Original Article
Comparison of the functions of glutathionylspermidine synthetase/amidase from *E. coli* and its predicted homologues YgiC and YjfC

Li Sui[1], John C Warren[2], Janelle PN Russell[2], Nina V Stourman[1,2]

[1]*Center for Applied Chemical Biology, Youngstown State University, One University Plaza, Youngstown, OH, 44555, USA;* [2]*Department of Chemistry, Youngstown State University, One University Plaza, Youngstown, OH, 44555, USA*

Abstract: Protein function prediction is very important in establishing the roles of various proteins in bacteria; however, some proteins in the *E. coli* genome have their function assigned based on low percent sequence homology that does not provide reliable assignments. We have made an attempt to verify the prediction that *E. coli* genes *ygiC* and *yjfC* encode proteins with the same function as glutathionylspermidine synthetase/amidase (GspSA). GspSA is a bifunctional enzyme that catalyzes the ATP-dependent formation and hydrolysis of glutathionylspermidine (G-Sp), a conjugate of glutathione (GSH) and spermidine. YgiC and YjfC proteins show 51% identity between themselves and 28% identity to the synthetase domain of the GspSA enzyme. YgiC and YjfC proteins were expressed and purified, and the properties of GspSA, YgiC, and YjfC were compared. In contrast to GspSA, proteins YgiC and YjfC did not bind to G-Sp immobilized on the affinity matrix. We demonstrated that all three proteins (GspSA, YgiC and YjfC) catalyze the hydrolysis of ATP; however, YgiC and YjfC cannot synthesize G-Sp, GSH, or GSH intermediates. *gsp, ygiC,* and *yjfC* genes were eliminated from the *E. coli* genome to test the ability of mutant strains to synthesize G-Sp conjugate. *E. coli* cells deficient in GspSA do not produce G-Sp while synthesis of the conjugate is not affected in *ΔygiC* and *ΔyjfC* mutants. All together our results indicate that YgiC and YjfC are not glutathionylspermidine synthetases as predicted from the amino acid sequence analysis.

Keywords: Glutathione, glutathionylspermidine, glutathionylspermidine synthetase/amidase, ATPase, ATP-grasp domain

Introduction

The complete *E. coli* genome was sequenced about fifteen years ago; however, despite significant efforts, many proteins encoded in the *E. coli* genome do not have function assignments, or have functions assigned based on the sequence analysis but have not been proven experimentally [1]. Protein function prediction is important in establishing the roles of various proteins in bacteria, and different approaches are undertaken to make the protein function assignments more accurate [2]. Nevertheless, the function prediction is often done based on low percent sequence similarities which makes it unreliable, or even misleading, in establishing more specific function of proteins. In many cases, the protein fold serves as a foundation for the assignment; however, multiple examples suggest that proteins with similar structures could participate in very different processes [3].

Here, we describe an attempt to verify the function prediction for the proteins encoded by two *E. coli* genes - *ygiC* and *yjfC.* YgiC and YjfC proteins are annotated as homologues of glutathionylspermidine synthetase/amidase (GspSA). This enzyme catalyzes the formation and hydrolysis of glutathionylspermidine (G-Sp), a conjugate of the tripeptide glutathione (GSH) and polyamine spermidine [4]. Both proteins YgiC and YjfC have a similar size of 45 kD and show 51% identity between themselves and 28% identity to the synthetase domain of glutahionylspermidine synthetase/amidase [5]. These

sequence similarities served as a base for the prediction that all three proteins are involved in the G-Sp synthesis, however until now, this prediction was not tested experimentally.

The existence of the unusual glutathione derivative - glutathionylspermidine conjugate in *E. coli* was detected several decades ago [6]. Later, the enzyme involved in its synthesis was identified and it was shown that this enzyme has two types of activities – synthetase, or amido-ligase, and amidase [4]. The GspSA protein is a homodimer in which each monomer consists of two distinct domains. The C-terminal domain is responsible for the synthesis of G-Sp while the N-terminal domain catalyzes the hydrolysis of the conjugate [5]. The synthetase domain of GspSA belongs to the class of ATP-grasp structural domains [7]. Proteins within the ATP-grasp superfamily demonstrate a very broad spectrum of physiological functions and reactions that they catalyze [8]. The common feature of these enzymes is the use of ATP to activate carboxyl groups for the reaction with nucleophiles via acylphosphate intermediate [9]. The members of this superfamily possess the structural elements which allow them to specifically bind a molecule of ATP in an orientation that promotes the reaction. Glutathione synthetase, the enzyme catalyzing the second step of GSH synthesis, condensation of γ-glutamylcysteine with glycine, is also a member of the ATP-grasp superfamily [10].

In our study we cloned, expressed, and purified YgiC and YjfC proteins. These proteins were tested for their capability to catalyze the formation of G-Sp, GSH, and its intermediates. An affinity column with immobilized G-Sp was used to check whether YgiC and YjfC proteins are capable of binding the conjugate. The ability of *E. coli* cells lacking *ygiC* and *yjfC* genes to synthesize G-Sp was also examined. Our results indicate that even though all three proteins (GspSA, YgiC, and YjfC) possess ATPase activity, YgiC and YjfC cannot synthesize G-Sp, GSH, or GSH intermediates. Strains lacking the *gsp* gene do not produce G-Sp conjugate while synthesis of the conjugate is not affected in *ygiC* and *yjfC* knockout strains.

Materials and methods

Materials

E. coli K-12 genomic DNA was from ATCC (Manassas). Primers were custom ordered from Invitrogen (Carlsbad, CA). Restriction enzymes were from New England Biolabs (Ipswich, MA). The pET20b(+) vector was from EMB Chemicals, Inc. (Gibbson, NJ). BL21(DE3) competent cells were from Stratagene (La Jolla, CA). Luria Bertani (LB) broth, ampicillin, ATP, chloramphenicol, dithiothreitol (DTT), glycine, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), isopropyl β-D-1-thiogalactopyranoside (IPTG), kanamycin, nicotinamide adenine dinucleotide (reduced form) (NADH), piperazine-N,N′-bis(2-ethanesulfonic acid) (PIPES), sodium chloride, sodium acetate, streptomycin sulfate, and bovine serum albumin (BSA) were from Amresco (Solon, OH). Glutahione (GSH), γ-glutamylcysteine, spermidine, 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB), lactate dehydrogenase (LDH), phosphoenolpyruvate, pyruvate kinase (PK), and formic acid were purchased from Sigma Aldrich (St. Louis, MO). Cysteine, glutamic acid, acetonitrile (HPLC grade), β-mercaptoethanol, and magnesium chloride were from Fisher Scientific (Pittsburgh, PA). DEAE-Cellulose, SP-Sephadex, Sephacryl S-200, and Epoxy-activated Sepharose 6B were from GE Healthcare (Piscataway, NJ). pET20b-*gsp* and pET20b-*gspC59A* vectors for expression of GspSA protein and its C59A mutant, respectively, *E. coli* BW25113-pKD46 and BW25113-*Δgsp::cmp* cells, and pKD3 and pKD13 vectors were generous gift from Dr. Richard Armstrong (Vanderbilt University, Nashville, TN).

DNA manipulations

ygiC gene was amplified from genomic *E. coli* DNA by PCR using the following primers: forward AAAA<u>GGATCC</u> ATGGAAAGAGTCAGTATTA-CCGAGCG and reverse AACCC<u>AAGCTT</u>ATTCA-ACAAAAATATGTGGATAAAACCG. The recognition sites for *BamH1* and *Hind III* restriction enzymes are underlined. The PCR product was digested with *BamH1* and *Hind III* restriction enzymes and subcloned into pET20b vector digested with the same enzymes. The resulting plasmid pET20b-*ygiC* was transformed into BL21(DE3) cells for protein expression. *yjfC* gene was amplified from genomic *E. coli* DNA by PCR using the following primers: forward AAAAAAAA<u>CATATG</u>CTGAGACACAACGTTCCTGTG and reverse AAA<u>GAATTC</u>TTATCCAGCAATGTAA-TGCGGAACG. The recognition sites for *NdeI* and *EcoRI* restriction enzymes are underlined.

The PCR product was digested with *NdeI* and *EcoRI* enzymes and ligated into pET20b vector digested with the same enzymes. The resulting pET20b-*yjfC* plasmid was transformed into BL21(DE3) cells for protein expression.

Protein expression

Overnight culture of BL21(DE3) cells containing either pET20b-*ygiC* or pET20b-*yjfC* grown at 37 $^{\circ}$C in LB media containing ampicillin at concentration 100 μg/mL was diluted 100 times into fresh LB/ampicillin media and grown with shaking at 37 $^{\circ}$C until OD_{600} of 0.6. At this time 0.3 mM IPTG was added to induce protein expression and the culture was allowed to grow for additional 4 hrs. The cells were harvested by centrifugation at 6000 x g for 10 min, the pellets were frozen and then resuspended in 50 mM HEPES buffer, pH 7.5 containing 1 mM EDTA and 1 mM DTT. The cells were disrupted by sonication (six 30 sec cycles with 1 min interval in between), and the cell debris were removed by centrifugation.

Protein purification

All buffers used for protein purification contained 1 mM DTT. The supernatant obtained after removal of cell debris was treated with streptomycin sulfate to a final concentration of 1% (w/v), and the precipitate was removed by centrifugation at 12,000 x g for 20 min. The supernatant was treated with ammonium sulfate to reach 75% saturation, and the precipitate was collected by centrifugation at 12,000 x g for 20 min. The pellet was dissolved in small amount of 50 mM potassium phosphate buffer, pH 7.5, dialyzed overnight against the same buffer, and loaded on the DEAE-cellulose column (11 x 3.0 cm) equilibrated with the same buffer. The column was washed with the buffer and then proteins were eluted with an NaCl gradient (0 – 0.4 M) in the same buffer. Fractions were analyzed by absorbance at 280 nm and SDS-PAGE. Fractions containing a strong band at 45 kD were pooled together, concentrated, and further purified by gel filtration chromatography using Sephacryl S-200 column (60 x 2 cm). Final concentration of proteins was determined by Bradford assay with BSA as a standard.

Synthesis of glutathionylspermidine

Glutathionylspermidine was synthesized enzymatically using mutant glutathionylspermidine synthetase/amidase (GspSA C59A). GspSA C59A protein was expressed in BL21(DE3) containing pET20b-*gspC59A* and purified as previously described [4]. The reaction mixture for the synthesis of G-Sp contained 15 μM GspSA C59A, 10 mM GSH, 10 mM spermidine, 15 mM ATP, and 15 mM $MgCl_2$ in 40 mM NaPIPES buffer, pH 6.8. The reaction mixture was incubated at 37 $^{\circ}$C for 2-3 hrs. GspSA C59A protein was denatured by heating at 90 $^{\circ}$C for 3 min and removed by centrifugation, while the rest of the mixture was loaded onto a SP-Sephadex column (10 x 1.7 cm) equilibrated with 50 mM potassium phosphate buffer, pH 7.0. The SP-Sepharose column was washed with 50 mM potassium phosphate buffer, pH 7.0 first to remove GSH and then the conjugate was eluted with a step-gradient of ammonium acetate (150, 300, 600 mM), pH 5.0. The collected fractions were tested with 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB) derivatizing agent for the presence of thiols, and derivatives were analyzed by HPLC. Fractions containing G-Sp were pooled together and lyophilized.

HPLC analysis of TNB-thiol derivatives

Thiols derivatized with DTNB were analyzed by HPLC using a reverse phase column (Atlantis T3, 3 μm, 4.6 x 50 mm, Waters Corporations, Milford, MA). Buffer A consisted of 0.9% (v/v) formic acid and buffer B was 100% acetonitrile. The concentration of buffer B was first increased from 5% to 30% over 9 minutes, then to 80% over 7 minutes, and held at 80% for 5 minutes. Buffer B concentration was then decreased to 5% over 1 minute and held constant for 10 minutes. The flow rate was set at 0.8 mL/min through the entire cycle which lasted 33 minutes. 50 μL of the sample was injected for analysis and the products were detected at 330 nm.

G-Sp immobilized affinity matrix

A solution containing about 1 mM of G-Sp conjugate was prepared in 0.1 M $NaHCO_3$ buffer, pH 10. 0.5 g of Epoxy-Activated Sepharose 6B was allowed to swell in water for 15 min at room temperature then washed several times with water and 0.1 M $NaHCO_3$ buffer, pH 10. The G-Sp solution was incubated with the resin at room temperature with constant rotation for 24 hrs and then removed from the resin by centrifugation. Unreacted epoxy-groups of the

resin were blocked by incubation with a 1 M solution of 2-mercaptoethanol for 24 hrs. Prepared modified G-Sp-Sepharose (about 3 mL) was washed several times with buffers of various pH (4–10) to remove all unbound material.

Affinity chromatography with G-Sp-Sepharose

G-Sp-sepharose was equilibrated with 50 mM potassium phosphate buffer, pH 7.5 and 200 µL of protein solution containing GspSA, YgiC, or YjfC was applied. The column was washed with 10 mL of phosphate buffer, then with solutions of NaCl with increasing concentrations ranging from 0.2 to 1 M prepared in the same buffer. Finally, the column was washed with a 1% solution of 2-mercaptoethanol. 1 mL fractions were collected and analyzed by SDS-PAGE.

ATPase activity assay

1 mL reaction mixture was prepared in 50 mM HEPES (pH 7.8) which contained 20 mM KCl, 10 mM $MgCl_2$, 1 mM phosphoenolpyruvate, 0.2 mM NADH, 20 units of pyruvate kinase (PK), 20 units of lactate dehydrogenase (LDH), 10 mM ATP and various amounts of GspSA, YgiC, and YjfC to a final concentration of 0.01 – 5 µM. The reactions were carried out at room temperature. All components except GspSA, YgiC, or YjfC were mixed in an Eppendorf tube with ATP added last. The reaction mixture was incubated at room temperature for 5 minutes, transferred to a quartz cuvette and used to blank the spectrophotometer at 340 nm. GspSA, YgiC, or YjfC was added to the reaction mixture, quickly mixed, and absorbance values were recorded at 10 second intervals for either 10 or 20 min using Agilent 8453 Diode Array Spectrophotometer (Agilent Technologies, Santa Clara, CA). The rate of ATP hydrolysis was calculated from the change of absorbance at 340 nm caused by the disappearance of NADH using the extinction coefficient of NADH 6220 M^{-1} cm^{-1}.

Ligase activity assays

The ligase activity of GspSA C59A, YgiC, and YjfC was tested with several combinations of substrates. The composition of each reaction mixture was as follows. *GSH and spermidine:* The reaction mixture contained 10 mM GSH,

10 mM spermidine, 15 mM ATP, 15 mM $MgCl_2$, and 5 µM of protein (GspSA C59A, YgiC, or YjfC) in a total reaction volume of 1 mL. *Cysteine and glutamate:* The reaction mixture contained 10 mM cysteine, 10 mM glutamate, 15 mM ATP, 15 mM $MgCl_2$, and 5 µM of protein (GspSA C59A, YgiC, or YjfC) in a total reaction volume of 1 mL. *Cysteine, glutamate, and glycine:* The reaction mixture contained 2 mM cysteine, 50 mM glutamate, 20 mM glycine, 15 mM ATP, 15 mM $MgCl_2$, and 5 µM of protein (GspSA C59A, YgiC, or YjfC) in a total reaction volume of 1 mL. γ-*Glutamylcysteine and glycine:* The reaction mixture contained 20 mM γ-Glu-Cys, 20 mM glycine, 10 mM ATP, 20 mM $MgCl_2$, and 10 µM of protein (GspSA C59A, YgiC, or YjfC) in a total reaction volume of 0.5 mL. γ-*Glutamylcysteine and spermidine:* The reaction mixture contained 20 mM γ-Glu-Cys, 20 mM spermidine, 10 mM ATP, 20 mM $MgCl_2$, and 10 µM of protein (GspSA C59A, YgiC, or YjfC) in a total reaction volume of 0.5 mL. All reactions were carried out at 37 °C in 40 mM PIPES buffer, pH 6.8. The reactions were initiated by the addition of ATP. 100 µL aliquots were removed from the reaction at several time points for analysis. Each aliquot was heated at 90 °C for 3 min and the protein precipitate was removed by centrifugation for 5 min. 10 µL of supernatant was mixed with 60 µL of 200 mM Tris buffer, pH 8.0 and 10 µL of 5 mM DTNB to produce thiol derivatives which were analyzed by HPLC as described above.

Creating gene disruptions

E. coli gene knockouts were performed according to the Wanner [11] protocol as follows. *E. coli* BW25113-pKD46 cells were made electrocompetent by growing in 100 mL of SOB media with 1 mM arabinose at 30 °C to OD_{600} of 0.6 and washing several times with decreasing volumes of 10% sterile glycerol. After the final wash, cells were resuspended in 150 µL 10% (v/v) glycerol solution, flash-frozen, and stored at -80 °C. To generate gene disruptions, first, PCR fragments (~1.1 kb) were prepared using template plasmids pKD3 or pKD13 to create antibiotic resistant cassettes for chloramphenicol or kanamycin respectively. Primers used to create the *yjfC*/chloramphenicol fragments included the forward primer: ACAACGTTCCT-GTGCGACGGGATCTGGACCAGATCGCCGC<u>GTG-TAGGCTGGGCTGCTTC</u> and the reverse primer:

ATGTAATGCGGAACGAAACGTGAGGTGTCTTTG-GTGATCACATATGAATATCCTCCTTG. Primers used to create the *ygiC*/kanamycin fragments included forward primer: GAAAGAGTCAG-TATTACCGAGCGCCCGGACTGGCGTGAGGT-GTGTAGGCTGGAGCTGCTTC and the reverse primer: AACCCAAGCTTATTCAACAAAAATATGTG-GATAAAACCGATTCCGGGGATCCGTCGACC.

Each primer consisted of 40 bases homologous to the disrupted gene and the 20 bases of the priming sequence of a template plasmid that are underlined.

Electrocompetent cells were transformed with the gel-purified PCR fragments by electroporation. 100 μL of electrocompetent cells along with 1.5 μL (about 90 ng) PCR fragment were placed in the electroporation cuvettes with a 0.1 cm gap (BTX Cuvettes, BTX Harvard Apparatus, Holliston, MA). Electroporation was performed using Bio-Rad Gene Pulser II (Bio-Rad Laboratories, Inc., Hercules, CA) with charging voltage 1.4 kV for 5 ms. Immediately after electroporation, the cells were suspended in 900 μL SOC media and incubated at 37 °C for about 1 hour. 400 μL of the culture were plated on selective LB agar and grown overnight at 30 °C. Success of gene elimination was confirmed by colony PCR using the following primers: *ygiC:* forward: GCCTGAAAGTGGTGGAAAAA and reverse GTGAGATGCAGCAACTGGAA; *yjfC:* forward: GCGCGGTAAAAGATGAAGAG and reverse: GTGTCCTGACAGCAGCAAAA.

To eliminate antibiotic resistance cassette, cells that had successful gene disruptions were made electrocompetent by the method described earlier. 100 μL of electrocompetent cells were transformed with the 1.5 μL of the helper plasmid pCP20 (20 ng/μL) by electroporation in cuvettes with a 0.2 cm gap (Bio-Rad). The cells were grown at 30 °C on LB agar selective medium. Cell colonies were then grown once nonselectively at either 37 °C or 43 °C overnight on LB agar medium to promote thermal induction. Individual colonies were then plated on LB agar nonselective and LB agar selective medium and grown at 30 °C to confirm the successful elimination of antibiotic resistance.

Test for G-Sp formation in E. coli

Wild type *E. coli* and knockout strains were grown in LB media overnight at 37 °C, diluted 100 fold into 3 mL fresh media in 15 mL Falcon tubes and grown to OD_{600} about 1. At this time the culture tubes were filled with the LB media to the top and tightly capped, the cultures were incubated with shaking at 37 °C for 24 hrs. Cells were harvested by centrifugation at 6000 x *g* for 10 min in the same tubes, the pellets were resuspended in 100 μL of lysozyme (400 μg/ mL) and incubated at room temperature for 5 min. 10 μL of 5 mM DTNB was added to the mixture and allowed to react for 2 min. The mixtures were then heated at 90 °C for 5 min to precipitate the proteins, and the precipitate was removed by 10 min centrifugation using a table-top centrifuge at maximum speed. The supernatant was analyzed by HPLC for thiol content as described previously.

Results and discussion

In our studies we cloned the genes and expressed and purified *E. coli* proteins YgiC and YjfC to compare them with the properties of glutathionylspermidine synthetase/amidase. *ygiC* and *yjfC* genes were PCR- amplified from *E. coli* genomic DNA and cloned into pET20b vector. The proteins were expressed in BL21(DE3) cells upon induction with IPTG. Corresponding proteins were purified using two chromatographic columns - anion-exchange DEAE-cellulose and gel-filtration. SDS-PAGE analysis demonstrated the presence of proteins with apparent molecular mass of 45 kDa. These proteins were used in further studies of enzymatic activities and interaction with G-Sp.

The sequence alignment

The synthetase domain of GspSA and both YgiC and YjfC proteins are assigned to the ATP-grasp superfamily of proteins. The enzymes of this family catalyze the ATP-dependent ligation of a carboxyl group and a nucleophile, usually an amino group or thiol. Even though proteins with ATP-grasp domain possess a great variety of substrate specificity; they share a number of structural motifs important for binding ATP. The similarities of GspSA with human glutathione synthetase allowed identifying the signature amino acid residues of ATP-grasp domain that are responsible for the ATP binding [7]. We mapped the amino acid residues that were found in the active site of the synthetase domain of GspSA [7] and compared them with the sequences of YgiC and YjfC proteins. The

```
GSP_ECOLI    DDTTILGWMIQTEDTEYSLPQPEIAGELLKISGARLENKGQFDGKWLDEKDPLQNAYVQA 240
YGIC_ECOLI   ----------------------------MERVSITERPDWREKAHEYG----FNFHTM 26
YJFC_ECOLI   ----------------------------MLRHNVPVRRDLDQIAADNG----FDFHII 26
                                      :     :  : :       :        :

GSP_ECOLI    NGQVINQDPYHYYTITESAEQELIKATNELHLMYLHATDKVLKDDNLLALFDIPKILWPR 300
YGIC_ECOLI   YGEPYWCEDAYYKLTLAQVEK-LEEVTAELHQMCLKVVEKVIASDELMTKFRIPKHTWSF 85
YJFC_ECOLI   DNEIYWDESRAYRFTLRQIEEQIEKPTAELHQMCLEVVDRAVKDEEILTQLAIPPLYWDV 86
              .:     :    *     .*: : : * *** * *...::.: .::::: : **   *

GSP_ECOLI    LRLSWQRRRHHMITGRMDFCMDERG-LKVYEYNADSASCHTEAGLILERWAEQGYKGNGF 359
YGIC_ECOLI   VRQSWLTHQPSLYS-RLDLAWDGTGEPKLLENNADTPTSLYEAAFFQWIWLEDQLNAGNL 144
YJFC_ECOLI   IAESWRARDPSLYG-RMDFAWCGNAPVKLLEYNADTPTSLYESAYFQWLWLEDARRSGII 145
              :  **   :   :    *:*:.    . *: * ***:.:.  *:. :   * *: ... :

GSP_ECOLI    NP--------AEGLINELAGAWKHSRARPFVHIMQDKDIEENYHAQFMEQALHQAGFETR 411
YGIC_ECOLI   PEGSDQFNSLQEKLIDRFVELREQYGFQLLHLTCCRDTVEDRGTIQYLQDCATEAEIATE 204
YJFC_ECOLI   PRDADQYNAIQERLISRFSEL---YSREPFYFCCCQDTDEDRSTVLYLQDCAQQAGQESR 202
                      * **..:          . :       . *:.   ::::. :*    :.

GSP_ECOLI    ILRGLDELGWDAAGQLIDGEGRLVNCVWKTWAWETAFDQIREVSDREFAAVPIRTGHPQN 471
YGIC_ECOLI   FLY-IDDIGLGEKGQFTDLQDQVISNLFKLYPWEFMLREMFST----------------- 246
YJFC_ECOLI   FIY-IEDLGLGVGGVLTDLDDNVIQRAFKLYPLEWMMRDDNGP----------------- 244
              :: ::::*  .  * : * :..::. :* :. *  : :

GSP_ECOLI    EVRLIDVLLRPEVLVFEPLWTVIPGNKAILPILWSLFPHHRYLLDTDFTVNDE--LVKTG 529
YGIC_ECOLI   ------KLEDAGVRWLEPAWKSIISNKALLPLLWEMFPNHPNLLPAYFAEDDH--PQMEK 298
YJFC_ECOLI   ------LLRKRREQWVEPLWKSILSNKGLMPLLWRFFPGHPNLLASWFDGEKPQIAAGES 298
                    *         .** *. * .**.::*:** :** *  ** : *   :.

GSP_ECOLI    YAVKPIAGRCGSNIDLVSHHEEVLDKTSGKFAEQKNIYQQLWCLPKVDGKYIQVCTFTVG 589
YGIC_ECOLI   YVVKPIFSREGANVSIIE-NGKTIEAAEGPYGEEGMIVQQFHPLPKFGDSYMLIGSWLVN 357
YJFC_ECOLI   YVRKPIYSREGGNVTIFDGKNNVVDHADGDYADEPMIYQAFQPLPRFGDSYTLIGSWIVD 358
             *. *** .* *.*: :.. : :.:: :.* :.:: * * : **:...* : :: *.

GSP_ECOLI    GNYGGTCLRGDESLVIKKESDIEPLIVVKK 619
YGIC_ECOLI   DQPAGIGIREDRALITQDMSRFYPHIFVE- 386
YJFC_ECOLI   DEACGMGIREDNTLITKDTSRFVPHYIAG- 387
              .:  *   :* *.:*: :. * : *   ..
```

Figure 1. Sequence alignment of GspSA synthetase domain, YgiC, and YjfC. The star (*) indicates fully conserved residue, colon (:) indicates conservation between groups of strongly similar properties, single dot (•) indicates conservation between groups of weakly similar properties. The shaded amino acid residues were found to participate in the binding of ATP by GspSA [7]. The amino acid residues in boxes represent conserved residues of ATP-grasp fold, R316, K498, K533. The amino acid residues highlighted in yellow are essential residues for substrates binding by GspSA as determined in [7].

sequence alignment of GspSA with YgiC and YjfC indicates that residues involved in the binding of triphosphate group and the formation of hydrophobic pocket for binding adenine of ATP are quite conserved in all three proteins (**Figure 1**). Distinct ATP-grasp domain residues Arg 316, Lys 498, and Lys 533 of GspSA are present in the structures of YgiC and YjfC proteins. Yet, there are more deviations in the nature of the amino acid residues that form the rest of the active site suggesting that YgiC and YjfC proteins may not use GSH and spermidine as substrates.

ATPase activity

The ATPase activity of GspSA, YgiC and YjfC, or their ability to hydrolyze ATP to ADP and phosphate, was tested by coupled assay. In this assay the hydrolysis of ATP to ADP by an enzyme is linked to two other enzymatic reactions. In the first reaction, ADP is utilized by pyruvate kinase which converts phosphoenolpyruvate to pyruvate. In the next step, LDH uses pyruvate to make lactate with concurrent oxidation of NADH to NAD$^+$. The reaction is followed by the decrease of the absorbance at 340 nm due to
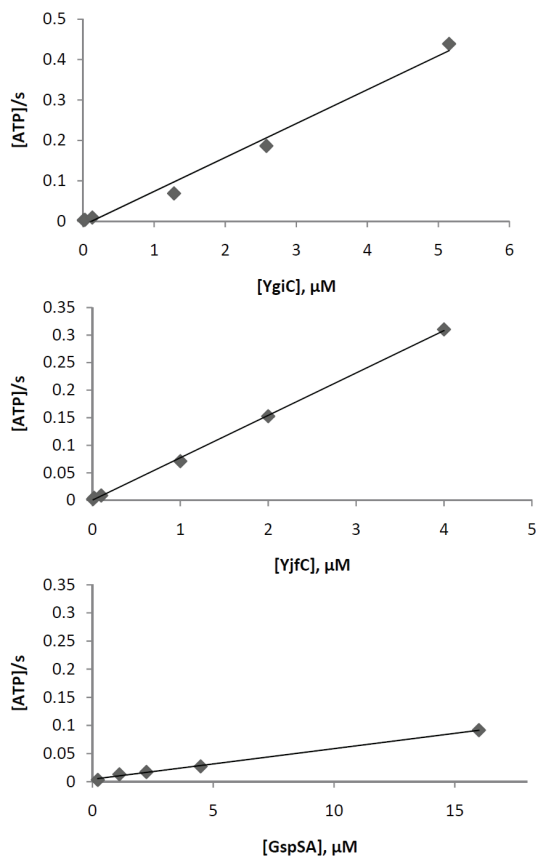
Figure 2. ATPase activity assay. The ATPase activities of GspSA, YgiC, and YjfC were determined by enzyme PK/LDH coupled assay. The change of absorbance at 340 nm due to oxidation of NADH was used to calculate the rate of hydrolysis of ATP. The initial concentration of ATP was 10 mM, the concentrations of YgiC and YjfC were varied between 0.01 and 5 µM, GspSA was varied between 0.2 and 16 µM.

the loss of NADH. First, we tested the ATPase activity of GspSA in the presence and absence of GSH. In the formation of a G-Sp conjugate, GspSA utilizes ATP to activate the carboxyl group of the glycine residue of GSH for the reaction with the N1-amino group of spermidine. In our experiments, the initial concentration of ATP was maintained constant at 10 mM. The concentration of YgiC and YjfC proteins was varied between 0.01 and 5 µM while GspSA concentration was varied between 0.2 and 16 µM. GspSA showed modest ATPase activity in the absence of other substrates. The availability and the amount of glutathione had little effect on the rate of the hydrolysis of ATP. However, when spermidine was added in the presence of GSH, the rate of the reaction significantly increased (data not shown). YgiC and YjfC proteins demonstrated comparable ability to

hydrolyze ATP to ADP. The ATPase activity of both Y-proteins was greater than the activity of GspSA under conditions used in our assays. **Figure 2** shows the linear relationship between the rate of ATP hydrolysis, which correlates with the rate of NADH oxidation, and protein concentration. Addition of glutathione, spermidine, or both to the reaction mixture did not change the rate of ATP hydrolysis for either YgiC or YjfC.

Ligase activity

Both Y-proteins are shorter than GspSA as they lack the amidase domain. Due to this, we tested only their synthetase potential. GspSA catalyzes the formation of a peptide bond between GSH and spermidine. The non-ribosomal formation of a peptide bond is an essential physiological reaction that is catalyzed by carbon-nitrogen ligases, also known as amido-ligases [12].

Several amido-ligases participate in the metabolism of glutathione and polyamines. Bacteria such as *E. coli* and *Pseudomonas* spp. utilize γ-glutamylation pathway for polyamine catabolism [13, 14]. In this pathway the enzymes with the amido-ligase activity conjugate polyamines, including spermidine, with the γ-carboxyl group of glutamate to prepare them for further deamination. In most organisms two amido-ligases contribute to the synthesis of GSH catalyzing two consecutive ATP-dependent condensations. First, γ-glutamyl-cysteine synthetase joins the γ-carboxyl group of glutamate with the α-amino group of cysteine. In the next step, the dipeptide is linked to the α-amino group of glycine by glutathione synthetase. Despite visible similarities in the reaction mechanisms of these two steps, the enzymes that catalyze them possess quite different structural architecture. γ-glutamyl-cysteine synthetase has a glutamine synthetase-like domain, while glutathione synthetase belongs to the ATP-grasp superfamily. Interestingly, in some bacteria these two functions are combined within one polypeptide chain capable of catalyzing both synthetic steps [15]. These polypeptides are organized into two domains responsible for individual reactions. It was shown that the domains involved in the catalysis of the ultimate step of GSH synthesis usually have an ATP-grasp fold [16].

There is very little sequence similarity between *E. coli* glutathione synthetase and YgiC, YjfC,
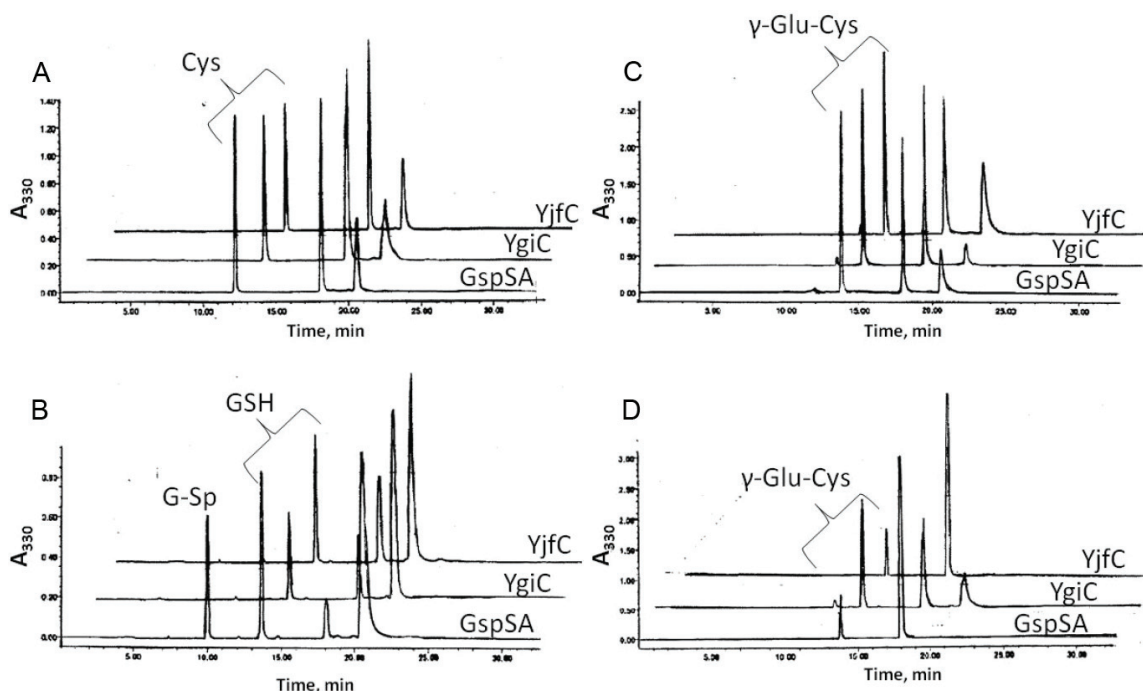
Figure 3. Ligase activity assay of GspSA, YgiC, and YjfC. The substrate composition for the reactions was: A. cysteine and glutamate, B. GSH and spermidine, C. γ-glutamylcysteine and spermidine, D. γ-glutamylcysteine and glycine. The reactions were stopped 3 hours after the addition of ATP by brief heating, derivatized with DTNB, and the thiol content of the reaction mixtures was analyzed by reverse-phase HPLC. Labeled peaks represent DTNB derivatives of corresponding thiols. Peaks at 18 min and 21 min are 2-nitro-5-thiobenzoate and the excess of DTNB correspondently.

and the synthetase domain of GspSA although all four proteins are assigned to ATP-grasp superfamily [17]. There are known examples of proteins with a similar structural fold, but quite different sequences, being capable of catalyzing the same reactions. *E. coli* protein YbdK is only remotely similar to γ-glutamyl-cysteine synthetase by sequence but belongs to the same structural family [18]. This protein has amido-ligase activity to couple the γ-carboxyl group of glutamate with the amino group of cysteine. The rate of this reaction was fairly slow and it remains unclear whether this activity of YbdK is physiologically relevant.

Considering possible involvement of proteins in glutathione metabolism, amido-ligase activities of GspSA, YgiC, and YjfC were analyzed using several combinations of nucleophiles and carboxyl group donors. The proteins were tested for their ability to conjugate GSH with spermidine, glutamate with cysteine in the presence or absence of glycine, and γ-glutamylcysteine with glycine or spermidine. Each protein was incubated with different com-

binations of the potential substrates in the presence of ATP. The reactions were stopped at various time points by brief heating, and the mixtures were examined for product formation. Each reaction included a substrate that contains a thiol group. This allowed us to examine the composition of the reaction mixtures using a C-18 reverse-phase HPLC column after derivatization of thiols with DTNB. This reagent reacts with sulfhydryl groups producing compounds that absorb light at 330 nm and could be easily observed by a UV detector. The only new product was formed in the reaction of GSH and spermidine in the presence of GspSA. Other combinations of proteins and substrates did not yield any new compounds based on the absence of the new peaks on the HPLC chromatograms (**Figure 3**).

Binding to G-Sp immobilized column

Interactions of YgiC and YjfC with G-Sp were tested using an affinity matrix with immobilized conjugate. To create this matrix, the glutathione-spermidine conjugate was synthesized
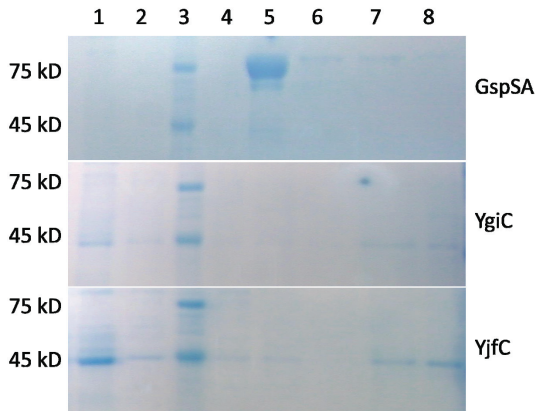
Figure 4. SDS-PAGE analysis of fractions collected from G-Sp affinity column. Lanes: 1 – fraction containing proteins that did not bind to the column; 2 – initial wash with 50 mM potassium phosphate buffer, pH 7.5; 3 – molecular marker; 4 – 0.2 M NaCl, 50 mM potassium phosphate buffer, pH 7.5; 5 – 0.6 M NaCl, 50 mM potassium phosphate buffer, pH 7.5; 6 - 1 M NaCl, 50 mM potassium phosphate buffer, pH 7.5; 7, 8 - 1% 2-mercaptoethanol, 50 mM potassium phosphate buffer, pH 7.5.

enzymatically by GspSA C59A. In order to increase the yield of G-Sp, we used a mutant enzyme in which amidase activity is inhibited by the replacement of an important catalytic residue, cysteine-59, with alanine [19]. The conjugate was purified by ion-exchange chromatography column and coupled with the activated epoxy-sepharose at pH 10. At this pH both the thiol and amino group of the conjugate are able to react with epoxide. Binding through different functional groups exposes various parts of the G-Sp molecule for contact with proteins and provides a greater variety of protein-ligand interactions.

Diluted solutions of GspSA, YgiC, and YjfC proteins were loaded on the G-Sp-sephadex column, which was washed with phosphate buffer to remove all unbound material. Bound proteins were eluted from the matrix with increasing concentration of sodium chloride up to 1 M. Finally, the column was washed with a reducing agent, 2-mercaptoethanol, to remove any protein that formed disulfide bonds with the immobilized conjugate. Fractions collected during elution were analyzed by SDS-PAGE (**Figure 4**). GspSA binds to immobilized G-Sp with relatively high affinity and elutes with about 0.6 M sodium chloride solution. YgiC protein did not bind significantly to the affinity matrix as most of it

was eluted within the initial wash. A considerable amount of YjfC was also eluted with the initial wash. Both Y- proteins formed disulfide bonds with the immobilized conjugate since small quantities of the protein stayed bound to the matrix at high concentrations of sodium chloride but were eluted by 2-mercaptoethanol.

Ability of bacterial strains to synthesize G-Sp

Attempts to assess the influence of the elimination of *ygiC* and *yjfC* genes on G-Sp production were described recently [20]. Both genes were eliminated from *E. coli* genome together with neighboring genes. Since in that study G-Sp was not detected even for wild type cells under growth conditions that were used, the effect of the elimination of corresponding genes on G-Sp production was not clear. To form the conjugate, bacteria should be grown to the stationary phase under anaerobic conditions [21]. We grew the cell cultures to the stationary phase in tubes completely filled with media which causes exhaustion of the oxygen available to bacteria, and analyzed the thiol content of the cell lysates. Under these growth conditions we were able to detect the presence of the glutathione-spermidine conjugate in the wild type *E. coli* cells. *gsp*, *ygiC*, and *yjfC* knockout *E. coli* strains were created using the method of Wanner [11] and antibiotic cassettes were eliminated as described in Materials and Methods. There were no visible differences between wild type and mutant strains growth in LB media. Δ*gsp*, Δ*ygiC,* and Δ*yjfC* strains were tested for their ability to produce G-Sp (**Figure 5**). Our results demonstrate that *E. coli* cells lacking *ygiC* and *yjfC* genes produce the conjugate in similar quantities as the wild type cells grown under the same conditions. G-Sp conjugate was not detected in the *gsp*-knockout strain of *E. coli*. This means that GspSA is required for the G-Sp synthesis and neither YgiC nor YjfC is capable of replacing this enzyme.

In the *E. coli* genome the *ygiC* gene is located next to *tolC.* The *tolC* gene product is an outer membrane protein involved in the efflux of some hydrophobic and amphiphilic molecules [22]. It was shown that *ygiC* and *tolC* genes are co-transcribed [20]. Based on these findings, it was hypothesized that the functions of YgiC and TolC could be related. Additionally, *ygiB*, a
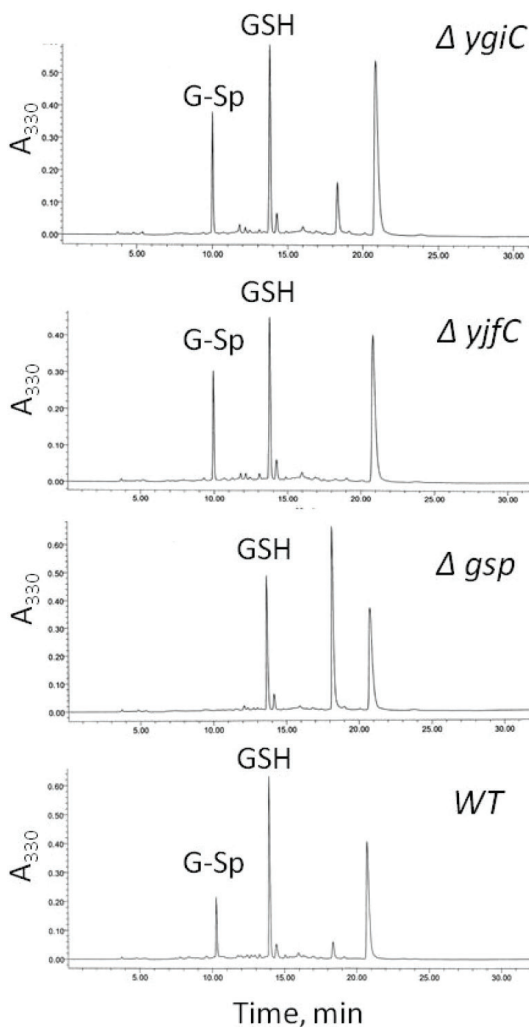
Figure 5. Analysis of thiol content in *E. coli* cells. Analysis of the thiol content of wild type and *gsp*, *yjfC,* and *ygiC* gene knockout *E. coli* strains grown in LB media to stationary phase under anaerobic conditions was performed using HPLC. Labeled peaks represent DTNB derivatives of G-Sp and GSH. Peak immediately followed GSH (14 min) corresponds to DTNB derivative of γ-glutamylcysteine. Peaks at 18 min and 21 min are 2-nitro-5-thiobenzoate and the excess of DTNB correspondently.

neighboring gene that is also co-transcribed with *ygiC* and *tolC*, is a predicted periplasmic protein. All together, these observations led to conclusions about the involvement of YgiC and YgiB proteins in the protection of the inner membrane of bacteria [20]. Our study indicates that since YgiC cannot synthesize G-Sp or even bind the conjugate, it is unlikely, therefore, that this protein participates in glutathionylspermidine synthesis. However, the ATPase activity of

the protein suggests that it may be involved in the transport system associated with TolC.

In conclusion, we demonstrated that similar to GspSA, proteins YgiC and YjfC possess ATPase activity and are capable of catalyzing the hydrolysis of ATP to ADP and phosphate. This activity does not require the presence of additional substrates for the ligation reaction. In contrast to GspSA, neither YgiC nor YjfC is able to produce the conjugate of glutathione and spermidine in the reactions carried out under the same conditions. Strains deficient in GspSA do not synthesize G-Sp under anaerobic conditions; however, the elimination of *ygiC* or *yjfC* genes does not affect the formation of G-Sp. All together, our results indicate that YgiC and YjfC are not glytathionylspermidine synthetases as predicted from their amino acid sequence analysis.

Acknowledgement

Address correspondence to: Dr. Nina V Stourman, Department of Chemistry, Youngstown State University, One University Plaza, Youngstown, OH, 44555, USA. E-mail: nvstourman@ysu.edu

References

[1]   Karp PD, Keseler IM, Shearer A, Latendresse M, Krummenacker M, Paley SM, Paulsen I, Collado-Vides J, Gama-Castro S, Peralta-Gil M, Santos-Zavaleta A, Peñaloza-Spínola MI, Bonavides-Martinez C, Ingraham J. Multidimensional annotation of the *Escherichia coli* K-12 genome. Nucleic Acids Res 2007; 35: 7577-90.

[2]   Keseler IM, Collado-Vides J, Santos-Zavaleta A, Peralta-Gil M, Gama-Castro S, Muñiz-Rascado L, Bonavides-Martinez C, Paley S, Krummenacker M, Altman T, Kaipa P, Spaulding A, Pacheco J, Latendresse M, Fulcher C, Sarker M, Shearer AG, Mackie A, Paulsen I, Gunsalus RP, Karp PD. EcoCyc: a comprehensive database of *Escherichia coli* biology. Nucleic Acids Res 2011; 39: D583-90.

[3]   Galperin MY, Koonin EV. Divergence and convergence in enzyme evolution. J Biol Chem 2012; 287: 21-8.

[4]   Bollinger JM Jr, Kwon DS, Huisman GW, Kolter R, Walsh CT. Glutathionylspermidine metabo-

lism in *Escherichia coli*. Purification, cloning, overproduction, and characterization of a bifunctional glutathionylspermidine synthetase/amidase. J Biol Chem 1995; 270: 14031-41.

[5] Kwon DS, Lin CH, Chen S, Coward JK, Walsh CT, Bollinger JM Jr. Dissection of glutathionylspermidine synthetase/amidase from *Escherichia coli* into autonomously folding and functional synthetase and amidase domains. J Biol Chem 1997; 272: 2429-36.

[6] Tabor H, Tabor CW. Isolation, characterization, and turnover of glutathionylspermidine from *Escherichia coli*. J Biol Chem 1975; 250: 2648-54.

[7] Pai CH, Chiang BY, Ko TP, Chou CC, Chong CM, Yen FJ, Chen S, Coward JK, Wang AH, Lin CH. Dual binding sites for translocation catalysis by *Escherichia coli* glutathionylspermidine synthetase. EMBO J 2006; 25: 5970-82.

[8] Galperin MY, Koonin EV. A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. Protein Sci 1997; 6: 2639-43.

[9] Fawaz MV, Topper ME, Firestine SM. The ATP-grasp enzymes. Bioorg Chem 2011; 39: 185-91.

[10] Polekhina G, Board PG, Gali RR, Rossjohn J, Parker MW. Molecular basis of glutathione synthetase deficiency and a rare gene permutation event. EMBO J 1999; 18: 3204-13.

[11] Datsenko KA, Wanner BL. One-step Inactivation of Chromosomal Genes in *Escherichia coli* K-12 Using PCR Products. Proc Natl Acad Sci USA 2000; 97: 6640-6645.

[12] Iyer LM, Abhiman S, Maxwell Burroughs A, Aravind L. Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. Mol Biosyst 2009; 5: 1636-60.

[13] Kurihara S, Oda S, Kato K, Kim HG, Koyanagi T, Kumagai H, Suzuki H. A novel putrescine utilization pathway involves gamma-glutamylated intermediates of *Escherichia coli K-12*. J Biol Chem 2005; 280: 4602-8.

[14] Yao X, He W, Lu CD. Functional characterization of seven γ-Glutamylpolyamine synthetase genes and the bauRABCD locus for polyamine and β-Alanine utilization in Pseudomonas aeruginosa PAO1. J Bacteriol 2011; 193: 3923-30.

[15] Janowiak BE, Hayward MA, Peterson FC, Volkman BF, Griffith OW. Gamma-glutamylcysteine synthetase-glutathione synthetase: domain structure and identification of residues important in substrate and glutathione binding. Biochemistry 2006; 45: 10461-73.

[16] Stout J, De Vos D, Vergauwen B, Savvides SN. Glutathione biosynthesis in bacteria by bifunctional GshF is driven by a modular structure featuring a novel hybrid ATP-grasp fold. J Mol Biol 2012; 416: 486-94.

[17] Fan C, Moews PC, Shi Y, Walsh CT, Knox JR. A common fold for peptide synthetases cleaving ATP to ADP: glutathione synthetase and D-alanine:d-alanine ligase of *Escherichia coli.* Proc Natl Acad Sci USA 1995; 92: 1172-6.

[18] Lehmann C, Doseeva V, Pullalarevu S, Krajewski W, Howard A, Herzberg O. YbdK is a carboxylate-amine ligase with a gamma-glutamyl:cysteine ligase activity: crystal structure and enzymatic assays. Proteins 2004; 56: 376-83.

[19] Lin CH, Kwon DS, Bollinger JM Jr, Walsh CT. Evidence for a glutathionyl-enzyme intermediate in the amidase activity of the bifunctional glutathionylspermidine synthetase/amidase from *Escherichia coli*. Biochemistry 1997; 36: 14930-8.

[20] Dhamdhere G, Zgurskaya HI. Metabolic shutdown in *Escherichia coli* cells lacking the outer membrane channel TolC. Mol Microbiol 2010; 77: 743-54.

[21] Smith K, Borges A, Ariyanayagam MR, Fairlamb AH. Glutathionylspermidine metabolism in *Escherichia coli.* Biochem J 1995; 312: 465-9.

[22] Sharff A, Fanutti C, Shi J, Calladine C, Luisi B. The role of the TolC family in protein transport and multidrug efflux. From stereochemical certainty to mechanistic hypothesis. Eur J Biochem 2001; 268: 5011-26.