TÜBİTAK

Turk J Biol
36 (2012) 513-526

doi:10.3906/biy-1110-2

Characterization of affinity tag features of recombinant *Tetrahymena thermophila* glutathione-S-transferase zeta for *Tetrahymena* protein expression vectors

Cem ÖZİÇ[1], Muhittin ARSLANYOLU[2]

[1]Department of Biology, Faculty of Sciences, Kafkas University, 36100 Kars - TURKEY

[2]Department of Biology, Faculty of Sciences, Anadolu University, Yunusemre Campus, 26470 Eskisehir - TURKEY

Received: 02.10.2011 ● Accepted: 03.04.2012

Abstract: Glutathione S-transferase (GST) is the one of most widely used affinity tags in biotechnology applications. The present study named a GST gene as the TtGSTz1 gene, characterized with the conserved glutathione binding motif (SSTSWRVRIAL) under a GST zeta subfamily from *Tetrahymena thermophila*. Phylogenetic analysis of the TtGstz1p protein sequence, with its orthologs from different GST classes, showed that it is a member of the unicellular GSTz monophyletic clade, which is positioned close to the bacterial GSTz clade. Seven codons of the TtGSTz1 gene were first engineered by introducing silent mutations (TAA > CAA or TAG > CAG) to the code for glutamine instead of stop signals in *E. coli*. The recombinant TtGSTz1 gene, with its 6XHis tag, was expressed using the pET16b expression plasmid and *E. coli* BL21(DE3). SDS-PAGE analysis of the 6XHis-TtGstz1p protein confirmed its expression and purification by nickel agarose and glutathione-sepharose 4B. Additionally, both anti-His and anti-GST antibodies recognized the purified 6XHis-TtGSTz in the western blot analysis. In conclusion, the results presented here suggest that the 6XHis-TtGstz1p fusion protein could be used as a dual tag in *Tetrahymena* expression vectors with a sequential 2-step protein purification procedure.

Key words: GSTzeta, *Tetrahymena thermophila*, recombinant, *E. coli*, protein expression, tag

Introduction

The characterization of gene function at the protein level requires the production and purification of large amounts of proteins. However, not all proteins of eukaryotic origin can be expressed in a fully functional form in *E. coli*. Therefore, investigations to find a more suitable eukaryotic host with protein expression constructs have persisted for decades, including the recent recombinant DNA studies in *Tetrahymena thermophila*. The construction of protein expression vectors based on rDNA origins (1) and their transformation with either electroporation (2) or microparticle bombardment in *T. thermophila* were achieved (3). Recently, a number of heterologous recombinant protein expression products were reported in *T. thermophila*, including the i-surface glycoprotein of *Ichthyophthirius multifiliis* (4), a circumsporozoite protein of *Plasmodium falciparum* (5), and a DNAse-I glycoprotein from *H. sapiens* (6). *T. thermophila* is a unicellular organism with biological features suitable for biotechnological use; for example, it has a large cell size (approximately 50 μm), a fast growth rate (a doubling time of approximately 2 h), and an ability to reach high cell densities (a few million cells/mL) under simple and inexpensive culture conditions.

Furthermore, *T. thermophila* offers simple long-term strain storage, ease of genetic manipulations and gene targeting due to its nuclear dimorphism, existence of a transcriptionally silent diploid (2n = 10) micronucleus, and a haploid somatic macronucleus (7). All of these studies suggest an opportunity for the use of *T. thermophila* as an alternative protein production host.

Target proteins are fused to affinity tags to facilitate stability, solubility, easy detection, and purification in different expression hosts (8). Two commonly used affinity tags for gene expression are glutathione *S*-transferase (GST) purified with glutathione affinity chromatography (9) and a multiple-histidine tag purified with immobilized nickel or cobalt ions (10). A dual tagged protein could be generated as a tripartite fusion protein by the addition of a GST or polyhistidine (6XHis) tag in either terminus for sequential 2-step purification (11,12). Alternatively, a single-piece dual tag could be placed at the C-terminus or N-terminus of the target protein to create a tripartite fusion protein (as proposed in this study). The recently sequenced macronuclear genome of *T. thermophila* revealed a number of GST genes that could be used as affinity tags for heterologous expression in *E. coli* or homologous protein expression in *T. thermophila* (13).

GSTs (EC 2.5.1.18) detoxify endobiotic and xenobiotic compounds by linking glutathione (GSH) covalently to a hydrophobic substrate, forming a less reactive and more polar glutathione *S*-conjugate (14,15). GST family proteins are characterized by similar tertiary structures and active site topology (16). Cytosolic GSTs are categorized into the alpha, beta, mu, pi, sigma, zeta, tau, delta, phi, omega, and theta classes. The zeta-class GSTs have a conserved N-terminal domain containing the GSH binding motif (SSCX(W/H)RVIAL), in which the first serine (Ser) helps the proper binding of GSH, and a less conserved C-terminal domain with a binding region for hydrophobic substrates (17).

In this study, a GSTz gene from *T. thermophila* (Tt*GSTz1*) was expressed in *E. coli* with a 6XHis tag after the introduction of 7 point mutations. Recombinant 6XHis-TtGstz1p was successfully affinity purified, not only with glutathione-sepharose beads, but also with Ni beads. The 6XHis-Gstz1p was recognizable by commercial anti-His and anti-GST antibodies. The characterized affinity features of the 6XHis-Gstz1p suggest that it could be used as a GST or dual tag in future *T. thermophila* protein expression vectors.

Materials and methods

Strains and materials used

The inbred strain *T. thermophila* SB210 was used and maintained in proteose peptone yeast extract medium (1% [w/v] Difco proteose peptone, 0.15% [w/v] yeast extract, 0.5 mM $FeCl_3$). All of the primers used in this study (Table) were from Bio Basic (Toronto, Canada).

Cloning of TtGSTz1 cDNA and its 3'UTR

Total RNA was isolated from exponentially growing cells (1 mL; approximately 1.5×10^5 cells/mL, overnight at 30 °C) using TRIzol Reagent (Sigma). Total RNA was treated with RQ1 DNAse I (Promega). Reverse transcription (RT) was performed according to the manufacturer's directions (Fermentas) using 1 unit of MMLV reverse transcriptase with 5 µg of total RNA and oligo dT_{22} primer. The TtGSTz1 protein-coding cDNA region was amplified with the F-GSTz and R-GSTz primers. The 3'UTR sequence of the TtGSTz cDNA was recovered with primers eF-Z6 and R-oligo dT_{22}. Amplification of these fragments was performed using Prime-Star DNA polymerase (Takara). Amplified cDNA fragments were cloned into pGEM-T Easy (Promega) and its sequence was confirmed with a CEQ 8000 Beckman-Coulter automated sequencer.

Construction of recombinant Tt*GSTz1* gene by site-directed PCR-mediated mutagenesis

Seven codons of the Tt*GSTz1* gene were first engineered by introducing silent mutations (TAA > CAA or TAG > CAG) to the code for glutamine instead of stop signals in *E. coli* (18). Briefly, primers containing point mutations (Table) were used to generate overlapping mutated DNA fragments of 10-18 bp, which were gel-purified and used as overlapping mega primers, joined in a second polymerase chain reaction (PCR) run with most 5' and 3' primers of 2 fragments. These steps were repeated until a full-length recombinant GSTz was formed. The introduced point mutations were confirmed using a CEQ8000

Table. All of the primer sequences used in this study. Complementary pairs of multiple-point-mutation primers for each position used in the first and second PCR run of site-directed mutagenesis of TtGSTz. Point mutations are capitalized. Bases with gray shadowing are the START and STOP codons. The underlined bases show the restriction enzyme position in the primers.

Primer name	Multiple-point-mutation primer set
2F	5′ cgctcgag**atg**gctggaaagctcaaagaaaattact 3′
eF-Z1	5′ agtgaa**C**aaacttctgaagaa 3′
eR-Z1	3′ gaaaacttctca**ctt**G**G**tttgaaga 5′
eF-Z2	5′ agtgaa**C**aaacttctgaagaa 3′
eR-Z2	3′ gaaaacttctca**ctt**G**G**tttgaaga 5′
eF-Z3	5′ ctttctttcct**C**aagatgcagttaaa 3′
eR-Z3	3′ ctcgtgggtgaagaagga**G**ttctacgt 5′
eF-Z4	5′ agagct**C**aaattagaggcttttgc 3′
eR-Z4	3′ cgtcaattttctcga**G**tttaatct 5′
eF-Z5	5′ caccctctc**C**aaatttgagggtgctt 3′
eR-Z5	3′ tgacggtaggtgggagag**G**tttaaac 5′
eF-Z6	5′ gaatatagc**C**aagacaagattcaatgg 3′
eR-Z6	3′ taacttttccttatatcg**G**ttctgttctaa 5′
2R	3′ gggtcgttc**G**ttcgactgttaggactattagttttctag **act**cacctaagc 5′
Protein expression primers	
FexpGSTzNdeI	5′ gcg**catatg**gctgaaagctcaaagaaaattact 3′
RexpGSTzBamHI	3′ ttaggactattagttttctagact**cctagg**cgc 5′
Cloning primers	
F-GSTz	5′ atggctgaaagctcaaagaaaattact 3′
R-GSTz	3′ ggactattagtttctagactgtcg 5′
R-oligo dT_{22}	3′ ttttttttttttttttttggatcctgcag 5′
mRNA expression control primers	
FTt17S	5′ cccatgcatgtgccagttcag 3′
RTt17S	3′ cagcattgttccatagacatccac 5′

Beckman-Coulter automated DNA sequencer. This clone was ligated into the NdeI and BamHI sites of pET-16b (+) (Novagene) and was transformed into *E. coli* strain BL21 (DE3) (Sigma).

Analysis of TtGSTz1 mRNA distribution by RT-PCR method

Expression levels of TtGSz1 mRNA were studied in *T. thermophila* after cold shock (30 °C to 4 °C) and sublethal oxidative stress (0.02% or 200 μM H_2O_2) treatment (19). Samples were taken at 15-min intervals for a duration of 60 min and analyzed using the reverse transcription (RT)-PCR method. Reverse transcription was executed with 5 μg of total RNA. PCR was performed with 1 μL of diluted cDNA (1:10) in a total volume of 25 μL using Speed STAR HS DNA Polymerase (Takara) enzyme for 27 cycles in the exponential range with FGSTz and RGSTz primers (Table). 17S rRNA primers (FTt17S

and RTt17S) (Table) were used to confirm that equal amounts of total RNA were used for each time point in the reverse transcription. The RT-PCR bands were gel-purified and confirmed with the CEQ 8000 DNA sequencer.

Western blot analysis of expressed and purified recombinant 6XHis-TtGstz1p

Expression of the recombinant protein was induced by the addition of 100 μM isopropyl β-D-thiogalactopyranoside (IPTG) in *E. coli* BL21 (DE3) carrying pET16b (+) 6XHis-TtGstz1p, and incubation was continued for a further 3 h at 37 °C. Purification of soluble 6XHis-TtGstz1p through Ni-NTA resin (QIAGEN) or glutathione affinity (Amersham Biosciences) column chromatography was performed according to the manufacturer's instructions. Analysis of the purified proteins on 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out with Coomassie brilliant blue staining. Proteins were transferred from the gel to a polyvinylidene fluoride filter (PVDF; Immobilon-P, Millipore). For western blotting, the ProteoQwest colorimetric kit with TMB substrate (Sigma) was used with mouse monoclonal anti-His_6 antibody (1:1000; Roche 135508) or mouse monoclonal anti-GST antibody (1:1000; Sigma G1160) and HRP-conjugated secondary antimouse antibody (1:3000; Sigma, A5225). Kaleidoscope prestained standards were used as molecular-mass markers (Bio-Rad).

Sequence analysis and homology search

The BLAST program was used for initial sequence retrieval (20). Multiple alignments of GST sequences were performed with ClustalW (version 1.75) (21), and the aligned sequences were shaded and analyzed for secondary protein structures using the ENDscript program (22). An amino acid alignment for phylogenetic tree was constructed with ClustalW and the phylogenetically acceptable characters were derived with the help of the Jalview program. Phylogenetic trees were plotted using neighbor-joining methods implemented in MEGA software, version 4.0, with 1000 bootstrap replicates (23). Homology modeling of putative TtGstz1p was constructed by using the SWISS-MODEL Server and visualized with the Swiss-PdbViewer, version 3.7 (24).

Results

Molecular cloning and primary structure of TtGSTzeta

The highly conserved GSH binding motif (SSCX(W/H)RVIAL) of the GSTz class helped to identify the GSTz gene in *T. thermophila* macronuclear genome (13). The GSTz mRNA sequence was constructed based on the 8 expressed sequence tags (ESTs) in the 5' region and 2 ESTs in the 3' region from *T. thermophila* SB210 (25). One of the ESTs in the 3' region of the mRNA was amplified with the eF-Z6 and R-oligo dT22 primer set as a 370-bp fragment in this study. The GSTz-coding mRNA sequence was named the Tt*GSTz1* gene (mRNA gene accession number: FJ175686, Genomic ID: XM_001009329.2, Genome Project ID: TTHERM_00575360) and its putative amino acid sequence was named TtGstz1p (Figure 1) based on the presence of conserved features of the GSTz class defined in this study. Tt*GSTz1* mRNA has an 81-bp sequence in the 5' noncoding region, a 663-bp protein coding sequence, and 84 bp in the 3' noncoding region. Its mRNA encodes a 221-residues-long putative peptide that has a predicted molecular mass of 24 kDa. The deduced amino acid sequence of Tt*GSTz1* is more than 40% identical to the zeta-class GSTs isolated from different organisms (Figure 2a); for instance, it is 51% identical to *Drosophila melanogaster* (MAAI2), 48.1% to *Mus musculus* (MAAI), and 48.6% to *Arabidopsis thaliana* (GSTZ1) (Figure 2a). GST isoenzymes belonging to the different classes generally have around 20% sequence identity. As expected, TtGstz1p has 24.6% identity with ZmGSTtau and 21.5% with HsGSTomega. A multiple alignment of the TtGstz1p with all of the GST classes showed 4 conserved structural or functional motifs: GSH binding motif I ($_{16}$SSTSWRVRIAL$_{26}$), substrate selection motif II ($_{59}$VPAL$_{62}$), GSH binding motif III ($_{71}$ESSAILE$_{77}$), and substrate binding motif IV ($_{165}$GDEITLAD$_{172}$). Motif I is a GSH binding catalytic region where the first catalytic Ser is aligned with the catalytic Ser residue of the theta, delta, and kappa classes (data not shown). There are 2 putative substrate binding amino acid sites in the GSTz class, which forms part of the active site based on the crystal structure of maleylacetoacetate isomerase/glutathione transferase zeta (26). In TtGstz1p, the

```
  tatttatttatttatttatttcattattaaattaattaataaatatttatataaataa -60

tcttcaaaattattaatataataaagatatatacaataaacatatcaattaatattttaa -1
 M  A  E  S  S  K  K  I  T  L  Y  S  Y  F  R  S  S  T  S  W
ATGgctgaaagctcaaagaaaattactctttacagttatttccgctcctcaacttcatgg 60
 R  V  R  I  A  L  N  L  K  K  I  D  Y  N  I  I  P  I  N  L
agagtaagaatagctcttaatcttaagaaaatagactacaacatcattcctattaacctt 120
 L  K  S  E  Q  T  S  E  E  Y  T  K  I  N  P  N  Q  G  V  P
ttgaagagtgaataaacttctgaagaatatactaaaatcaatcctaactagggagtccct 180
 A  L  K  Y  G  D  E  V  I  I  E  S  S  A  I  L  E  F  L  E
gctttaaagtatggagacgaggttattattgaaagctctgctatcttagagttttggaa 240
 E  V  F  P  E  H  P  L  L  P  Q  D  A  V  K  R  A  Q  I  R
gaagtatttcctgagcacccacttcttccttaagatgcagttaaaagagcttaaattaga 300
 G  F  C  Q  V  I  N  T  A  I  H  P  L  Q  N  L  R  V  L  N
ggcttttgccaagtaattaacactgccatccaccctctctaaaatttgagggtgcttaat 360
 K  I  E  K  E  Y  S  Q  D  K  I  Q  W  L  K  F  W  V  T  K
aagattgaaaaggaatatagctaagacaagattcaatggctcaaattctgggttactaag 420
 G  L  T  A  I  E  E  L  L  K  N  S  H  G  K  Y  C  F  G  D
ggtttgacagcaattgaagaattattaaagaattctcatggaaaatattgctttggtgat 480
 E  I  T  L  A  D  L  F  L  V  P  Q  V  Q  G  V  V  D  R  F
gaaataacacttgctgacttattcttggtacctcaagttcaaggagttgtagatagattc 540
 Q  F  D  L  T  P  F  P  N  I  A  E  V  L  K  N  L  K  E  I
caatttgatttaacccctttccccaatattgcagaagttttaaaaaatttaaaggaaatt 600
 P  E  F  V  A  A  S  P  S  K  Q  A  D  N  P  D  N  Q  K  I
cctgagttcgttgctgcctcacccagcaagtaagctgacaatcctgataatcaaaagatc 660
 *
TGAaaataaaagcatatctttttaatatttttgtgttaaattataatttgtaactaaaat 720
ttattatttttacaattaaattcaaatcataatt                           780
```

Figure 1. The full length cDNA nucleotide and deduced amino acid sequences of Tt*GSTz1*. The predicted protein sequence is presented in single-letter code between the start (ATG) and stop (TGA) codons. The DNA sequence includes the 5′ and 3′ noncoding regions underlined and bolded. The conserved 4 motifs are colored with gray shadowing.

first one is a very conserved $_{116}$HPLQNL$_{122}$ sequence, located between helix α4 and α5, and the other is a less conserved $_{181}$GVVDR$_{185}$ sequence in the C-terminal end of helix α7 (Figure 2a).

An overall structural model of the TtGstz1p was constructed on the basis of the coordinates of the *Mus musculus* GSTz structure (2cz2A) with 46% identity and a 2.40e^{-42} e-value (Figure 2b). The prediction of the 3D model structuring of TtGstz1p clearly shows that it forms 2 spatially distinct domains; a smaller, well conserved GSH binding N-terminal domain (domain I) and a larger, less conserved substrate binding C-terminal domain (domain II) (Figure 2c). Domain I has a $\beta_1\alpha_1\beta_2\alpha_2\beta_3\beta_4\alpha_3$ thioredoxin fold, while domain II has an entirely helical structure. These 2 domains are joined by an 11-residue linker region that adopts a coil structure (Figures 2b and 2c).

The rooted phylogenetic tree of putative TtGstz1p, with all of the classes of the GST family, was constructed with the neighbor-joining method with EcGRX2 thioredoxin and GSTkappa as an outgroup (Figure 3a) (17). Overall, neighbor-joining tree topology was supported with high bootstrap values (>71%) in each node. The tree was divided into 2 major groups with a 71% bootstrap value; clade 1 was formed by lambda, zeta, omega, and tau and clade 2 was composed of delta, phi, theta, sigma, alpha, pi, and mu. Thus, phylogenetically, TtGstz1p

a

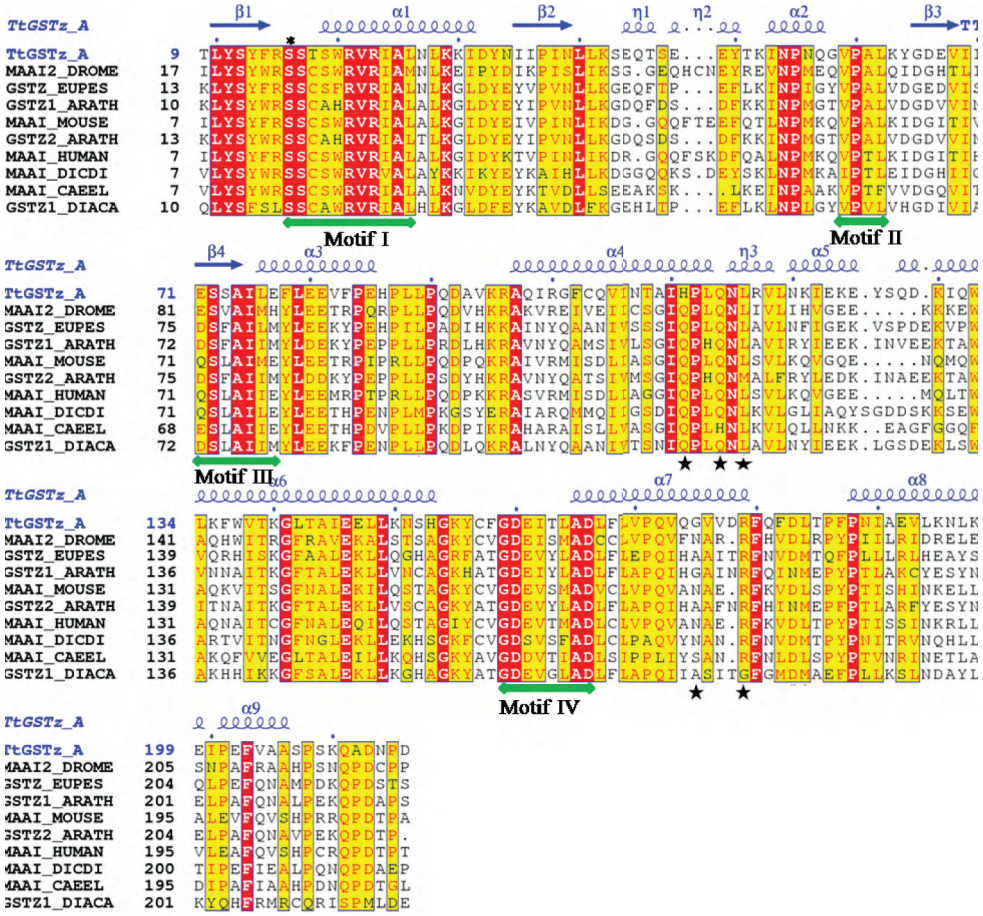


Figure 2. Sequence alignment of members of the GSTzeta class and the 3D structure prediction of TtGstz1p. **a)** Sequence alignment of GSTzeta class members compared with the secondary structure of TtGstz1p. The alignment was constructed using ClustalW and evaluated with ESPript. Alpha helices and beta strands are represented as helices and arrows, respectively, and strict beta turns are marked with TT. The proposed catalytic residue Ser_{16} (underlined) in the conserved region SSCX[W/H]RVIAL is marked with an asterisk. Periods indicate gaps introduced to optimize alignment. Similarities between the TtGSTz_A (blue) and homologous sequences aligned with CLUSTALW are rendered by colored shadowing. Residues with low similarity are written in black; they are in red and framed in yellow if they are highly similar, and residues are written in white on a red background in case of strict identity. The used orthologs were the following sequences: TtGSTz_A (*T. thermophila* **FJl75686**), MAAI2_DROME (*Drosophila melanogaster* **Q9VHD2**), GSTZ_EUPES (*Euphorbia esula* **P57108**), GSTZ1_ARATH (*Arabidopsis thaliana* **Q9ZVQ3**), MAAI_MOUSE (*Mus musculus* **Q9WVL0**), GSTZ2_ARATH (*A. thaliana* **Q9ZVQ4**), MAAI_HUMAN (*Homo sapiens* **O43708**), MAAI_DICDI (*Dictyostelium discoideum* **Q54YN2**), MAAI_CAEEL (*Caenorhabditis elegans* **Q18938**), and GSTZ1_DIACA (*Dianthus caryophyllus* **P28342**). ⋆ marks the putative substrate binding amino acid sites of MAAI (26). Green horizontal bar shows the positionally conserved structural or functional 4 motifs in all GSTs. **b)** Superposition of TtGstz1p (in green; residue range of 6 to 217) and *Mus musculus* GSTz (2cz2A; in black; residue range of 4 to 215) backbone α-carbon structure. **c)** A molecular model representation of TtGstz1p; the ribbon model of TtGstz1p, where α-helices are drawn as ribbons, β-strands as arrows, and coils as black loops. The red arrow indicates the position of the catalytic residue Ser16 in motif I of the GSH binding domain. The blue arrow shows the linker sequence of domain I and domain II. The green arrow points to the substrate binding region. All of the secondary structures are numbered. N and C represent the N and the C terminal ends.

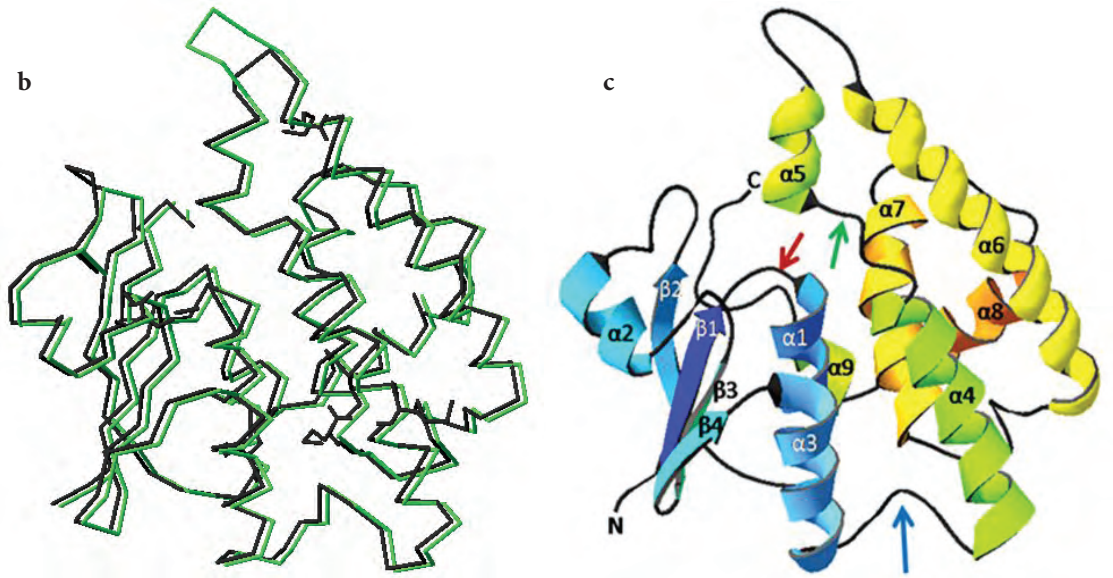


Figure 2. Continued.

was clearly a member of the zeta clade composed of human and mouse GSTz members with a 100% bootstrap value. In addition to this class determining analysis, the TtGstz1p was aligned with orthologous GSTz sequences from 61 different organisms. A resulting unrooted neighbor-joining tree showed that the phylogenetic relationship of the members of the GST zeta class appears to be based on the origin of the organism with unicellular, bacterial, plant, invertebrate, insect, and vertebrate (reptiles, birds, and mammalians) clades (Figure 3b). EcGRX2 thioredoxin (*E. coli*) was used as an outgroup to show the ancestral root in the tree. These major clades were internally supported with high bootstrap values but were not stable in the connecting nodes (data not shown). The unicellular zeta clade was composed of TtGstz1p, a putative maleylacetoacetate isomerase gene from *T. thermophila* (TtMAAI), and *Paramecium* zeta was closely positioned between outgroup EcGRX2 and the bacterial zeta clades.

mRNA distribution of *TtGSTz1*

The mRNA expression level of Tt*GSTz1* was investigated by using the RT-PCR method in response to cold shock (shift from 30 °C to 4 °C) and sublethal H_2O_2 oxidative stress. *Tetrahymena* cells, grown overnight and used as the untreated cell control, showed the level of TtGSTz1 mRNA expression before the stress treatment (Figure 4, lane 1). The oxidative stress-treated cells had almost the same amounts of mRNA at all of the time intervals (Figure 4, central gel, lanes 1-5). In the cold-shocked cells, however, the mRNA level of Tt*GSTz1* showed an increase at 30 and 45 min but decreased to the control level at 60 min (Figure 4, upper gel, lanes 1-5). The level of 17S rRNA transcript was equal at all of the time intervals (Figure 4, bottom gel, lanes 1-5). Based on this observation, Tt*GSTz1* mRNA transcript seems to be induced by a cold shock but not by sublethal oxidative stress conditions. This result may imply that Tt*GSTz1* is transiently necessary in increased amounts under cold stress. Further understanding of the physiology of Tt*GSTz1* requires comprehensive repetitive studies on the stress responsive changes of activity, protein, and mRNA under various biological conditions.

Codon adaptation, recombinant protein expression, and purification of TtGSTz

The site-directed mutagenesis strategy based on the codon usage table of *T. thermophila* and *E. coli* to alter the first base T to C in TAA and TAG glutamine codons of Tt*GSTz* as a silent mutation is given in Figure 5a, and the PCR products and their extension results are shown in Figure 5b. The DNA fragment carrying all of the mutations was cloned and transferred to the pET16b and transformed into *E.*

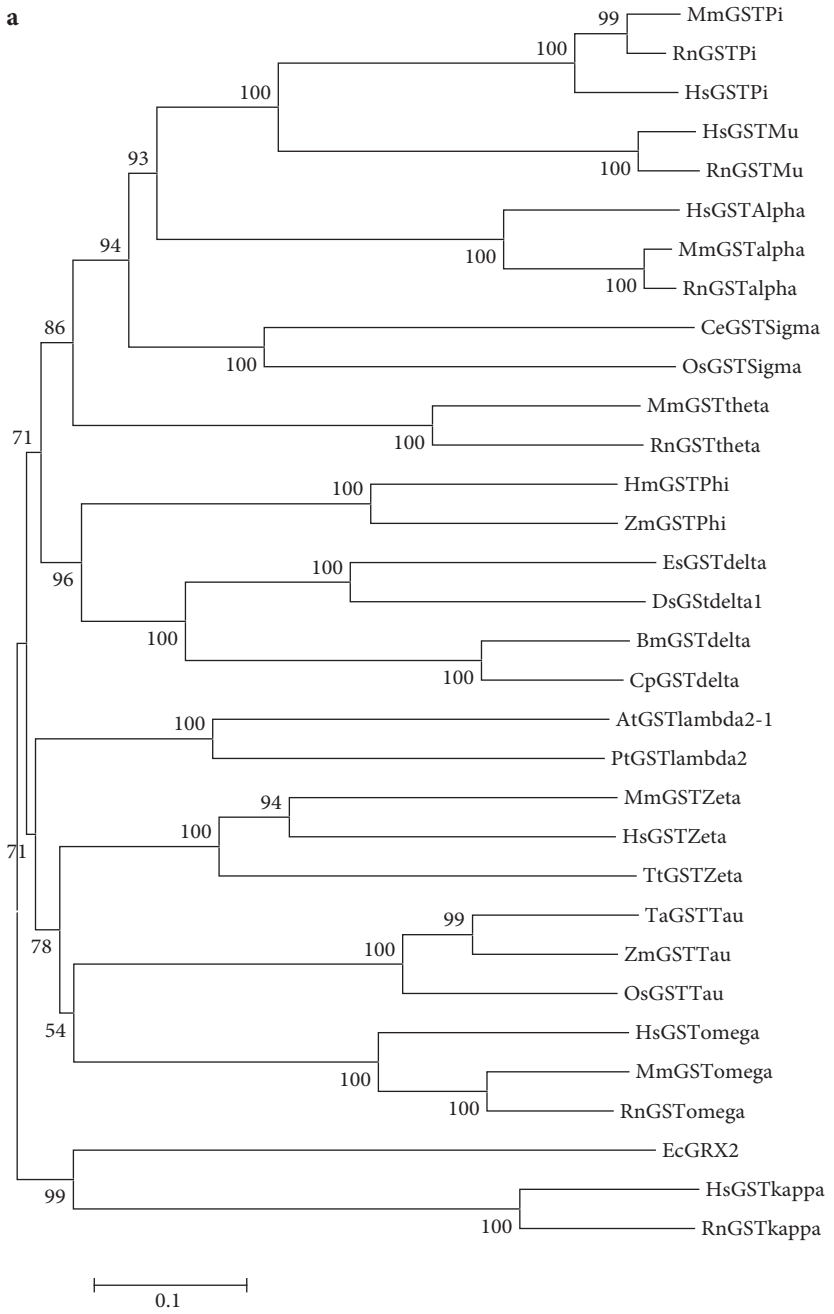
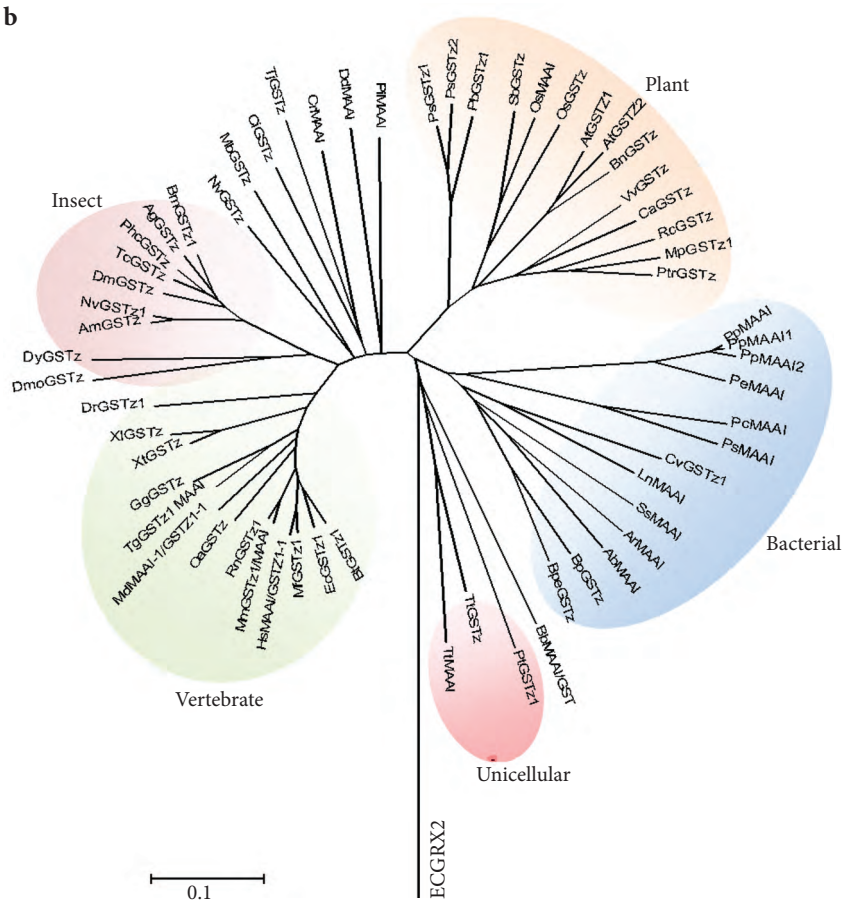


Figure 3. Evolutionary relationships of TtGstz1p with all of the GST classes and GSTz orthologs. **a)** A rooted phylogenetic tree analysis of TtGstz1p with all of the GST classes. The phylogenetic analyses were conducted using the neighbor-joining method in MEGA4. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. All of the positions containing gaps and missing data were eliminated from the dataset using Jalview and the complete deletion option in MEGA4. The substitution model was the p-distance in the amino acid level. There were a total of 285 positions in the final dataset coming from a total of 32 different GSTs - Alveolata/Ciliate: TtGSTzeta (*Tetrahymena thermophila*, **FJ175686**); Bacteria: EcGRX2 (*Escherichia coli*, **NP_287198.1**); Mammalian, Animal: MmGSTomega (*Mus musculus*, **NP_034492.1**), RnGSTomega (*Rattus norvegicus*, **Q9Z339.2**), HsGSTomega (*Homo sapiens*, **NP_004823.1**), HsGSTZeta (*H. sapiens*, **AAB96392.1**), MmGSTZeta (*M. musculus*, **EDL02925.1**), MmGSTtheta (*M. musculus*, **Q64471.3**), RnGSTtheta (*R. norvegicus*, **NP_445745.1**), HsGSTMu (*H. sapiens*, **NP_000840.2**), RnGSTMu (*R. norvegicus*, **NP_742035.1**), MmGSTPi (*M. musculus*, **NP_038569.1**), RnGSTPi (*R. norvegicus*, **NP_036709.1**), HsGSTPi (*H. sapiens*, **NP_000843.1**), MmGSTalpha (*M. musculus*, **NP_032207.3**), RnGSTalpha (*R. norvegicus*, **NP_058709.2**), HsGSTAlpha (*H. sapiens*, **NP_665683.1**), HsGSTkappa (*H. sapiens*, **NP_057001.1**), RnGSTkappa (*R. norvegicus*, **NP_852036.1**); Plant: AtGSTlambda2-1 (*Arabidopsis thaliana*, **NP_001119157.1**), PtGSTlambda2 (*Populus trichocarpa*, **ADB11342.1**), OsGSTTau (*Oryza sativa*, **A2XMN2.1**), TaGSTTau (*Triticum aestivum*, **CAC94002.1**), ZmGSTTau (*Zea mays*, **NP_001149886.1**), HmGSTPhi (*Hyoscyamus muticus*, **P46423.1**), ZmGSTPhi (*Z. mays*, **NP_001105412.1**); Arthropoda: EsGSTdelta (*Eriocheir sinensis*, **ACT78699.1**); Cephalopoda: OsGSTSigma (*Ommastrephes sloani*, **P4608**); Nematode: CeGSTSigma (*Caenorhabditis elegans*, **NP_501846.1**); Insecta, Lepidoptera: BmGSTdelta (*Bombyx mori*, **NP_001036974.1**), CpGSTdelta (*Cydia pomonella*, **ACG69436.1**); Fly: DsGSTdelta1 (*Drosophila simulans*, **AAK66764.1**). **b)** A rooted phylogenetic tree analysis of TtGstz1p with GSTz orthologs. The phylogenetic analyses were conducted using the neighbor-joining method in MEGA4, as described in the caption for Figure 4. There were a total of 175 positions in the final dataset coming from the 61 different GSTz - Animal, Mammalia: BtGSTz1 (*Bos taurus*, **NP_001069154.1**), EcGSTz1 (*Equus caballus*, **XP_001493411.1**), MfGSTz1 (*Macaca fascicularis*, **ABO21636.1**), HsMAAI/GSTZ1-1 (*Homo sapiens*, **O43708.3**), MmGSTz1/MAAI (*Mus musculus*, **EDL02925.1**), RnGSTz1 (*Rattus norvegicus*, **NP_001102915.1**), OaGSTz (*Ornithorhynchus anatinus*, **XP_001505837.1**), MdMAAI-1/GSTZ1-1 (*Monodelphis domestica*,



XP_001367277.1); Actinopterygii: DrGSTz1 (*Danio rerio*, **NP_001025442.1**); Aves: GgGSTz (*Gallus gallus*, **XP_001233621.**1), TgGSTz1(MAAI) (*Taeniopygia guttata*, **XP_002200574.1**); Amphibia: XtGSTz (*Xenopus tropicalis*, **AAI61329.1**), XlGSTz (*X. laevis*, **NP_001088856.1**); Insecta, Diptera: DyGSTz (*Drosophila yakuba*, **XP_002096943.1**), DmGSTz (*D. mojavensis*, **XP_001998510.1**), DmGSTz (*D. melanogaster*, **NP_996190.1**), AgGSTz (*Anopheles gambiae*, **XP_312009.2**); Hymenoptera: AmGSTz (*Apis mellifera*, **XP_394562.1**), NvGSTz1 (*Nasonia vitripennis*, **NP_001165931.1**); Coleoptera: TcGSTz (*Tribolium castaneum*, **XP_973541.2**); Paraneoptera: PhcGSTz (*Pediculus humanus corporis*, **XP_002426106.1**); Lepidoptera: BmGSTz1 (*Bombyx mori*, **NP_001037418.1**); Anthozoa: NvGSTz (*Nematostella vectensis*, **XP_001632220.1**); Choanoflagellida: MbGSTz (*Monosiga brevicollis*, **XP_001743166.1**); Tunicata: CiGStz (*Ciona intestinalis*, **XP_002122117.1**); Crustacea: TjGSTz (*Tigriopus japonicus*, **ACE81250.1**), CrMAAI (*Caligus rogercresseyi*, **ACO11269.1**); Amoebozoa: DdMAAI (*Dictyostelium discoideum*, **XP_642170.1**); Mycetozoa: PiMAAI (*Phytophthora infestans*, **EEY69157.1**); Plants, Spermatophyta, Coniferopsida: PsGSTz (*Picea sitchensis*, **ABK25945.1**), PsGSTz (*P. sitchensis*, **ABK24481.1**), PbGSTz (*Pinus brutia*, **ADB45877.1**); Liliopsida: SbGSTz (*Sorghum bicolor*, **XP_002450558.1**), OsMAAI (*Oryza sativa*, **AAX92930.1**), OsGSTz (*O. sativa*, **ABI17930.1**); Eudicotyledons: AtGSTZ1 (*Arabidopsis thaliana*, **NP_178344.1**), AtGSTZ2 (*A. thaliana*, **NP_178343.1**), BnGSTz (*Brassica napus*, **AAO60042.1**), VvGSTz (*Vitis vinifera*, **XP_002273077.1**), CaGSTz (*Capsicum annuum*, **ABQ88335.1**), RcGSTz (*Ricinus communis*, **XP_002519388.1**), MpGSTz1 (*Malva pusilla*, **AAO61856.1**), PtGSTz (*Populus trichocarpa*, **XP_002328824.1**); Bacteria, Gammaproteobacteria: PpMAAI (*Pseudomonas putida*, **NP_746728.1**), PpMAAI1 (*P. putida*, **YP_001670825.1**), PMAAI (*Psychrobacter* sp., **YP_001279637.1**), PeMAAI (*Pseudomonas entomophila*, **YP_610067.1**), PcMAAI (*Psychrobacter cryohalolentis*, **YP_580822.1**), PpMAAI2 (*Pseudomonas putida*, **AAO12529.1**), SMAAI (*Shewanella* sp.,**YP_870410.1**), AbMAAI (*Acinetobacter baumannii*, **ZP_04662105.1**), ArMAAI (*A. radioresistens*, **ZP_05361359.1**), as an outgroup EcGRX2 (*Escherichia coli*, GRX2,**NP_287198.1**); Betaproteobacteria: CvGSTz1 (*Chromobacterium violaceum*, **NP_900642.1**), LnMAAI (*Lutiella nitroferrum*, **ZP_03697904.1**), BpGSTz (*Bordetella parapertussis*, **NP_884118.1**), BpGSTz (*B. petrii*, **YP_001631756.1**); Deltaproteobacteria: BbMAAI/GST (*Bdellovibrio bacteriovorus*, **NP_967294.1**); Alveolata/ Ciliate: TtGSTz (*Tetrahymena thermophila*, **FJ175686**), TtMAAI (*T. thermophila*, **XP_001009332.1**), PtGSTz (*Paramecium tetraurelia*, **XP_001443850.1**).

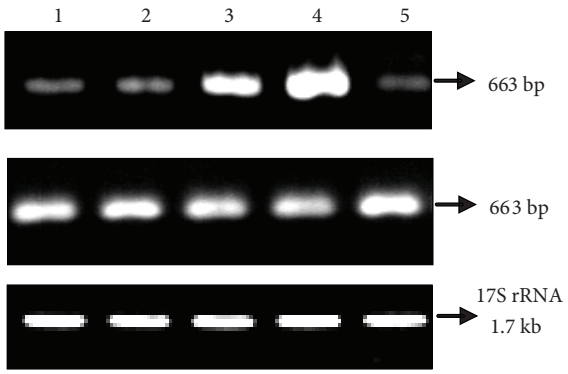
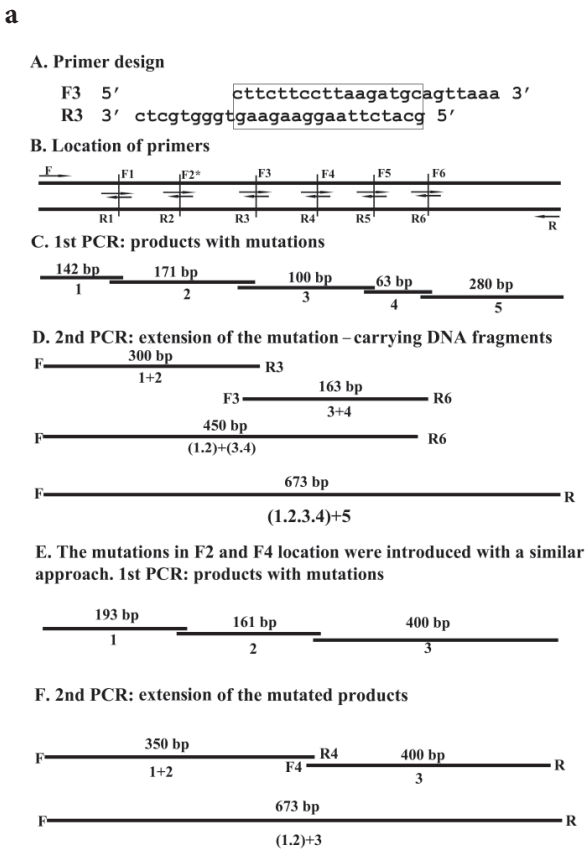


Figure 4. Tt*GSTz1* mRNA expression under cold shock and sublethal oxidative stress. Tt*GSTz1* mRNA expression was analyzed with RT-PCR on agarose gel. Total RNA was isolated from cells grown at 30 °C overnight and shifted to 4 °C for cold shock (upper) and to 0.020% H_2O_2 for sublethal oxidative stress (middle). 17S rRNA gene was used as a control to show the equal use of the total RNA samples in RT (lower). The control cells were grown at 30 °C overnight (lane 1). The cells were shifted to respective stresses for 15 min (lane 2), 30 min (lane 3), 45 min (lane 4), and 60 min (lane 5). After RT-PCR, cDNA products were analyzed with 1% agarose gel including ethidium bromide



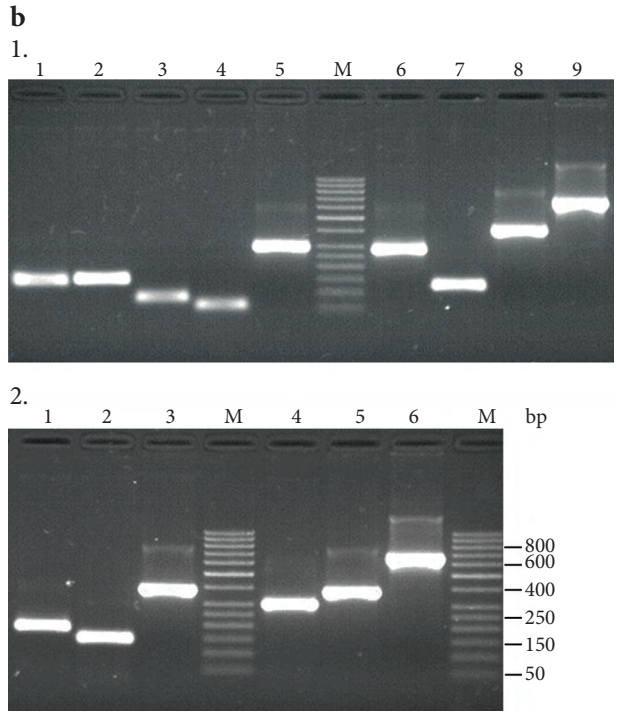


Figure 5. The site-directed mutagenesis of the Tt*GSTz1* gene. **a)** The site-directed mutagenesis strategy of Tt*GSTz1* gene using overextension PCR: **A** - an example for a complementary primer pair carrying mutation, **B** - the location of needed point mutations of Tt*GSTz1* gene, **C** - the strategy to introduce 5 point mutations to the protein coding region and the expected sizes of mutation-carrying PCR products, **D** - the strategy to extend the first PCR products in the second PCR run to yield full-length optimized Tt*GSTz1*, and **E** - and **F** - the mutations in the F2 and F4 locations were introduced with a similar approach. **b)** The codon optimization of Tt*GSTz1* gene: **1** - The product sizes of the mutation-carrying amplified 5 fragments in the first PCR were obtained with the primer set and their respective Tm in PCR; 2F/R1 at 52 °C (lane 1), F1/R3 at 56 °C (lane 2), F3/R5 at 57 °C (lane 3), F5/R6 at 55 °C (lane 4), and F6/2R at 58 °C (lane 5). The purified first PCR fragments were used as a template to extend in the second PCR with most of the 5' and 3' end primer set F/R3 at 55 °C (lane 6), F3/R6 at 52 °C (lane 7), F3/2R at 56 °C (lane 8), and 2F/2R at 55 °C (lane 9). **2** - The mutations in the F2 and F4 locations were introduced with a similar approach. The fragments produced in the first PCR were gained with 2F/R2 at 51 °C (lane 1), F2/R4 at 57 °C (lane 2), and F4/2R at 57 °C (lane 3). The products of the second extension PCR with the most 5' and 3' end primers of the used fragments were obtained with 2F/R4 at 51 °C (lane 4), F4/2R at 51 °C (lane 5), and 2F/2R at 55 °C (lane 6). M: 50-1000 bp ladder (Bio Basic GM345). The DNA fragments were separated in 2% agarose gel and visualized with ethidium bromide.

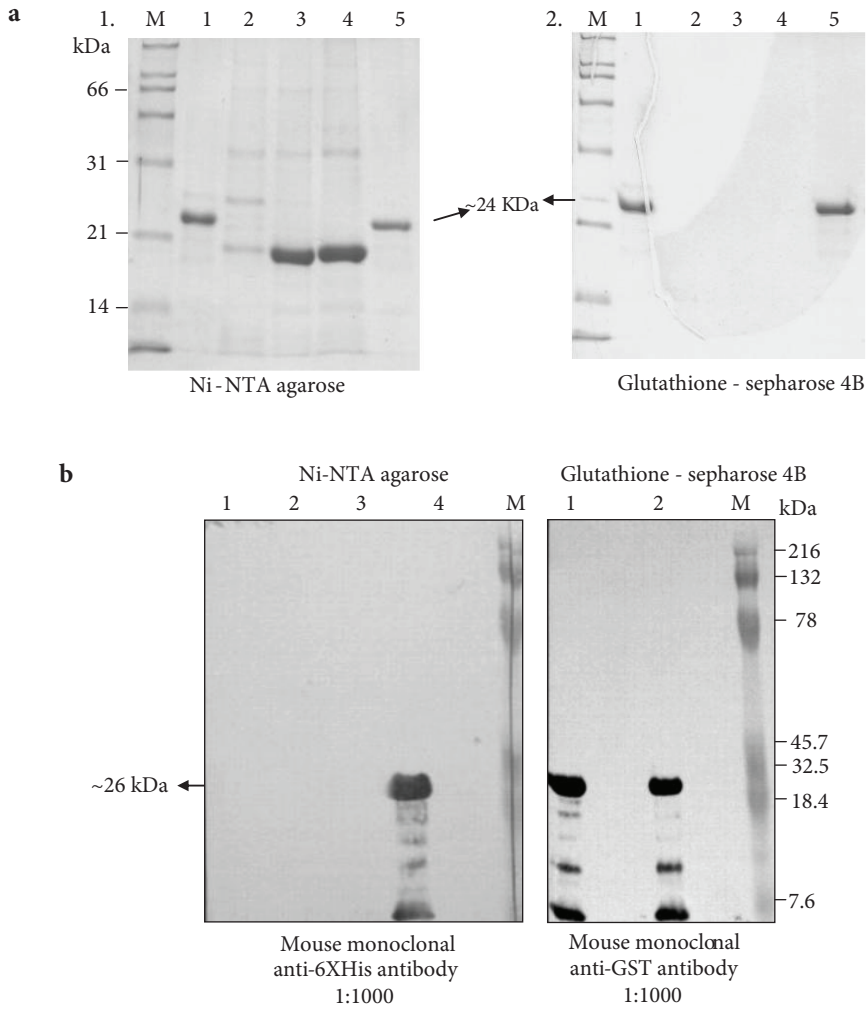


Figure 6. SDS-PAGE and western blot analysis of affinity purified recombinant 6XHis-TtGstz1p proteins. **a)** SDS-PAGE analysis: the recombinant proteins were purified with Ni-NTA agarose and glutathione-sepharose 4B. Lane M: SDS-PAGE standards-broad range (Bio-Rad). Lane 1: positive control - GST protein produced from pGEX 4T-1 isolated with glutathione-sepharose 4B; lane 2: negative control - purified proteins from uninduced cells; lane 3: proteins from 1-h IPTG-induced cells; lane 4: proteins from 2-h IPTG-induced cells; lane 5: proteins from 3-h IPTG-induced cells. Purified recombinant proteins were analyzed with 10% SDS-PAGE and visualized with Coomassie Blue R-250 staining. **b)** Western blot analysis: 6X-TtGstz1p fusion proteins were purified with Ni-NTA agarose and glutathione-sepharose 4B. Proteins were separated with 10% SDS-PAGE and blotted to PVDF membranes. The membranes were probed with a mouse monoclonal anti-6XHis antibody, 1:1000 (Ni-NTA agarose), and a mouse monoclonal anti-GST antibody, 1:1000 (glutathione-sepharose 4B). Lane M is kaleidoscope prestained standards (Bio-Rad). Left: lane 1, positive control - GST protein produced from pGEX 4T-1 isolated with glutathione-sepharose 4B; lane 2, 6X-TtGstz1p proteins from 1-h IPTG-induced cells; lane 3, proteins from 2-h IPTG-induced cells; lane 4: proteins from 3-h IPTG-induced cells. Right: lane 1. positive control - GST protein produced from pGEX 4T-1; lane 2, proteins from 3-h IPTG-induced cells.

coli BL21 (DE3). SDS-PAGE analysis with Coomassie blue staining showed that the affinity purified recombinant protein was expressed as a 24-kDa protein at 3 h after IPTG induction (Figure 6a, lane 5), but it is absent in the uninduced cells (Figure 6a, lane 2). The 6XHis-TtGstz1p protein was purifiable using both the nickel agarose and glutathione-sepharose 4B methods (Figure 6b, lane 5). Western blot analysis using anti-His antibodies showed that the 24-kDa protein contains the His tag (Figure 6a,

lane 4), but the negative control was not recognized (Figure 6b, left gel, lane 1). Anti-GST antibodies in the western blot analysis also recognized not only the recombinant purified 24-kDa protein (Figure 6b, right gel, lane 2) but also the GST positive control in lane 1. Overall, the 6XHis-TtGstz1p (24 kDa) was produced in *E. coli* and purified by both purification methods.

Discussion

The recent *T. thermophila* genome project identified a large number of putative genes in the GST family (13). However, in the literature, there is only one paper about *Tetrahymena thermophila* GST enzymes, in which an increasing level of GST activity in the early stationary phase of the growth cycle was reported (27). Searching the *Tetrahymena* macronuclear genome project data with the tools of bioinformatics resulted in the identification of about putative 62 members belonging to the mu, theta, omega, and zeta classes (unpublished results). In this study, one member of the *Tetrahymena* GSTzeta class was classified as Tt*GSTz1* based on the presence of zeta-class conserved motif I (SSTSWRVRIAL) where the conserved cysteine (C) was replaced by threonine (T), which has similar amino acid properties in the conserved consensus motif I (SSCX[WH]RVRIAL) of the GSTz class.

The TtGstz1p had a position within the zeta clade when compared with all of the known GST classes in this study, which were very well conserved over a considerable evolutionary period (17,28). This neighbor-joining rooted tree (Figure 3a) analysis suggested that the most ancient classes of cytosolic GSTs were lambda for clade 1 and delta for clade 2, rather than theta and omega as previously reported (29). This major division of the GST family reported here was not in agreement with the maximum-likelihood tree (28), in which the omega, tau, and lambda classes were absent; moreover, their tree topology had low bootstrap values. However, when TtGstz1p was analyzed with zeta orthologs from different taxa (Figure 3b), it seemed to be phylogenetically localized between the outgroup *E. coli* GRX2 (17) and the prokaryotic GSTzeta clade. In the unicellular zeta clade, there was a Deltaproteobacteria member *Bdellovibrio* GSTz sequence with a low bootstrap value of 49%. The positional phylogenetic proximity of TtGstz1p to the prokaryotic zeta clade and outgroup may suggest that the *Tetrahymena* GSTz gene was acquired from bacteria by horizontal gene transfer, a relatively common occurrence in ciliates (30). A similar conclusion was suggested for a bacterial-like citrate synthase gene from *T. thermophila* (31). Consequently, the phylogenetically bacteria-like unicellular TtGstz1p is highly recognizable by the mouse monoclonal anti-GST antibody, which is produced against a GSTmu protein from a platyhelminth parasite, *Schistosoma japonicum* (9). Accordingly, this antibody is not able to recognize higher eukaryotic GSTs from rat, rabbit, porcine, bovine liver, or human placenta when tested by ELISA (32). It would be very interesting to see if the other commercial GST antibodies could recognize bacteria-like TtGstz1p.

Knowledge of the stress-dependent distributions of TtGSTz1 mRNA could be used to enhance the state of gene expression related to its biotechnological use. Using RT-PCR analysis, we demonstrated that TtGSTz1 mRNA is present under all of the studied conditions, including untreated vegetative cells grown at 30 °C. Previously, microarray (33) and EST (Unigene ID: Tth.75) profile (34) studies of *Tetrahymena thermophila* showed similar results of Tt*GSTz1* mRNA being highly present in vegetative cells; however, it drops to the basal level in conjugative and starved cells (33,34) and is absent in heavy metal (11 mM $CdCl_2$ and 500 mM $CuSO_4$ for 1 h)-stressed cells (34). Therefore, it is reasonable to say that the TtGSTz1 gene is a ubiquitously expressing gene in vegetative cells, but its mRNA expression levels at some time points during starvation and conjugation are 10-fold higher than the background or totally absent. The presence of OoGSTz1 mRNA from *Oryza sativa* (rice) in cold and H_2O_2 oxidative stress was reported by Tsuchiya et al. (35). Transgenic rice plants overexpressing the OoGSTz1 cDNA showed enhancement of chilling tolerance in germination and seedling stages (36). The reason for differential mRNA expression of Tt*GSTz1* gene response against cold shock still remains unknown in *T. thermophila*.

Homologous and heterologous protein expression in *T. thermophila* has a number of difficulties to be addressed, such as the need for a highly efficient

transformation method, better-designed protein expression vectors, and a *Tetrahymena* protein expression host supplemented with inducible rare tRNA genes for human gene expression (6). To obtain a large amount of proteins in *T. thermophila*, fusion protein techniques with a GST tag may help stabilize and purify heterologous protein (8). The results of this study suggest that the basic affinity characteristics of TtGstz1p offer a biotechnological use as a C- or N-terminal dual tag directly in *E. coli* or indirectly in *Tetrahymena* protein expression vectors for sequential 2-step protein purification. The use of homologous TtGSTz1 instead of *Schistosoma japonicum* GSTmu protein as a tag in *Tetrahymena* protein expression vectors could be advantageous due to the stability of the mRNA and the codon usage of *T. thermophila*. However, endogenic TtGSTz1 mRNA expression at 30 °C (Figure 4, lane 1) might mean the presence of its protein expression. Generally, recombinant protein expression in *T. thermophila* is executed in cells grown at 30 °C; therefore, GSH affinity purification of a recombinant protein with a TtGSTz1 tag could lead to some endogenic contamination of TtGstz1p as well as other GST homologs. This problem could be resolved with the use of a dual tag of 6XHis-TtGstz1p with Ni^{2+}-NTA columns or a sequential 2-step purification procedure on glutathione-agarose and Ni^{2+}-NTA columns in *Tetrahymena* studies (10).

In conclusion, we report evidence that a novel member of the GSTzeta class (Tt*GSTz1*) is present in the unicellular *T. thermophila*. Based on the observations presented in this study, we suggest that 6XHis-TtGSTz can be used as a dual tag in *Tetrahymena* expression vectors to improve the high level purification of recombinant full-size fusion proteins by a sequential 2-step procedure on glutathione-agarose and Ni^{2+}-NTA columns. Presently, its use is still unknown in *T. thermophila*; nevertheless, the information from this study will certainly provide further insight into the test or discovery of a better GST tag candidate in *T. thermophila*

Acknowledgments

We would like to thank Dr Eduardo Orias for strain *T. thermophila* SB210.

Corresponding author:

Muhittin ARSLANYOLU

Department of Biology,

Faculty of Sciences,

Anadolu University, Yunusemre Campus,

26470 Eskisehir - TURKEY

E-mail: marslanyolu@anadolu.edu.tr

References

1. Yu GL, Blackburn EH. Transformation of *Tetrahymena thermophila* with a mutated circular ribosomal DNA plasmid vector. Proc Natl Acad Sci USA 86: 8487-8491, 1989.
2. Gaertig J, Gorovsky MA. Efficient mass transformation of Tetrahymena thermophila by electroporation of conjugants. PNAS 89: 9196-9200, 1992.
3. Bruns PJ, Cassidy-Hanley D. Biolistic transformation of macro- and micronuclei. Methods Cell Biol 62: 501-512, 2000.
4. Gaertig J, Gao Y, Tishgarten T et al. Surface display of a parasite antigen in the ciliate *Tetrahymena thermophila*. Nat Biotechnol 17: 462-465, 1999.
5. Peterson DS, Gao Y, Asokan K et al. The circumsporozoite protein of *Plasmodium falciparum* is expressed and localized to the cell surface in the free-living ciliate *Tetrahymena thermophila*. Mol Biochem Parasitol 122: 119-126, 2002.
6. Weide T, Herrmann L, Bockau U et al. Secretion of functional human enzymes by *Tetrahymena thermophila*. BMC Biotechnol 6: 19, 2006.
7. Orias E, Hamilton EP, Orias JD. Tetrahymena as a laboratory organism: useful strains, cell culture, and cell line maintenance. Methods Cell Biol 62: 189-211, 2000.
8. Davies AH, Jowett JBM, Jones IM. Recombinant baculovirus vectors expressing glutathione *S*-transferase fusion proteins. Biotechnology 11: 933-936, 1993.
9. Smith DB, Johnson KS. Single-step purification of polypeptides expressed in *Escherichia coli* as fusion with glutathione *S*-transferase. Gene 67: 31-40, 1988.
10. Nilsson J, Stahl S, Lundeberg J et al. Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. Protein Expr Purif 11: 1-16, 1997.
11. Panagiotidis CA, Silverstein SJ. pALEX, a dual-tag prokaryotic expression vector for the purification of full length proteins. Gene 164: 45-47, 1995.
12. Truong K, Khorchid A, Ikura M. A fluorescent cassette-based strategy for engineering multiple domain fusion proteins. BMC Biotechnology 3: 8, 2003.

13. Eisen JA, Coyne RS, Wu M et al. Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. PLoS Biol 4: 286, 2006.

14. Bock KW, Lilienblum W, Fischer G et al. The role of conjugation reactions in detoxication. Arch Toxicol 60: 22-29, 1987.

15. Wilce MCJ, Parker MW. Structure and function of glutathione S-transferases. Biochim Biophys Acta 1205: 1-18, 1994.

16. Habig WH, Pabst MJ, Jacoby WB. Glutathione S-transferases. J Biol Chem 249: 7130-7139, 1974.

17. Frova C. Glutathione transferases in the genomics era: new insights and perspectives. Biomol Eng 23: 149-169, 2006.

18. Herlitze S, Koenen M. A general and rapid mutagenesis method using polymerase chain reaction. Gene 91: 143-147, 1990.

19. Ferreira PM, Costa V, Guerreiro P et al. Heat shock effect on glutathione and superoxide dismutase in *Tetrahymena pyriformis*. Cell Biol Int Rep 16: 19-26, 1992.

20. Altschul SF, Madden TL, Schaffer AA et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402, 1997.

21. Higgins D, Thompson J, Gibson T et al. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680, 1994.

22. Gouet P, Courcelle E, Stuart DI et al. ESPript: multiple sequence alignments in PostScript. Bioinformatics 15: 305-308, 1999.

23. Tamura K, Dudley J, Nei M et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599, 2007.

24. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18: 2714-2723, 1997.

25. Stover N, Krieger C, Binkley G et al. *Tetrahymena* Genome Database (TGD): a new genomic resource for *Tetrahymena thermophila* research. Nucleic Acids Res 34: 500-503, 2006.

26. Polekhina G, Board PG, Blackburn AC et al. Crystal structure of maleylacetoacetate isomerase/glutathione transferase zeta reveals the molecular basis for its remarkable catalytic promiscuity. Biochemistry 40: 1567-1576, 2001.

27. Overbaugh M, Lau EP, Marino VA et al. Purification and preliminary characterization of monomeric glutathione S-transferase from *Tetrahymena thermophila*. Arch Biochem Biophys 261: 227-234, 1988.

28. Board PG, Baker RT, Chelvanayagam G et al. Zeta, a novel class of glutathione transferases in a range of species from plants to humans. Biochem J 328: 929-935, 1997.

29. Fonseca RR, Johnson WE, O'Brien SJ et al. Molecular evolution and the role of oxidative stress in the expansion and functional diversification of cytosolic glutathione transferases. BMC Evol Biol 10: 281, 2010.

30. Ricard G, McEwan NR, Dutilh BE et al. Horizontal gene transfer from Bacteria to rumen Ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment. BMC Genomics 7: 22, 2006.

31. Mukai A, Endoh H. Presence of a bacterial-like citrate synthase gene in *Tetrahymena thermophila*: recent lateral gene transfers (LGT) or multiple gene losses subsequent to a single ancient LGT? J Mol Evol 58: 540-549, 2004.

32. Campbell MJ, McFall P, Niederhuber JE. Production and characterization of a monoclonal antibody against *Schistosoma japonicum* glutathione S transferase. J Immunol Methods 18: 73-78, 1995.

33. Miao W, Xiong J, Bowen J et al. Microarray analyses of gene expression during the *Tetrahymena thermophila* life cycle. PloS ONE 4: e4429, 2009.

34. Pontius JU, Wagner L, Schuler GD. UniGene: a unified view of the transcriptome. In: McEntyre, J, Ostell J. eds. The NCBI handbook. Bethesda (MD): National Center for Biotechnology Information; 2003: pp. 21/1-21/12.

35. Tsuchiya T, Takesawa T, Kanzaki H et al. Genomic structure and differential expression of two tandem-arranged GSTZ genes in rice. Gene 335: 141-149, 2004.

36. Takesawa T, Ito M, Kanzaki H et al. Overexpressing of zeta glutathione S-transferase in transgenic rice enhances germination and growth at low temperature. Mol Breed 9: 93-101, 2002.