of the *sn*-glycerol 3-phosphate regulon of *E. coli* [17]. Recently, for the latter proteins rhodanese activity was demonstrated.

It could be shown that the promoter of the gene *Sen1* encoding AAA80303 contains sequence motifs resembling conserved motifs among stress-inducible genes. Examination of the expression patterns of the *Sen1* gene under various senescing conditions revealed that *Sen1* and *ST1* gene expressions are associated with *Arabidopsis* leaf senescence [13,18] and *Sen1* is also strongly induced in leaves by abscisic acid or ethephon [18]. These lines of evidence have to be followed systematically using gene-specific probes.

Recent investigations have shown that the rhodanese signatures are structurally related to the catalytic subunit of Cdc25 phosphatase enzymes [22] and that the two enzyme families

are likely to share a common evolutionary origin. An increasing number of reports indicate that rhodanese modules are versatile sulfur carriers that have adapted their function to fulfill the need for reactive sulfane-sulfur in distinct metabolic and regulatory pathways [2]. The high number of expressed ST in *Arabidopsis* indicates their important role in the organism.

Acknowledgements: We would like to thank Pamela von Trzebiatowski and Julia Volker for their excellent technical assistance. We are grateful to Prof. A. Schmidt, University of Hannover, for helpful discussions. The work was financially supported by the Deutsche Forschungsgemeinschaft (PA 764/1-3, SCHM 307/15-1) and the Fonds der Chemischen Industrie.

References

- [1] Westley, J. (1973) *Adv. Enzymol.* 39, 327–368.
- [2] Bordo, D. and Bork, P. (2002) *EMBO Rep.* 3, 741–746.
- [3] Spallarossa, A., Donahue, J.L., Larson, T.J., Bolognesi, M. and Bordo, D. (2001) *Structure* 9, 1117–1125.
- [4] Vennesland, B., Castric, P.A., Conn, E.E., Solomonson, L.P., Volini, M. and Westley, J. (1982) *Fed. Proc.* 41, 2639–2648.
- [5] Donadio, S., Shafiee, A. and Hutchinson, R. (1990) *J. Bacteriol.* 172, 350–360.
- [6] Bonomi, F., Pagani, S., Cerletti, P. and Canella, C. (1977) *Eur. J. Biochem.* 72, 17–24.
- [7] Wood, J.L. and Fiedler, H. (1953) *J. Biol. Chem.* 205, 231–234.
- [8] Nagahara, N., Okazaki, T. and Nishino, T. (1995) *J. Biol. Chem.* 270, 16230–16235.
- [9] Colnaghi, R., Cassinelli, G., Drummond, M., Forlani, F. and Pagani, S. (2001) *FEBS Lett.* 500, 153–156.
- [10] Hatzfeld, Y. and Saito, K. (2000) *FEBS Lett.* 470, 147–150.
- [11] Nakamura, T., Yamaguchi, Y. and Sano, H. (2000) *Eur. J. Biochem.* 267, 5621–5630.
- [12] Papenbrock, J. and Schmidt, A. (2000) *Eur. J. Biochem.* 267, 145–154.
- [13] Papenbrock, J. and Schmidt, A. (2000) *Eur. J. Biochem.* 267, 5571–5579.
- [14] Laemmli, U.-K. (1970) *Nature* 227, 680–685.
- [15] Burow, M., Kessler, D. and Papenbrock, J. (2002) *Biol. Chem.* 383, 1363–1372.
- [16] Bradford, M.M. (1976) *Anal. Biochem.* 72, 248–254.
- [17] Ray, W.K., Zeng, G., Potters, M.B., Mansuri, A.M. and Larson, T.J. (2000) *J. Bacteriol.* 182, 2277–2284.
- [18] Oh, S.A., Lee, S.Y., Chung, I.K., Lee, C.H. and Nam, H.G. (1996) *Plant Mol. Biol.* 30, 739–754.
- [19] Shimada, Y., Wu, G.J. and Watanabe, A. (1998) *Plant Cell Physiol.* 39, 139–143.
- [20] Klimmek, O., Kreis, V., Klein, C., Simon, J., Wittershagen, A. and Kroger, A. (1998) *Eur. J. Biochem.* 253, 263–269.
- [21] Adams, H., Teertstra, W., Koster, M. and Tommassen, J. (2002) *FEBS Lett.* 518, 173–176.
- [22] Hofmann, K., Bucher, P. and Kajava, A.V. (1998) *J. Mol. Biol.* 282, 195–208.