Misler, S., & Hurlbut, W. P. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 991-995.
Nicholls, D. G., Rugolo, M., Scott, I. G., & Meldolesi, J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7924-7928.
Robello, M., Rolandi, R., Alemã, S., & Grasso, A. (1984) *Proc. R. Soc. London, Ser. B 220*, 477-487.
Tzeng, M. C., & Siekevitz, P. (1979) *J. Neurochem. 33*, 263-274.
Valtorta, F., Madeddu, L., Meldolesi, J., & Ceccarelli, B. (1984) *J. Cell Biol. 99*, 124-132.
Vicentini, L. M., & Meldolesi, J. (1984) *Biochem. Biophys. Res. Commun. 121*, 538-544.

Glutathione Reductase from *Escherichia coli*: Cloning and Sequence Analysis of the Gene and Relationship to Other Flavoprotein Disulfide Oxidoreductases†

Shaun Greer and Richard N. Perham*
Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, England
Received October 17, 1985; Revised Manuscript Received January 9, 1986

ABSTRACT: A glutathione reductase negative strain of *Escherichia coli* K-12 was isolated as a thermoresistant survivor when a *gor*::Mu*cts*Ap lysogen was subjected to elevated temperature. It was found that in addition to being ampicillin sensitive this mutant was hypersensitive to arsenate, which may be connected with the fact that the *gor* gene maps between 77 and 78 min on the *E. coli* genome, close to the *pit* locus encoding the major arsenate transport system of *E. coli*. A derivative of this mutant was used as the recipient in a screen of the Clarke and Carbon hybrid plasmid bank of *E. coli* DNA. A plasmid, pGR, was isolated that encodes both an arsenate-resistance element and glutathione reductase. Restriction mapping of this plasmid showed that the insert DNA is approximately 10 kilobase pairs in length, and a fragment of the *gor* gene was identified that allowed the *gor* gene to be accurately mapped on pGR by a combination of restriction analysis and Southern blotting. The DNA sequence of the *gor* gene was determined and found to encode a protein of 450 amino acid residues. The glutathione reductase of *E. coli* is very homologous to the human enzyme and is also related (though less closely) to other flavoprotein disulfide oxidoreductases whose sequences are available. These enzymes have retained a common mechanism while evolving different specificities.

Glutathione reductase (EC 1.6.4.2) is a widespread enzyme that catalyzes the reduction of oxidized glutathione by NADPH:

$$GSSG + NADPH + H^+ \rightarrow NADP^+ + 2GSH$$

It is a member of an important class of flavoprotein enzymes, the disulfide oxidoreductases, which appear to share a common catalytic mechanism. Other members of the class are dihydrolipoamide dehydrogenase (EC 1.6.4.3) (Reed, 1974; Packman et al., 1984), thioredoxin reductase (EC 1.6.4.5) (Holmgren, 1980), and mercuric reductase (Fox & Walsh, 1982). Studies of their amino acid sequences reveal homologies between dihydrolipoamide dehydrogenase, glutathione reductase, and mercuric reductase, implying that they have arisen by divergent evolution from a common ancestor (Perham et al., 1978; Williams et al., 1982; Fox & Walsh, 1983). On the other hand, thioredoxin reductase is sufficiently different for it to be likely that this enzyme has arisen by convergent evolution toward a common mechanism (Perham et al., 1978).

The *lpd* gene of *Escherichia coli*, encoding dihydrolipoamide dehydrogenase (Stephens et al., 1983), and the *mer*A gene of transposon Tn 501 from *Pseudomonas aeruginosa*, encoding mercuric reductase (Brown et al., 1983), have been cloned and their nucleotide sequences determined. This in turn has enabled their primary structures to be inferred. The complete amino acid sequence (Krauth-Siegel et al., 1982) and X-ray crystallographic structure at 2-Å resolution (Thieme et al., 1981) of human glutathione reductase are also known, which has shed considerable light on the reaction mechanism (Pai & Schulz, 1983).

Glutathione is not an essential metabolite in *E. coli* (Apontoweil & Berends, 1975; Fuchs & Warner, 1975). Mutants (*gor*) that lack glutathione reductase have been isolated from *E. coli* mutated by bacteriophage Mu insertion and shown to be similarly unimpaired in growth (Davis et al., 1982). This phenotype has placed considerable difficulty in the way of attempts to clone the *gor* gene, but we now describe further experiments with the *gor* mutation that have allowed the isolation of a plasmid carrying the *E. coli gor* gene from the Clarke and Carbon (1976) bank of recombinant plasmids. The determination of the DNA sequence of this gene was undertaken to add another flavoprotein oxidoreductase sequence to those few already known. Apart from advancing our understanding of the mechanism and evolutionary history of this important group of enzymes, this should also permit more stringent tests of mechanism to be applied by means of in vitro mutagenesis.

Materials and Methods

Materials. Complex bacteriological media were from Difco Laboratories Ltd., East Molesey, Surrey. NADPH, 5,5′-dithiobis(2-nitrobenzoic acid), and oxidized glutathione were supplied by Sigma Chemical Co. Ltd., Poole, Dorset. Agarose and low melting point agarose (Seaplaque) were from Miles

† This work was supported by a grant from the Science and Engineering Research Council to R.N.P. S.G. was supported by a studentship from the Department of Education for Northern Ireland.

0006-2960/86/0425-2736$01.50/0

Scientific, Stoke Poges, Slough, Berkshire. Restriction enzymes were obtained from New England Biolabs (U.K. distributors CP Laboratories Ltd., Bishops Stortford, Hertfordshire). NACS minicolumns were supplied by the manufacturer, BRL Ltd., Trident House, Paisley, Renfrewshire. Dithiothreitol, IPTG, and BCIG were also from BRL. Calf intestinal alkaline phosphatase (special quality for molecular biology), T4 DNA ligase, rATP, and the Klenow fragment of *E. coli* DNA polymerase I were from BCL, Boehringer Mannheim House, Lewes, East Sussex. T4 DNA polymerase was from Pharmacia P-L Biochemicals Ltd., Pharmacia House, Milton Keynes, Buckinghamshire, as were all nucleotides (except for radiolabeled dATP) used in DNA sequencing. [α^{35}S]dATP(S) (410 Ci/mmol) was obtained from Amersham International plc, Amersham, Buckinghamshire.

Bacterial Strains and Media. JM2267 is *gor*::Mu*cts*Ap (Davis et al., 1982) derived from JM2137, which is F$^-$ Δ(*his gnd*) Δ*lac ara*D Str^R. JA200 is F$^+$ *trp*E5 *rec*A *thr leu lac*Y (Clarke & Carbon, 1976). K10 is a wild-type *E. coli* K10 strain; such strains are known to be *pit* (Sprague et al., 1975). Nutrient broth contained 10 g/L Bacto tryptone, 5 g/L yeast extract, and 5 g/L NaCl, solidified if necessary with 1.5% (w/v) Bacto agar. Ampicillin, if required, was added at 10 mg/L. Other supplements are specified in the text. The Tris minimal medium used for tests of arsenate sensitivity contained 0.1 M Tris [tris(hydroxymethyl)aminomethane], 0.8 mM $MgSO_4$, 6.8 mM sodium citrate, 18.7 mM NH_4Cl, and 0.1 mM K_2HPO_4, final pH 7.5, supplemented with glucose (2 g/L) and, when necessary, amino acids (20 mg/L).

Glutathione Reductase Assays. Colonies on plates were screened for glutathione reductase activity after exposure to $CHCl_3$ vapor by assay with NADPH and 5,5′-dithiobis(2-nitrobenzoic acid) (Nbs_2) in situ, as described by Davis et al. (1982). In cell-free extracts (Davis et al., 1982), glutathione reductase activity was assayed by following the oxidation of NADPH (0.1 mM) in the presence of oxidized glutathione (1.2 mM), as described by Worthington and Rosemeyer (1976).

Selection of pit–gor Deletions. An overnight culture (0.25 mL) of strain JM2267 was spread on each of nine plates of nutrient agar supplemented with 20 mM sodium arsenate, 1 mM potassium phosphate, and 25 mM sodium citrate. The plates were incubated at 42 °C. As colonies arose, they were picked onto nutrient agar supplemented with 25 mM sodium citrate alone and tested for ampicillin resistance by replica plating onto nutrient agar containing ampicillin at a concentration of 10 mg/L. Strains apparently sensitive to ampicillin were streak-purified on nutrient–citrate medium and tested again.

Isolation of Strain SG8. An overnight culture (0.25 mL) of strain SG5 was plated onto nutrient agar supplemented with nalidixic acid (10 mg/L). After overnight growth at 37 °C, a resistant colony was selected and purified. This strain was designated SG8.

Screening of Clarke and Carbon Bank. This was performed by a simple replica mating technique. A lawn of mixed clones from the Clarke and Carbon (1976) bank of *E. coli* DNA was replicated onto a nutrient agar plate. After growth for 7 h, this fresh lawn was in turn replicated onto a plate of Tris minimal medium supplemented with 0.2 mM sodium arsenate and nalidixic acid (10 mg/L), spread with 0.1 mL of an overnight culture of strain SG8 concentrated 10-fold in 0.9% saline. Nalidixic acid was included to select against growth of the donor strain JA200. After incubation for 2 days at 30 °C, the plate was checked for the growth of colonies.

Isolation and Characterization of Plasmid DNA. Bacterial transformations, the rapid screening of plasmid-containing strains, and the isolation of large amounts of plasmid were carried out as described by Maniatis et al. (1982). Small amounts of plasmid were prepared by a quick-boiling method (Holmes & Quigley, 1981). Restriction enzyme digests were also prepared as described for the relevant enzyme by Maniatis et al. (1982), generally under conditions aimed at a 10–20-fold overdigestion. Digests were analyzed by electrophoresis on 0.8% agarose gels (Maniatis et al., 1982). Prime cut probes were made by the method of Hudson and Davidson (1984); Southern blotting and subsequent hybridizations were carried out as described by Maniatis et al. (1982).

DNA Sequence Determination and Analysis. NACS minicolumns were used to purify DNA from low melting point agarose as described by the manufacturer. DNA eluted from these columns was extracted twice with phenol and then precipitated with ethanol before use. The circularization of DNA and its subsequent sonication, the fractionation of sonicated DNA, and the construction of m13 clones were all carried out as described by Deininger (1983) with the exception that a NACS minicolumn was used. DNA from m13 clones was prepared and sequenced by standard techniques (Biggin et al., 1983; Sanger et al., 1980), and sequence data were assembled and analyzed with the DB system of Staden (1983).

The protein sequence of human glutathione reductase was aligned with that of *E. coli* glutathione reductase by means of a small specially written Algol 68C program implementing a simple dynamic programming algorithm (Sankoff & Kruskal, 1983) and using the score matrix described by Staden (1982). This is a modification of the MDM_{78} matrix (Schwartz & Dayhoff, 1978), which allows a numerical weight to be placed on the simple substitution of one amino acid by another. Deletions were dealt with by scoring each individual deletion element with a value of 5 (equivalent to an unfavored substitution); in addition, all deletions except those running through the ends of proteins were each given a block penalty of −25 (since all deletions are unlikely events, but those at the ends of a protein are less likely to have a significant effect on function).

Results

Construction of the Δgor Strain SG5. *E. coli* strain JM2267 is *gor*::Mu*cts*Ap (Davis et al., 1982) and is thus of limited value in cloning because it is temperature sensitive and ampicillin resistant. An attempt was therefore made to isolate a strain of JM2267 in which the *gor* gene and the Mu phage inserted in it had been deleted from the chromosome; this strain would be neither temperature sensitive nor ampicillin resistant. The *gor* gene is known to map between 77 and 78 min on the *E. coli* genome (Davis et al., 1982), close to the *pit* locus (Bachmann, 1983), which is one of the two *E. coli* transport systems for inorganic phosphate and in particular is the system largely responsible for the uptake of arsenate (Sprague et al., 1975). Strains in which the *pit* locus is mutated are thus relatively insensitive to arsenate in the growth medium. Growth of strain JM2267 on arsenate-containing medium at 42 °C will presumably select for cells that are neither temperature sensitive nor arsenate sensitive; these should include *pit–gor* deletion mutants.

The nutrient–arsenate medium used for the screen (see Materials and Methods) was found to provide only a weak counterselection against *pit*$^+$ cells. Strains isolated from this screen were tested on ampicillin-containing medium for the loss of the Mu phage. One strain was ampicillin-sensitive, but

Table I: Arsenate Sensitivity of Strains K10, JM2137, JM2267, and SG5[a]

arsenate concn (mM)	*E. coli* strain			
	K10	JM2137	JM2267	SG5
0	5	5	5	5
0.02	5	5	5	5
0.05	5	5	5	4
0.1	5	5	5	0
0.2	5	4	5	0
0.5	5	3	3	0
1.0	4	1	0	0

[a] Patches of the four strains were grown overnight at 30 °C on nutrient agar plates. Next day they were replica-plated onto plates containing Tris minimal medium supplemented with various concentrations of arsenate. After incubation at 30 °C for 48 h, the growth of each strain was assessed on an arbitrary scale from 0 to 5 (0 = no growth; 5 = maximum growth). The *pit* (and hence arsenate-resistant) strain K10 was used as a control.

when tested again on the original selection medium, it proved unexpectedly to be more (not less) arsenate sensitive than its parent *E. coli* JM2267. However, cell-free extracts of this strain were found to lack glutathione reductase activity. It presumably contains a deletion of the *gor* gene and neighboring DNA, which also somehow renders the cells sensitive to arsenate. The isolation of this strain, designated SG5, opened the way to a simple method for cloning the *gor* gene.

Characterization of Arsenate Sensitivity of Strain SG5. The *gor* mutation in *E. coli* has no known phenotype that could be utilized to select for cells containing glutathione reductase activity (Davis et al., 1982) and thereby to clone the *gor* gene from a gene bank. However, if sections of DNA conferring a selectable arsenate resistance could be cloned back into strain SG5, it was argued that the *gor* gene might be cloned too provided that the linkage between it and the arsenate-resistance element was close enough. For such experiments, a medium with a better defined arsenate:phosphate ratio was required, and it was found (Table I) that growth on a Tris minimal medium containing optimally 0.2 mM sodium arsenate would reliably select against strain SG5 in favor of strains JM2267 or JM2137, the parent of JM2267.

It was feared that reversion of the mutation conferring arsenate sensitivity on strain SG5 might interfere with such a selection. However, further experiments demonstrated that reversion was undetectable during growth of this strain on test plates of Tris minimal medium containing 0.2 mM sodium arsenate, designed to simulate a screen of the Clarke and Carbon (1976) gene bank. This ruled out any interference from revertants in the use of strain SG5 to screen the Clarke and Carbon (1976) gene bank for the putative arsenate-resistance gene(s) and the nearby *gor* gene. Prolonged growth of strain SG5 on this medium did lead to the appearance of revertants, but these grew much more slowly than strains JM2267 and JM2137 and were phenotypically distinct from them under these conditions.

Screening the Clarke and Carbon Bank for Plasmids Conferring Arsenate Resistance on Strain SG5. To facilitate selection against growth of the donor strain, JA200, during a screen of the Clarke and Carbon (1976) bank of recombinant plasmids, it is desirable to include an antibiotic, such as nalidixic acid, in the growth medium. Strain SG8, a nalidixic acid resistant derivative of strain SG5, was therefore isolated (as described under Materials and Methods) and used as the recipient. Of the arsenate-resistant clones isolated during the screen, nine were tested by direct assay of colonies on agar plates for the presence of glutathione reductase activity. All of the strains appeared to be glutathione reductase positive.

Table II: Glutathione Reductase Activity in Cell-Free Extracts of *E. coli* Strains JM2137, SG5, and SG5As^R

strain	protein concn (mg/mL)	glutathione reductase activity [nmol min^{-1} (mg of protein)$^{-1}$]
JM2137	8.8	172
SG5	8.9	0.1
SG5As^R	7.4	487

The cells were screened for plasmids (Maniatis et al., 1982), and the strain containing what seemed to be the smallest plasmid was used for the isolation of plasmid DNA (Holmes & Quigley, 1981). This DNA preparation transformed *E. coli* SG5 to arsenate resistance, and one of these arsenate-resistant transformants (designated SG5As^R) was purified. Assay of a cell-free extract revealed that the cells had become glutathione reductase positive, with enzymic activities some 2 or 3 times that of the original parent strain JM2137 (Table II).

Strain SG5 is not *rec*A and the glutathione reductase bearing plasmid could therefore have become altered during growth of the cells on arsenate-containing selective media. To isolate a *rec*A strain containing the plasmid, the area of the Clarke and Carbon bank plate that had given rise to the original arsenate-resistant strain on the first screen was now used as the plasmid bank stock for a second screen. This process was repeated until a single strain was obtained that contained a plasmid conferring arsenate resistance and glutathione reductase activity on strain SG5. This Clarke and Carbon strain was designated JA200/pGR.

Mapping the gor Gene on Plasmid pGR. The DNA of plasmid pGR was prepared from strain JA200/pGR, and as judged by agarose gel electrophoresis, it appeared homogeneous. It could transform arsenate-sensitive *gor* strains to arsenate-resistant strains, and cell-free extracts of these arsenate-resistant transformants were found to contain glutathione reductase activity. By summing the sizes of fragments produced by digestion with restriction endonucleases *Sph*I and *Pst*I, the plasmid was found to be about 16.6 kbp in length. Since the ColE1 plasmid on its own is approximately 6.6 kbp (Chan et al., 1985), the insert in pGR must be roughly 10 kbp.

To facilitate DNA sequencing, the insert was mapped with restriction endonucleases, particularly those that are predicted to cleave ColE1 DNA close to the site of insertion. Inspection of the restriction map revealed that no convenient enzymes existed that would cleanly excise the whole insert DNA. Therefore, an attempt was made to map the *gor* gene on the plasmid so that a fragment on which it was located could be sequenced first. The amino acid sequence of human glutathione reductase is known (Krauth-Siegel et al., 1982), and it was a reasonable supposition that human glutathione reductase and *E. coli* glutathione reductase would be homologous, especially in those regions of the protein involved in the catalytic activity. With this in mind, 60 clones of random-sheared pGR DNA were constructed in the filamentous bacteriophage m13mp9 and their DNA sequences determined. When these were translated, one of them appeared to encode an amino acid sequence from the C-terminal region of glutathione reductase incorporating a histidine residue implicated in the catalytic mechanism (Pai & Schulz, 1983). The DNA sequence extended well beyond the putative termination codon, and in this region downstream of the gene was a recognition site for the restriction enzyme *Nsi*I.

This made it possible to map accurately both the position and the orientation of the *gor* gene on pGR. A prime cut probe made from the above m13 clone by using *Nsi*I as the cutting

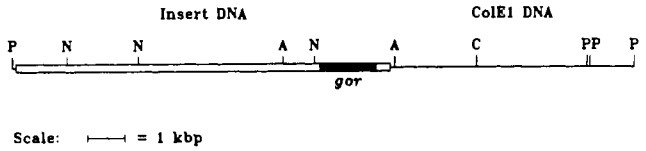

FIGURE 1: Restriction map of ColE1 plasmid (pGR) carrying *E. coli gor* gene: A, *Ava*II; C, *Cla*I; N, *Nsi*I; P, *Pst*I. Only the two *Ava*II sites relevant to excising the *gor* gene from pGR are shown; several more exist on the plasmid. The *gor* gene is about 1.4 kbp long and was mapped to the position shown; transcription proceeds from right to left.

enzyme was hybridized to various restriction digests of pGR. Obviously, any fragment to which the probe would hybridize contained DNA sequence between the 3′-end of the *gor* gene and the *Nsi*I cleavage site. Thus, by hybridization to both single enzyme digests and double digests with *Nsi*I, it was possible to locate the *gor* gene on pGR and to deduce the direction of transcription on the plasmid (Figure 1).

DNA Sequence of gor Gene. As shown in Figure 1, the *gor* gene could be excised from pGR on a 3-kbp *Ava*II fragment of DNA. A preparative *Ava*II digest of 141 μg of pGR DNA was resolved by electrophoresis on a 0.4% low melting agarose gel. The 3-kbp band was excised from the gel, and the DNA was purified from the agarose in a NACS minicolumn. The fragment was then self-ligated and randomly sheared by sonication, and clones of the resulting DNA were constructed in bacteriophage m13mp9 and sequenced. Sequence data were assembled until the whole of the glutathione reductase gene was covered on both strands. The DNA sequence of the *gor* gene and the predicted amino acid sequence are shown in Figure 2.

Immediately upstream of the gene is a sequence homologous to known *E. coli* promoters (Rosenberg & Court, 1979); the −35 region of this sequence is not close to the consensus TTGACA and, hence, is difficult to assign absolutely (one possibility is shown in Figure 2), but the Pribnow box is almost the canonical TATAAT except for a single substitution in one of the less well conserved positions. Immediately downstream of the gene is a potential terminator of transcription (Rosenberg & Court, 1979), although terminators can be difficult to assign [see, for example, Platt (1981)]. This is a hairpin structure containing a 12-bp inverted repeat that is immedi-

```
                    -35                                 PB
GGAGTAATTGCAGCCATTGCTGGCACCTATTACGTCTCGCGCTACAATCGCGGTAATCAA
      10        20        30        40        50        60

   SD
          M  T  K  H  Y  D  Y  I  A  I  G  G  G  S
CGATAAGGACACTTTGTCATGACTAAACACTATGATTACATCGCCATCGGCGGCGGCAGC
      70        80        90       100       110       120

 G  G  I  A  S  I  N  R  A  A  M  Y  G  Q  K  C  A  L  I  E
GGCGGTATCGCCTCCATCAACCGCGCGGCTATGTACGGCCAGAAATGTGCGCTGATTGAA
     130       140       150       160       170       180

 A  K  E  L  G  G  T  C  V  N  V  G  C  V  P  K  K  V  M  W
GCCAAAGAGCTGGGCGGCACCTGCGTAAATGTTGGCTGTGTGCCGAAAAAAGTGATGTGG
     190       200       210       220       230       240

 H  A  A  Q  I  R  E  A  I  H  M  Y  G  P  D  Y  G  F  D  T
CACGCGGCGCAAATCCGTGAAGCGATCCATATGTACGGCCCGGATTATGGTTTTGATACC
     250       260       270       280       290       300

 T  I  N  K  F  N  W  E  T  L  I  A  S  R  T  A  Y  I  D  R
ACTATCAATAAATTCAACTGGGAAACGTTGATCGCCAGCCGTACCGCCTATATCGACCGT
     310       320       330       340       350       360

 I  H  T  S  Y  E  N  V  L  G  K  N  N  V  D  V  I  K  G  F
ATTCATACTTCCTATGAAAACGTGCTCGGTAAAAATAACGTTGATGTAATCAAAGGCTTT
     370       380       390       400       410       420

 A  R  F  V  D  A  K  T  L  E  V  N  G  E  T  I  T  A  D  H
GCCCGCTTCGTTGATGCCAAAACGCTGGAGGTAAACGGCGAAACCATCACGGCCGATCAT
     430       440       450       460       470       480

 I  L  I  A  T  G  G  R  P  S  H  P  D  I  P  G  V  E  Y  G
ATTCTGATCGCCACAGGCGGTCGTCCGAGCCACCCGGATATTCCGGGCGTGGAATACGGT
     490       500       510       520       530       540

 I  D  S  D  G  F  F  A  L  P  A  L  P  E  R  V  A  V  V  G
ATTGATTCTGATGGCTTCTTCGCCCTTCCTGCTTTGCCAGAGCGCGTGGCGGTTGTTGGC
     550       560       570       580       590       600

 A  G  Y  I  A  V  E  L  A  G  V  I  N  G  L  G  A  K  T  H
GCGGGTTACATCGCCGTTGAGCTGGCGGGCGTGATTAACGGCCTCGGCGCGAAAACGCAT
     610       620       630       640       650       660

 L  F  V  R  K  H  A  P  L  R  S  F  D  P  M  I  S  E  T  L
CTGTTTGTGCGTAAACATGCGCCGCTGCGCAGCTTCGACCCGATGATTTCCGAAACGCTG
     670       680       690       700       710       720

 V  E  V  M  N  A  E  G  P  Q  L  H  T  N  A  I  P  K  A  V
GTCGAAGTGATGAACGCCGAAGGCCCGCAGCTGCACACCAACGCCATCCCGAAAGCGGTA
     730       740       750       760       770       780

 V  K  N  T  D  G  S  L  T  L  E  L  E  D  G  R  S  E  T  V
GTGAAAAATACCGATGGTAGCCTGACGCTGGAGCTGGAAGATGGTCGCAGTGAAACGGTG
     790       800       810       820       830       840

 D  C  L  I  W  A  I  G  R  E  P  A  N  D  N  I  N  L  E  A
GATTGCCTGATTTGGGCGATTGGTCGCGAGCCTGCCAATGACAACATCAACCTGGAAGCC
     850       860       870       880       890       900

 A  G  V  K  T  N  E  K  G  Y  I  V  V  D  K  Y  Q  N  T  N
GCTGGCGTTAAAACTAACGAAAAAGGCTATATCGTCGTCGATAAATATCAAAACACCAAT
     910       920       930       940       950       960

 I  E  G  I  Y  A  V  G  D  N  T  G  A  V  E  L  T  P  V  A
ATTGAAGGTATTTACGCGGTGGGCGATAACACGGGTGCAGTGGAGCTGACACCGGTGGCA
     970       980       990      1000      1010      1020

 V  A  A  G  R  R  L  S  E  R  L  F  N  N  K  P  D  E  H  L
GTTGCAGCGGGTCGCCGTCTCTCTGAACGCCTGTTTAATAACAAGCCGGATGAGCATCTG
    1030      1040      1050      1060      1070      1080

 D  Y  S  N  I  P  T  V  V  F  S  H  P  P  I  G  T  V  G  L
GATTACAGCAACATTCCGACCGTGGTCTTCAGCCATCCGCCGATTGGTACTGTTGGTTTA
    1090      1100      1110      1120      1130      1140

 T  E  P  Q  A  R  E  Q  Y  G  D  D  Q  V  K  V  Y  K  S  S
ACGGAACCGCAGGCGCGCGAGCAGTATGGCGACGATCAGGTGAAAGTGTATAAATCCTCT
    1150      1160      1170      1180      1190      1200

 F  T  A  M  Y  T  A  V  T  T  H  R  Q  P  C  R  M  K  L  V
TTCACCGCGATGTATACCGCCGTCACCACTCACCGCCAGCCGTGCCGCATGAAGCTGGTG
    1210      1220      1230      1240      1250      1260

 C  V  G  S  E  E  K  I  V  G  I  H  G  I  G  F  G  M  D  E
TGCGTTGGATCGGAAGAGAAGATTGTCGGTATTCACGGCATTGGCTTTGGTATGGACGAA
    1270      1280      1290      1300      1310      1320

 M  L  Q  G  F  A  V  A  L  K  M  G  A  T  K  K  D  F  D  N
ATGTTGCAGGGCTTCGCGGTGGCGCTGAAGATGGGGGCAACCAAAAAAGACTTCGACAAT
    1330      1340      1350      1360      1370      1380

 T  V  A  I  H  P  T  A  A  E  E  F  V  T  M  R  *
ACCGTCGCCATTCACCCAACGGCGGCAGAAGAGTTCGTGACAATGCGTTAAATGTTAAAG
    1390      1400      1410      1420      1430      1440

                TER
GGCTAAGAGTAGTGTGGCTCTTAGCCCTTAATTACGTTTCCGCTATCAGTTCAGAAGCTGA
    1450      1460      1470      1480      1490      1500
```

FIGURE 2: DNA sequence of *E. coli gor* gene and predicted amino acid sequence of glutathione reductase. Potentially important control regions are shown underlined: −35, promoter −35 region; PB, Pribnow box; SD, Shine–Dalgarno sequence; TER, stem–loop structure of the putative transcription terminator.

	Human glutathione reductase	E.coli glutathione reductase	E.coli dihydrolipoamide dehydrogenase
Asp	21	24	25
Asn	17	23	14
Thr	31	32	26
Ser	31	16	14
Glu	29	30	39
Gln	11	9	8
Pro	24	20	21
Gly	43	44	51
Ala	42	44	50
Val	44	40	45
Met	15	12	9
Ile	29	33	39
Leu	34	27	34
Tyr	13	16	8
Phe	14	15	14
His	16	14	13
Lys	34	24	39
Arg	17	18	15
Cys	10	6	5
Trp	3	3	4
Total	478	450	473
M_r (apoenzyme monomer)	51609	48717	50554

FIGURE 3: Amino acid compositions of three flavoprotein disulfide oxidoreductases: human glutathione reductase (Krauth-Siegel et al., 1982), *E. coli* dihydrolipoamide dehydrogenase (Stephens et al., 1983), and *E. coli* glutathione reductase (this paper). The N-terminal methionine residue is included in the *E. coli* glutathione reductase tabulation.

F	TTT	5	S	TCT	3	Y	TAT	9	C	TGT	2
F	TTC	10	S	TCC	4	Y	TAC	7	C	TGC	4
L	TTA	1	S	TCA	0	*	TAA	1	*	TGA	0
L	TTG	3	S	TCG	1	*	TAG	0	W	TGG	3
L	CTT	1	P	CCT	2	H	CAT	7	R	CGT	7
L	CTC	3	P	CCC	0	H	CAC	7	R	CGC	11
L	CTA	0	P	CCA	2	Q	CAA	2	R	CGA	0
L	CTG	19	P	CCG	16	Q	CAG	7	R	CGG	0
I	ATT	17	T	ACT	6	N	AAT	8	S	AGT	1
I	ATC	16	T	ACC	13	N	AAC	15	S	AGC	7
I	ATA	0	T	ACA	3	K	AAA	20	R	AGA	0
M	ATG	12	T	ACG	10	K	AAG	4	R	AGG	0
V	GTT	10	A	GCT	3	D	GAT	17	G	GGT	16
V	GTC	7	A	GCC	17	D	GAC	7	G	GGC	26
V	GTA	4	A	GCA	5	E	GAA	19	G	GGA	1
V	GTG	19	A	GCG	19	E	GAG	11	G	GGG	1

Total codon count = 451.

FIGURE 4: Codon totals for the *E. coli* glutathione reductase gene. There are 451 codons in all, including initiation and termination codons.

ately followed by 10 bases of sequence, 80% of which are A or T. Whether these putative control sequences upstream and downstream of the *gor* gene are functional in vivo can only be ascertained by transcript mapping, but their existence suggests that the *gor* gene is transcribed independently and is not part of an operon.

The initiator codon of the *gor* gene was tentatively identified as the methionine codon at the 5′-end of the open-reading frame. This identification was supported by the presence just downstream of this codon of amino acid sequence homology to the FAD binding domain of human glutathione reductase and by the presence 9 bp upstream of a 6-bp sequence complementary to the 3′-end of *E. coli* 16S ribosomal RNA (Shine & Dalgarno, 1975). Also, just upstream of the initiator codon and in the same reading frame is a stop codon (TAA), so the protein cannot start upstream of this point.

The protein predicted from the sequence of the *gor* gene consists of 450 amino acids (including the initiating methionine residue), which makes it smaller than either human glutathione reductase (478 residues) (Krauth-Siegel et al., 1982), *E. coli* dihydrolipoamide dehydrogenase (473 residues) (Stephens et al., 1983), or *P. aeruginosa* mercuric reductase (561 residues) (Brown et al., 1983). The amino acid composition (Figure 3) and the codon usage (Figure 4) are shown, as are the amino acid compositions of human glutathione reductase and *E. coli* dihydrolipoamide dehydrogenase (Figure 3). Comparison of the data shown in Figure 4 with the codon usage in strongly and less strongly expressed *E. coli* genes (Grosjean & Fiers, 1982) shows that although individual codon blocks do seem to fall into either the strongly or less strongly expressed classes, the codon-usage pattern of the protein as a whole shows no such unambiguous tendency. This intermediate codon-usage pattern may reflect the fact that glutathione reductase occupies an intermediate position in the expression levels of proteins in the cell.

Discussion

The cloning strategy that we used was predicated on the supposed proximity on the *E. coli* genome of the *gor* gene (Davis et al., 1982) and the *pit* locus (Sprague et al., 1975). Of particular importance was the isolation of the arsenate-sensitive strain SG5. The mutation conferring sensitivity to arsenate was suppressible by plasmid-encoded determinants, suggesting that it does not arise from a chromosomal rearrangement but is more probably due to the loss of an unknown gene or genes in this region. This unknown gene may simply code for a product that renders arsenate less toxic within the cell, but it could be connected somehow with the *pit* system. The exact connection, if any, that arsenate sensitivity has with mutations in or near the *pit* locus is unclear, but it may become apparent when the complete nucleotide sequence of the insert DNA (some 10 kilobase pairs) of plasmid pGR is determined and its encoded products are identified. At all events, the isolation of strain SG5, though fortuitous, has revealed a novel and interesting arsenate-sensitive phenotype in *E. coli* and enabled us to follow a simple route to the cloning of the *gor* gene.

Perhaps the most interesting aspect of the work at the present stage is a comparison of the *E. coli* glutathione reductase with other related enzymes. It is immediately obvious that the primary structures of the glutathione reductases from *E. coli* and human red blood cells bear a very great resemblance to each other, implying close similarities in three-dimensional structure as well (Figure 5). The sequence homology is especially marked at the C-terminal ends of the proteins; this region in human glutathione reductase embodies the interface (glutathione-binding) domain of the molecule (Thieme et al., 1981) and presumably serves the same purpose in the *E. coli* enzyme. Apart from simple substitutions, the differences in primary structure between the two proteins are mostly due to deletions from the *E. coli* enzyme relative to the human enzyme. In the one case where the deletion is made from the human amino acid sequence (following residue 82), it consists of a single residue (Figure 5).

There are four deletions from the *E. coli* enzyme relative to the human enzyme. No sequence homologous to the N-terminal 16 amino acids of the human enzyme exists in the *E. coli* protein. The N-terminal 18 amino acids of human glutathione reductase have no crystal structure (Thieme et al., 1981) and have not been implicated in catalysis (Pai & Schulz,

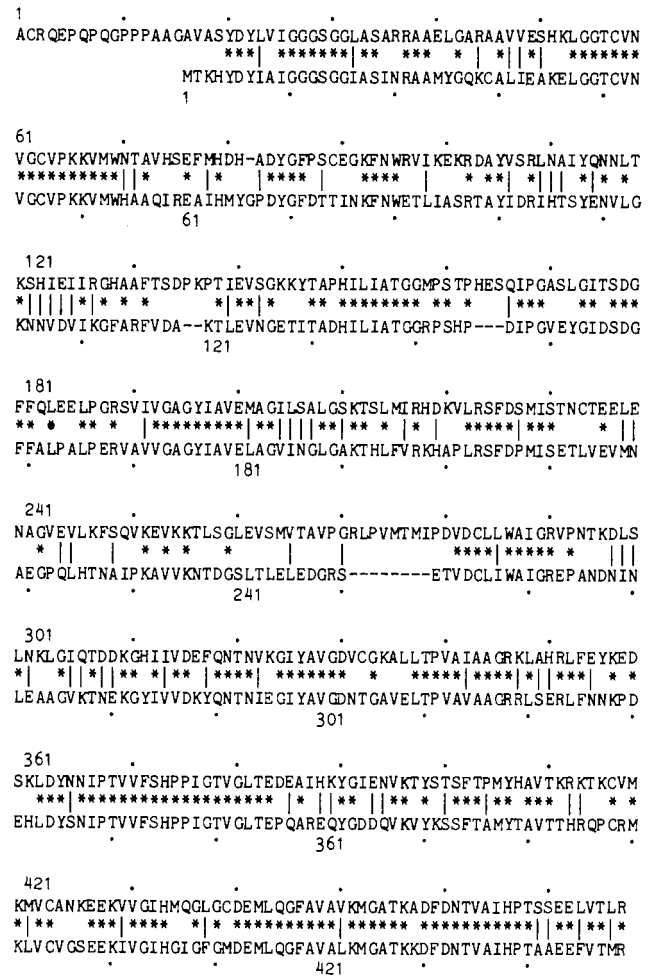

FIGURE 5: Alignment of the primary structure of human glutathione reductase (Krauth-Siegel et al., 1982; upper strand) and *E. coli* glutathione reductase (lower strand). Alignment codes: (*) identity; (|) weight for substitution from the scoring matrix greater than 10, corresponding to a conservative substitution. Insertions in one protein relative to the other are indicated by dashes. Numbering of amino acids refers to the residue adjacent to the first digit of the number.

1983). This would be consistent with the deletion of a large part of this region from the *E. coli* enzyme, though it is conceivable that a flexible arm at the N-terminus could have some part to play in other properties of the enzyme, e.g., regulation, in the human red blood cell (Thieme et al., 1981). Another deletion from the *E. coli* enzyme is the loss of those residues corresponding to residues 164–166 in the human enzyme, where they form part of the loop connecting the FAD and NADP domains. This slight shortening of an interdomain loop is not necessarily of great overall structural importance. The remaining two deletions correspond to residues 136–137 and residues 272–279 in the human enzyme and are actually within the FAD and NADP domains, respectively. The FAD and NADP domains are structurally homologous in the human enzyme (Krauth-Siegel et al., 1982), and the two deletions in the *E. coli* enzyme appear from sequence matching to occur in homologous positions inside these two domains, within areas of β-sheet. The exact structural placing of these deletions must of course await crystallographic studies of *E. coli* glutathione reductase since estimates based on sequence homology matching may well not be accurate in three dimensions. But it is of interest to note that previous comparisons between human glutathione reductase, *p*-hydroxybenzoate reductase (Wierenga et al., 1983), and *E. coli* dihydrolipoamide dehydrogenase (Rice et al., 1984) have also found deletions in these regions; it seems that these areas of the FAD and NAD(P) domains are favored positions for any such deletions to occur.

Conservation of amino acid residues implicated in the catalytic mechanism of human glutathione reductase (Pai & Schulz, 1983) is almost perfect; only one such residue is substituted in the *E. coli* enzyme. Even this substitution, of His-219 in the human protein by a lysine residue in the *E. coli* enzyme (Figure 5), is conservative, given that His-219 has been assigned a part in binding the 2′-phosphate group of NADP and thereby discriminating against NAD (Krauth-Siegel et al., 1982). One other substitution to note is that Cys-90 in the human enzyme, which forms an unusual intersubunit disulfide bridge (Thieme et al., 1981), has been replaced by threonine. Thus, it is unlikely that *E. coli* glutathione reductase will resemble the human enzyme in being a covalently linked dimer.

Until recently, knowledge of the primary structures of flavoprotein disulfide oxidoreductases was limited to that obtained from direct determination of amino acid sequences [for reviews, see Williams et al. (1984) and Greer & Perham (1984)], and a complete sequence was available only for human glutathione reductase (Krauth-Siegel et al., 1982). In the past few years, the entire sequences of two more of these important enzymes, *E. coli* dihydrolipoamide dehydrogenase (Stephens et al., 1983) and *P. aeruginosa* mercuric reductase (Brown et al., 1983), have been predicted from the DNA sequences of the relevant cloned genes. We have presented here for the first time the complete DNA sequence of a gene encoding glutathione reductase, from which the amino acid sequence of the *E. coli* enzyme could be predicted. *E. coli* glutathione reductase turns out to be more closely related to its human homologue than it is to *E. coli* dihydrolipoamide dehydrogenase or to *P. aeruginosa* mercuric reductase. Crystallographic work on human glutathione reductase (Thieme et al., 1981; Pai & Schulz, 1983) has given an insight into the molecular mechanism of this enzyme and, by inference, into the mechanism of those enzymes related to it. Because of the very considerable sequence similarity between human glutathione reductase and the *E. coli* enzyme, it is certain that the three-dimensional structures of these enzymes will also be similar. Thus, it should now be possible to test some of the predictions made about the catalytic mechanism of the human enzyme by site-directed mutagenesis of the *E. coli* protein.

ACKNOWLEDGMENTS

We are grateful to Dr. M. Jones-Mortimer and Dr. P. Alefounder for helpful discussion and to Professor J. Guest for the gift of the Clarke and Carbon gene bank.

Registry No. DNA (*Escherichia coli* gene *gor*), 101077-52-5; glutathione reductase (*Escherichia coli* reduced), 101077-54-7; glutathione reductase, 9001-48-3; arsenate, 15584-04-0; flavoprotein disulfide oxidoreductase, 9029-19-0.

REFERENCES

Apontoweil, P., & Berends, W. (1975) *Biochim. Biophys. Acta 399*, 10–22.

Bachmann, B. J. (1983) *Microbiol. Rev. 47*, 180–230.

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Brown, N. L., Ford, S. J., Pridmore, D., & Fritzinger, D. C. (1983) *Biochemistry 22*, 4089–4095.

Chan, P. T., Ohmori, H., Tomizawa, J., & Lebowitz, J. (1985) *J. Biol. Chem. 260*, 8925–8935.

Clarke, L., & Carbon, J. (1976) *Cell (Cambridge, Mass.) 9*, 91–99.

Davis, N. K., Greer, S., Jones-Mortimer, M. C., & Perham, R. N. (1982) *J. Gen. Microbiol. 128*, 1631–1634.
Deininger, P. L. (1983) *Anal. Biochem. 129*, 216–223.
Fox, B., & Walsh, C. T. (1982) *J. Biol. Chem. 257*, 2498–2503.
Fox, B., & Walsh, C. T. (1983) *Biochemistry 22*, 4082–4088.
Fuchs, J. A., & Warner, H. R. (1975) *J. Bacteriol. 124*, 140–148.
Greer, J. A. G., & Perham, R. N. (1984) in *Flavins and Flavoproteins* (Bray, R. C., Engel, P. C., & Mayhew, S. G., Eds.) pp 125–137, de Gruyter, Berlin.
Grosjean, H., & Fiers, W. (1982) *Gene 18*, 199–209.
Holmes, D. S., & Quigley, M. (1981) *Anal. Biochem. 114*, 193–197.
Holmgren, A. (1980) *Experientia, Suppl. 36*, 149–180.
Hudson, G. S., & Davidson, B. E. (1984) *J. Mol. Biol. 180*, 1023–1051.
Krauth-Siegel, R. L., Blatterspiel, R., Saleh, M., Schiltz, E., Schirmer, R. H., & Untucht-Grau, R. (1982) *Eur. J. Biochem. 121*, 259–267.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Packman, L. C., Hale, G., & Perham, R. N. (1984) *EMBO J. 3*, 1315–1319.
Pai, E. F., & Schulz, G. E. (1983) *J. Biol. Chem. 258*, 1752–1757.
Perham, R. N., Harrison, R. A., & Brown, J. P. (1978) *Biochem. Soc. Trans. 6*, 47–50.
Platt, T. (1981) *Cell (Cambridge, Mass.) 24*, 10–23.
Reed, L. J. (1974) *Acc. Chem. Res. 7*, 40–46.
Rosenberg, M., & Court, D. (1979) *Annu. Rev. Genet. 13*, 319–353.
Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., & Roe, B. A. (1980) *J. Mol. Biol. 143*, 161–178.
Schwartz, R. M., & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, Vol. 5, pp 353–358, National Biomedical Research Foundation, Washington, DC.
Shine, J., & Dalgarno, L. (1975) *Nature (London) 254*, 34–38.
Sprague, G. F., Jr., Bell, R. M., & Cronan, J. E., Jr. (1975) *Mol. Gen. Genet. 143*, 71–77.
Staden, R. (1982) *Nucleic Acids Res. 10*, 2951–2961.
Staden, R. (1983) in *DNA Sequencing, Laboratory Techniques in Biochemistry and Molecular Biology* (Hindley, J., Ed.) Vol. 10, pp 311–368, Elsevier, Amsterdam.
Stephens, P. E., Lewis, H. M., Darlison, M. G., & Guest, J. R. (1983) *Eur. J. Biochem. 135*, 519–527.
Thieme, R., Pai, E. F., Schirmer, R. H., & Schulz, G. E. (1981) *J. Mol. Biol. 152*, 763–782.
Williams, C. H., Jr., Arscott, L. D., & Schulz, G. E. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2199–2201.
Williams, C. H., Jr., Arscott, L. D., & Swenson, R. P. (1984) in *Flavins and Flavoproteins* (Bray, R. C., Engel, P. C., & Mayhew, S. G., Eds.) pp 95–109, de Gruyter, Berlin.
Worthington, D. J., & Rosemeyer, M. A. (1976) *Eur. J. Biochem. 67*, 231–138.

CORRECTION

Spectra, Membrane Binding, and Potentiometric Responses of New Charge Shift Probes, by Eric Fluhler, Valerie G. Burnham, and Leslie M. Loew*, Volume 24, Number 21, October 8, 1985, pages 5749–5755.

Page 5749. In the title, Reponses should read Responses.

Page 5750. The equation should read

$$F_C = \frac{(F_\infty - F_0)C}{C + K_d/n} + F_0$$

Page 5754. In Table III, column 2, all the exponents of 10 should have negative signs.