Glutathione synthetase: similarities of the proteins from *Schizosaccharomyces pombe* and *Arabidopsis thaliana*

Chang-Lin WANG and David J. OLIVER

Department of Botany, Iowa State University, Ames, IA 50011, U.S.A.

Glutathione synthetase predicted from the reported gene sequence from *Schizosaccharomyces pombe* is substantially smaller than the equivalent protein predicted from the cDNAs sequenced from *Arabidopsis thaliana*, *Saccharomyces cerevisiae* and other eukaryotes. Sequence alignments of the proteins encoded by the cDNA clones for glutathione synthetase from *Arabidopsis* and *S. pombe* show that the *Arabidopsis* protein contains 200 extra amino acids at the N-terminus. In order to test if this sequence is essential in the function of the protein, the full-length *Arabidopsis* protein and as two N-terminal deletions ($\Delta 67-71$ and $\Delta 67-200$) were expressed in *S. pombe* mutant MN101, which lacks endogenous glutathione synthetase activity. Although the wild-type plant cDNA could complement the yeast mutation, neither deletion mutant was able to restore glutathione-de-

pendent cadmium resistance. When the three proteins were expressed as fusion proteins in *Escherichia coli*, they accumulated to the same level, but only the plasmid containing the full-length cDNA, pFLAG222, produced detectable enzyme activity *in vitro*. These results suggested that the N-terminus of the *Arabidopsis* glutathione synthetase is essential for its function and opened up the possibility that there was a sequencing error in the reported *S. pombe* sequence. Therefore the *gsh2* sequence from wild-type *S. pombe* and the mutant strain MN101 were determined. The wild-type *S. pombe* *gsh2* encodes a protein that is about the same length as that found in *Arabidopsis*, and the MN101 mutation involves a frameshift mutation early in the glutathione synthetase reading frame.

INTRODUCTION

Glutathione fills a number of roles in all organisms, with its major function directed towards protecting them from environmental insults [1,2]. GSH-dependent reactions are important in protecting plants from oxidative stress [3,4], from xenobiotic compounds [5], and from heavy metals [6]. In the latter case, GSH serves as a substrate for the synthesis of phytochelatin. Mutations that inhibit GSH synthesis, or its conversion into phytochelatin, result in increased Cd^{2+} sensitivity in plants [7,8] and in the fission yeast, *Schizosaccharomyces pombe* [9].

Glutathione is synthesized from common amino acids in a two-step reaction. γ -Glutamylcysteine synthetase catalyses the ATP-dependent ligation of glutamate and cysteine to form γ -glutamylcysteine. This dipeptide is then condensed with glycine in an ATP-dependent reaction catalysed by GSH synthetase to form GSH. cDNA and genomic clones for GSH synthetase (*gsh2*) have recently been isolated from *Arabidopsis* [10–12]. Translation of these clones predicts a protein of 478 amino acid residues. This sequence shares substantial similarity to the 474-residue proteins from rat [13], humans [14], and *Xenopus* [15]. The protein from *Saccharomyces cerevisiae* contains 491 residues. Surprisingly, the protein reported for *S. pombe* is substantially smaller [16]. The sequence predicted from the gene from fission yeast is only 285 amino acids with the N-terminal region found in the other eukaryotes completely missing.

The ability of the *Arabidopsis* cDNA for *gsh2* to restore GSH synthetase activity in an *S. pombe* mutant [12] suggested that these two proteins share a common function and probably a similar structure. This has prompted us to investigate whether the N-terminal region of the *Arabidopsis* protein is essential for enzyme activity when expressed in *S. pombe* and *Escherichia coli* or if it serves some function that is unique to higher eukaryotes. Once N-terminal deletions had demonstrated that this region was essential, we focused our attention on the sequence of the *S. pombe* gene. In this paper we show that the protein predicted by the *S. pombe* gene is very similar to that of other eukaryotes.

MATERIALS AND METHODS

The two deletion mutations, $\Delta 67-71$ (missing amino acids Ala⁶⁷ to Gln⁷¹) and $\Delta 67-200$ (where amino acids Ala⁶⁷ to Lys²⁰⁰ are deleted), were generated from the *Arabidopsis* cDNA clone in the *S. pombe* expression vector pREP3XN [12,17,18] by partial digestion with *Hind*III and ligation of the resulting plasmids. The two plasmids (named pGSH222 Δ^{67-71} and pGSH222 Δ^{67-200}) containing the correct deletion fragments were identified by restriction digests and confirmed by DNA sequencing. The two clones, along with the full-length cDNA, plasmid pGSH222 [18], were independently transformed into the GSH synthetase-deficient *S. pombe* mutant, MN101 [9]. The presence of the plasmids was confirmed by colony PCR and Southern-blot hybridization. The ability of these constructs to encode a functional GSH synthetase protein was tested by the growth of the resulting transformants in the presence of 0.5 mM Cd^{2+} . The wild-type yeast strain (ATCC 38399) grows normally under these conditions, the MN101 mutant completely fails to grow, and the MN101 mutant expressing the *Arabidopsis* *gsh2* cDNA grows like the wild-type [12].

The three *Arabidopsis* cDNA clones (full length, $\Delta 67-71$ and $\Delta 67-200$) were subcloned using PCR and appropriate primers (GCGAATTCATGGAATCACAGAAACC and GTTGTCG-ACTGATCAAGATGGTTGTGAA) into the *Eco*RI and *Sal*I sites of the *E. coli* expression vector pFLAG-1 (Science Imaging System, New Haven, CT, U.S.A.) to yield plasmids pFLAG222, pFLAG222 Δ^{67-71} and pFLAG222 Δ^{67-200} . In this vector the cDNAs were expressed as a translation fusion with an N-terminal *ompA* signal sequence and the FLAG-1 epitope. These constructs were transformed into *E. coli* DH5 α [19]. GSH synthetase expressed by the FLAG-1 vector was isolated from the periplasmic space after isopropyl β -D-thiogalactoside (0.5 mM) induction, and the crude periplasmic extract was assayed as ATP-dependent synthesis of [¹⁴C]GSH from γ -glutamylcysteine and [¹⁻¹⁴C]glycine as described previously [12,18].

The *gsh2* gene from the MN101 mutant as well as wild-type *S. pombe* were isolated by PCR amplification from *S. pombe* DNA and sequenced.

RESULTS

An alignment of the proteins predicted by the cDNA clones of *gsh2* from *Arabidopsis* and the fission yeast, *S. pombe*, suggests that the plant protein, and in fact the protein from all other eukaryotes [12], is about 200 amino acid residues longer than the protein predicted from the fission yeast gene (Figure 1). Most of this additional sequence is at the N-terminus. This raises questions about how essential this region of the *Arabidopsis* protein might be. We tested the importance of this domain by deleting portions of the N-terminus of the plant GSH synthetase and testing the activity of these deletion proteins.

The cDNA encoding the *Arabidopsis* GSH synthetase contains three *Hind*III sites that could be used to delete fragments of the N-terminal region of the protein without disrupting the resulting reading frame (Figure 1). Partial *Hind*III digestions resulted in two deleted cDNA clones. The smaller deletion, $\Delta 67-71$, encoded a protein with only five amino acids missing (Ala⁶⁷ to Gln⁷¹). The larger deletion $\Delta 67-200$, resulted in a protein missing 134 residues, Ala⁶⁷ to Lys²⁰⁰. This protein is missing most of the

Figure 1 Graphic representation of the cDNA clones for GSH synthetase from *Arabidopsis* and the original clone from *S. pombe*

The *Arabidopsis* cDNA is shown with the start and stop codon indicated and the *Hind*III sites used to generate the deletion mutants designated by the letter H. The two deletion mutants that resulted from *Hind*III digestion and ligation are indicated. The original *S. pombe* GSH synthetase sequence is also shown.

Table 1 Expression of *Arabidopsis* GSH synthetase cDNA and $\Delta 67-71$ and $\Delta 67-200$ deletion in *S. pombe*

The Wild-type yeast (ATCC 38399), the GSH synthetase-deficient mutant, MN101, transformed with the control plasmid (pREP3XN), and MN101 transformed with the wild-type cDNA and the two different deletion clones were tested for their ability to grow in the presence of increasing concentrations of Cd²⁺ (0–0.5 mM). The *in vitro* enzyme activity for these constructs was also determined.

Strain	Growth					GSH synthetase activity (nmol/h per mg of protein)
	[Cd ²⁺]	0	0.005	0.05	0.5	
Wild-type	++	++	++	++	++	4.8 ± 0.6
MN101/pREP3XN	++	—	—	—	—	0
MN101/intact cDNA	++	++	++	++	++	2.5 ± 0.2
MN101/ $\Delta 67-71$	++	—	—	—	—	0
MN101/ $\Delta 67-200$	++	—	—	—	—	0

Figure 2 Western blot of the recombinant GSH synthetase produced in *E. coli* using anti-FLAG antibody after separation by SDS/PAGE

Lane 1, $\Delta 67-71$; lane 2, $\Delta 67-200$; lane 3, control FLAG-1; lane 4, wild-type *Arabidopsis* GSH synthetase.

Table 2 Expression of *Arabidopsis* cDNA and $\Delta 67-71$ and $\Delta 67-200$ deletion clones in *E. coli*

Expression of the transgenes in the *E. coli* expression vector, pFLAG-1, was induced with 0.05 mM isopropyl β -D-thiogalactoside for 2 h at 37 °C. The enzyme was assayed from a periplasmic space fraction and the amount of enzyme was verified by Western blotting using the anti-FLAG-1 antibody.

Strain	Glutathione synthetase activity (nmol/h per mg of protein)
DH5 α /pFLAG-1	0
DH5 α /intact cDNA	85.7 ± 6.5
DH5 α / $\Delta 67-71$	0
DH5 α / $\Delta 67-200$	0

	1	136	326	878
<i>gsh2</i> (new)	ATG GAA ATT	CCT -ATT	GAA GCT	GGT
<i>gsh2</i> (old)	ATG GAA ATT	CCT TAT T	GAA G-T	GGT
MN101	ATG GAA -TTG	CTT -ATT	GAA GCT	GAT

Figure 3 Sequence changes in the new *gsh2* sequence (accession no. Z69369) as compared with the original reported sequence (accession no. M85179/S67467) and the mutations found in MN101

additional N-terminal sequence not predicted in the *S. pombe* sequence (Figure 1).

These two deletion constructs, as well as the full-length *gsh2* gene, were placed in the *S. pombe* expression vector pREP3XN under control of the *nmt1* promoter [17]. These plasmids were then transformed and their proteins expressed in the *S. pombe* mutant, MN101, which lacks endogenous GSH synthetase activity. The plasmid pGSH222 containing the full-length *Arabidopsis* cDNA clone was able to complement the missing GSH synthetase gene. This allowed the yeast to produce phytochelatin and thereby restore its ability to grow in the presence of 0.5 mM CdCl₂ like the wild-type yeast [12]. Neither of the two deletion clones, however, were able to complement the GSH synthetase mutation and, like the MN101 strain, neither transformant was able to grow at 0.005 mM CdCl₂ (Table 1). When GSH synthetase activities were measured in extracts from these yeast strains, pGSH222 returned the enzyme level in MN101 to half of the wild-type levels. Neither of the deletions appeared to restore any enzyme activity to this mutant (Table 1).

It has not been possible to detect the transgenic plant protein in *S. pombe* by non-enzymic means. In order to be sure that the deletion proteins are stable, we expressed the full-length and the two deleted proteins with a FLAG-1 epitope tag in *E. coli*, thus

<i>Arabidopsis</i>	MESQKPIFDLEKLDDEFVQKLVYDALVWSSSLHGLVVGDKSYQKSGNVPGVGLMHAPIALL	60
<i>S. pombe</i>	MEIEKYTP EQIEELGKG ARDFAFAHG VVTELSVSK EGRN IA TQ IPIT LF	50
<i>S. cerevisiae</i>	MAHYPPSKD QLNELIQ EVNQ W AIT-NGLSMYP PKFEENPSNASVSP VTIY	49
<i>Arabidopsis</i>	PTAFPEAYWKQACNV TPL FNELID RVSLDGK FLQDSLS RTK KVDV FTSRLLDIH SKMLER	120
<i>S. pombe</i>	PSVIPHGAF VEAVSVQ KAYNKLYAKIANDY EFLRLHLQ SI TKYDEF MNKLWNLYQKHREA	110
<i>S. cerevisiae</i>	PTPIPRKCFDEAVQIQPV FNELYARITQDMAQ PDSYLHKTTE ALALSDSEFTGKLWSLYL	109
<i>Arabidopsis</i>	NKKEDIRLGLHR FDYMLDEETNSLLQ IEMN TISCSF PGLSRL VSQ LHQ SLLRSY GDQIGI	180
<i>S. pombe</i>	VAHLKENQF- QPLSLGVFRSDY MVHQDD SF IGCK QVEFNTISV SF GGVSKAVSNL HAYCS	169
<i>S. cerevisiae</i>	ATLKSAQYKK QNFRLGIFRSDY LIDKKKGTE QIKQVEFNTVS VS FAGLSEK VDR LH SYLN	169
<i>Arabidopsis</i>	DSERPINT STIQFADALAKAW LEYSNPRA VVMVIVQ PEERNMYD QHLLSSI LREKHN IV	240
<i>S. pombe</i>	QSGLYRKPLTTNYLT VNTSVSGICTG ISNAVDAYRDY VKNIT SKMNIASD NTK PIVLFV	229
<i>S. cerevisiae</i>	RANKYDPKGPIYNDQNMVISDSGYLLSKALAK AVESYKSQ ----- QSSSTSDPIVAFIV	224
<i>Arabidopsis</i>	VIRKTLAEVEKEG SVQ ED ETLIVGG ----- QAVAVVY	274
<i>S. pombe</i>	KGGERNIT DQRTLEYEL LNRFH VISKR IDIAELNSLIHDKSSNKLYMKTS FTTYE AVAVVY	289
<i>S. cerevisiae</i>	QRNERNV FDQKVLE LN LLEK FGTKSVRLTFDDVNDKLFIDDKTGKLFIRD- TEQEIAVVY	283
<i>Arabidopsis</i>	FRSGYTPNDHPSESE WNARLLIESSAVK CP SIAYHLTG SKKIQQELAK PGVLER F LDNK	332
<i>S. pombe</i>	YRVGYALDD-PSQEA WDMRLTIENT LAIKCPSISTH LAG SKKIQQVLA ESNALER F LEGD	349
<i>S. cerevisiae</i>	YRTGYTTDYTSEK DWEAR LFL EKS FAIKAPDLLTQLSGSKKIQQL LTDE GV LKGK YISDA	343
<i>Arabidopsis</i>	EDIAKL RKCFAGLWS L DDS---- EIVQ AIEK PGLFVMK PQREGGGNN Y GDDVREN LLR	388
<i>S. pombe</i>	ELQAVRS-TFAD MYPLDDTPRG KEG IK LAF EK PEDFVLK PQREGGGNN TYK DIPGLLSK	408
<i>S. cerevisiae</i>	EKKSSLLKTFV KIYPLDDTK L REGKRL AL SEPSKY VLK PQREGGGNN VYK ENIPN FL KG	403
<i>Arabidopsis</i>	LQKEGEEGNA AYIL MQRIFPKVSNMFLVREGVY HKHQA --- I SELGVY GAYL RSK DEVIV	448
<i>S. pombe</i>	MPQ EEWD --- SYIL MRYINAVPSQ NYIL KGERPEKFDV VEIGILGTIV WNIKT- DEVVQ	467
<i>S. cerevisiae</i>	IEER HWD --- AYIL MELIEPELNENN II LRDNKSYNEPI-- I SELGIY GC VLFNDEQ VLS	461
<i>Arabidopsis</i>	NEQSGYLMRTK IAS SDEGGVAAGFGVLD SIYLI-	478
<i>S. pombe</i>	NGQSGFICRTKPKKT NEGGVATGYAS LS SIELSE	498
<i>S. cerevisiae</i>	NEFSGSLLRSK FNT SNEGGVAAGFGCLDS IILY-	491

Figure 4. Sequence alignment for the GSH synthetase proteins from *Arabidopsis thaliana*, *S. pombe* and *Sac. cerevisiae*

Amino acid residues that are shared by two or three sequences are shown in bold. The methionine originally assumed to begin the *S. pombe* protein is underlined.

allowing detection with the anti-FLAG-1 monoclonal antibody [18]. The proteins proved very unstable, but we were able to detect full-length protein in the periplasmic space. This protein was isolated and assayed as a crude periplasmic extract. Western blotting showed that pFLAG222 produced a protein with an apparent mass of 54 kDa. This agrees with the size noted earlier for the protein analysed *in vitro* [18]. The $\Delta 67-71$ clone produced a protein of the same apparent size, and the $\Delta 67-200$ construct resulted in a protein with an apparent size of 40 kDa (Figure 2). *E. coli* expressing the plasmid pFLAG222 encoding the full-length cDNA produced an active GSH synthetase (85.7 nmol/h per mg of protein) when measured *in vitro*. The N-terminal-deleted proteins produced by pFLAG222 $\Delta 67-71$ and pFLAG222 $\Delta 67-200$ both lacked detectable GSH synthetase activity even when the same amount of immunodetectable protein was present (Table 2).

These results suggested that the N-terminus of the *Arabidopsis* GSH synthetase is essential and opened up the possibility that there was a sequencing error in the reported *S. pombe* gene. We therefore amplified and sequenced a section of the gene from wild-type *S. pombe* (ATCC 38399). The sequence for this gene was simultaneously determined by another group (accession no. Z69369). This sequence shows that there were two frameshift errors in the original (accession no. M85179/S67467) sequence (Figure 3) for the gene, which resulted in a shortened reading frame beginning at an ATG that was located 213 amino acids (639 bp) downstream from the correct start sequence (Figure 1). The protein encoded by the full-length gene contains 498 amino acids and has an estimated molecular mass of 56.2 kDa. The *Arabidopsis* protein is 53.9 kDa. A revised amino acid alignment for the proteins predicted by genes from *S. pombe*, *Sac. cerevisiae* and *Arabidopsis* is shown in Figure 4. This alignment shows that

the regions of sequence identity are shared throughout the length of the proteins. One highly conserved sequence (amino acids 364–376 in the *Arabidopsis* sequence) has been implicated in forming a flexible loop that covers the active site of the enzyme [18].

We also isolated the *gsh2* gene from the *S. pombe* mutant MN101 and sequenced it. Surprisingly, the gene contained two mutations, a deletion in the third codon of the reading frame (ATG GAA-TT instead of ATG GAA ATT) and a G → A mutation at base 879 in the reading frame. The frameshift results in the protein terminating at a TAG at 79 bp. The glycine to aspartate mutation at 879 bp is not expressed because of the premature termination of the protein.

DISCUSSION

The N-terminal region of the *Arabidopsis* GSH synthetase is clearly necessary for enzymic activity. Deletion of even a small portion of the protein resulted in the complete loss of activity. This loss was documented *in vivo* because the deleted enzymes were unable to restore any Cd²⁺ resistance when expressed in the MN101 mutant of *S. pombe*. We had earlier demonstrated that Cd²⁺ sensitivity in this organism is based on the lack of GSH synthetase activity and that expression of the *Arabidopsis* cDNA complemented the missing yeast gene [12,18]. Since the protein produced by neither pGSH222Δ^{67–71} nor pGSH222Δ^{67–200} was capable of restoring any level of Cd²⁺ resistance, it seems clear that these enzymes were not functioning *in vivo*. The mutant enzymes expressed in *S. pombe* were not detectable by *in vitro* measures of enzyme activity. The sensitivity of this assay is about 5% of the activity of the unmodified *Arabidopsis* protein. It was possible that these modified proteins were extremely unstable in *S. pombe* and that accounted for the lack of enzyme activity. We therefore expressed them in our *E. coli* system where we could estimate the amount of protein present immunologically. Although the deleted proteins were isolated at the same levels as the wild-type protein, they had no detectable (less than 3% of the wild-type) enzyme activity.

This raised an interesting question. If both the plant and the yeast enzymes have the same catalytic function and share so much sequence similarity throughout the C-terminal two-thirds of the protein, what is the function of the N-terminal third of the plant proteins and why is it needed for the enzyme to function in yeast if the endogenous yeast enzyme lacks this portion of the protein? We addressed this question by cloning and sequencing the *gsh2* from *S. pombe*. A re-evaluation of the *S. pombe* gene revealed that the protein predicted from this sequence was actually slightly larger than the proteins found in the higher eukaryotes and that there is substantial sequence conservation throughout the protein (Figure 4). This suggests that the *S. pombe* protein is actually quite similar to that found in higher eukaryotes and that the N-terminal region of the GSH synthetase

has an essential function in determining the activity of the protein.

There are still some difficult questions about the *S. pombe* GSH synthetase enzyme that need to be resolved. The isolated protein has been reported to be an α₂β₂ heterotetramer composed of two 33 kDa and two 26 kDa proteins as estimated by SDS/PAGE [16]. The other eukaryotic proteins are homodimers of the 52–56 kDa subunit. In both cases the intact protein would have the measured molecular mass of 110–120 kDa. The *S. pombe* protein presented in Figure 4 has a molecular mass of 56.2 kDa. This is sufficient mass that the 33 kDa and the 26 kDa proteins reported could be derived from it by proteolysis. Partial peptide sequence of the larger subunit shows regions predicted by the C-terminal half of the *S. pombe* *gsh2* gene. Some indirect support for the proteolytic source of the two peptides observed in the *S. pombe* protein is noted in a more recent report by Nakagawa et al. [20] showing that the purified enzyme contains a significant amount of a 56 kDa protein that might represent the primary translation product. Further experimentation will be necessary to resolve this issue.

REFERENCES

- Bergmann, L. and Rennenberg, H. (1993) in Sulfur Nutrition and Assimilation in Higher Plants, (De Kok, L. J., Stolen, I., Rennenberg, H., Brunold, C. and Rauser, W. E., eds.), pp. 109–123, SPB Academic Publishing, The Hague
- Alscher, R. G. (1989) *Physiol. Plant.* **77**, 457–464
- Foyer, C. H., Descourvieres, P. and Kunert, K. J. (1994) *Plant Cell Environ.* **17**, 507–523
- Foyer, C. H., Lelandais, M. and Kunert, K. J. (1994) *Physiol. Plant.* **92**, 696–717
- Lamoureux, G. L., Shimabukuro, R. H. and Frear, D. S. (1994) in Herbicide Resistance in Weeds and Crops, (Caseley, J. C., Cussans, G. W. and Atkin, R. K., eds.), pp. 227–261, Butterworth-Heinemann, Oxford
- Rauser, W. E. (1995) *Plant Physiol.* **109**, 1141–1149
- Howden, R., Andersen, C. R., Goldsbrough, P. B. and Cobbett, C. S. (1995) *Plant Physiol.* **107**, 1067–1073
- Howden, R., Goldsbrough, P. B., Andersen, C. R. and Cobbett, C. S. (1995) *Plant Physiol.* **107**, 1059–1066
- Mutoh, M. and Hayashi, Y. (1988) *Biochem. Biophys. Res. Commun.* **151**, 32–39
- Rawlins, M. R., Leaver, C. J. and May, M. J. (1995) *FEBS Lett.* **376**, 81–86
- Ullmann, P., Gondet, L., Potier, S. and Bach, T. J. (1996) *Eur. J. Biochem.* **236**, 662–669
- Wang, C.-L. and Oliver, D. J. (1996) *Plant Mol. Biol.* **31**, 1093–1104
- Huang, C.-S., He, W., Meister, A. and Anderson, M. E. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1232–1236
- Gali, R. R. and Board, P. G. (1995) *Biochem. J.* **310**, 353–358
- Habenicht, A., Hille, S. and Knochel, W. (1993) *Biochim. Biophys. Acta* **1174**, 295–298
- Mutoh, N., Nakagawa, C. W., Ando, S., Tanabe, K. and Hayashi, Y. (1991) *Biochem. Biophys. Res. Commun.* **181**, 430–436
- Forsburg, S. L. (1993) *Nucleic Acids Res.* **21**, 2955–2956
- Wang, C.-L. and Oliver, D. J. (1997) *Biochem. J.* **322**, 241–244
- Sambrook, J., Fritsch, T. E. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Nakagawa, C. W., Mutoh, N. and Hayashi, Y. (1993) *Biochem. Cell Biol.* **71**, 447–453