BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

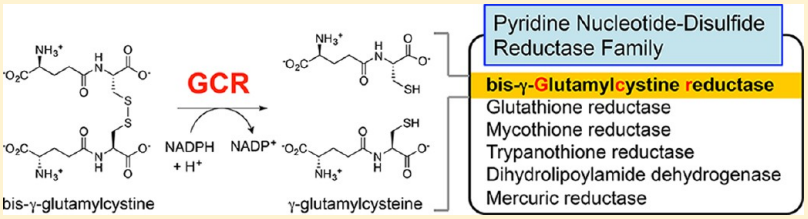
pubs.acs.org/biochemistry

The Orphan Protein Bis-γ-glutamylcystine Reductase Joins the Pyridine Nucleotide Disulfide Reductase Family

Juhan Kim and Shelley D. Copley*

Department of Molecular, Cellular and Developmental Biology and Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado 80309, United States

Supporting Information



ABSTRACT: Facile DNA sequencing became possible decades after many enzymes had been purified and characterized. Consequently, there are still "orphan" enyzmes for which activities are known but for which encoding genes have not been identified. Identification of the genes encoding orphan enzymes is important because it allows correct annotation of genes of unknown function or with misassigned function. Bis-γ-glutamylcystine reductase (GCR) is an orphan protein that was purified in 1988. This enzyme catalyzes the reduction of bis-γ-glutamylcystine. γ-Glutamylcysteine is the major low-molecular weight thiol in halobacteria. We purified GCR from *Halobacterium* sp. NRC-1 and identified the sequence of 23 tryptic peptides by nano-liquid chromatography electrospray ionization tandem mass spectrometry. These peptides cover 62% of the protein predicted to be encoded by a gene in *Halobacterium* sp. NRC-1 that is annotated as mercuric reductase. GCR and mercuric reductase activities were assayed using enzyme that was expressed in *Escherichia coli* and refolded from inclusion bodies. The enzyme had robust GCR activity but no mercuric reductase activity. The genomes of most, but not all, halobacteria for which whole genome sequences are available have close homologues of GCR, suggesting that there is more to be learned about the low-molecular weight thiols used in halobacteria.

Massive genome sequencing efforts in recent years have contributed millions of sequences to genomic databases. Functions for the vast majority of these sequences have been predicted computationally on the basis of similarities to the sequences of other proteins and a variety of other genomic clues such as genome context and phylogenetic profiling.[1−3] Computational annotations are usually accurate at the superfamily level. However, predictions of specific functions are often wrong. As a result of misannotation and the subsequent transfer of erroneous annotations, the database is littered with incorrect assignments of function.[4]

On the other side of the picture, there are a number of "orphan" proteins for which functions are known but for which the corresponding genes have not been identified.[5−8] Bis-γ-glutamylcystine reductase (GCR) is one of these orphan proteins. GCR from *Halobacterium halobium* was purified and characterized by Sundquist and Fahey in 1988.[9] The enzyme catalyzes the reaction shown in Figure 1, which is similar to that catalyzed by the well-studied enzyme glutathione reductase. Bis-γ-glutamylcystine (γ-Glu-Cys), which lacks the glycine moiety of glutathione, is a major intracellular thiol in halobacteria, Archaea that are adapted for life in high-salt environments. Maintenance of reduced γ-Glu-Cys in halobacteria requires GCR. Here we report the identification of the gene encoding GCR in *Halobacterium* sp. NRC-1. The enzyme is misannotated as a mercuric reductase. GCR belongs to the pyridine nucleotide disulfide reductase family and is found only in halobacteria. However, some halobacteria lack GCR, suggesting that there is diversity with respect to mechanisms for maintaining the redox state of the cytoplasm and protection against oxidative damage even within the *Halobacterium* clade.

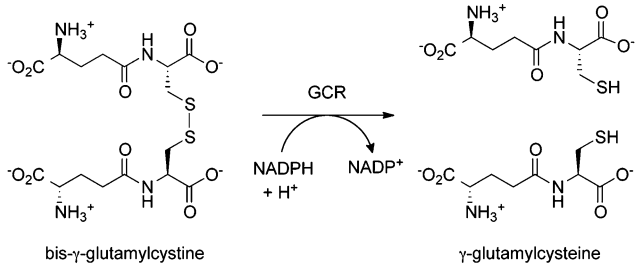


Figure 1. Reaction catalyzed by bis-γ-glutamylcystine reductase (GCR).

Received: March 14, 2013
Revised: April 1, 2013
Published: April 5, 2013

EXPERIMENTAL PROCEDURES

Growth of *Halobacterium* sp. NRC-1. *Halobacterium* sp. NRC-1 and its genomic DNA were generous gifts from N. Baliga (Institute for Systems Biology, Seattle, WA). A single colony grown for 1 week on *H. halobium* complex agar medium at 42 °C was inoculated into 5 mL of *H. halobium* complex medium (CM)[10] in a 15 mL culture tube. After being incubated for 4 days at 42 °C while being shaken at 250 rpm, the culture was added to 100 mL of CM in a 500 mL Erlenmeyer flask and incubated for an additional 4 days at 42 °C while being shaken. At that point, 10 mL aliquots of the culture were used to inoculate 1 L of *H. halobium* complex medium in each of seven 4 L Erlenmeyer flasks. Cultures were incubated for 4 days, and the cells were harvested by centrifugation at 4000*g* and room temperature for 40 min. Cell pellets were stored at −80 °C before use.

Chemicals and Other Materials. Bis-γ-glutamylcystine was prepared by passing O_2 through an aqueous solution of γ-Glu-Cys (94 mg dissolved in 3.0 mL of deionized water). The pH of the solution was adjusted to 8.0 with NH_4OH before the oxidation reaction.[11] The purity of the lyophilized bis-γ-glutamylcystine was assessed by ^1H and ^{13}C NMR in D_2O. The product was more than 99% pure, and no remaining γ-Glu-Cys was detectable: ^1H NMR (400 MHz, D_2O) δ 4.48 (dd, J = 4.0, 9.2 Hz, 1 H), 3.76 (dd, J = 5.2, 6.8 Hz, 1 H), 3.23 (dd, J = 4.0, 14 Hz, 1 H), 2.93 (dd, J = 9.2, 14 Hz, 1 H), 2.47 (m, 2 H), 2.16 (m, 2H); ^{13}C NMR (75 MHz, D_2O) δ 176.9, 174.3, 174.1, 54.3, 54.2, 39.8, 31.7, 26.5.

Butyl-Sepharose FF, HiTrap chelating HP, and HisTrap HP (immobilized Ni^{2+}) resins were purchased from GE Healthcare Biosciences (Pittsburgh, PA). Immobilized Cu^{2+} resin was prepared from HiTrap chelating HP resin using 0.1 M $CuCl_2$ following the manufacturer's instructions.

GCR Activity Assay. GCR activity was detected as described by Sundquist and Fahey.[12] One unit of enzyme activity is defined as the amount of enzyme that catalyzes conversion of 1 μmol of substrate per minute with 1 mM bis-γ-glutamylcystine and 0.42 mM NADPH. For reactions with varying concentrations of bis-γ-glutamylcystine, the concentration of NADPH was held constant at 1.7 mM.

Mercuric Reductase Activity Assay. Mercuric reductase activity was assayed by following the oxidation of NADPH at 340 nm and room temperature.[13] Assays were conducted in 50 mM sodium phosphate (pH 6.7) containing 3 M KCl, 1.3 M NaCl, 1 mM EDTA, 0.34 mM NADPH, and up to 1 mM $HgCl_2$.

Purification of GCR from *Halobacterium* sp. NRC-1. GCR was partially purified from 5 g cell pellets by the method of Sundquist and Fahey,[9] except that a butyl-Sepharose FF column was used instead of a Sepharose 4B column. Protein concentrations were determined by the method of Bradford.[14] Sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) was conducted using 4 to 20% gradient gels. Protein bands were visualized using a SilverQuest silver staining kit (Life Technologies, Grand Island, NY).

Mass Spectroscopic Analysis of GCR. A protein band obtained after SDS−PAGE of a sample obtained after purification of GCR using a column of immobilized Ni^{2+} resin was analyzed by nano-liquid chromatography (NanoLC) electrospray ionization tandem mass spectrometry (ESI-MS/MS) by ProtTech Inc. (Norristown, PA). The protein gel slice was treated with dithiothreitol (20 mM) and iodoacetamide (55 mM), successively, to reduce and alkylate cysteine residues. In-gel digestion of the protein sample was performed with sequencing-grade modified trypsin (Promega) in 100 mM ammonium bicarbonate (pH 8.5). The tryptic digest was analyzed using a high-pressure liquid chromatography system (Agilent) with a reverse phase C18 column (8 cm, 75 μm inside diameter) packed with 3 μm particles (pore size of 300 Å). Eluted peptides were analyzed with an ion trap mass spectrometer (LCQ-DECA XP PLUS, Thermo Scientific). The MS/MS data were used to search the nonredundant protein database RefSeq (http://www.ncbi.nlm.nih.gov/RefSeq) with Protech's ProQuest software suite.

Cloning of the Gene Encoding GCR. The gene encoding GCR (accession number NP_279293.1) was amplified by polymerase chain reaction from *Halobacterium* sp. NRC-1 genomic DNA with LA Taq polymerase in GC-I buffer provided by the manufacturer (Takara Bio, Inc., Otsu, Shiga, Japan) using the following primers: 5′-primer, 5′-GAC GAC GAC AAG ATG ACT ACC GAG CAA CCA CAC-3′; and 3′-primer, 5′-GAG GAG AAG CCC GGT TAC AGC TCG GCC GCG GCG TC-3′. The amplified gene was cloned into pET46 (EMD Millipore) by ligation-independent cloning (following the manufacturer's protocol) under control of an isopropyl β-D-thiogalactopyranoside (IPTG)-inducible T7 promoter, resulting in incorporation of a His_6 tag at the N-terminus of the protein.

Overproduction of *Halobacterium* GCR in *Escherichia coli*. *Halobacterium* sp. NRC-1 GCR was overproduced from pET46 in *E. coli* ArcticExpress (DE3) RP (Agilent Technologies). Terrific broth[15] (15 mL) containing 100 μg/mL ampicillin in a 50 mL Erlenmeyer flask was inoculated with a single colony carrying the expression plasmid obtained after overnight growth on LB agar medium with 100 μg/mL ampicillin at 37 °C. The culture was incubated while being shaken at 37 °C and 200 rpm until the OD_{600} reached 0.5. IPTG was added to give a final concentration of 0.5 mM, and the culture was shaken for 4 h at 37 °C and 200 rpm. Cells were harvested by centrifugation at 3500*g* for 10 min at 4 °C. Cell pellets were stored at −80 °C before being used.

Refolding and Reconstitution of Overproduced GCR. A 30 mg portion of a cell pellet from *E. coli* ArcticExpress (DE3) RP was resuspended in 1 mL of phosphate-buffered saline (PBS, pH 7) containing 1 mg/mL lysozyme and protease inhibitor mixture [used to give 1.2 mM 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF), 0.1 mM Bestatin, 15 μM E-64, 15 μM Pepstatin A, and 5 mM EDTA (Research Products International)]. After incubation for 10 min at room temperature, the cells were disrupted by sonication (2 × 4 min on ice) using a Virsonic Sonicator Cell Disruptor 600 (SP Scientific Co.). Insoluble fractions containing GCR were recovered by centrifugation at 16000*g* and 4 °C for 10 min. Protein refolding and reconstitution were performed according to the procedure used to refold and reconstitute *Haloferax volcanii* dihydrolipoamide dehydrogenase overproduced in *E. coli*.[16] The insoluble proteins were dissolved in 1 mL of solubilization buffer containing 2 mM EDTA, 50 mM DTT, and 8 M urea in 20 mM Tris-HCl (pH 8.0). The resulting protein solution was slowly diluted in 20 mL of refolding buffer containing 3 M KCl, 1.3 M NaCl, 35 μM FAD, 1 mM NAD, 0.3 mM glutathione disulfide, and 3 mM glutathione in 20 mM Tris-HCl (pH 8.0).

Purification of Refolded GCR. Refolded GCR was purified using a 1 mL immobilized Cu^{2+} column equilibrated with 50

mM sodium phosphate (pH 6.7) (buffer A) containing 1.23 M $(NH_4)_2SO_4$. A 1 mL HiTrap chelating HP column was connected to the distal end of the immobilized Cu^{2+} column to prevent elution of free Cu^{2+} into the collected fractions. The column was washed with 20 mL of buffer A containing 1.23 M $(NH_4)_2SO_4$. Fractions (1 mL) were collected during elution with a linear gradient from 0 to 500 mM imidazole in buffer A containing 1.23 M $(NH_4)_2SO_4$ (20 mL, total). Fractions were analyzed by SDS–PAGE on 12% polyacrylamide gels to identify fractions containing GCR.

Sequence Analysis. InterProScan version 4.8[17] at the European Bioinformatics Institute (EBI)[18] was used to identify conserved sequence domains and their functional annotations in GCR. Multiple sequence alignments were generated using Muscle.[19] Pairwise sequence identities were calculated using *needle* from the EMBOSS package[20] using the BLOSUM35 matrix with a gap-opening penalty of 10 and a gap-extension penalty of 0.5.

RESULTS

Identification of the Gene Encoding GCR from *Halobacterium* sp. NRC-1. We purified a protein with GCR activity from extracts of *Halobacterium* sp. NRC-1 following the method used by Sundquist and Fahey to purify GCR from *H. halobium*[9] (Table S1 of the Supporting Information). After four steps of column purification, one protein band observed after SDS–PAGE matched the size of the previously purified GCR from *H. halobium* (Figure S1 of the Supporting Information). NanoLC–ESI-MS/MS analysis of a tryptic digest of this gel band identified 23 peptide sequences (Table S2 of the Supporting Information). A search against the nonredundant RefSeq database found exact sequence matches for all 23 peptides in a protein from *Halobacterium* sp. NRC-1. Sixty-two percent of the matching protein sequence was covered by the peptide fragments (Figure 2). To our surprise, this *Halobacterium* sp. NRC-1 protein is encoded by a gene named *merA* and annotated as a mercury(II) reductase (accession number NP_279293). This annotation seemed unlikely to be correct, as the protein lacks the two consecutive cysteine residues found at the C-terminus of other mercuric reductases that are required for binding Hg(II) at the active site.[21]

```
         10         20         30         40         50
---------|---------|---------|---------|---------|
MTTEQPHVVIVGAYGSAGVAAAEELVDEDVTLTLIDDGEPGGGLCILRGC   50

MPSKEVLSAGAHRYQARHDDRLTGAVPDVDLDAVVDTKDDHVLGFAEHRR  100

AAVHDMAEREHVTFHHDTARFVDDRTLVVDGERIEADYVVLGTGSALNVP  150

DLPGIADTDWMGSRDVLDATEFPDSGVVMGFGYVGVELVPYLSEAADMDL  200

TVIEHDDRPLDSFHPVFGEEILAQYREEFDVDIRTEVREEAVEETDDGGV  250

EVRLDDGTTAEGEQLELFTGRTPSYPAGIEETRLTPDPEWVDPATMQARD  300

DDHVEVVGDAMGERMLLHNAKEEGYAAGRNALAVERGGDLETYDPFSHQV  350

MFSGLGVFPFASLGLTAEAARAAGHDVVTSQRDASSDGVFKTKDAARGAA  400

RLVVDADDGTVLGYHGLHYHADVMAKTMQVVLAAGMDVREIPDRAYHPTT  450

PEVLDGLFGDAAAEL                                     465
```

Figure 2. Mapping of peptides derived from tryptic digestion of *Halobacterium* sp. NRC-1 GCR onto the sequence of a protein annotated as mercuric reductase (accession number NP_279293). Fragments detected from ESI-MS/MS analysis are shown as dotted bars under the corresponding sequences. Fragments detected by ESI-MS/MS covered 287 of 465 amino acid residues.

Heterologous Expression, Refolding, and Purification of Active GCR from *E. coli*. To obtain larger quantities of the pure protein for kinetic characterization, we expressed GCR in *E. coli*. The gene annotated as *Halobacterium* sp. NRC-1 *merA* was cloned into pET46 in frame with a sequence encoding an N-terminal His_6 tag. The protein was well-expressed in several *E. coli* strains [*E. coli* BL21(DE3), BL21 Codon Plus (DE3) RP, Tuner(DE3), and Arctic Express (DE3) RP] under a variety of conditions, including concentrations of IPTG ranging from 10 μM to 0.5 mM, induction times ranging from 3 h to overnight, and temperatures ranging from 10 to 37 °C. However, the protein was insoluble in every case. This is a common phenomenon when proteins from halophiles are expressed in *E. coli*; halophilic proteins have evolved to be soluble and active under high-salt conditions and do not necessarily fold properly under the conditions of the *E. coli* cytoplasm.[22,23]

We refolded and reconstituted GCR from inclusion bodies using a protocol that was successful in refolding a dihydrolipoamide reductase from *Ha. volcanii* that had been expressed in *E. coli*.[16] Inclusion bodies containing GCR were dissolved in 8 M urea and then slowly diluted into a refolding buffer containing FAD and NAD at room temperature. GCR activity increased and then leveled off within 4 h. The reconstituted GCR was purified using an immobilized Cu^{2+} column (Figure 3A and Figure S2B and Table S3 of the

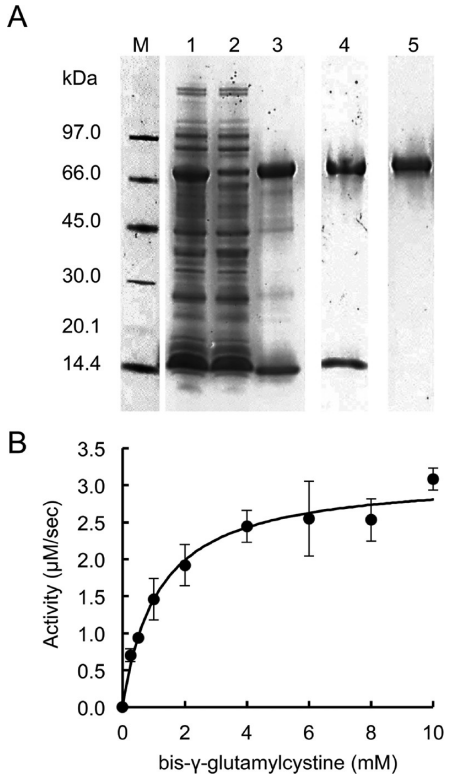


Figure 3. (A) SDS–PAGE analysis of *Halobacterium* sp. NRC-1 N-His_6-GCR overproduced in *E. coli* Arctic (DE3) RP: lane 1, lysate; lane 2, soluble supernatant; lane 3, insoluble precipitate; lane 4, refolded protein; lane 5, protein obtained after purification using an immobilized Cu^{2+} resin. (B) GCR activity of the purified protein as a function of bis-γ-glutamylcystine concentration.

dx.doi.org/10.1021/bi4003343 | *Biochemistry* 2013, 52, 2905–2913

seq no.	ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	DLDH_AZOVI	100																											
2	DLDH_MYCTU	37	100																										
3	DLDH_NEIME	40	36	100																									
4	DLDH_PSEFL	84	37	39	100																								
5	DLDH_GEOST	42	38	43	44	100																							
6	DLDH_YEAST	44	37	39	44	43	100																						
7	GR_ECOLI	27	28	28	28	29	29	100																					
8	GR_BARHE	25	28	27	26	29	25	35	100																				
9	GR_CHRGR	30	29	27	31	30	28	51	38	100																			
10	GR_YEAST	26	27	25	27	29	28	49	32	44	100																		
11	GCR_HNRC1	24	28	23	24	26	24	27	24	27	24	100																	
12	GCR_HALWA	22	27	24	24	29	23	27	23	25	24	58	100																
13	GCR_HALXA	24	27	25	25	28	25	25	23	29	24	57	60	100															
14	GCR_HALMU	26	29	25	28	29	27	26	23	28	25	57	62	59	100														
15	MYTR_MYCTU	28	30	29	27	29	26	31	28	28	27	25	26	24	26	100													
16	MYTR_CORGL	30	30	29	31	30	27	32	30	32	29	27	26	22	28	52	100												
17	MYTR_CORDI	29	33	27	29	29	30	31	29	30	24	27	25	23	27	55	66	100											
18	MYTR_RHOER	30	31	29	31	29	0	34	28	32	26	25	24	24	26	65	55	57	100										
19	MYTR_AMYME	31	32	31	30	30	29	32	29	31	28	26	23	24	27	56	50	53	57	100									
20	TYTR_TRYCR	30	28	26	29	28	25	36	38	35	33	23	22	20	23	27	25	24	25	27	100								
21	TYTR_TRYBB	28	28	26	29	29	26	37	38	34	34	23	22	21	24	25	26	24	26	27	82	100							
22	TYTR_CRIFA	28	28	28	28	27	26	36	38	35	32	24	23	23	24	26	26	25	26	29	69	69	100						
23	TYTR_LEIIN	27	27	27	26	25	24	34	39	32	31	22	22	20	23	25	24	24	25	27	64	65	75	100					
24	MERA_TN501	24	25	25	25	28	25	26	25	27	22	22	24	22	24	24	24	24	23	24	21	25	25	22	100				
25	MERA_BACCE	23	22	22	25	24	26	23	24	23	22	21	20	22	21	21	22	23	24	23	20	20	20	20	37	100			
26	MERA_STRLI	29	32	32	31	31	31	29	30	32	26	27	27	27	28	34	31	30	35	32	25	26	25	27	39	36	100		
27	MERA_SULSO	28	30	28	30	32	30	29	27	28	25	27	25	27	28	26	29	29	29	27	26	29	27	27	30	28	35	100	
28	MERA_HALMA	28	32	32	28	31	30	28	26	29	26	28	28	29	28	28	29	31	31	31	26	25	27	25	29	28	42	37	100

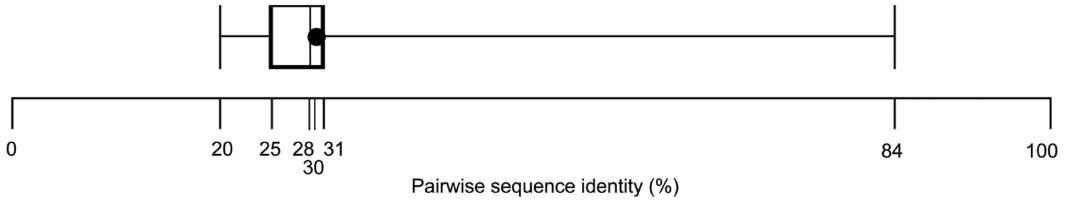


Figure 4. Pairwise sequence identities between selected proteins belonging to different families in the pyridine nucleotide disulfide reductase superfamily: dihydrolipoamide dehydrogenase (DLDH), glutathione reductase (GR), bis-γ-glutamylcystine reductase (GCR), mycothione reductase (MYTR), trypanothione reductase (TYTR), and mercuric reductase (MERA). The box-and-whisker plot shows the minimum, one-quarter quartile, median, average (●), three-quarter quartile, and maximal pairwise sequence identities among these sequences. The accession numbers for the proteins shown are listed in Table S4 of the Supporting Information.

Supporting Information). The His_6-tagged GCR bound more tightly to this column than the native enzyme (Figure S2 of the Supporting Information), probably because of the binding of the N-terminal His_6 tag to the resin. The purified protein reduced bis-γ-glutamylcystine effectively, with a k_{cat} of 54 ± 8 s^{-1}, a K_M of 1.1 ± 0.1 mM, and a k_{cat}/K_M of (4.9 ± 0.9) × 10^4 M^{-1} s^{-1} (Figure 3B). These kinetic parameters agree well with those reported by Sundquist and Fahey (k_{cat} of 28 s^{-1}, K_M of 0.81 mM, and k_{cat}/K_M of 3.5 × 10^4 M^{-1} s^{-1}).[12]

Purified GCR Does Not Have Mercuric Reductase Activity. Because the gene encoding GCR is currently annotated as *merA*, we measured the mercuric reductase activity of the protein by following the oxidation of NADPH at 340 nm and room temperature.[13] Assays were conducted in 50 mM sodium phosphate (pH 6.7) containing 3 M KCl, 1.3 M NaCl, 1 mM EDTA, 0.34 mM NADPH, and up to 1 mM $HgCl_2$. No activity was observed over 5 min in the presence of 0.6 μM enzyme, whereas GCR reductase activity was easily detectable over 30 s in the presence of 0.06 μM enzyme. Further, GCR activity was completely inhibited by addition of 1 mM $HgCl_2$ (Figure S3 of the Supporting Information). This finding is consistent with previous reports showing that GCR is inhibited by many divalent metal ions, including Cu^{2+}, Co^{2+}, and Hg^{2+}.[9]

GCR Belongs to the Pyridine Nucleotide Disulfide Oxidoreductase Family. The sequence of GCR has highly significant matches to the FAD/NAD(P) binding domain (PFAM, PF07992) and the dimerization domain (PFAM, PF02582) of the pyridine nucleotide disulfide oxidoreductase family; the *E* values are 8.3 × 10^{-19} and 3.43 × 10^{-13}, respectively. PROSITE[24] recognized a pattern for the class I pyridine nucleotide disulfide oxidoreductase active site, and PRINTS[25] reported a set of motifs as a grouped signature for the class I pyridine nucleotide disulfide reductases.

Proteins in the pyridine nucleotide disulfide oxidoreductase family catalyze reduction of a wide range of disulfide substrates, and their sequences are very divergent (Figure 4). However, all members of the family share a common mechanism, which is initiated by the transfer of a hydride from a pyridine nucleotide cofactor to flavin adenine dinucleotide (FAD), followed by delivery of reducing equivalents to a cysteine of the active site disulfide and ultimately to the substrate disulfide or, in the case of mercuric reductase, Hg^{2+}.[26]

Figure 5 shows a multiple-sequence alignment of *Halobacterium* sp. NRC-1 GCR and closely related putative GCRs from other halobacteria with sequences of known pyridine nucleotide disulfide oxidoreductase family members, including glutathione reductases, mycothione reductases, trypanothione reductases, dihydrolipoylamide dehydrogenases, and mercuric reductases. (All of these proteins belong to PFAM family PF07992.) Conserved sequence motifs known to interact with the two cofactors, FAD and NADPH, are highlighted. Most of the

dx.doi.org/10.1021/bi4003343 | *Biochemistry* 2013, 52, 2905−2913

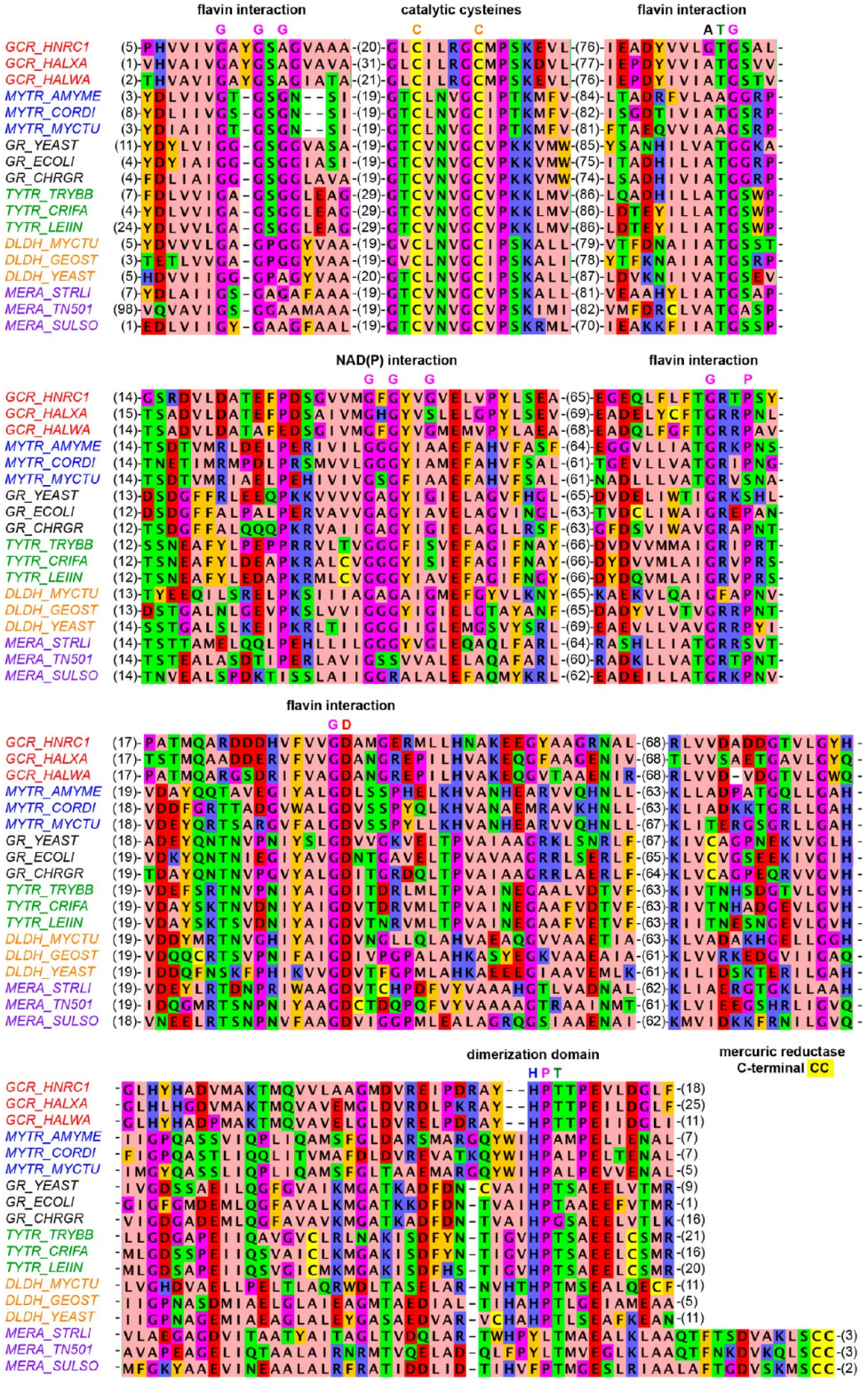


Figure 5. Multiple-sequence alignment of GCRs from several halobacteria with sequences from other subfamilies of the pyridine nucleotide disulfide reductase family, including dihydrolipoamide reductases (LPDA), glutathione reductases (GR), mycothione reductases (MYTR), trypanothione reductases (TYTR), and mercuric reductases (MERA). Residues found in conserved motifs involved in interactions with the flavin or NAD(P), catalytic cysteines, and the C-terminal two-cysteine motif of the mercuric reductase are indicated above the alignment.[26,43] Accession numbers for the proteins shown are listed in Table S4 of the Supporting Information.

sequences also share the C-terminal dimerization domain with a signature HPT sequence. The exception is the group of mercuric reductases, which have a distinctive C-terminal domain containing two cysteine residues that are involved in binding Hg(II) at the active site. The multiple-sequence alignment and the conservation of several motifs in GCR support its inclusion in the pyridine nucleotide disulfide oxidoreductase family.

dx.doi.org/10.1021/bi4003343 | *Biochemistry* 2013, 52, 2905−2913

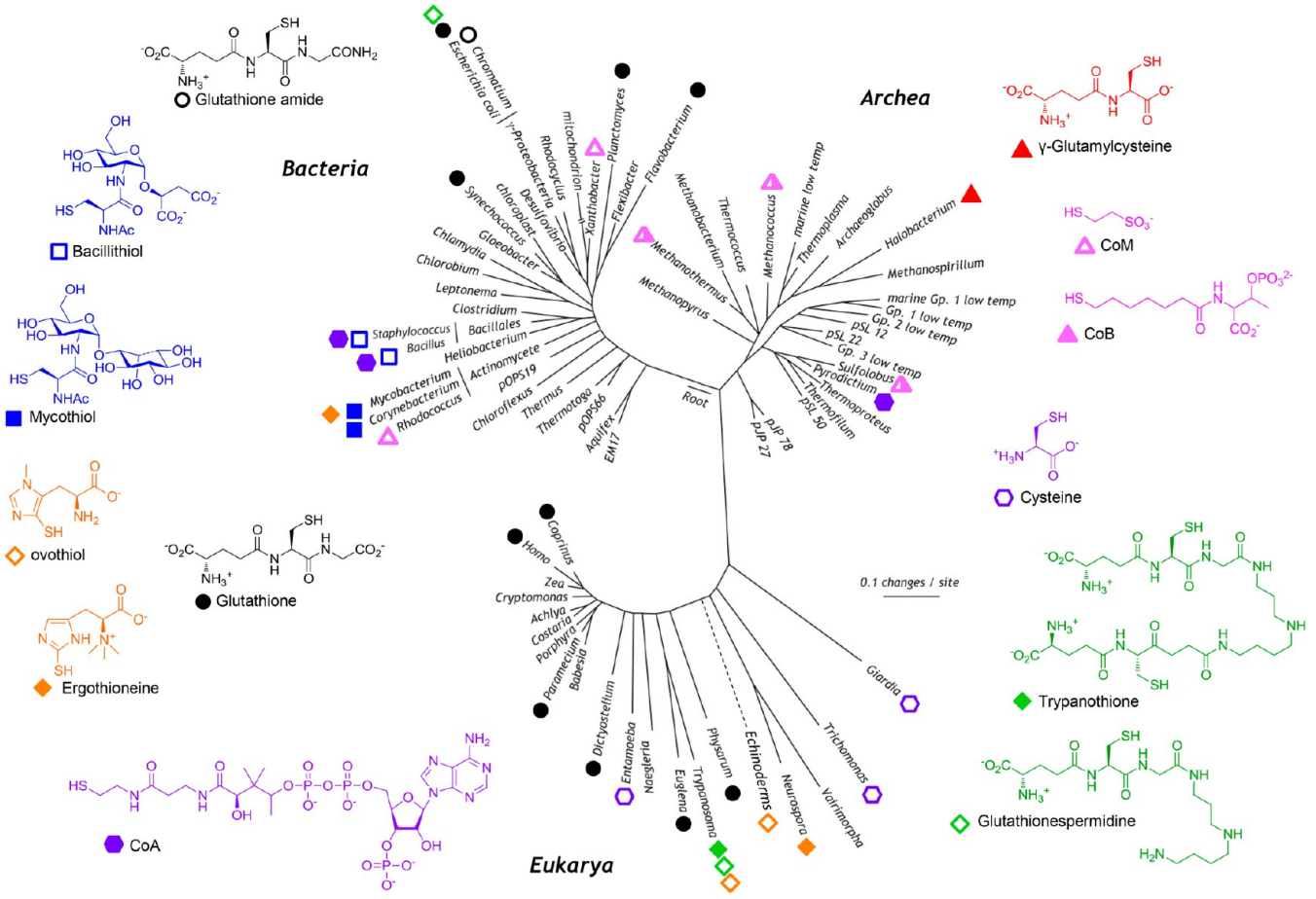


Figure 6. Major low-molecular weight thiols used by diverse lineages mapped onto a phylogenetic tree constructed from 16S rRNA sequences (courtesy of N. Pace). Note that CoA is used as a cofactor in all domains of life but also serves the roles of typical low-molecular weight thiols in *Staphylococcus*, *Bacillus*,[44,45] and *Pyrodictium*.[46] Similarly, cysteine serves as the primary low-molecular weight thiol in *Giardia*, *Entamoeba*, and *Trichomonas*.[35] References for the occurrence of other thiols: glutathione amide,[47] mycothiol,[48,49] bacillithiol,[50,51] trypanothione and glutathionylspermidine,[52] glutathionylspermidine in *E. coli*,[53] ergothioneine,[54−56] CoM-SS-CoB,[57,58] and CoM.[59,60]

DISCUSSION

Low-molecular weight thiols serve many important roles in cells. They act as redox buffers for the maintenance of the redox state of molecules in the cell. They reduce disulfide bonds caused by oxidation of cellular thiols and react with alkylating reagents, thus protecting DNA and proteins.[27,28] Thiols can serve as substrates in enzymatic reactions[29,30] and participate in the regulation of protein function and cell signaling.[31−33] Although the use of low-molecular weight thiols for such purposes is common, there is extraordinary diversity among the structures used by different evolutionary lineages (see Figure 6).[31,32,34,35] Further diversity is found in the enzymes that regenerate the thiols after they are oxidized. Most characterized thiol disulfide reductases, including glutathione reductase, trypanothione reductase, and mycothione reductase, belong to the pyridine nucleotide disulfide oxidoreductase family within the two-dinucleotide binding domain flavoprotein (tDBDF) superfamily[26] and use either NADPH or NADH as a hydride donor. In the case of ovothiol, which is found in sea urchin eggs,[36] the corresponding disulfide is reduced by glutathione rather than a reductase protein. In protozoan parasites, ovothiol disulfide can be reduced by trypanothione.[37] Thus, various systems for using thiols to protect against oxidative damage appear to have evolved convergently in different lineages long after the divergence of the LUCA into the bacterial, archaeal, and eukaryal domains.

Halobacteria are unique in their use of γ-Glu-Cys as a major low-molecular weight thiol.[38] We have previously postulated that the ability to make γ-Glu-Cys arose in halobacteria via horizontal gene transfer of a gene encoding γ-glutamyl cysteine ligase (GshA) from a cyanobacterium.[39] Typically, γ-Glu-Cys is converted to glutathione, the major thiol found in eukaryotes and Gram-negative bacteria, by glutathione synthetase. γ-Glu-Cys lacks the glycine residue that is present in glutathione. This discrepancy may be related to the high salt content of the *Halobacterium* cytoplasm. Cysteine residues are susceptible to autoxidation, which is catalyzed by heavy metal ions complexed by the thiol, amino, and carboxylate groups.[40] In glutathione, the amino and carboxylate groups of cysteine are involved in amide bonds with glutamate and glycine, which substantially decreases the rate of autoxidation. The presence of a high salt concentration decreases the rate of autoxidation of Cys, so formation of amide bonds to glutamate and glycine is less critical. Curiously, γ-Glu-Cys is actually more stable than glutathione in the presence of a high salt concentration.[12] Thus, the simpler thiol serves perfectly well in the halobacteria, and there has apparently been no selective pressure to expend energy and carbon to add an additional glycine residue.[39]

dx.doi.org/10.1021/bi4003343 | *Biochemistry* 2013, 52, 2905−2913

Genes encoding closely related homologues of *Halobacterium* sp. NRC-1 GCR are found in the genomes of 12 of the 18 halobacteria for which full genome sequences are available (Figure 7). Surprisingly, we could not detect homologues of

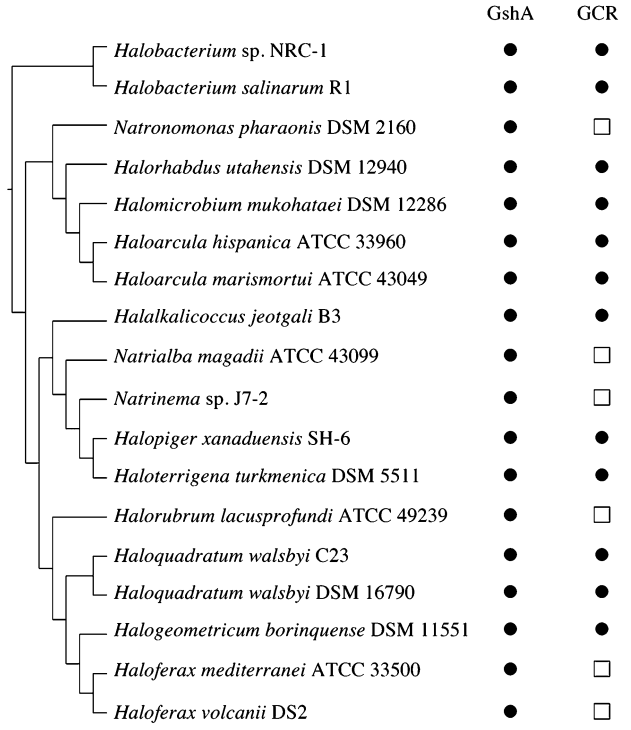


Figure 7. Occurrence of γ-glutamylcysteine synthase (GshA) and bis-γ-glutamylcystine reductase (GCR) among halobacteria. The presence (●) or absence (□) of each protein is indicated for each species on the phylogenetic cladogram inferred from whole genome information using CVTree version 2.[61]

GCR from six halobacteria. Each of these species has a homologue of GshA that is 60−70% identical to *Halobacterium* sp. NRC-1 GshA, so presumably all are capable of making γ-Glu-Cys. Halobacteria that lack a homologue of GCR may have a nonhomologous enzyme that serves this function. Alternatively, these Archaea may use a different low-molecular weight thiol, possibly one derived from γ-Glu-Cys. It is intriguing that there is such diversity even within the *Halobacterium* clade.

Halobacterium sp. NRC-1 GCR belongs to the pyridine nucleotide disulfide oxidoreductase family. This makes a great deal of sense, given the ability of all enzymes in the family to reduce a disulfide bond using electrons derived from NADPH that are passed through a flavin and a disulfide on the enzyme before reaching the substrate. What is surprising is the high level of sequence divergence among the family members (Figure 4), which suggests that this family has been evolving for a very long period of time and makes phylogenetic analysis difficult. Enzymes involved in the synthesis of low-molecular weight thiols and the reduction of the corresponding disulfides likely evolved at the time O_2 began to appear in the atmosphere[39] more than 2.5 billion years ago.[41] Dihydrolipoamide dehydrogenase, which is a component of pyruvate dehydrogenase, the α-ketoglutarate dehydrogenase complex, and the glycine cleavage system, was likely present in the last universal common ancestor[42] and may have been the progenitor of the family of pyridine nucleotide disulfide reductases that now includes glutathione reductase, trypanothione reductase, mycothione reductase, mercuric reductase, and now GCR.

■ ASSOCIATED CONTENT

Supporting Information

Purification of GCR from *Halobacterium* sp. NRC-1, purification of native GCR from *Halobacterium* sp. NRC-1 (Table S1), peptides identified by ESI-MS/MS analysis of GCR (Table S2), purification of recombinant N-His_6-GCR (Table S3), accession numbers for sequences used in the multiple-sequence alignment shown in Figure 5 (Table S4), SDS−PAGE analysis of samples taken during purification of GCR from *Halobacterium* sp. NRC-1 (Figure S1), elution profiles of native GCR and N-His_6-GCR expressed in *E. coli* on immobilized Cu^{2+} resin (Figure S2), and inhibition of GCR activity by Hg^{2+} (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Department of Molecular, Cellular and Developmental Biology, University of Colorado Boulder, Boulder, CO 80309. Telephone: (303) 492-6328. Fax: (303) 492-1149. E-mail: shelley.copley@colorado.edu.

Funding

This work was supported by National Institutes of Health Grant RO1 GM083285 to S.D.C.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

GCR, bis-γ-glutamylcystine reductase; γ-Glu-Cys, γ-glutamylcysteine; NMR, nuclear magnetic resonance spectroscopy; NADPH, nicotinamide adenine dinucleotide phosphate; NAD, nicotinamide adenine dinucleotide; ESI-MS/MS, electrospray ionization tandem mass spectrometry; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; AEBSF, 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride; EDTA, ethylenediaminetetraacetic acid; IPTG, isopropyl β-D-thiogalactopyranoside; tDBDF, two-dinucleotide binding domain flavoproteins.

■ REFERENCES

(1) Doerks, T., van Noort, V., Minguez, P., and Bork, P. (2012) Annotation of the *M. tuberculosis* hypothetical orfeome: Adding functional information to more than half of the uncharacterized proteins. *PLoS One 7*, e34302.

(2) Green, M. L., and Karp, P. D. (2007) Using genome-context data to identify specific types of functional associations in pathway/genome databases. *Bioinformatics 23*, i205−i211.

(3) Baughman, J. M., Perocchi, F., Girgis, H. S., Plovanich, M., Belcher-Timme, C. A., Sancak, Y., Bao, X. R., Strittmatter, L., Goldberger, O., Bogorad, R. L., Koteliansky, V., and Mootha, V. K. (2011) Integrative genomics identifies MCU as an essential component of the mitochondrial calcium uniporter. *Nature 476*, 341−345.

(4) Schnoes, A. M., Brown, S. D., Dodevski, I., and Babbitt, P. C. (2009) Annotation error in public databases: Misannotation of molecular function in enzyme superfamilies. *PLoS Comput. Biol. 5*, e1000605.

(5) Lespinet, O., and Labedan, B. (2006) Puzzling over orphan enzymes. *Cell. Mol. Life Sci. 63*, 517−523.

(6) Pouliot, Y., and Karp, P. D. (2007) A survey of orphan enzyme activities. *BMC Bioinf. 8*, 244.
(7) Hitchcock, D. S., Fedorov, A. A., Fedorov, E. V., Dangott, L. J., Almo, S. C., and Raushel, F. M. (2011) Rescue of the orphan enzyme isoguanine deaminase. *Biochemistry 50*, 5555−5557.
(8) Hall, R. S., Agarwal, R., Hitchcock, D., Sauder, J. M., Burley, S. K., Swaminathan, S., and Raushel, F. M. (2010) Discovery and structure determination of the orphan enzyme isoxanthopterin deaminase. *Biochemistry 49*, 4374−4382.
(9) Sundquist, A. R., and Fahey, R. C. (1988) The novel disulfide reductase bis-γ-glytamylcystine reductase and dihydrolipoamide dehydrogenase from *Halobacterium halobium*: Purification by immobilized-metal-ion affinity chromatography and properties of the enzymes. *J. Bacteriol. 170*, 3459−3467.
(10) DasSarma, S., Fleischmann, E. M., and Rodriguez-Valera, F. (1995) Media for halophiles. In *Archaea: A Laboratory Manual—Halophiles* (Robb, F. T., Place, A. R., Sowers, K. R., Schreier, H. J., Fleischmann, E. M., DasSarma, S., and Fleischmann, E. M., Eds.) pp 225−230, Cold Spring Harbor Laboratory Press, Plainview, NY.
(11) Strumeyer, D., and Bloch, K. (1962) γ-L-Glutamyl-L-cysteine from glutathione. *Biochem. Prep. 9*, 52−55.
(12) Sundquist, A. R., and Fahey, R. C. (1989) The function of γ-glutamylcysteine and bis-γ-glutamylcystine reductase in *Halobacterium halobium*. *J. Biol. Chem. 264*, 719−725.
(13) Fox, B., and Walsh, C. T. (1982) Mercuric reductase. Purification and characterization of a transposon-encoded flavoprotein containing an oxidation-reduction-active disulfide. *J. Biol. Chem. 257*, 2498−2503.
(14) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*, 248−254.
(15) Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
(16) Connaris, H., Chaudhuri, J. B., Danson, M. J., and Hough, D. W. (1999) Expression, reactivation, and purification of enzymes from *Haloferax volcanii* in *Escherichia coli*. *Biotechnol. Bioeng. 64*, 38−45.
(17) Zdobnov, E. M., and Apweiler, R. (2001) InterProScan: An integration platform for the signature-recognition methods in InterPro. *Bioinformatics 17*, 847−848.
(18) Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res. 38*, W695−W699.
(19) Edgar, R. C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res. 32*, 1792−1797.
(20) Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet. 16*, 276−277.
(21) Moore, M. J., Miller, S. M., and Walsh, C. T. (1992) C-terminal cysteines of Tn501 mercuric ion reductase. *Biochemistry 31*, 1677−1685.
(22) Allers, T. (2010) Overexpression and purification of halophilic proteins in *Haloferax volcanii*. *Bioengineered Bugs 1*, 288−290.
(23) Sakakibara, N., Han, M., Rollor, C. R., Gilson, R. C., Busch, C., Heo, G., and Kelman, Z. (2008) Cloning, Purification, and Partial Characterization of the *Halobacterium sp*. NRC-1 minichromosome Maintenance (MCM) Helicase. *Open Microbiol. J. 2*, 13−17.
(24) Sigrist, C. J., Cerutti, L., de Castro, E., Langendijk-Genevaux, P. S., Bulliard, V., Bairoch, A., and Hulo, N. (2010) PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res. 38*, D161−D166.
(25) Attwood, T. K., Beck, M. E., Bleasby, A. J., and Parry-Smith, D. J. (1994) PRINTS: A database of protein motif fingerprints. *Nucleic Acids Res. 22*, 3590−3596.
(26) Ojha, S., Meng, E. C., and Babbitt, P. C. (2007) Evolution of function in the "two dinucleotide binding domains" flavoproteins. *PLoS Comput. Biol. 3*, e121.
(27) Chan, J. Y., Ruchirawat, M., Lapeyre, J. N., and Becker, F. F. (1983) The protective role of thiol reducing agents in the in vitro inhibition of rat liver DNA methylase by direct acting carcinogens. *Carcinogenesis 4*, 1097−1100.
(28) Shan, X. Q., Aw, T. Y., and Jones, D. P. (1990) Glutathione-dependent protection against oxidative injury. *Pharmacol. Ther. 47*, 61−71.
(29) Warner, J. R., Lawson, S. L., and Copley, S. D. (2005) A mechanistic investigation of the thiol-disulfide exchange step in the reductive dehalogenation catalyzed by tetrachlorohydroquinone dehalogenase. *Biochemistry 44*, 10360−10368.
(30) Oakley, A. (2011) Glutathione transferases: A structural perspective. *Drug Metab. Rev. 43*, 138−151.
(31) Fahey, R. C. (2013) Glutathione analogs in prokaryotes. *Biochim. Biophys. Acta 1830*, 3182−3198.
(32) Smirnova, G. V., and Oktyabrsky, O. N. (2005) Glutathione in bacteria. *Biochemistry (Moscow, Russ. Fed.) 70*, 1199−1211.
(33) Dalle-Donne, I., Rossi, R., Colombo, G., Giustarini, D., and Milzani, A. (2009) Protein S-glutathionylation: A regulatory device from bacteria to humans. *Trends Biochem. Sci. 34*, 85−96.
(34) Masip, L., Veeravalli, K., and Georgiou, G. (2006) The many faces of glutathione in bacteria. *Antioxid. Redox Signaling 8*, 753−762.
(35) Van Laer, K., Hamilton, C. J., and Messens, J. (2013) Low molecular weight thiols in thiol disulfide exchange. *Antioxid. Redox Signaling 18*, 1642−1653.
(36) Turner, E., Hager, L. J., and Shapiro, B. M. (1988) Ovothiol replaces glutathione peroxidase as a hydrogen peroxide scavenger in sea urchin eggs. *Science 242*, 939−941.
(37) Ariyanayagam, M. R., and Fairlamb, A. H. (2001) Ovothiol and trypanothione as antioxidants in trypanosomatids. *Mol. Biochem. Parasitol. 115*, 189−198.
(38) Newton, G. L., and Javor, B. (1985) γ-Glutamylcysteine and thiosulfate are the major low-molecular-weight thiols in halobacteria. *J. Bacteriol. 161*, 438−441.
(39) Copley, S. D., and Dhillon, J. K. (2002) Lateral gene transfer and parallel evolution in the history of glutathione biosynthesis genes. *Genome Biol. 3*, research0025.1−research0025.16.
(40) Fahey, R. C., and Sundquist, A. R. (1991) Evolution of glutathione metabolism. *Adv. Enzymol. Relat. Areas Mol. Biol. 64*, 1−53.
(41) Anbar, A. D., Duan, Y., Lyons, T. W., Arnold, G. L., Kendall, B., Creaser, R. A., Kaufman, A. J., Gordon, G. W., Scott, C., Garvin, J., and Buick, R. (2007) A whiff of oxygen before the great oxidation event? *Science 317*, 1903−1906.
(42) Ouzounis, C. A., Kunin, V., Darzentas, N., and Goldovsky, L. (2006) A minimal estimate for the gene content of the last universal common ancestor: Exobiology from a terrestrial perspective. *Res. Microbiol. 157*, 57−68.
(43) Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R., and Bateman, A. (2010) The Pfam protein families database. *Nucleic Acids Res. 38*, D211−D222.
(44) Nicely, N. I., Parsonage, D., Paige, C., Newton, G. L., Fahey, R. C., Leonardi, R., Jackowski, S., Mallett, T. C., and Claiborne, A. (2007) Structure of the type III pantothenate kinase from *Bacillus anthracis* at 2.0 Å resolution: Implications for coenzyme A-dependent redox biology. *Biochemistry 46*, 3234−3245.
(45) Mallett, T. C., Wallen, J. R., Karplus, P. A., Sakai, H., Tsukihara, T., and Claiborne, A. (2006) Structure of coenzyme A-disulfide reductase from *Staphylococcus aureus* at 1.54 Å resolution. *Biochemistry 45*, 11278−11289.
(46) Harris, D. R., Ward, D. E., Feasel, J. M., Lancaster, K. M., Murphy, R. D., Mallet, T. C., and Crane, E. J., III (2005) Discovery and characterization of a Coenzyme A disulfide reductase from *Pyrococcus horikoshii*. Implications for this disulfide metabolism of anaerobic hyperthermophiles. *FEBS J. 272*, 1189−1200.
(47) Vergauwen, B., Pauwels, F., Jacquemotte, F., Meyer, T. E., Cusanovich, M. A., Bartsch, R. G., and Van Beeumen, J. J. (2001) Characterization of glutathione amide reductase from *Chromatium*

gracile. Identification of a novel thiol peroxidase (Prx/Grx) fueled by glutathione amide redox cycling. *J. Biol. Chem. 276*, 20890−20897.
(48) Newton, G. L., Arnold, K., Price, M. S., Sherrill, C., Delcardayre, S. B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R. C., and Davis, C. (1996) Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes. *J. Bacteriol. 178*, 1990−1995.
(49) Patel, M. P., and Blanchard, J. S. (1999) Expression, purification, and characterization of *Mycobacterium tuberculosis* mycothione reductase. *Biochemistry 38*, 11827−11833.
(50) Newton, G. L., Rawat, M., La Clair, J. J., Jothivasan, V. K., Budiarto, T., Hamilton, C. J., Claiborne, A., Helmann, J. D., and Fahey, R. C. (2009) Bacillithiol is an antioxidant thiol produced in Bacilli. *Nat. Chem. Biol. 5*, 625−627.
(51) Gaballa, A., Newton, G. L., Antelmann, H., Parsonage, D., Upton, H., Rawat, M., Claiborne, A., Fahey, R. C., and Helmann, J. D. (2010) Biosynthesis and functions of bacillithiol, a major low-molecular-weight thiol in Bacilli. *Proc. Natl. Acad. Sci. U.S.A. 107*, 6482−6486.
(52) Lantwin, C. B., Schlichting, I., Kabsch, W., Pai, E. F., and Krauth-Siegel, R. L. (1994) The structure of *Trypanosoma cruzi* trypanothione reductase in the oxidized and NADPH reduced state. *Proteins 18*, 161−173.
(53) Smith, K., Borges, A., Ariyanayagam, M. R., and Fairlamb, A. H. (1995) Glutathionylspermidine metabolism in *Escherichia coli*. *Biochem. J. 312* (Part2), 465−469.
(54) Ta, P., Buchmeier, N., Newton, G. L., Rawat, M., and Fahey, R. C. (2011) Organic hydroperoxide resistance protein and ergothioneine compensate for loss of mycothiol in *Mycobacterium smegmatis* mutants. *J. Bacteriol. 193*, 1981−1990.
(55) Seebeck, F. P. (2010) In vitro reconstitution of mycobacterial ergothioneine biosynthesis. *J. Am. Chem. Soc. 132*, 6632−6633.
(56) Bello, M. H., Barrera-Perez, V., Morin, D., and Epstein, L. (2012) The *Neurospora crassa* mutant NcΔEgt-1 identifies an ergothioneine biosynthetic gene and demonstrates that ergothioneine enhances conidial survival and protects against peroxide toxicity during conidial germination. *Fungal Genet. Biol. 49*, 160−172.
(57) Bennati, M., Weiden, N., Dinse, K. P., and Hedderich, R. (2004) ^{57}Fe ENDOR spectroscopy on the iron-sulfur cluster involved in substrate reduction of heterodisulfide reductase. *J. Am. Chem. Soc. 126*, 8378−8379.
(58) Mander, G. J., Duin, E. C., Linder, D., Stetter, K. O., and Hedderich, R. (2002) Purification and characterization of a membrane-bound enzyme complex from the sulfate-reducing archaeon *Archaeoglobus fulgidus* related to heterodisulfide reductase from methanogenic archaea. *Eur. J. Biochem. 269*, 1895−1904.
(59) Kofoed, M. A., Wampler, D. A., Pandey, A. S., Peters, J. W., and Ensign, S. A. (2011) Roles of the redox-active disulfide and histidine residues forming a catalytic dyad in reactions catalyzed by 2-ketopropyl coenzyme M oxidoreductase/carboxylase. *J. Bacteriol. 193*, 4904−4913.
(60) Krum, J. G., and Ensign, S. A. (2000) Heterologous expression of bacterial Epoxyalkane:Coenzyme M transferase and inducible coenzyme M biosynthesis in *Xanthobacter* strain Py2 and *Rhodococcus rhodochrous* B276. *J. Bacteriol. 182*, 2629−2634.
(61) Xu, Z., and Hao, B. (2009) CVTree update: A newly designed phylogenetic study platform using composition vectors and whole genomes. *Nucleic Acids Res. 37*, W174−W178.