Molecular Basis of Formaldehyde Detoxification

CHARACTERIZATION OF TWO S-FORMYLGLUTATHIONE HYDROLASES FROM ESCHERICHIA COLI, FrmB AND YeiG

Received for publication, February 1, 2006, and in revised form, March 22, 2006. Published, JBC Papers in Press, March 27, 2006, DOI 10.1074/jbc.M600996200

Claudio F. Gonzalez, Michael Proudfoot, Greg Brown, Yuriy Korniyenko, Hirotada Mori, Alexei V. Savchenko, and Alexander F. Yakunin

From the 1 Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario M5G 1L6, Canada and the 2 Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan

The *Escherichia coli* genes *frmB* (yaiM) and *yeiG* encode two uncharacterized proteins that share 54% sequence identity and contain a serine esterase motif. We demonstrated that purified FrmB and YeIG have high carboxylesterase activity against the model substrates, *p*-nitrophenyl esters of fatty acids (C2–C6) and *α*-naphthyl acetate. However, both proteins had the highest hydrolytic activity toward S-formylglutathione, an intermediate of the glutathione-dependent pathway of formaldehyde detoxification. With this substrate, both proteins had similar affinity (Km = 0.41–0.43 mM), but FrmB was almost 5 times more active. Alanine replacement mutagenesis of YeIG demonstrated that Ser145, Asp233, and His256 are absolutely required for activity, indicating that these residues represent a serine hydrolase catalytic triad in this protein and in other S-formylglutathione hydrolases. This was confirmed by inspecting the crystal structure of the *Saccharomyces cerevisiae* S-formylglutathione hydrolase YJG8 (Protein Data Bank code 1pv1), which has 45% sequence identity to YeIG. In *E. coli* cells, the expression of *frmB* was stimulated 45–75 times by the addition of formaldehyde to the growth medium, whereas YeIG was found to be a constitutive enzyme. The simultaneous deletion of both *frmB* and *yeiG* genes was required to increase the sensitivity of the growth of *E. coli* cells to formaldehyde, suggesting that both FrmB and YeIG contribute to the detoxification of formaldehyde. Thus, FrmB and YeIG are S-formylglutathione hydrolases with a Ser-His-Asp catalytic triad involved in the detoxification of formaldehyde in *E. coli*.

Formaldehyde is an extremely reactive chemical producing covalently cross-linked complexes with proteins and nucleic acids (1–3). It is a common environmental contaminant found in many industrial and medical products, as well as being endogenously produced in all living organisms as a result of metabolism (methionine, histamine, methanol, and methylamine), spontaneous dissociation of 5,10-methylene tetrahydrofolate, or oxidative demethylation of DNA and RNA (4–9). To prevent the lethal and mutagenic effects of formaldehyde, several repair mechanisms have evolved. Detoxification of formaldehyde can be carried out by enzymes like formaldehyde dismutase, methylformate synthase, or glutathione-independent formaldehyde dehydrogenase (10–12). However, these enzymes have been found in only a limited number of organisms, whereas a glutathione-dependent repair system appears to be widespread in nature and has been found in most prokaryotes (except for archaea) and all eukaryotes (7, 13–15).

In this process (shown in Scheme 1), formaldehyde spontaneously reacts with GSH to produce S-hydroxymethylglutathione, which is then oxidized by formaldehyde dehydrogenase to S-formylglutathione. Finally, this compound is hydrolyzed to formate and GSH by S-formylglutathione hydrolase (FGH).2 These functionally related enzymes, formaldehyde dehydrogenase and FGH, also show genetic linkage, since their genes are adjacent in many bacterial genomes or even fused in some eukaryotes (16).

Formaldehyde dehydrogenases (class III alcohol dehydrogenases) were extensively characterized both structurally and biochemically (17–20), whereas FGHs have received far less attention. Only three eukaryotic FGHs (from the human liver, *Saccharomyces cerevisiae*, and *Arabidopsis thaliana*) (21–23) have been purified and partially characterized, but the homologous enzyme from bacteria has not yet been purified. The FGH enzymes from yeasts and *A. thaliana* are not strictly specific to S-formylglutathione and also show significant carboxylesterase activity against model substrates *α*-naphthyl acetate and *p*-nitrophenyl acetate (22, 23). Although the sequence analysis suggested that FGHs are serine hydroxylases, it has been proposed that FGHs are cysteinyl hydroxylases, because the *A. thaliana* FGH was highly sensitive to the inhibition by *N*-ethylmaleimide and heavy metals (23). In humans, this enzyme (identical to esterase D) was highly expressed in liver and kidney, and its expression was stimulated by phenobarbital treatment (24). Genetic polymorphism of esterase D was discovered in various human populations, and the reduced enzymatic activity of this enzyme was found to be associated with the susceptibility to several pathological conditions (tumor, liver cirrhosis, retinoblastoma, obesity, and autism) (25–28).

Much less information is available about the prokaryotic FGHs. In *Paracoccus denitrificans*, the gene encoding FGH (*fghA*) has been identified and shown to be homologous to the human esterase D (15). The recent work on *Escherichia coli* suggested that the three-gene operon *frmR, frmA, frmB* is likely to encode a complete pathway for degradation of formaldehyde (29). The expression of *frmA* and *frmB* in *E. coli* cells was induced by formaldehyde 20–100-fold over the level of uninduced cells. The *frmR* gene is likely to encode the transcriptional repressor of this operon, whereas *frmA* and *frmB* were predicted to encode S-hydroxymethylglutathione dehydrogenase and S-formylglutathione hydrolase, respectively (29).

The abbreviations used are: FGH, S-formylglutathione hydrolase; *p*NP, *p*-nitrophenyl; MOPS, 4-morpholinopropanesulfonic acid; r.m.s., root mean square.

8 This work was supported by Genome Canada through the Ontario Genomics Institute, National Institutes of Health Grant GM62414-01, CREST (Japan Science and Technology), and the Inamori Foundation (to H. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
9 The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1 and S2.
1 To whom correspondence should be addressed. Tel.: 416-946-0075; Fax: 416-978-8528; E-mail: a.iakounine@utoronto.ca.

2 The abbreviations used are: FGH, S-formylglutathione hydrolase; *p*NP, *p*-nitrophenyl; MOPS, 4-morpholinopropanesulfonic acid; r.m.s., root mean square.
In this work, we have purified and biochemically characterized the *E. coli* FrmB and its paralog, YeiG, which shares 54% amino acid sequence identity with FrmB. Both proteins were shown to be active FGHs with a Ser-His-Asp catalytic triad. Our results demonstrated that YeiG is a constitutive enzyme in *E. coli*, whereas the expression of FrmB is greatly stimulated by formaldehyde and that both proteins are involved in formaldehyde detoxification in vivo.

**MATERIALS AND METHODS**

**Chemicals and Strains**—All chemicals were purchased from Sigma. S-Formylglutathione was synthesized from formylthioglycolate and glutathione as described previously (30). Nitrosogluthathione was prepared according to Schapiro et al. (31). The primers, strains, and plasmids used in this work are listed in supplemental Table S1.

**Cloning, Overexpression, and Purification of FrmB and YeiG**—The *frmB* open reading frame was PCR-amplified using chromosomal DNA of the *E. coli* BW25113 strain and the cloning primers *frmB*-Clon-F and *frmB*-Clon-R (supplemental Table S2). The restriction sites BamHI and Ndel were added to the primers and used to directionally clone the PCR product into pET15b (AmpR) as previously described (32). The recombinant plasmid was transformed into the *E. coli* BL21 (DE3) strain for overexpression. The YeiG open reading frame was obtained from the Genobase collection (33). The cells were grown on LB medium at 37 °C to an *A*_{600} of ~0.6, and the protein expression was induced by the addition of 0.7–1 mM isopropyl 1-thio-β-d-galactopyranoside. After induction, the cells were harvested at 16 °C overnight. The cells were collected by centrifugation and resuspended in binding buffer (500 mM NaCl, 50 mM Tris-HCl (20 mM), pH 8.0). When FrmB was dialyzed, 0.5 mM of TCEP was added to the buffer to prevent protein precipitation.

**Enzymatic Assays—**Esterase activity was measured spectrophotometrically using 2–10 μg/ml of purified enzyme at 37 °C. Carboxylesterase activity was measured using *p*-nitrophenyl esters of various fatty acids (C2–C6) (34). The amount of *p*-nitrophenol released was quantified at *A*_{400} using the extinction coefficient ε = 16,300 M^{-1} cm^{-1}. Stock solutions of *p*NP-esters were prepared using acetonitrile as a solvent. Aryl esterase activity was determined by measuring the *α*-naphthol released from *α*-naphthyl acetate and *α*-naphthyl butyrate at 310 nm (ε = 3000 M^{-1} cm^{-1}). Stock solutions (100 mM) of *α*-naphthyl esters were prepared in acetonitrile. Thioesterase activity was assayed using palmitoyl-CoA as previously described (35). Enzymatic activity against the glutathione derivatives (S-formylglutathione, S-lactoylglutathione, S-(1,2-dicarboxyethyl)glutathione, S-hexylglutathione, and S-(4-nitrobenzyl)glutathione) was followed continuously at 240 nm (21). The reaction rate was calculated using the following extinction coefficients: ε = 3,000 M^{-1} cm^{-1} for S-formylglutathione, ε = 3,100 M^{-1} cm^{-1} for S-lactoylglutathione.

**Construction of frmB and yeiG Deletion Mutants**—Deletion mutants were generated by the methods described by Datsenko and Wanner (36). The *Km* cassette was PCR-amplified from pKD4 by using specific primers for the *frmB* gene (*frmB*-KOF plus frmB-KOR) and for the *yeiG* gene (yeiG-KOF plus yeiG-KOR). The PCR product was purified and used to transform competent cells of the *E. coli* BW25113 strain containing pKD46. To induce the ARED recombination proteins, the cells were grown on LB medium containing 1 mM arabinose. The competent cells were transformed by electroporation using a Gene Pulser (pulse controller at 200 ohms, capacitance at 250 microfarads, and voltage at 25 kV). Kanamycin-resistant transformants were recovered on LB agar plates containing 25 μg/ml of kanamycin. The insertion of the recombinant cassette was confirmed by PCR using a set of primers complementary to a region upstream and downstream of each target gene. The sets of primers used were *frmB*-CF/*frmB*-CR and yeiG-CF/yeiG-CL for *frmB* and *yeiG*, respectively. Plasmid pKD46 was eliminated by growing the cells on LB plates at 37 °C. The mutagenesis process was completed by deletion of kanamycin cassette using pCP20 as described by Datsenko and Wanner (36).

**Quantitative Reverse Transcription-PCR**—Bacterial cells were cultured in MOPS minimal medium (37) with 0.05% glucose as a carbon source with shaking at 37 °C. When the cells reached *A*_{600} = 0.3, the cultures were split into three aliquots. Individual subcultures contained the following: no inducer, 0.25 mM formaldehyde, and 250 mM methanol. Cells were collected by centrifugation 45 min later. Total RNA was subsequently isolated with the RNeasy Mini Kit (Qiagen) in accordance with the manufacturer’s protocol. Residual DNA present in the DNA preparations was removed by RNase-free DNase (Fermentas). cDNAs were synthesized with the superscript first-strand synthesis kit (Invitrogen) in accordance with the manufacturer’s instructions and stored at −20 °C prior to use. Real-time quantitative PCR was carried out on the Applied Biosystems 7300 apparatus (Applied Biosystems) using Platinum SYBR Green qPCR SuperMix UDG (Invitrogen) in accordance with the manufacturer’s recommended protocol.

**Site-directed Mutagenesis**—Site-directed mutagenesis was performed using the QuickChange™ site-directed mutagenesis kit (Stratagene) according to the manufacturer’s protocol. The *yeiG* open reading frame cloned on pCA24N expression vector was used as a template. Primers carrying specific mutations were 41–42 nucleotides in length; the target codon was flanked by ~20 bp each side. The amino acids selected to be mutated were all changed to alanine. The standard PCR mixture contained 50–100 ng of template DNA and 150–250 ng of each primer. The methylated plasmid was digested with DpnI (twice), and 4 μl of each reaction were used to transform competent cells. Kanamycin-resistant transformants were recovered on LB agar plates containing 25 μg/ml of kanamycin. The insertion of the recombinant cassette was confirmed by PCR using a set of primers complementary to a region upstream and downstream of each target gene. The sets of primers used were *frmB*-CF/*frmB*-CR and *yeiG*-CF/*yeiG*-CL for *frmB* and *yeiG*, respectively. Plasmid pKD46 was eliminated by growing the cells on LB plates at 37 °C. The mutagenesis process was completed by deletion of kanamycin cassette using pCP20 as described by Datsenko and Wanner (36).
RESULTS AND DISCUSSION

E. coli Has Two Genes Encoding Putative S-Formylglutathione Hydrolases—In E. coli, the frmB (yaiM) gene encodes a putative FGH and is part of an operon with frmR (encoding the putative transcriptional repressor for this operon) and frmA (encoding an S-hydroxymethylglutathione dehydrogenase). This operon is probably involved in the detoxification of formaldehyde (29). The BLAST search of the E. coli genome with the FrmB sequence revealed the presence of a paralog, the YeiG protein, which has 54% sequence identity to FrmB (Fig. 1).

In contrast to frmB, the yeiG gene is a single gene operon located between folE and cirA and transcribed in the opposite direction to these genes. No paralogs of FrmR or FrmA were found in the E. coli genome. YeiG is annotated as an esterase (Swiss-Prot or EcoCyc data base), because in our recent work we demonstrated that YeiG possesses esterase activity in general enzymatic screens with model substrates (pNP-butyrate and palmitoyl-CoA) (38). Analysis of amino acid sequences of FGHs from sequenced genomes demonstrated that this enzyme is well conserved with sequences from bacteria and eukaryotes having 44–80% identity (Fig. 1). Interestingly, the BLAST analysis of archaeal and many Gram-positive bacterial genomes revealed no presence of either FGH homolog, suggesting that these organisms use other enzymes for the detoxification of formaldehyde.

This scattered phylogenetic distribution of FGHs and rather high sequence similarity among FGHs from different organisms suggest a relatively recent evolutionary origin of this enzyme, which was probably spread mainly by the lateral gene transfer (39). Analysis of sequenced genomes revealed that besides three E. coli strains (K12, O157:H7, and O6), two FGH genes were only found in three environmental bacteria: Rhizobium (Sinorhizobium) meliloti, Shewanella oneidensis, and Photobacterium profundum. In contrast to E. coli, in R. meliloti and S. oneidensis genomes, both putative FGH genes show an operon association with genes encoding putative glutathione-dependent formaldehyde dehydrogenases, suggesting that both operons might be involved in formaldehyde detoxification.

Enzymatic Activity of FrmB and YeiG—Purified FrmB and YeiG were screened for catalytic activity against a panel of 15 esterase substrates (carboxyl- and thioesters). YeiG showed detectable activity (17.6 nmol/min mg protein) against palmitoyl-CoA (C–S bond hydrolysis), high activity against phenyl-acetate and several pNP-esters of short chain fatty acids (C–O bond hydrolysis) (Fig. 2). FrmB did not hydrolyze palmitoyl-CoA and showed significant esterase activity only toward two pNP-esters (pNP-acetate and pNP-propionate).
and α-naphthyl acetate (Fig. 2). With pNP-propionate as substrate, both proteins had maximal activity at pH 7.5–8.8 (data not shown).

Our experiments with purified FrmB and YeiG revealed high formylglutathione hydrolase activity in both proteins. This activity involves the hydrolysis of the C–S bond, indicating that these proteins are promiscuous esterases capable of hydrolysis of two types of bonds: C–O and C–S. With S-formylglutathione as substrate, both proteins showed saturation kinetics with a sigmoidal saturation curve (Hill coefficient \( n = 2.2 \pm 0.44 \)), indicating positive cooperativity in S-formylglutathione binding. With other substrates (Table 1), both proteins exhibited hyperbolic saturation curves. High concentrations of S-formylglutathione (>1 mM) inhibited the activity of both proteins. FrmB and YeiG showed similar affinity to S-formylglutathione, but FrmB had 4–5 times higher specific activity (Table 1). For both proteins, S-formylglutathione was the best substrate, indicating that this metabolite might be their natural substrate. The kinetic parameters of the E. coli FrmB and YeiG (Table 1) were similar to those of eukaryotic FGHs from yeast, A. thaliana, and human liver (21–23). Like for the A. thaliana FGH (23), the activity of both FrmB and YeiG was inhibited by the sulfhydryl inhibitors (N-ethylmaleimide, iodoacetate, ZnCl_{2}, and CuCl_{2} (1 mM; data not shown)), suggesting the importance of SH groups for catalysis in both proteins. In the available sequences of known and putative FGHs, there are two conserved Cys residues (Cys^{26} and Cys^{54} in YeiG; Fig. 1) that might be targeted by these sulfhydryl inhibitors.

YeiG also showed significant activity against lactoylglutathione (Fig. 2). With this substrate, it had 20 times higher hydrolyase activity than FrmB. No hydrolyase activity was found toward S-(1,2-dicarboxyethyl)glutathione, S-hexylglutathione, or S-(4-nitrobenzyl)glutathione. S-Lactoylglutathione is an intermediate of the detoxification of another aldehyde, methylglyoxal. In all organisms, methylglyoxal is produced spontaneously from the glycolytic intermediates (dihydroxyacetone phosphate or glyceraldehyde 3-phosphate) or by enzymes (methylglyoxal synthase, amino oxidases) (40–43). Like formaldehyde, methylglyoxal is very toxic, and its elevated production results in rapid modification of proteins, nucleotides, and other substrates generating advanced glycation end products (44). The methylglyoxal-induced modifications may regulate gene expression and produce multiple negative effects in diabetes and other pathological conditions (45, 46).

Methylglyoxal can be detoxified by the glutathione-independent glyoxalase III, by the glutathione-dependent system comprising glyoxalase I and glyoxalase II, or by various aldo and keto reductases (47–50). Glyoxalase I produces S-lactoylglutathione, which is then hydrolyzed by glyoxalase II to D-lactate and glutathione (51). Although the presence of glyoxalase II (GloB) in E. coli has been reported, the biochemical characterization of this enzyme has not been published (52). However, the affinity and activity of the E. coli YeiG toward S-lactoylglutathione are within the range of these parameters of known glyoxalases from other organisms (0.05–3.0 mM; 1.34–2007 μmol/min/mg of protein) (BRENDA data base). The glyoxalase system of Gram-negative bacteria is thought to operate in the periplasmic space (42). YeiG may represent a cytoplasmic equivalent of glyoxalase II, specialized for detoxification of endogenous rather than extracellular methylglyoxal. In addition, high activity of YeiG toward α-naphthyl acetate might indicate that this enzyme is also involved in the degradation of a presently unknown naphthyl-containing metabolite in E. coli cells.

### Table 1

| Substrate                  | \( K_m \) (mM) | \( V_{max} \) (units/mg s) | \( k_{cat} \) (s\(^{-1}\)) | \( k_{cat}/K_m \) (s\(^{-1}\) M\(^{-1}\)) |
|----------------------------|---------------|----------------------------|---------------------------|---------------------------------|
| FrmB pNP-acetate (C2)      | 0.29 ± 0.02   | 0.58 ± 0.02                | 0.30 ± 0.01               | 1.03 × 10\(^5\)                |
| pNP-propionate (C3)        | 0.83 ± 0.04   | 0.27 ± 0.03                | 0.14 ± 0.02               | 0.17 × 10\(^4\)                |
| S-Lactoylglutathione       | 0.60 ± 0.21   | 0.09 ± 0.02                | 0.05 ± 0.01               | 0.08 × 10\(^4\)                |
| S-Formylglutathione        | 0.41 ± 0.03   | 55.0 ± 4.0                 | 28.5 ± 2.07               | 6.90 × 10\(^4\)                |
| YeiG pNP-acetate (C2)      | 0.45 ± 0.05   | 0.50 ± 0.08                | 0.26 ± 0.04               | 0.57 × 10\(^5\)                |
| pNP-propionate (C3)        | 0.48 ± 0.03   | 0.71 ± 0.01                | 0.37 ± 0.05               | 0.76 × 10\(^4\)                |
| pNP-butyrate (C4)          | 0.70 ± 0.03   | 0.70 ± 0.05                | 0.36 ± 0.03               | 0.51 × 10\(^4\)                |
| pNP-caproate (C6)          | 0.95 ± 0.01   | 0.64 ± 0.08                | 0.49 ± 0.04               | 0.52 × 10\(^4\)                |
| α-Naphthyl acetate         | 1.03 ± 0.13   | 9.79 ± 0.56                | 5.09 ± 0.29               | 4.94 × 10\(^4\)                |
| S-Formylglutathione        | 0.58 ± 0.18   | 1.97 ± 0.29                | 1.02 ± 0.15               | 1.75 × 10\(^4\)                |
| S-Formylglutathione        | 0.43 ± 0.05   | 12.6 ± 1.30                | 6.51 ± 0.67               | 15.1 × 10\(^4\)                |

*Units = μmol/min.*
of this catalytic triad. Like in typical serine hydrolases (55), the catalytic Ser145 of E. coli YeiG is located on the predicted tight turn between a β-strand and an α-helix. Likewise, Asp218 and His254 are also located on the loops connecting predicted α-helices and β-strands of YeiG. The lack of esterase activity in the S145A mutant was also confirmed using S-lactoylglutathione as a substrate. The alanine replacement of Asp80, Asp199, and Asp218 reduced the catalytic activity of YeiG without a strong effect on the substrate affinity (Table 2). Surprisingly, the D255A mutant showed ∼2 times lower activity but 2 times higher affinity to α-naphthyl acetate.

Crystal Structure of FGH: S. cerevisiae YJG8 Protein—The only available three-dimensional structure of FGH, that of the S. cerevisiae YJG8 protein, was solved in the course of a structural genomics project (Protein Data Bank code 1pv1). Although YJG8 is annotated as a putative serine esterase, previous work identified the presence of S-formylglutathione hydrolase activity in this protein (22). YJG8 shows 45 and 41% of sequence identity to the E. coli YeiG and FrmB, respectively, and their sequence alignment demonstrated an absolute conservation of the serine hydrolase catalytic triad (Fig. 1). Because of this high sequence conservation between FGHs, we used the YJG8 structure to describe the FGH catalytic site.

The crystal structure of YJG8 (Protein Data Bank code 1pv1) revealed a dimer of dimers in the crystallographic unit cell (Fig. 3). This is consistent with the tetrameric form of the native E. coli YeiG demonstrated using analytical gel filtration (121.2 kDa). Extensive dimer/dimer contacts contain comparable amounts of hydrophobic (5) and hydrophilic (6) residues. The two times smaller dimer/dimer contacts (seven residues) are more hydrophilic (Glu99, Ser101, Asp103, and Lys200). The structure of the YJG8 monomer shows the general topology of the α/β hydrolase fold, with a central nine-stranded β-sheet flanked by three helices on one side and by 10 helices on the other side (Fig. 3A). The canonical α/β hydrolase fold has an eight-stranded, mostly parallel β-sheet with the strand β2 antiparallel to the rest (strand order 12435678) (56, 57). The YJG8 monomer structure shows that it contains a canonical eight-stranded β-sheet to which an additional anti-parallel strand (β1) has been added, creating a nine-stranded β-sheet with the strand order 123546789 (Fig. 3B). The YJG8 β-sheet is highly twisted with the first and last strands oriented almost perpendicularly to one another.

Using the Dali search (58), we compared the structure of YJG8 with other available protein structures. The five closest matches include dipeptidyl peptidase IV (Protein Data Bank code 1r9m; Z score 18.0; r.m.s. deviation 3.3 Å), prolyl endopeptidase (Protein Data Bank code 1o6g; Z score 16.2; r.m.s. deviation 3.8 Å), epoxide hydrolase 2 (Protein Data Bank code 1ij5; Z score 16.0; r.m.s. deviation 2.9 Å), triacylglycerol lipase (Protein Data Bank code 1k8g; Z score 14.2; r.m.s. deviation 3.5 Å), and serine carboxypeptidase KEX1 from yeast (Protein Data Bank code 1ac5; Z score 13.9; r.m.s. deviation 3.5 Å). These proteins all have hydrolase activity against various substrates and show only 9–16% of total sequence identity to YJG8 but with the conservation of all three residues of the serine hydrolase catalytic triad (except for the catalytic serine in epoxide hydrolase 2). The overall topology of YJG8 (dimer of homodimers) is very similar to that of the tetrameric dipeptidyl peptidase IV (Protein Data Bank code 1r9m).

![Diagram of the yeast YJG8 Protein](Image)

FIGURE 3. Structure of the yeast YJG8 (Protein Data Bank 1pv1). A, ribbon diagram of one subunit of YJG8. The monomer is composed of a single domain with an α/β-hydrolase fold. α-Helices are shown in red, and β-strands are yellow. Also shown as a thick wire model are the residues of the catalytic triad: Ser161, His276, Asp241. B, the coordination of the side chains of the residues of the catalytic triad of YJG8. The pictures were drawn based on the Protein Data Bank entry 1pv1 using the PyMOL program (available on the World Wide Web at www.pymol.org).
likely active site with the three residues of the catalytic triad arranged in a classical way of serine hydrolases (Fig. 3B). The active site is formed by parts of helices α1, α6, α8, and α13; strands β4 and β6; and several flexible loops (β4–α1, β5–α3, β6–α6, β8–α12, and β9–α13). The walls of the catalytic cavity are formed by the conserved residues from the sequence blocks 1 and 4 (containing the serine nucleophile of the catalytic triad) and 5 (containing the histidine base of the catalytic triad) (Fig. 1). Most of the residues from two other blocks of conserved residues are located on the flexible loops and seem not to be directly involved in substrate coordination or protein oligomerization. Additional conserved residues identified in the catalytic cavity are Asn64, Lys68, Tyr102, Asp103, Tyr197, and Phe243. These residues are likely to be involved in substrate binding. The surface location of the FGH active site provides relatively easy access for substrates and is consistent with the promiscuous character of these enzymes. The length of the cavity is ~30.9 Å (from Gly105 to Val271), which fits the size of the glutathione molecule (~25 Å) and explains the preference of the FGH enzymes to short acyl chain substrates (Fig. 2). The catalytic Ser161 is located in the middle of this cavity (14.6 Å from the Gly105).

The conserved Cys60, located in the center of the active site pocket, close to the catalytic Ser161 (8.5 Å), presumably accounts for the inhibition by thiol-specific reagents. Alanine mutagenesis of analogous Cys54 in the E. coli YeiG produced no effect on catalytic activity but slightly reduced the substrate affinity of this protein and greatly increased its resistance to thiol inhibitors. As shown in Fig. 4, both the wild type and C26A (another conserved Cys) YeiG proteins were strongly inhibited by 1 mM N-ethylmaleimide (K_i = 101.5 and 91.3 μM, respectively), whereas the C54A mutant retained almost full activity. Phenylmethysulfonyl fluoride (PMSF) was used as a control to rule out the possibility of nonspecific inhibition. The inhibition of the C54A mutant by PMSF was negligible, indicating that the inhibition by N-ethylmaleimide is specific for thiol groups. Additionally, the C54A mutant was not inhibited by iodoacetate, further confirming the specificity of the inhibition by N-ethylmaleimide.

**Figure 4.** Effect of N-ethylmaleimide (A) and phenylmethysulfonyl fluoride (B) on the hydrolysis of α-naphthyl acetate by the E. coli wild type and mutant YeiG. Reaction mixtures contained 1 mM α-naphthyl acetate and 3 μg of enzyme. Other reaction conditions were as described under "Materials and Methods." PMSF, phenylmethysulfonyl fluoride.

**Figure 5.** Possible reaction mechanism of the hydrolysis of S-formylglutathione by FGHs. The reaction proceeds through the following steps: 1) nucleophilic attack of Ser161 to form the first tetrahedral intermediate; 2) decomposition of the tetrahedral intermediate-1 and formation of the formyl enzyme intermediate; 3) release of reduced glutathione; 4) nucleophilic attack of water and the formation of the second tetrahdral intermediate; and 5) release of the second product (formate) and formation of free enzyme.
fluoride, which is known to react with the catalytic serine in hydrolases (59), produced similar levels of inhibition in all three YeIG proteins (Fig. 4). Thus, the inhibiting effect of thiol inhibitors on the activity of FGHs can be explained by the binding of the inhibitor molecule to the conserved cysteine residue (Cys54 in YeIG or Cys56 in YJG8) located in the active site of this enzyme, consequently blocking the binding of substrate to the active site of FGHs.

The sequential and three-dimensional arrangement of the catalytic residues Ser161, His276, and Asp241 corresponds to that of related α/β-hydrolases (54–56, 60). Ser161 is embedded in the sequence Gly159–Ser161–Met162–Gly163–Gly164, conserved in all characterized FGHs (Fig. 1) and in many predicted esterases. This residue is located on a sharp turn (nucleophilic elbow), which connects strand β6 and helix α6. His276 is located at the end of an eight-residue loop situated between the strand β9 and helix α13 (Fig. 1). Its imidazole group is in the proper position to make hydrogen bonds with the catalytic Ser161 (3.65 Å) and Asp241 (2.73 Å) (Fig. 3B). Asp241 is located on the long loop (14 residues) between the strand β8 and helix α12 (Fig. 1). Its carboxylate groups form hydrogen bonds with the catalytic His276 (Fig. 3B). Analysis of the YJG8 active site suggests that like in the chymotrypsin-type esterases (60), its oxyanion hole is formed by the main chain NH groups of the substrate to the active site of FGHs.

The sequential and three-dimensional arrangement of the catalytic residues Ser161, His276, and Asp241 corresponds to that of related α/β-hydrolases (54–56, 60). Ser161 is embedded in the sequence Gly159–Ser161–Met162–Gly163–Gly164, conserved in all characterized FGHs (Fig. 1) and in many predicted esterases. This residue is located on a sharp turn (nucleophilic elbow), which connects strand β6 and helix α6. His276 is located at the end of an eight-residue loop situated between the strand β9 and helix α13 (Fig. 1). Its imidazole group is in the proper position to make hydrogen bonds with the catalytic Ser161 (3.65 Å) and Asp241 (2.73 Å) (Fig. 3B). Asp241 is located on the long loop (14 residues) between the strand β8 and helix α12 (Fig. 1). Its carboxylate groups form hydrogen bonds with the catalytic His276 (Fig. 3B). Analysis of the YJG8 active site suggests that like in the chymotrypsin-type α/β-hydrolases (60), its oxyanion hole is formed by the main chain NH groups of the catalytic Ser161 and Gly159. The lack of a side chain at position 159 (YJG8 numbering) allows correct positioning of the amido hydrogen to form the required hydrogen bond with the oxyanion substrate.

The present work allows a putative picture of the catalytic mechanism of the S-formylglutathione hydrolysis catalyzed by FGHs (Fig. 5). Consistent with previous mechanistic studies on serine hydrolases (61), Ser161 of YJG8 is properly positioned for nucleophilic attack of the carbonyl carbon of the bound S-formylglutathione molecule. The importance of Ser161 during catalysis is supported by mutagenic studies of the E. coli YeIG showing that a S145A mutation results in an inactive enzyme (Table 2). His276 and Asp241, the other residues of the catalytic triad in the yeast YJG8, are positioned to assist in the general base-catalyzed attack of the serine nucleophile, and their significance for catalysis was also confirmed by mutational studies on YeIG (Table 2). The nucleophilic attack by the Ser161 hydroxyl group on the carbonyl carbon atom of S-formylglutathione leads to the formation of a tetrahedral intermediate that is stabilized through interactions with an oxyanion hole formed by the backbone NH groups of the catalytic Ser161 and Gly159 (Fig. 5). The tetrahedral intermediate breaks down by general acid catalysis to a formyl-enzyme intermediate with the formyl moiety covalently attached to Ser161. During the formaylation step, the imidazole group of His276 transfers the proton of the serine hydroxyl to the sulfydryl leaving group regenerating the reduced glutathione. The formyl-FGH is then deformaolated through the reverse reaction pathway of formaylation with formation of a second tetrahedral intermediate releasing free formate, but in this reaction, a water molecule instead of the serine residue is the attacking nucleophile (Fig. 5). Deformaolation produces the regenerated enzyme and the reaction product, formic acid.

The molecular basis of genetic polymorphism of the human FGH (esterase D) has been characterized recently (62). The rarer polymorphic allele ESD*2 was found to contain Glu190 instead of the Gly190 present in the most common allele ESD*1. The ESD*7 allele (originated from ESD*2) has a D231G mutation, whereas the ESD*5 allele (originated from ESD*1) has a G257D mutation. The structure of YJG8 shows that Gly257 (Gly257 in humans) and Glu246 (Asp231 in humans) are located far away from the protein active site and therefore should have no significant influence on activity. However, Gly273 (Gly277 in humans) is a conserved residue present in all known FGHs and located just 2 residues before the catalytic His276. The replacement of this Gly to Asp (as in ESD*5 allele) will introduce strong negative charge close to the catalytic His278 and Asp241 and therefore can potentially affect the interaction of these residues. Moreover, in YJG8, Gly273 is one of 14 residues involved in the dimer intersubunit interaction (hydrogen-bonded to Ser8), and its replacement by Asp might also perturb the oligomerization of enzyme.

Gene Expression Studies (Real Time PCR)—To study the expression of the two FGHs in E. coli and their role in the detoxification of formaldehyde, we designed two single deletion mutants (ΔfrmB and ΔyeiG) and a double mutant (ΔfrmB-ΔyeiG) (see “Materials and Methods”). Using real time PCR, we characterized the effect of different growth conditions and gene deletions on the expression of both FGHs in E. coli (Fig. 6).

In the wild type strain, the level of uninduced expression of frmB was ~5 times higher than that of yeiG. These basally expressed FGHs are perhaps involved in the detoxification of endogenous formaldehyde continuously produced by normal cellular metabolism or demethylolation reactions. The addition of formaldehyde (as well as methanol, nitrosoglutathione, methylviologen, and H2O2) to the growth medium had no effect on the expression of yeiG in both the wild type and ΔfrmB strains (Fig. 6B). However, formaldehyde greatly stimulated the expression of frmB: 45-fold in wild type and 70-fold in ΔyeiG over the level of uninduced cells (Fig. 6A). This is consistent with previous studies that reported a 10–207-fold increase in the expression of the E. coli frmAB in response to formaldehyde addition (29, 63). Methylviologen or H2O2 did not affect the expression of frmB, indicating that this gene is not regulated by oxidative stress. The higher level of the frmB induction by formaldehyde in the yeiG deletion mutant (Fig. 6A) might indicate that E. coli cells compensate for the absence of YeIG by increased synthesis of FrmB and suggests that these two enzymes have overlapping functions in this organism. Thus, in E. coli, the expression of frmB is stimu-
lulated by exogenous formaldehyde or by deletion of yeiG, whereas the yeiG expression is not regulated by exogenous formaldehyde or methanol or by deletion of frmB. The uninduced (basal) expression of both FGHs in E. coli cells is perhaps associated with the degradation of endogenous formaldehyde produced by normal cellular metabolism or demethylation reactions, whereas exogenous formaldehyde induce additional synthesis of more active FGH (FrmB) in E. coli.

E. coli Resistance to Formaldehyde—Both single deletion mutants (ΔfrmB and ΔyeiG) and a double mutant (ΔfrmB–ΔyeiG) behaved as the wild type strain under normal aerobic growth conditions on the minimal MOPS-glucose medium (wild type strain under normal aerobic growth conditions on the minimal MOPS-glucose medium (Table 3). These results suggest that single mutants, the remaining FGH can compensate for the lack of another enzyme. Therefore, simultaneous inactivation of both the frmB and yeiG genes is required to increase the sensitivity of E. coli cells to formaldehyde. 

Table 3

| Formaldehyde | Growth rate (μmax) |
|--------------|--------------------|
|              | WT                 | ΔfrmB | ΔyeiG | ΔfrmB–ΔyeiG |
|              | μm/s               |        |       |             |
| No addition  | 0.51 ± 0.01        | 0.49 ± 0.02 | 0.51 ± 0.01 | 0.51 ± 0.02 |
| 0.10         | 0.48 ± 0.02        | 0.49 ± 0.02 | 0.50 ± 0.02 | 0.47 ± 0.01 |
| 0.35         | 0.46 ± 0.02        | 0.47 ± 0.01 | 0.50 ± 0.02 | 0.36 ± 0.02 |
| 0.40         | 0.47 ± 0.03        | 0.43 ± 0.05 | 0.50 ± 0.01 | 0.22 ± 0.05 |

Effect of formaldehyde on the growth rate (μmax) of the E. coli wild type (WT) and mutant strains

Acknowledgments—We thank all members of the Ontario Center for Structural Proteomics for help in conducting experiments.

References

1. Bolt, H. M. (1987) J. Cancer Res. Clin. Oncol. 113, 305–309
2. Chaw, Y. F., Crane, L. E., Lange, P., and Shapiro, R. (1980) Biochemistry 19, 5525–5531
3. Heck, H. D., Casanova, M., and Starr, T. R. (1990) Crit. Rev. Toxicol. 20, 397–426
4. Ma, T. H., and Harris, M. M. (1988) Mutat. Res. 196, 37–59
5. Precious, E., Gunn, C. E., and Lyles, G. A. (1988) Biochem. Pharmacol. 37, 707–713
6. Fall, R., and Benson, A. A. (1996) Trends Plant Sci. 1, 296–310
7. Hanson, A. D., Gage, D. A., and Shachar-Hill, Y. (2000) Trends Plant Sci. 5, 206–213
8. Falnes, P. O., Johansen, R. F., and Seeberg, E. (2002) Nature 419, 178–182
9. Aas, P. A., Otterlei, M., Falnes, P. O., Vagbo, C. B., Skorpen, F., Akbari, M., Sundheim, O., Bjorås, M., Slupphaug, G., Seeberg, E. R., and Krokan, H. E. (2003) Nature 421, 859–863
10. Bystryn, L. V., Govorushkina, N. I., Van Ophem, P. W., Hektor, H. J., Dijkhuizen, L., and Duine, J. A. (1993) J. Gen. Microbiol. 139, 1979–1985
11. Sakai, Y., Muramuto, A. P., Sembiring, L. T., and Yato, N. (1995) FEMS Microbiol. Lett. 127, 229–234
12. Ito, K., Takahashi, M., Yoshimoto, T., and Tsuru, D. (1994) J. Bacteriol. 176, 2483–2491
13. Uotila, L., and Koivusalo, M. (1979) Arch. Biochem. Biophys. 196, 33–45
14. Lee, W.-H., Wheatley, W., Benedict, W. F., Huang, C. M., and Lee, E.-Y.-H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 6790–6794
15. Harms, N., Ras, J., Reijnders, W. N., van Spanning, R. J., and Stouthamer, A. H. (1996) J. Bacteriol. 178, 6296–6299
16. Stover, N. A., Cavalcanti, A. R., Li, A. J., Richardson, B. C., and Landweber, L. F. (2005) Mol. Biol. Evol. 22, 1539–1542
17. Mouliès, J.-M., Holmquist, B., and Valle, B. I. (1991) Biochemistry 30, 5743–5749
18. Jensen, D. E., Belka, G. K., and Du Bois, G. C. (1998) Biochem. J. 331, 659–668
19. Liu, L., Hausladen, A., Zeng, M., Que, L., Heitman, J., and Stamler, J. S. (2001) Nature 410, 490–494
20. Sanghni, P. C., Bosron, W. F., and Hurley, T. D. (2002) Biochemistry 41, 15118–15119
21. Uotila, L., and Koivusalo, M. (1974) J. Biol. Chem. 249, 7664–7672
22. Degrassi, G., Uotila, L., Klima, R., and Venturi, V. (1999) Appl. Environ. Microbiol. 65, 3470–3472
23. Kordic, S., Cummins, I., and Edwards, R. (2002) Arch. Biochem. Biophys. 399, 232–238
24. Lee, Y.-E. H., and Lee, W.-H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 6337–6341
25. Cowell, J. K., Rutland, P., Jay, M., and Hungerford, J. F. (1986) Hum. Genet. 74, 298–301
26. Rito, E. R., Mason-Brothers, A., Menkes, J. H., and Sparks, R. S. (1988) Arch. Gen. Psychiatry 45, 600
27. Spitsyn, V. A., Nafikova, A. K., Spitsyna, N. K., and Afanas'eva, I. S. (2001) Genetika 37, 698–707
28. Irizarry, K., Hu, G., Wong, M. L., Licino, J., and Lee, C. J. (2001) Pharmacogenomics J. 1, 193–203
29. Herzig, C. and Blattner, F. R. (2004) J. Bacteriol. 186, 6714–6720
30. Uotila, L. (1981) Methods Enzymol. 77, 424–430
31. Schapiron, J. M., Libby, S. J., and Fang, F. C. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 8496–8501
32. Zhang, R. G., Skaran, T., Katz, J. E., Beasley, S., Khachaturyan, A., Yyas, S., Arrow-smith, C. H., Clarke, S., Edwards, A., Joachimiak, A., and Savchenko, A. (2001) Structure 9, 1095–1106
33. Kitagawa, M., Stover, N. A., Cavalcanti, A. R., Li, A. J., Richardson, B. C., and Landweber, L. F. (2005) Mol. Biol. Evol. 22, 1539–1542
34. Cooper, R. A. (1984) Adv. Enzymol. 44, 381–424
35. Vorderwülbecke, T., Kieslich, K., and Erdmann, H. (1992) Enzyme Microb. Technol. 14, 631–639
36. Berge, R. K., and Farstad, M. (1981) Methods Enzymol. 71, 234–242
37. Datsenko, K. A., and Wanner, B. L. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 6640–6645
38. Neidhart, F. C., Bloch, P. L., and Smith, D. F. (1974) J. Bacteriol. 114, 196–213
39. Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000) Nature 405, 299–304
40. Freedberg, W. B., Kistler, W. S., and Lin, E. C. (1971) J. Bacteriol. 108, 137–144
41. Shahani, P. C., Bosron, W. F., and Hurley, T. D. (2002) Biochemistry 41, 15118–15119
42. Shinozaki, K., Thorvalds, R. J., Giardino, I., Beisswenger, P., Thorpe, S. R., Onorato, J., and Brownlee, M. (1998) J. Clin. Invest. 101, 1142–1147
E. coli S-Formylglutathione Hydrolases

45. Ramasamy, R., Yan, S. F., and Schmidt, A. M. (2006) Cell 124, 258–260
46. Yao, D., Taguchi, T., Matsumura, T., Pestell, R., Edelstein, D., Giardino, I., Suske, G., Ahmed, N., Thornalley, P. J., Sarthy, V. P., Hammes, H.-P., and Brownlee, M. (2006) Cell 124, 275–286
47. Thornalley, P. J. (1990) Biochem. J. 269, 1–11
48. Misra, K., Banerjee, A. B., Ray, S., and Ray, M. (1995) Biochem. J. 305, 999–1003
49. Misra, K., Banerjee, A. B., Ray, S., and Ray, M. (1996) Mol. Cell. Biochem. 156, 117–124
50. Inoue, Y., and Kimura, A. (1995) Adv. Microbiol. Physiol. 37, 177–227
51. Thornalley, P. J. (1995) Crit. Rev. Oncol. Hematol. 20, 99–128
52. Ferguson, G. P., Totemeyer, S., MacLean, M. J., and Booth, I. R. (1998) Arch. Microbiol. 170, 209–218
53. Brenner, S. (1988) Nature 334, 528–530
54. Cygler, M., Schrag, J. D., Sussman, J. L., Harel, M., Silman, I., Gentry, M. K., and Doctor, B. P. (1993) Protein Sci. 2, 366–382
55. Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolov, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., and Schrag, J. (1992) Protein Eng. 5, 197–211
56. Heikinheimo, P., Goldman, A., Jeffries, C., and Ollis, D. L. (1999) Structure 7, 141–146
57. Holmquist, M. (2000) Curr. Protein Pept. Sci. 1, 209–235
58. Holm, L., and Sander, C. (1995) Trends Biochem. Sci. 20, 478–480
59. Gold, A. M. (1995) Biochemistry 4, 897–901
60. Polgar, L. (2005) Cell. Mol. Life Sci. 62, 2161–2172
61. Powers, J. C., Odake, S., Oleksyszyn, J., Hori, H., Ueda, T., Boduszek, B., and Kam, C. (1993) Agents Actions Suppl. 42, 3–18
62. Yuasa, I., Umetsu, K., Tsuchida, S., Irizawa, Y., Henke, L., and Henke, J. (2004) Hum. Biol. 76, 479–488
63. Gutbeil, W. G., Kasimoglu, E., and Nicholson, P. C. (1997) Biochem. Biophys. Res. Commun. 238, 693–696