JOURNAL OF BACTERIOLOGY, Nov. 1992, p. 7080–7089
0021-9193/92/227080-10$02.00/0

Vol. 174, No. 22

Clostridium sticklandii Glycine Reductase Selenoprotein A Gene: Cloning, Sequencing, and Expression in *Escherichia coli*

GREGORY E. GARCIA† AND THRESSA C. STADTMAN*

Laboratory of Biochemistry, National Heart, Lung, and Blood Institute, National Institutes of Health, Building 3, Room 108, Bethesda, Maryland 20892

Received 10 July 1992/Accepted 8 September 1992

Gene *grdA*, which encodes selenoprotein A of the glycine reductase complex from *Clostridium sticklandii*, was identified and characterized. This gene encodes a protein of 158 amino acids with a calculated M_r of 17,142. The known sequence of 15 amino acids around the selenocysteine residue and the known carboxy terminus of the protein are correctly predicted by the nucleotide sequence. An opal termination codon (TGA) corresponding to the location of the single selenocysteine residue in the polypeptide was found in frame at position 130. The *C. sticklandii grdA* gene was inserted behind the *tac* promotor of an *Escherichia coli* expression vector. An *E. coli* strain transformed with this vector produced an 18-kDa polypeptide that was not detected in extracts of nontransformed cells. Affinity-purified anti-*C. sticklandii* selenoprotein A immunoglobulin G reacted specifically with this polypeptide, which was indistinguishable from authentic *C. sticklandii* selenoprotein A by immunological analysis. Addition of the purified expressed protein to glycine reductase protein components B and C reconstituted the active glycine reductase complex. Although synthesis of enzymically active protein A depended on the presence of selenium in the growth medium, formation of immunologically reactive protein did not. Moreover, synthesis of enzymically active protein in a transformed *E. coli selD* mutant strain indicated that there is a nonspecific mechanism of selenocysteine incorporation. These findings imply that mRNA secondary structures of *C. sticklandii grdA* are not functional for UGA-directed selenocysteine insertion in the *E. coli* expression system.

The glycine reductase present in several clostridia catalyzes the reductive deamination of glycine to acetylphosphate and ammonia (1, 28). Protein A, one of the components of the glycine reductase complex, is a small, acidic, heat-stable glycoprotein that contains selenium in the form of a selenocysteine residue (8, 26, 31). Attempts to determine the complete amino acid sequence of protein A from *Clostridium sticklandii* were unsuccessful because of a blocked amino terminus (9). Although the sequence of an internal selenopeptide, -Cys-Phe-Val-Sec-Thr-Ala-Ala-Gly-Ala-Met-Asp-Leu-Glu-Asn-Gln-Lys-, was known (24), oligonucleotide probes based on the genetic code for this sequence were not sufficiently specific for the isolation of the gene from genomic libraries of *C. sticklandii* (12a). However, additional amino acid sequence information was obtained for an analogous selenoprotein A from *Clostridium purinolyticum* that does not have a blocked amino terminus, and this information allowed isolation of the gene, which was cloned and sequenced (13). The *C. purinolyticum* gene has an in-frame TGA codon at position 127, and the amino acid sequence of an Endo-Glu-C selenocysteine peptide from the gene product is identical to the sequence of the corresponding *C. sticklandii* peptide. On the basis of this degree of similarity and the cross-reactivity of the *C. purinolyticum* protein with anti-*C. sticklandii* selenoprotein A antibodies (25), a strategy of selecting the gene from *C. sticklandii* by using a probe consisting of the entire gene from *C. purinolyticum* seemed reasonable. Since most of the known chemical properties of selenoprotein A had been determined for the protein isolated from *C. sticklandii*, it was especially desirable to have the corresponding gene available for further analysis. A description of the successful isolation and sequencing of the gene for selenoprotein A from *C. sticklandii* and of studies on the expression of the gene in *Escherichia coli* is presented in this paper.

MATERIALS AND METHODS

Materials. Competent *E. coli* DH5α and DH5αMCR were obtained from Bethesda Research Laboratories. Transformations were performed as recommended by the manufacturer. *E. coli* MB08 (F⁻ *selD proC23 trp-30 his-51 lac-28 rpsL101* λ⁺) was a gift from Marie-Andre Mandrand-Berthelot. T4 DNA ligase, T4 DNA kinase, *Taq* DNA polymerase, *Eco*RI methylase, restriction endonucleases, isopropylthio-β-galactoside (IPTG), and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside were purchased from Bethesda Research Laboratories. Modified T7 DNA polymerase (Sequenase) was obtained from United States Biochemical Corp. α-^{35}S-labeled dATP (500 Ci/mmol) and [α-^{32}P]ATP (3,000 Ci/mmol) were purchased from New England Nuclear. $^{75}SeO_3H_2$ (1,000 Ci/mol) was obtained from the Research Reactor Facility at the University of Missouri. A DNA nonisotopic labeling system, positively charged nylon membranes (Sureblot), membrane blocking solution (MBS), nitroblue tetrazolium, and 5-bromo-4-chloro-3-indolyl phosphate were obtained from ONCOR. Polyvinylidene difluoride membranes (Immobilon) were purchased from Millipore Corp. Polyacrylamide (16%) gels, gradient (10 to 27%) polyacrylamide gels, and a Western blot module were obtained from Novex, San Diego, Calif. Rabbit anti-sheep immunoglobulin G (IgG) conjugated to alkaline phosphatase

* Corresponding author.
† Present address: Department of Biochemistry, Walter Reed Army Institute of Research, Washington, DC 20307.

and rabbit nonspecific serum were obtained from Pierce Chemical Co. Alkaline phosphatase chemiluminescent reagent AMPPD was obtained from Tropix, Inc. Other reagents were of the highest grade available.

DNA isolation and recombinant DNA methods. Genomic DNA was prepared from *C. sticklandii* ATCC 12662 by the method of Saito and Miura (22). Plasmid pGG7 containing the *grdA* gene from *C. purinolyticum* has been described previously (13). Plasmids pUC13 and pUC19, *E. coli* expression vector plasmid pKK223-3, and M13 universal and reverse sequencing primers were purchased from Pharmacia. All other oligonucleotide primers were synthesized in our laboratory by using an Applied Biosystems model 380B DNA synthesizer and were purified on polyacrylamide gels.

Molecular biology procedures were performed as described by Maniatis et al. (20). Small-scale plasmid isolations were performed by the method of Birnboim and Doly (5), as modified by Thompson (30). Supercoiled plasmid DNA was purified by CsCl ultracentrifugation and was sequenced by the dideoxy chain termination method of Sanger et al. (23), modified for use with α-^{35}S-labeled dATP and modified T7 DNA polymerase (U.S. Biochemical Corp.) as described in the Sequenase kit instructions.

Probe generation. The *grdA* gene from *C. purinolyticum* was amplified by the polymerase chain reaction, which was performed with a Perkin-Elmer thermal cycler as described by the manufacturer. Oligonucleotide primers corresponding to the amino and carboxyl termini were used to amplify the full-length *grdA* gene, using plasmid pGG7 as the source. The amplified DNA was analyzed with a Phastsystem apparatus using a homogeneous 20% polyacrylamide gel (Pharmacia) as described previously (13). A 450-bp product was detected, which corresponds to the size predicted for the *grdA* gene of *C. purinolyticum*. The 450-bp product was purified on agarose gels and labeled with [α-^{32}P]dCTP by nick translation, as described in the Bethesda Research Laboratories kit instructions.

The *C. sticklandii grdA* gene (isolated as described below) in plasmid pGG9 was amplified and purified like plasmid pGG7. The amplified DNA product was labeled nonisotopically with biotin-modified nucleotides by nick translation, as described in the ONCOR kit instructions.

Plasmid construction for the expression of the selenoprotein A gene. The *grdA* gene of *C. sticklandii* was amplified from plasmid pGG9 by the polymerase chain reaction method as described above. Oligonucleotide primers corresponding to the N terminus and C terminus were used to amplify the full-length gene. The two internal restriction endonuclease *Eco*RI sites were methylated with *S*-adenosylmethionine and *Eco*RI methylase. *Eco*RI linkers were phosphorylated with T4 kinase and ligated to the methylated DNA. The linkered DNA was digested with restriction endonuclease *Eco*RI and purified on agarose gels. The purified product was ligated to prokaryotic expression vector pKK223-3 digested with restriction endonuclease *Eco*RI. *E. coli* DH5αMCR was transformed, and the initial selection was by growth on Luria broth agar plates containing 200 μg of ampicillin per ml (LBA medium). Ampicillin-resistant clones were picked and grown overnight aerobically in 5 ml of LBA medium. Recombinants were selected by hybridization to the *C. sticklandii grdA* probe on Southern blots of plasmids isolated from the selected clones. A positive recombinant plasmid, pGG12, was sequenced to determine the orientation of the insertion and to confirm the gene sequence. Both strands of the gene insertion of plasmid pGG12 were sequenced.

Expression of the selenoprotein A gene in *E. coli* for immunological analysis. *E. coli* DH5αMCR transformed with pGG12 was grown overnight in 5 ml of LBA medium at 37°C. The entire culture was used to inoculate 100 ml of LBA medium supplemented with 2 μM $MnCl_2$, 2 μM Na_2MoO_4, and 1 μM Na_2SeO_3 in a 100-ml volumetric flask. The cells were induced by adding IPTG to a final concentration of 0.4 mM and were incubated anaerobically for 14 h. ^{75}Se-labeled cells were prepared by supplementing the induction medium with $^{75}SeO_3Na_2$ (0.5 μCi/ml). Following induction, cells were harvested by centrifugation and kept on ice. The cells were washed with ice-cold Tris-buffered saline (0.01 M Tris-HCl [pH 7.2], 0.15 M NaCl). The final cell pellet was resuspended in 1 ml of Tris-buffered saline and stored at −20°C. Untransformed cultures of *E. coli* DH5αMCR were prepared as described above, except that the antibiotic was omitted.

Partial purification of recombinant selenoprotein A from *E. coli*. *E. coli* transformed with pGG12 was grown either aerobically or anaerobically at 37°C in a 10-liter fermentor containing double-strength Luria broth (2% tryptone, 1% yeast extract, 1% NaCl) or anaerobically in a glucose minimal medium (27); both media were supplemented with 0.2 mg of ampicillin per ml (see Table 2). Cells were grown to an optical density at 550 nm of 0.6 and then induced by adding lactose (final concentration, 0.2%) and IPTG (final concentration, 0.4 mM). Na_2SeO_3 was added to a final concentration of 1 μM. For aerobic cultures the induction was allowed to proceed for 3 h, at which time the optical density at 600 nm was 5. The harvested cells were frozen by dropping them into liquid N_2. A 3-g portion of cells was thawed in 2 volumes of buffer A (50 mM Tricine-HCl [pH 7.5], 2 mM dithiothreitol, 10 mM EDTA) supplemented with 0.1% phenylmethylsulfonyl fluoride, and the cells were broken by sonication, using a Branson model 450 sonicator. After centrifugation at 26,000 × *g* for 20 min, the supernatant was applied to a DE-52 column (2.6 by 2.8 cm) equilibrated in buffer A. The column was washed with buffer A to remove nonabsorbed proteins and then was eluted with 0.2 M potassium phosphate (pH 7.1)–2 mM dithiothreitol–1 mM EDTA. The acidic protein fraction remaining on the column was eluted with 1 M NaCl in 10 mM potassium phosphate (pH 7.1)–2 mM dithiothreitol–1 mM EDTA and was precipitated by adding 0.56 g of ammonium sulfate per ml. The protein pellet was dissolved in a small volume of 50 mM Tricine-HCl (pH 7.5)–1 mM dithiothreitol–1 mM EDTA and was desalted over a PD-10 column (Pharmacia). When necessary, protein fractions were concentrated with a Centricon-10 microconcentrator (Amicon).

Preparation of source antibody-antigen complex. *C. sticklandii* selenoprotein A was partially purified through the preparative DEAE high-performance liquid chromatography step as described previously (24), except that the protein was not alkylated. To facilitate detection, a tracer amount of labeled selenoprotein A, which was present in a crude ^{75}Se-labeled cell extract (171,000 cpm), was added to 3 mg of the protein preparation. After fractionation on an analytical DEAE high-performance liquid chromatography column, the selenoprotein A peak fractions were pooled, desalted, and concentrated on a Centricon-10 microconcentrator. The concentrated sample (40 μl) contained 26,000 cpm of ^{75}Se. Aliquots of this solution were subjected to sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis in a 10 to 27% gradient polyacrylamide gel (18). Following electrophoresis the gel was equilibrated in transfer buffer (10 mM cyclohexylaminopropanesulfonic acid-NaOH [pH 11], 10% methanol), and the proteins were transferred to a polyvinyli-

dene difluoride membrane in a Novex Western blot module at 35 A for 8 h at 22°C. After the membrane was stained with Coomassie brilliant blue and the blot was exposed to X-ray film (Kodak type SB5 film) overnight, the selenoprotein A band was excised and destained with 100% methanol. The destained membrane was treated with MBS (ONCOR, Inc.) for 3 h at 37°C, dried, and stored desiccated at −20°C. On the day of use the source antigen membranes were wetted with methanol, rinsed with water, and placed in separate bullet tubes containing 0.2 ml of MBS. To each tube was added 20 μl of a 1/200 dilution of sheep anti-*C. sticklandii* selenoprotein A immune serum that had been preabsorbed with an *E. coli* cell extract (15). After incubation for 3 h at 37°C to allow antibody absorption to the antigen, the membranes were washed four times for 5 min with 1 ml of MBS at room temperature and then used immediately for antibody exchange.

Immunological analysis by antibody exchange. Synthesis of *C. sticklandii* selenoprotein A in transformed *E. coli* strains was monitored by using a modified antibody exchange reaction (14). The proteins in 50-μl cell suspensions or in DEAE-cellulose eluate fractions were solubilized by heating the preparations in SDS buffer and were separated electrophoretically in SDS by using 16% polyacrylamide gels. After the separated proteins were transferred to Immobilon P membranes as described above, the membranes were rinsed with water and stained with 0.1% Ponceau SS (Sigma) in 5% acetic acid. Strips containing individual protein lanes were cut, destained with water, and blocked with MBS for 3 h at 37°C. These target strips, which were rinsed two times with MBS, were then combined, two to a tube, and incubated with a single freshly prepared antigen-antibody source strip in MBS for 3 h at 37°C. Following the antibody exchange reaction, the target strips were washed for 5 min with MBS four times and were incubated for 1.5 h at 36°C in a secondary antibody solution consisting of rabbit anti-sheep IgG conjugated to alkaline phosphatase (diluted to a titer of 1/5,000 in 10 ml of MBS). The strips treated with the second antibody were washed, and the alkaline phosphatase substrate was added for visualization of the selenoprotein A band, as described in the ONCOR kit instructions. The high level of background staining observed in the initial experiments was reduced by adding 0.5% nonspecific rabbit serum to the secondary antibody solution.

Chemiluminescent detection of alkaline phosphatase linked to rabbit anti-sheep IgG was performed as described by the manufacturer (Tropix, Inc.). Densitometric scans of X-ray exposures were performed with a model 2202 Ultrascan Densitometer (LKB, Inc.).

Miscellaneous assays. The activity of selenoprotein A was measured by glycine reductase assays performed as described previously (25, 29). The reaction mixtures contained complementary proteins B and C of the glycine reductase complexes from either *C. sticklandii* or *C. purinolyticum*. Protein concentrations were determined by using the bicinchonic acid protein assay (Pierce Chemical Co., Rockford, Ill.) as described by the manufacturer. Bovine gamma globulin was the standard protein.

Nucleotide sequence accession number. The sequence reported in this paper was deposited in EMBL/GenBank data base under accession number M60399.

RESULTS

Cloning and nucleotide sequence of the gene (*grdA*) encoding the selenoprotein component of the glycine reductase complex. The selenoprotein components of the glycine reductase complexes from *C. sticklandii* and *C. purinolyticum* are small, acidic, heat-stable proteins that share cross-reactivity with sheep anti-*C. sticklandii* selenoprotein A IgG and are partially interchangeable in heterologous enzyme preparations (25). The selenocysteine peptides derived from the two proteins by cleavage with endoproteinase Endo-Glu-C are identical (13, 24). Since the gene for the *C. purinolyticum* selenoprotein A already had been isolated (13), it was used to locate the corresponding gene in *C. sticklandii*. A DNA probe produced by amplifying the *C. purinolyticum grdA* gene from plasmid pGG7 with the polymerase chain reaction technique was used to screen a Southern blot of *C. sticklandii* genomic DNA digested to completion with restriction endonuclease *Hin*dIII. The probe hybridized to a DNA fragment of approximately 1,500 bp. The probe was used to screen a library of 1,000- to 2,000-bp size-selected *C. sticklandii* DNAs digested to completion with *Hin*dIII and cloned into pUC13 digested with *Hin*dIII. The probe hybridized to a clone containing a recombinant plasmid that had a DNA insertion of approximately 1,500 bp. The plasmid was designated pGG9. To generate a smaller fragment for sequencing, plasmid pGG9 was digested with *Sau*3A1, and a 550-bp fragment that hybridized to the probe on Southern blots was subcloned into plasmid pUC19 digested with *Bam*HI to produce plasmid pGG10.

On the basis of the insertion sequence determined from pGG10, new oligonucleotide primers were synthesized, and the remaining portion of the gene in pGG9 was sequenced. Both strands of the gene were sequenced. The entire selenoprotein A gene sequence was contained within the 1,500-bp fragment. The nucleotide sequence of the *grdA* gene is shown in Fig. 1 together with the predicted amino acid sequence. The deduced amino acid sequence of the corresponding *grdA* gene product from *C. purinolyticum* (13) is also indicated to show amino acid replacements and omissions. The *grdA* gene from *C. sticklandii* encodes a protein of 158 amino acids with a calculated molecular weight of 17,142. A putative ribosome binding site (GGAGG) is present 6 bases upstream from the initiator ATG. An in-frame termination codon, TGA, is located at position 130 in a nucleotide context that corresponds exactly to the amino acid sequence of the selenocysteine-containing peptide produced by Endo-Glu-C protease digestion (24). This provides one more example of the usage of an in-frame TGA codon to specify selenocysteine insertion into the gene product rather than to signal termination. The C-terminal aspartate, previously determined by hydrazinolysis of purified protein (9), is correctly predicted by the gene sequence. There are two tandem TAA translation termination codons. A second putative ribosome binding site (GGAGG) was found within 7 bases of the translation termination site for selenoprotein A. This site is located 9 bases upstream from a potential open reading frame (ORF 2) at base 499.

The codon usage (Table 1) for the selenoprotein A gene from *C. sticklandii* indicates that the G+C content is 39.3 mol%. The amino acids alanine, glycine, and proline, for which the first two codon positions are occupied by G or C, account for 16.7 mol% of the G+C content of the gene. An almost identical G+C content (39.7 mol%) was found for the gene from *C. purinolyticum* (13), and this was attributable to a high alanine-plus-glycine content. Similarly, the "wobble" position bases for all of the amino acid codons in the messages for both proteins are primarily A or U. In the *C. sticklandii* gene the exceptions are asparagine and serine, with 75 and 40 mol% G+C, respectively. The single histidine

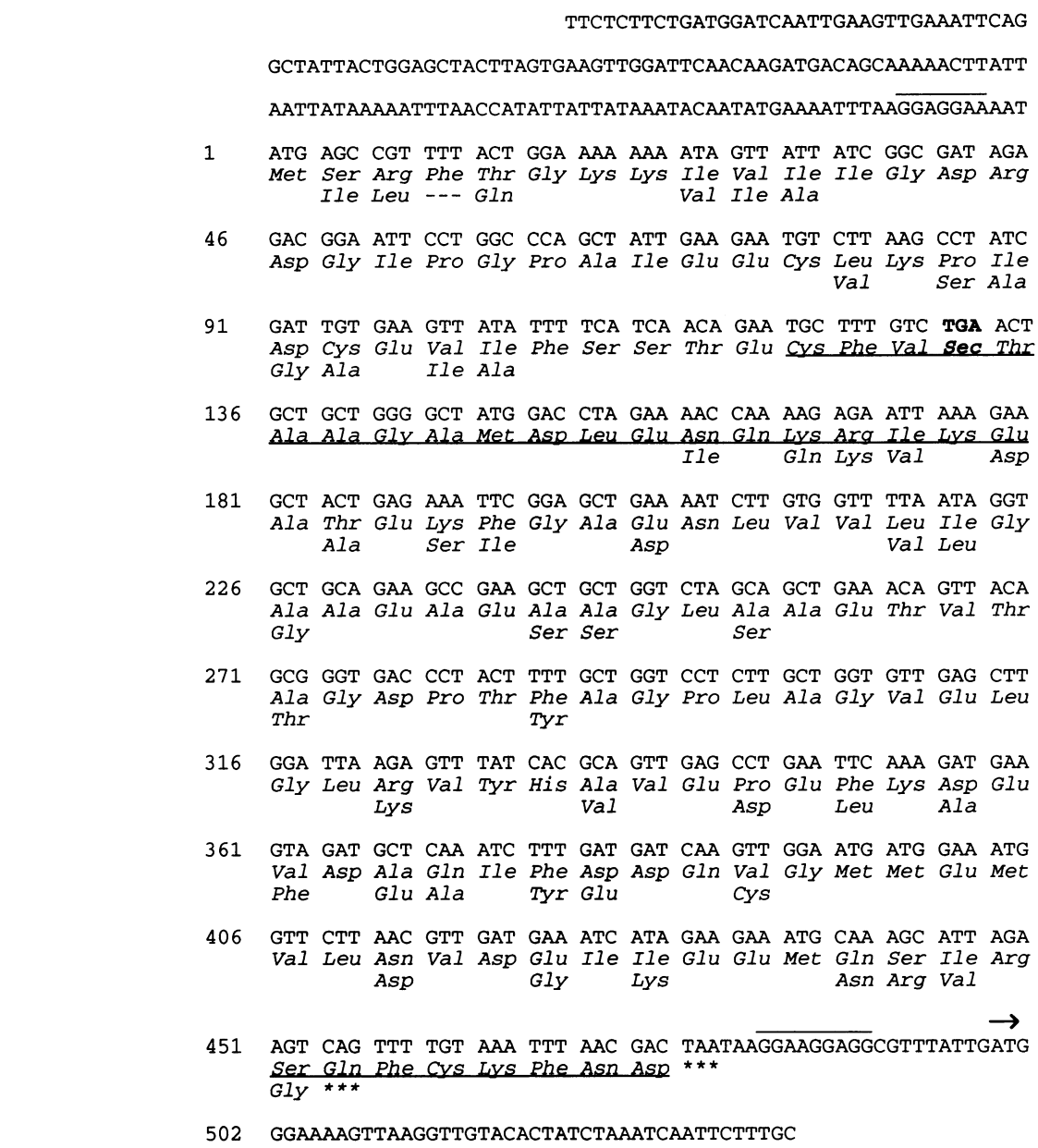

FIG. 1. Nucleotide sequence and predicted amino acid sequence of the *grdA* gene from *C. sticklandii*. The site of translation initiation is designated base 1. The putative ribosome binding sites for the *grdA* gene and for a second open reading frame are indicated by overlining. The codon corresponding to selenocysteine is found at base 130 and is shown in boldface. The potential second open reading frame site for initiation of translation is indicated by the arrow. Underlining indicates that the predicted sequence and the known peptide amino acid sequence are identical. The third line of sequences shows the predicted amino acid sequence of selenoprotein A from *C. purinolyticum* (13). Only the amino acid differences are shown. A single gap was inserted at amino acid position 4 and is indicated by dashes. The asterisks indicate the translation termination codon sites.

codon has a C at the wobble position. For leucine, with two possible bases (U and C) in the first position of the codon, 7 of the 9 codons start with C (77%); 100% of the wobble position bases are U or A.

Expression of *C. sticklandii* selenoprotein A gene in *E. coli*. It has been shown that efficient read-through of the UGA codon in the *E. coli fdhF* message requires the presence of a 40-base mRNA segment immediately downstream from the UGA codon, and a stem-loop structure has been proposed to explain this observation (35). This proposed stem-loop structure and a similar one consisting of 48 bases (2, 3) in the *E. coli fdnG* mRNA have calculated stability energies of −17.6 and −15.5 kcal/mol (-7.36×10^4 and -6.49×10^4 J/mol) respectively (Fig. 2A and B). The structure predicted for the *fdnG* gene product has a 5-base loop on the 3′ side of the stem, which results in a stem length similar to that of the *fdhF* structure. Additional evidence that a specific stem-loop structure downstream from UGA is required for read-through of UGA and insertion of selenocysteine came from attempts to express a mutated formate dehydrogenase gene

TABLE 1. Codon usage for *C. sticklandii* selenoprotein A gene

Amino acid	Codon	Usage
Phe	UUU	7
	UUC	2
Leu	UUA	2
	UUG	0
	CUU	5
	CUC	0
	CUA	2
	CUG	0
Ile	AUU	5
	AUC	4
	AUA	4
Met	AUG	6
Val	GUU	10
	GUC	1
	GUA	1
	GUG	1
Ser	UCU	0
	UCC	0
	UCA	2
	UCG	0
	AGU	1
	AGC	2
Pro	CCU	5
	CCC	0
	CCA	1
	CCG	0
Thr	ACU	4
	ACC	0
	ACA	3
	ACG	0
Ala	GCU	13
	GCC	1
	GCA	3
	GCG	1
Tyr	UAU	1
	UAC	0
Ter	UAA	1
His	CAU	0
	CAC	1
Gln	CAA	4
	CAG	1
Asn	AAU	1
	AAC	3
Lys	AAA	6
	AAG	2
Asp	GAU	7
	GAC	4
Cys	UGU	3
	UGC	1
Trp	UGG	0
Arg	CGU	1
	CGC	0
	CGA	0
	CGG	0
	AGA	4
	AGG	0
Gly	GGU	5
	GGC	2
	GGA	5
	GGG	1
Glu	GAA	16
	GAG	3
Sec	UGA	1

(*fdhA*) from *Methanobacterium formicicum* in *E. coli* (16). A segment of this gene and a corresponding segment of the *E. coli fdhF* gene are homologous, except that a cysteine codon, TGC, occurs in place of TGA. However, simple mutation of the TGC codon to TGA proved insufficient to allow synthesis of a selenocysteine-containing *M. formicicum* gene product in *E. coli*. Instead, additional mutations that generated a base sequence in the loop and a portion of the adjacent stem identical to that in the *E. coli fdhF* mRNA were required. This apparent specificity of a unique sequence requirement for a stem-loop structure indicates that factors other than structural stability alone are involved in UGA translation.

A computer-assisted analysis of the *C. purinolyticum grdA* sequence around the UGA site (13) demonstrated that there was a potential stem-loop secondary structure (Fig. 2C) with an energy of stability of −19.4 kcal/mol (-8.12×10^4 J/mol), suggesting a potential for regulation of selenocysteine insertion. However, when an mRNA secondary structure analysis was performed on the corresponding sequence of the *C. sticklandii grdA* gene product, a different result was obtained (Fig. 2D). For an equivalent sequence length, the predicted stem-loop structure is energetically weak (−8.0 kcal/mol [-3.35×10^4 J/mol]) and has two loops of 7 bases each and one loop of 13 bases. Such a stem-loop structure encompassing a 53-base sequence presumably would not be formed. Only when very large structures, such as those shown in Fig. 2E, were analyzed was evidence of a potential stem-loop structure of sufficient stability detected. If the *E. coli* mRNA secondary structure model for the specific insertion of selenocysteine is applicable to clostridia, then a segment of the *grdA* mRNA translated sequence should have this function. Alternatively, in at least one eukaryotic system, a segment of the 3′ untranslated region of the mRNA located downstream from the translation termination site is required for UGA read-through (4). This region of the tetraiodothyronine 5′-deiodinase message, termed a selenocysteine insertion element, has the potential to form an mRNA stem-loop secondary structure. When a similar segment derived from the 3′ untranslated region of the glutathione peroxidase message was substituted for the deleted deiodinase mRNA segment, it restored UGA read-through and synthesis of the deiodinase gene product. This implies that a similar mechanism is operative for selenocysteine insertion at the UGA codon for glutathione peroxidase gene expression.

To test the possibility that there is an mRNA structure in the clostridial selenoprotein A message that can function in an *E. coli* expression system to direct selenocysteine incorporation at the UGA codon, a *grdA* gene construction was used to transform *E. coli*. For this purpose the *grdA* gene of *C. sticklandii* was inserted behind the strong *E. coli tac* promoter of prokaryotic expression vector pKK223-3 to create plasmid pGG12. Following induction of the transformant, cell extract proteins were fractionated by SDS-polyacrylamide gel electrophoresis, transferred to a polyvinylidene difluoride membrane, and subjected to immunological analysis with affinity-purified sheep anti-*C. sticklandii* selenoprotein A antibodies. Extracts from cells transformed with plasmid pGG12 contained a protein that reacted specifically with the purified antibodies (Fig. 3A, lanes 4, and Fig. 4, lane A); no protein A was detected in a nontransformed cell extract (Fig. 3A, lanes 1, and Fig. 4, lane B) or in cells transformed with vector alone (Fig. 3A, lanes 2). The 18-kDa antibody-specific protein also was detected in extracts of cells containing pGG12 that were not induced with ITPG, but the amount was less (Fig. 3A, lanes 3), indicating that the level of synthesis under noninducing conditions was low. The expressed protein behaved in a manner identical to authentic *C. sticklandii* selenoprotein A on Western blots (Fig. 3B). The protein produced in recombinant cells was

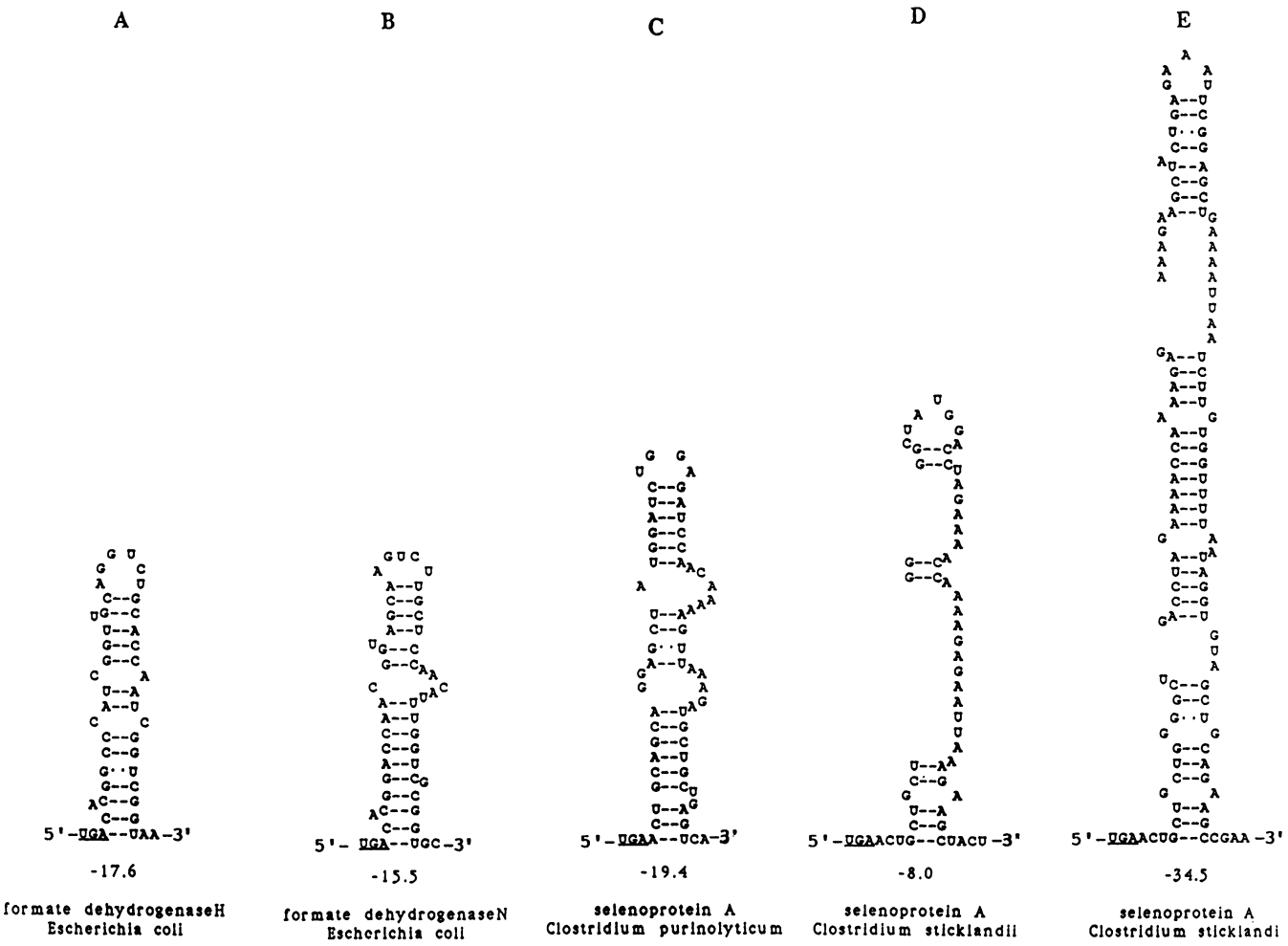

FIG. 2. Potential secondary structure analysis of prokaryotic selenoprotein mRNAs downstream from the selenocysteine UGA codons. The orientations of the sequences are indicated. The UGA codons are underlined. The calculated free energy values (in kilocalories per mole [1 kcal = 4,184 J]) are shown beneath the sequences. The structures and stability energies were predicted by using the RNAFOLD program of Zucker and Steigler (36) from the PC/GENE program package (IntelliGenetics, Inc.), using the energy parameters of Freier et al. (12). The sequences were determined by Zinoni et al. (35) (A), by Berg et al. (2) (B), by us previously (13) (recalculated by using the RNAFOLD program) (C), and by us in this study (D and E). Panels D and E show alternative structures of the UGA codon downstream sequence for *C. sticklandii*.

tested for its ability to complement glycine reductase preparations that had been freed of selenoprotein A. Protein fractions from wild-type cells transformed with pGG12 were able to reconstitute active glycine reductase when they were added to protein components B and C (Table 2), but no protein A activity was detected in comparable preparations of nontransformed cells (data not shown). The level of enzyme activity was dependent on the amount of selenium in the culture medium during growth (Table 2), whereas synthesis of immunologically reactive protein was selenium independent. The same amounts of 18-kDa antibody-specific protein were detected in strain DH5αMCR(pGG12) cells cultured in a defined low-sulfur medium in the absence of selenium as in the presence of 5 μM selenocysteine (data not shown). However, the protein fractions from cells grown in the absence of selenium exhibited no detectable selenoprotein A activity. In contrast, the enzyme from cells cultured in the presence of 5 μM selenocysteine had a specific activity of 1.14 μmol/mg (Table 2). The specific enzyme activities based on the amount of immunologically reactive protein in the various preparations from selenium-supplemented or selenium-deficient cells clearly show the dependence of catalytic activity on selenium availability (Table 3). Enzymically active protein was produced under both aerobic and anaerobic growth conditions in selenium-supplemented media, but the amount of active protein was greater in cells grown anaerobically. The lower level of activity in aerobically grown cells was not caused by depletion of selenium because of the high cell density attained, because increasing the amount of added selenite by fivefold did not increase enzyme activity (Table 2). When the activities of partially purified preparations of recombinant protein were compared with the activity of authentic selenoprotein A of comparable purity, the activity of the best *E. coli* product obtained was only about 1/10th that expected (1.5 versus 18.9 μmol of glycine reduced per mg of protein in 90 min) (Table 2).

In previous studies it had been shown that a *Salmonella typhimurium selA1* mutant (17) and *E. coli* MBO8 (19) are incapable of specific incorporation of selenocysteine into formate dehydrogenase directed by the UGA codon, yet

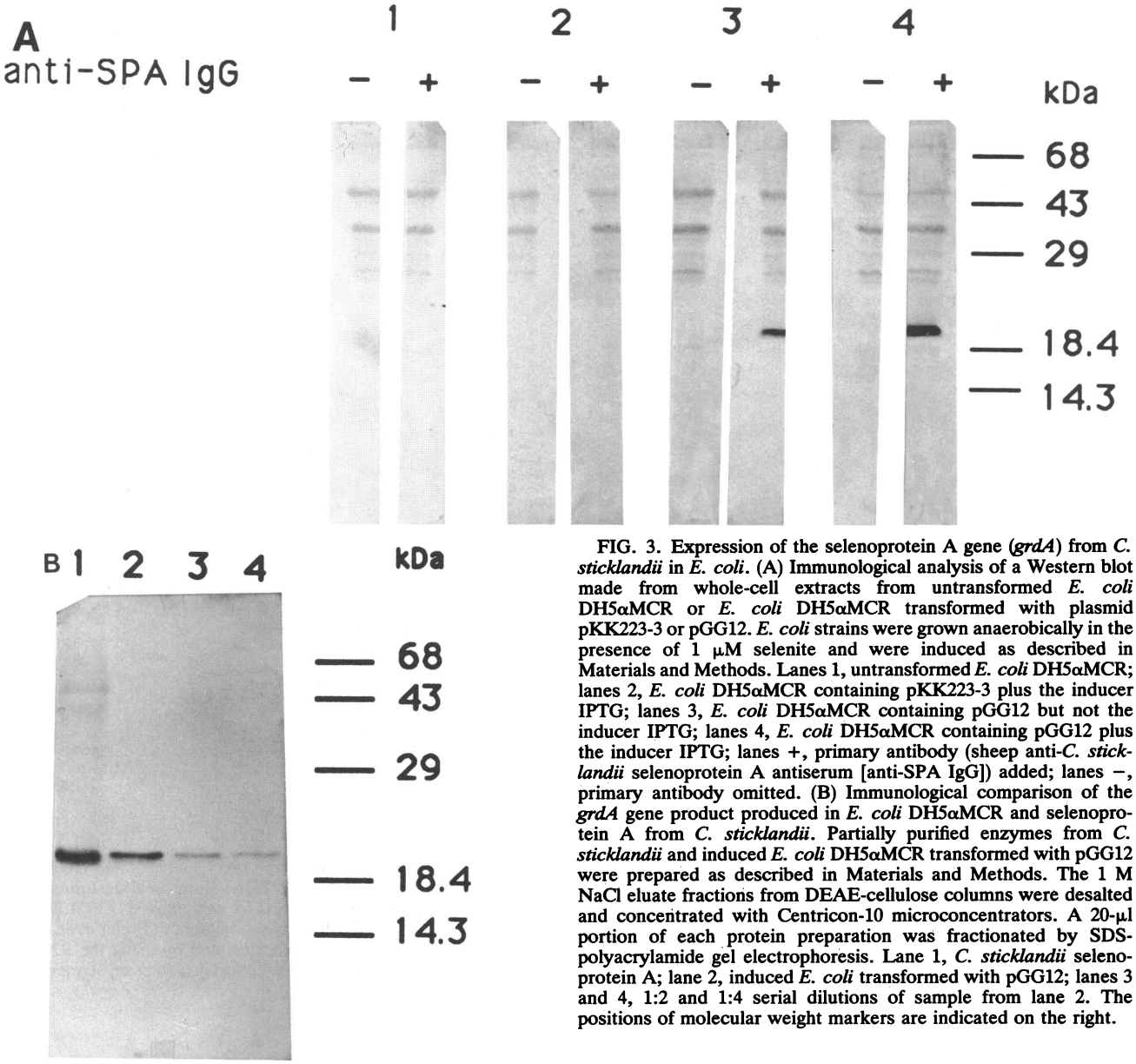

FIG. 3. Expression of the selenoprotein A gene (*grdA*) from *C. sticklandii* in *E. coli*. (A) Immunological analysis of a Western blot made from whole-cell extracts from untransformed *E. coli* DH5αMCR or *E. coli* DH5αMCR transformed with plasmid pKK223-3 or pGG12. *E. coli* strains were grown anaerobically in the presence of 1 μM selenite and were induced as described in Materials and Methods. Lanes 1, untransformed *E. coli* DH5αMCR; lanes 2, *E. coli* DH5αMCR containing pKK223-3 plus the inducer IPTG; lanes 3, *E. coli* DH5αMCR containing pGG12 but not the inducer IPTG; lanes 4, *E. coli* DH5αMCR containing pGG12 plus the inducer IPTG; lanes +, primary antibody (sheep anti-*C. sticklandii* selenoprotein A antiserum [anti-SPA IgG]) added; lanes −, primary antibody omitted. (B) Immunological comparison of the *grdA* gene product produced in *E. coli* DH5αMCR and selenoprotein A from *C. sticklandii*. Partially purified enzymes from *C. sticklandii* and induced *E. coli* DH5αMCR transformed with pGG12 were prepared as described in Materials and Methods. The 1 M NaCl eluate fractions from DEAE-cellulose columns were desalted and concentrated with Centricon-10 microconcentrators. A 20-μl portion of each protein preparation was fractionated by SDS-polyacrylamide gel electrophoresis. Lane 1, *C. sticklandii* selenoprotein A; lane 2, induced *E. coli* transformed with pGG12; lanes 3 and 4, 1:2 and 1:4 serial dilutions of sample from lane 2. The positions of molecular weight markers are indicated on the right.

both strains synthesize free selenocysteine and insert it nonspecifically into numerous cellular proteins (17, 27). In these mutants random substitution of selenocysteine for cysteine during protein synthesis appears to occur. Use of a cysteine isoacceptor tRNA mischarged occasionally with selenocysteine would be a plausible explanation for this (34). Because of a defective *selD* gene in the *Salmonella selA1* and *E. coli* MBO8 mutants (27), selenophosphate is not formed (32), and in the absence of this reactive selenium donor compound aminoacrylyl-$tRNA^{Sec}_{UCA}$ cannot be converted to selenocysteyl-$tRNA^{Sec}_{UCA}$, a component that is essential for UGA-directed specific incorporation of selenocysteine into protein (11). However, a cysteine tRNA (anticodon GCA) differing only in the "wobble" base of the anticodon might function as a UGA suppressor, thus allowing expression of the clostridial gene and insertion of cysteine (or occasionally selenocysteine). To test this possibility, *E. coli* mutant MBO8 was chosen as an alternate host for plasmid pGG12. As shown in Table 2, protein that exhibited detectable enzymic activity in the glycine reductase assay was synthesized in transformed strain MBO8 grown in Luria broth supplemented with selenite. An immunologically reactive protein of the expected molecular weight was present (Fig. 4, lane C), and the calculated specific activity of this protein was comparable to that of the protein produced by the transformed wild-type strain (Table 3). The inability of this strain to produce hydrogen from formate (in an in vivo assay for hydrogenase-linked formate dehydrogenase activity) indicated that no reversion of the *selD* mutation had occurred. A few attempts to increase selenium incorporation and thus induce synthesis of an enzyme exhibiting higher catalytic activity by growth in a low-sulfur medium supplemented with various levels of selenite or selenocysteine were not successful (Table 2). The tendency of strain MBO8 to favor selenomethionine formation over selenocysteine formation observed in previous studies (27) may be an important limiting factor under these conditions.

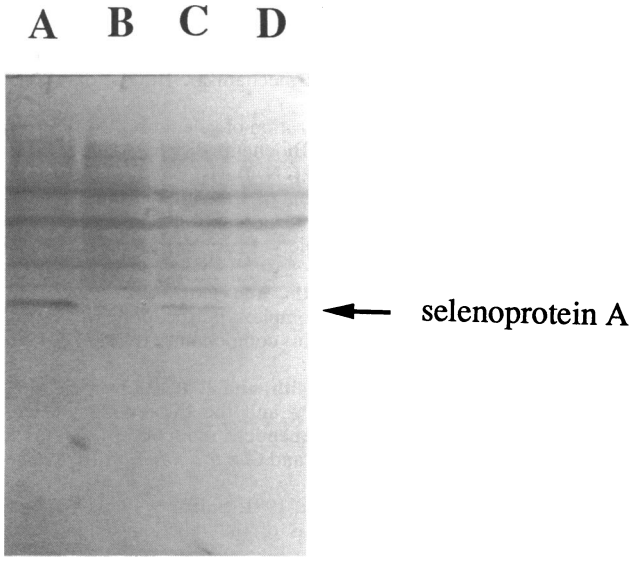

FIG. 4. Expression of immunologically reactive protein in *E. coli selD* mutant strain MBO8. Wild-type or mutant *E. coli* strains with or without plasmid pGG12 were induced with IPTG during anaerobic growth in Luria broth in the presence of 1 μM selenite. Immunological analysis was performed as described in Materials and Methods. Lane A, *E. coli* DH5αMCR transformed with pGG12; lane B, *E. coli* DH5αMCR; lane C, *E. coli* MBO8 transformed with pGG12; lane D, *E. coli* MBO8.

DISCUSSION

Native selenoprotein A isolated from *C. sticklandii* is an acidic, glycosylated protein species with a molecular mass of approximately 18 kDa based on its deduced amino acid content. A slightly larger apparent molecular mass based on its migration in SDS gels (Fig. 4) could be due both to its SDS-binding capacity and its glycosyl group content. Although the protein produced in *E. coli* has not been examined for the presence of glycosyl groups, it was indistinguishable electrophoretically from the enzyme isolated from *C. sticklandii* (Fig. 3B). The selenoprotein component of glycine reductase isolated from *Eubacterium acidaminophilum* also is a glycoprotein (10).

TABLE 2. Effect of selenium addition on biologically active selenoprotein A synthesis in *E. coli* host strains transformed with pGG12

Selenium supplement	Concn (μM)	Selenoprotein A sp act[a]		
		Wild-type strain[b] (aerobic conditions)	Wild-type strain[b] (anaerobic conditions)	Strain MBO8 (*selD*)[c] (anaerobic conditions)
None[d]		0.083	0.39	NT[e]
Selenite[d]	1	0.50	1.50	0.47
Selenite[d]	5	0.25	NT	NT
None[f]		NT	0	NT
Selenocysteine[f]	5	NT	1.14	NT
Selenocysteine[f]	10	NT	NT	0.3
Selenite[f]	1	NT	NT	0

[a] Expressed in micromoles of glycine reduced in 90 min per milligram of protein of comparable purity (acidic protein fractions eluted from DEAE-cellulose). The specific activity of *C. sticklandii* selenoprotein A of comparable purity was 18.9 μmol/mg, compared with 1.5 μmol/mg for selenoprotein A produced in wild-type *E. coli* cultures with optimal Se supplementation.
[b] *E. coli* DH5αMCR.
[c] *E. coli* strain that is not able to insert selenium into formate dehydrogenase and tRNAs (19) because of defective selenophosphate synthetase (32).
[d] Host cells were cultured in double-strength Luria broth.
[e] NT, not tested.
[f] Host cells were cultured in a glucose minimal medium (27) that was rendered low in sulfur concentration by replacing 0.8 mM $MgSO_4$ with 0.5 mM reduced glutathione.

TABLE 3. Specific activities of selenoprotein A synthesized in *E. coli* strains transformed with pGG12

Concn of selenium supplement (μM)	% Protein A sp act[a]		
	Wild-type strain[b] (aerobic conditions)	Wild-type strain[b] (anaerobic conditions)	Strain MBO8 (*selD*)[c] (anaerobic conditions)
0	9	34	
1	37	100	122
5	13		

[a] Specific activity is defined as follows: (micromoles of glycine reduced in 90 min)/(area units of selenoprotein A). The data were normalized to data for wild-type *E. coli* DH5αMCR(pGG12) grown under anaerobic conditions in medium supplemented with 1 μM selenite; 100% is defined as 1.32 μmol of glycine reduced in 90 min per 7.84 cm^2 (0.17). The area units were determined by calculation from densitometric scans of Kodak type SB5 X-ray film with an LKB Inc. model 2202 Ultrascan densitometer. The film was exposed for 1 min to the AMPP (Tropix Inc.) chemiluminescent product of alkaline phosphatase conjugated to rabbit anti-sheep IgG. For each sample 0.875 μg of protein was used.
[b] *E. coli* DH5αMCR.
[c] See Table 2, footnote *c*.

The codon usage for the selenoprotein A gene shows that A-ending codons for glutamate, lysine, and glutamine predominate and G-ending codons occur much less frequently (Table 1). This is of interest in view of the fact that in *C. sticklandii* the cognate tRNAs for these amino acids exist almost entirely as 2-selenouridine- rather than 2-thiouridine-containing species (7; unpublished data). Whereas a sulfur-modified uridine in the "wobble" position of the anticodons of these tRNAs results in a preference for A-ending codons, this bias is overcome by substitution of selenium for sulfur (21, 33). Since an adequate selenium supply in the culture medium is essential for selenoprotein A synthesis in *C. sticklandii*, this ensures that the 2-selenouridine forms of glutamate, lysine, and glutamine tRNA species are much more abundant than the corresponding sulfur-modified tRNAs. Thus, A-ending and G-ending codons for these amino acids should be translated at similar rates.

Expression of the *C. sticklandii* selenoprotein A gene in *E. coli* resulted in the synthesis of a full-length, immunologically reactive protein, indicating suppression and read-through of the in-frame UGA codon. However, this protein product exhibited only about 1/10th the expected catalytic activity of a protein A component of the glycine reductase complex. Upon further investigation it was found that synthesis of catalytically active protein depended on the level of selenium in the culture medium, but synthesis of immunologically reactive material was independent of the selenium level. Since a defined sulfur medium devoid of selenium supported synthesis of full-length, immunologically reactive protein equally well and this product exhibited no detectable catalytic activity, a nonspecific mechanism of UGA translation involving a cysteine tRNA species was suspected. Esterification of $tRNA^{Cys}$ with selenocysteine in place of cysteine is well documented in *E. coli* (34). Such a mechanism operating in lieu of the normal UGA-directed specific incorporation of selenocysteine would result in numerous

selenoprotein A molecules containing cysteine at the active site and only a few containing selenocysteine. Furthermore, in the absence of selenium in the medium, only cysteine would be incorporated, and an enzymically inactive protein would be formed. The ability of a *selD* mutant of *E. coli* (strain MBO8) that was transformed with the *grdA* plasmid to synthesize catalytically active selenoprotein A when selenium was provided showed that this premise was correct. The inability of this strain to synthesize selenophosphate (32) that is required for conversion of seryl-tRNA^{Sec} to selenocysteyl-tRNA^{Sec} prevents specific UGA-directed insertion of selenocysteine at an essential active site of an enzyme. However, in the absence of selenophosphate synthesis, selenocysteine is formed by an alternative route and is incorporated nonspecifically in conspicuous amounts in numerous proteins of the cell (17, 27). Under these conditions random insertion of selenocysteine in place of cysteine presumably occurred at positions corresponding to cysteine codons. In the selenoprotein A gene product synthesized in wild-type *E. coli* or in the *selD* mutant strain, cysteine and/or selenocysteine insertion must have occurred at the position corresponding to the UGA codon in addition to the four cysteine codons. Peptide maps and amino acid sequence data generated by using high-specific-activity ^{75}Se-labeled protein may provide direct information on these points. Even with ^{75}Se-labeling it still is necessary to distinguish between [^{75}Se]selenomethionine and [^{75}Se]selenocysteine in proteins that contain both methionine and cysteine.

The low catalytic activity of the *grdA* gene product synthesized in *E. coli* in the presence of selenium and the ability of the system to read through the in-frame UGA codon in the absence of selenium and synthesize the full-length inactive gene product suggest that the mRNA secondary structure required for specific UGA translational insertion of selenocysteine by *E. coli* (35) is not present in the clostridial message. In view of the present lack of detailed information concerning the factors required for expression of clostridial genes and the apparent incompatibility of the *E. coli* UGA translation system and the *C. sticklandii grdA* message, use of a compatible transformed clostridial strain appears to be an alternative.

ACKNOWLEDGMENTS

We thank Mike Nauman for technical assistance in the large-scale growth of *E. coli*, J. Nathan Davis for help with immunological assays, and Merry Peters for assistance in the preparation of the manuscript.

REFERENCES

1. **Arkowitz, R. H., and R. H. Abeles.** 1989. Identification of acetyl phosphate as the product of clostridial glycine reductase: evidence for an acyl enzyme intermediate. Biochemistry **28**:4639–4644.
2. **Berg, B. L., C. Baron, and V. Stewart.** 1991. Nitrate inducible formate dehydrogenase in *Escherichia coli* K-12. II. Evidence that a mRNA stem-loop is essential for decoding opal (UGA) as selenocysteine. J. Biol. Chem. **266**:22386–22391.
3. **Berg, B. L., J. Li, J. Heider, and V. Stewart.** 1991. Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. I. Nucleotide sequence of the *fdnGHI* operon and evidence that opal (UGA) encodes selenocysteine. J. Biol. Chem. **266**:22380–22385.
4. **Berry, J. M., L. Banu, Y. Y. Chen, S. J. Mandel, J. D. Kieffer, J. W. Harney, and P. R. Larsen.** 1991. Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3′ untranslated region. Nature (London) **353**:273–276.
5. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extract procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513.
6. **Bock, A., and T. C. Stadtman.** 1988. Selenocysteine, a highly specific component of certain enzymes, is incorporated by a UGA-directed co-translational mechanism. BioFactors **1**:245–250.
7. **Ching, W.-M.** 1986. Characterization of selenium tRNA^{Glu} from *Clostridium sticklandii*. Arch. Biochem. Biophys. **244**:137–146.
8. **Cone, J. E., R. Martin del Rio, J. N. Davis, and T. C. Stadtman.** 1976. Chemical characterization of the selenoprotein component of clostridial glycine reductase: identification of selenocysteine as the organoselenium moiety. Proc. Natl. Acad. Sci. USA **73**:2659–2663.
9. **Cone, J. E., R. Martin del Rio, and T. C. Stadtman.** 1977. Clostridial glycine reductase complex: purification and characterization of the selenoprotein component. J. Biol. Chem. **252**:5337–5344.
10. **Dietricks, D., M. Meyer, M. Reith, and J. R. Andreesen.** 1991. Interaction of selenoprotein P_A and the thioredoxin system, components of the NADPH-dependent reduction of glycine in *Eubacterium acidaminophilum* and *Clostridium litoralis*. J. Bacteriol. **173**:5883–5991.
11. **Forchhammer, K., and A. Böck.** 1991. Selenocysteine synthase from *Escherichia coli*. Analysis of the reaction sequence. J. Biol. Chem. **266**:6324–6328.
12. **Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**:9373–9377.

12a. **Garcia, G.** Unpublished data.

13. **Garcia, G. E., and T. C. Stadtman.** 1991. Selenoprotein A component of the glycine reductase complex from *Clostridium purinolyticum*: nucleotide sequence of the gene shows that selenocysteine is encoded by UGA. J. Bacteriol. **173**:2093–2098.
14. **Hammarback, J. A., and R. B. Vallee.** 1990. Antibody exchange immunochemistry. J. Biol. Chem. **265**:12763–12766.
15. **Harlow, E., and D. Lane.** 1988. Antibodies, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Heider, J., and A. Böck.** 1992. Targeted insertion of selenocysteine into the α subunit of formate dehydrogenase from *Methanobacterium formicicum*. J. Bacteriol. **174**:659–663.
17. **Kramer, G. F., and B. Ames.** 1988. Isolation and characterization of a selenium metabolism mutant of *Salmonella typhimurium*. J. Bacteriol. **170**:736–743.
18. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
19. **Leinfelder, W., K. Forchhammer, F. Zinoni, G. Sawers, M.-A. Mandrand-Berthelot, and A. Böck.** 1988. *Escherichia coli* genes whose products are involved in selenium metabolism. J. Bacteriol. **170**:540–546.
20. **Maniatis, T. E., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Ohashi, Z., M. Saneyoshi, F. Harada, H. Hara, and S. Nishimura.** 1970. Presumed anticodon structure of glutamic acid tRNA from *E. coli*: a possible location of a 2-thiorudine derivative in the first position of the anticodon. Biochem. Biophys. Res. Commun. **40**:866–872.
22. **Saito, H., and K.-I. Miura.** 1963. Preparation of transforming deoxyribonucleic acid by phenol treatment. Biochim. Biophys. Acta **72**:619–629.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
24. **Sliwkowski, M. X., and T. C. Stadtman.** 1988. Selenoprotein A of the clostridial glycine reductase complex: purification and amino acid sequence of the selenocysteine-containing peptide. Proc. Natl. Acad. Sci. USA **85**:368–371.
25. **Sliwkowski, M. X., and T. C. Stadtman.** 1988. Selenium-dependent glycine reductase: differences in physico-chemical properties and biological activities of selenoprotein A components isolated from *Clostridium sticklandii* and *Clostridium purinolyticum*. BioFactors **1**:293–296.

26. **Stadtman, T. C.** 1966. Glycine reduction to acetate and ammonia: identification of ferredoxin and another low molecular weight acidic protein as components of the reductase system. Arch. Biochem. Biophys. **113**:9–19.
27. **Stadtman, T. C., J. N. Davis, E. Zehelein, and A. Bock.** 1989. Biochemical and genetic analysis of *Salmonella typhimurium* and *Escherichia coli* mutants defective in specific incorporation of selenium into formate dehydrogenase and tRNAs. BioFactors **2**:35–44.
28. **Stadtman, T. C., P. Elliott, and L. Tiemann.** 1958. Studies on the enzymic reduction of amino acids. III. Phosphate esterification coupled with glycine reduction. J. Biol. Chem. **231**:961–973.
29. **Tanaka, H., and T. C. Stadtman.** 1979. Selenium-dependent clostridial glycine reductase: purification and characterization of the two membrane-associated protein components. J. Biol. Chem. **254**:447–452.
30. **Thompson, J. A.** 1986. A review of high-performance liquid chromatography in nucleic acids research. III. Isolation, purification, and analysis of supercoiled plasmid DNA. Biochromatography **1**:68–80.
31. **Turner, D. C., and T. C. Stadtman.** 1973. Purification of protein components of the clostridial glycine reductase system and characterization of protein A as a selenoprotein. Arch. Biochem. Biophys. **154**:366–381.
32. **Veres, Z., L. Tsai, T. D. Scholz, M. Politino, R. S. Balaban, and T. C. Stadtman.** 1992. Synthesis of 5-methylaminomethyl-2-selenouridine in tRNAs: ^{31}P NMR studies show the labile selenium donor synthesized by the *selD* gene product contains selenium bonded to phosphorus. Proc. Natl. Acad. Sci. USA **89**:2975–2979.
33. **Wittwer, A. J., and W.-M. Ching.** 1989. Selenium-containing $tRNA^{Glu}$ and $tRNA^{Lys}$ from *Escherichia coli*: purification, codon specificity, and translational activity. BioFactors **2**:27–34.
34. **Young, P. A., and I. I. Kaiser.** 1975. Aminoacylation of *Escherichia coli* cysteine tRNA by selenocysteine. Arch. Biochem. Biophys. **171**:483–489.
35. **Zinoni, F., J. Heider, and A. Böck.** 1990. Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. Proc. Natl. Acad. Sci. USA **87**:4660–4664.
36. **Zucker, M., and P. Steigler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.