JOURNAL OF BACTERIOLOGY, Mar. 1991, p. 2093–2098
0021-9193/91/062093-06$02.00/0

Vol. 173, No. 6

Selenoprotein A Component of the Glycine Reductase Complex from *Clostridium purinolyticum*: Nucleotide Sequence of the Gene Shows that Selenocysteine Is Encoded by UGA

GREGORY E. GARCIA* AND THRESSA C. STADTMAN

Laboratory of Biochemistry, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892

Received 1 October 1990/Accepted 9 January 1991

The gene encoding the selenoprotein A component of glycine reductase was isolated from *Clostridium purinolyticum*. The nucleotide sequence of this gene (*grdA*) was determined. The opal termination codon (TGA) was found in-frame at the position corresponding to the location of the selenocysteine residue in the gene product. A comparison of the nucleotide sequences and secondary mRNA structures corresponding to the selenoprotein A gene and the *fdhF* gene of *Escherichia coli* formate dehydrogenase shows that there is a similar potential for regulation of the specific insertion of selenocysteine at the UGA codon.

Selenium is specifically incorporated as selenocysteine into a few proteins of both eukaryotic and prokaryotic origin (for recent reviews, see references 16 and 18). Certain clostridial species contain a selenium-dependent glycine reductase complex composed of three protein components: selenoprotein A, protein B, and protein C (14, 19, 22). This complex, in the presence of acetate kinase, catalyzes the reductive deamination of glycine to acetate and ammonia with the concomitant synthesis of ATP (1, 17). The role of the selenoprotein component of the complex is especially interesting from the chemical point of view. A Schiff base intermediate formed between a reactive carbonyl group of protein B and glycine in the first step of the overall reaction (1, 19) is decomposed to give an Se-carboxymethyl derivative of the selenocysteine residue in selenoprotein A, presumably the result of a direct attack by the ionized selenol group (1). The selenoether derivative of protein A is then reductively cleaved in the presence of protein C, forming an acetyl ester of protein C. Phosphorolysis of this ester forms acetyl phosphate, the precursor of acetate and ATP.

The glycine reductase selenoprotein component of *Clostridium purinolyticum* is similar to the *Clostridium sticklandii* selenoprotein A (4, 5, 13, 14, 22). These are small, acidic, heat-stable proteins of about 12,000 molecular weight that contain a single selenocysteine (Sec) residue. Antibodies elicited to *C. sticklandii* protein A cross-react with *C. purinolyticum* protein A, and the two selenoproteins are partially interchangeable as complements of heterologous preparations of proteins B and C (13).

The amino acid sequence around the selenocysteine residue of protein A of *C. sticklandii* was determined (12) to be as follows: -CysPheValSecThrAlaAlaGlyAlaMetAspLeuGlu AsnGlnLys-. Early studies that used intact cells to investigate the mechanism of incorporation of selenocysteine into this protein had shown (15) that the selenium atom of added selenocysteine was rapidly incorporated whereas no incorporation of the carbon chain of the doubly labeled substrate was detected. A posttranslational modification of an existing amino acid residue (e.g., serine) in a preformed polypeptide seemed a viable possibility, although no immunologically reactive precursor could be detected. Since direct sequence analysis of the gene encoding this protein should give a more definitive clue to the origin of the selenocysteine residue, it was decided to approach the problem from this direction. Because of the lack of mutants in the glycine reductase pathway, the absence of practical methods for selection of such mutants, and the dearth of genetic information on the clostridial species that contain the glycine reductase complex, synthetic oligonucleotide probes were utilized to locate the selenoprotein A gene. In the case of *C. sticklandii* selenoprotein A, which has a blocked amino terminus, the available amino acid sequence data were limited to an internal region of the protein surrounding the selenocysteine residue. Although separate oligonucleotide probes corresponding to the predicted gene sequences of this area were synthesized, these probes proved to be insufficient to select the desired DNA fragment. The codon UGA was used for selenocysteine based on the finding that TGA corresponds to selenocysteine in mammalian glutathione peroxidase (3) and the presence of TGA in the formate dehydrogenase gene (*fdhF*) of *Escherichia coli* (25). In these initial studies, genomic *C. sticklandii* DNA digested to completion with restriction nucleases was used. A high degree of nonspecific binding of the probes to regions of DNA fragments rich in adenine and thymidine residues was observed, and this effect prevented selection of a specific DNA fragment on Southern blots of DNA digests (7a). An analogous selenoprotein A component of a glycine reductase produced by a purine-fermenting organism, *C. purinolyticum* (6, 13), was used to obtain more amino acid sequence data for generation of additional complementary oligonucleotide probes. This protein proved to have a free N terminus, and useful sequence data for the N-terminal region could be obtained by direct analysis of the intact protein. The additional amino acid sequence information and the successful isolation of the selenoprotein A gene from *C. purinolyticum*, together with its nucleotide sequence analysis, are described in this communication.

MATERIALS AND METHODS

Materials. T4 DNA ligase, Taq DNA polymerase, restriction enzymes, and *E. coli* DH5α competent cells were from

* Corresponding author.

Bethesda Research Laboratories. Modified T7 DNA polymerase (Sequenase) was from United States Biochemical Corp. Phage M13 universal and reverse sequencing primers, plasmids pUC-13 and pUC-19, were from Pharmacia. [γ-^{35}S]dATP (500 Ci/mmol), [α-^{32}P]dCTP (3,000 Ci/mmol), and [γ-^{32}P]ATP (3,000 Ci/mmol) were purchased from New England Nuclear. Iodoacetic acid was from Aldrich Chemical Co., and ampicillin was from Sigma Chemical Co. All other reagents were of the highest grade available.

Bacterial growth and selenoprotein A purification. *C. purinolyticum* ATCC 33906 was grown in a glycine-yeast extract-mineral salts medium (6) supplemented with 1 μM SeO_3^{2-} and 11 mM uric acid. For ^{75}Se-labeled cells, 2-liter cultures containing the same medium supplemented with $^{75}SeO_3^{2-}$ (1 to 2 mCi/μmol) were prepared. Cell lysates were prepared in a French pressure cell. ^{75}Se-labeled selenoprotein A was isolated essentially as described before (13) except that the protein was reduced and alkylated with iodoacetate prior to preparative DEAE–high-performance liquid chromatography (HPLC). As a final purification step, the protein was subjected to molecular sieve chromatography on a prototype Superdex 75 HR 10/30 column (10-μm beads) from Pharmacia LKB Biotechnology Inc., Uppsala, Sweden, equilibrated in 6 M guanidine hydrochloride–0.1% trifluoroacetic acid and developed with the same buffer at a flow rate of 0.3 ml/min. Guanidine was removed and the protein was concentrated on a Centricon-10 microconcentrator (Amicon, Inc.). The protein was stored in 10 mM sodium phosphate (pH 7.1)–1 mM EDTA.

Amino-terminal sequence analysis. Partially purified ^{75}Se-labeled selenoprotein A was fractionated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were transferred to a polyvinylidine difluoride membrane (Immobilon, Millipore Corp.) as described by the manufacturer. The blotted proteins were stained with Coomassie blue, and then the blot was exposed to Kodak X-OMAT X-ray film. A single ^{75}Se-labeled protein (apparent molecular weight, 18,000) that migrated as expected for the pure protein (M_r, 15,700, based on amino acid composition [13]) was found. The protein band was cut out and the amino-terminal sequence was determined on an ABI 477A protein sequencer at the Macromolecular Structure Facility at Michigan State University, East Lansing.

Endoproteinase Glu-C digest and purification of peptides by reverse-phase HPLC. One milligram of purified selenoprotein A was dialyzed versus 50 mM NH_4CO_3, pH 8, in Spectropor 3 tubing (3,500-molecular-weight cutoff) for 12 h at 4°C. To the dialyzed protein solution (4 ml) was added 0.5 ml of 0.5 M Tris hydrochloride, pH 8. Protease digestion was begun by the addition of 10 μl of a freshly prepared protease Endo-Glu-C (*Staphylococcus aureus* V8 protease) solution (10 μg/ml in 0.5 M Tris hydrochloride, pH 8). The digestion mixture was incubated at 25°C. After 3 h, 0.25 ml of 2% SDS was added along with another 10-μl aliquot of protease, and the digestion was continued for 12 h.

The reaction was terminated by the addition of 2 N HCl to pH 2, and acetonitrile was added to a final concentration of 5%. An aliquot (2 ml) of the sample was injected onto a Vydac C_{18} column attached to a Hewlett-Packard 1090 HPLC. The solvent system used consisted of buffer A (94.5% H_2O, 5% CH_3CN, 0.5% trifluoroacetic acid) and buffer B (CH_3CN–0.5% trifluoroacetic acid). The development program was as follows: 100% buffer A, 0 to 15 min; to 55% buffer B, 15 to 65 min; to 100% buffer B, 65 to 70 min. The flow rate was 1 ml/min. Fractions containing identical peptide peaks (from several runs) were combined and concentrated to dryness in a Savant Speed-Vac apparatus. The pooled peptides were dissolved in 6 M guanidine hydrochloride–0.5% trifluoroacetic acid and rerun on the C_{18} column, using a stepwise gradient. Stepwise gradient elution of the peptides was accomplished as follows. The sample was loaded onto a column equilibrated in buffer A. A linear gradient of buffer B was increased to within 1% of the elution concentration of the original HPLC run within 5 min, and the column was washed for another 5 min. Buffer B was increased to the elution concentration in a single step, and the column was washed for 5 min to elute the peptides. The appropriate peptide peaks were collected, the samples were dried, and the peptides were subjected to sequence analysis as described above.

DNA isolation and recombinant DNA methods. Total genomic DNA was isolated by the procedure of Saito and Miura (10). General genetic manipulations were performed as described before (9). Small-scale plasmid preparations were by the alkaline lysis method of Birnboim and Doly (2) as modified by Thompson (20). Plasmid isolated for sequencing was purified by CsCl ultracentrifugation. Supercoiled plasmid was sequenced by the dideoxy-chain termination method of Sanger et al. (11) modified for use with [^{35}S]dATP and modified T7 DNA polymerase (Sequenase kit instructions; U.S. Biochemical Corp.). Specific oligonucleotide primers and probes were synthesized on an ABI 380B DNA synthesizer and purified by electrophoresis on polyacrylamide gels.

Gene amplification for probe production. A modified polymerase chain reaction technique was used to amplify a portion of the selenoprotein A gene. Degenerative oligonucleotide primers corresponding to the amino-terminal portion (primer A) and the internal selenocysteine peptide sequence (primer B) that reflect every codon combination per the universal genetic code were synthesized. The amplification mix was essentially as described previously (8) except the deoxynucleoside triphosphate concentrations used were 1 mM. The polymerase chain reaction was performed in a Perkin-Elmer thermal cycler. The amplification mixture was heated to 94°C for 2 min and then cooled to room temperature. Taq DNA polymerase (2.5 U/μl) was added to a total of 1.25 U, and the final mixture was overlaid with mineral oil. The cycling reaction consisted of the following steps: denaturation at 94°C for 2 min, hybridization at 65°C for 2 min, and chain polymerization at 72°C for 3 min.

A. MetIleLeuGlnGlyLysLysValIleAlaIleGlyAspAspAspGlyIleProGlyPro

B. AspGlnCysGlyMetMet

C. IleAlaPheSerSerThrGlu

D. AspGluLeuLysAlaGlu

E. CysValLysSerAla

F. PheAspGluAlaIleTyrGlu

G. ThrValThrThrGlyAspProThrTyrAlaGlyProLeuAlaGlyValGlu

H. CysPheValSecThrAlaAlaGlyAlaMetAspLeuGlu

I. MetAsnArgValArg

FIG. 1. Partial protein sequence of the selenoprotein component of the glycine reductase complex from *C. purinolyticum*. A, Amino-terminal peptide; B through I, Endo-Glu-C peptides listed in their order of elution from a C_{18} reverse-phase HPLC column with increasing acetonitrile concentrations of 8, 15, 24, 24.5, 24.7, 25, 29, and 39%, respectively. Selenocysteine is represented by Sec.

A. EcoRI *GlnGlyLysLysValIleAlaIleGly*
5′-GAATTCCAPGGNAAPAAPGTNAT3GCNAT3GG-3′

B. SacI *AspMetAlaGlyAlaAlaThrSecValPheCys*
5′-GAGCTCQTCCATNGCNCCNGCNGCNGTTCANACPAAPCA-3′

FIG. 2. Oligonucleotides for the polymerase chain reaction amplification of the selenoprotein A gene from *C. purinolyticum* genomic DNA. Degenerative oligonucleotide primers were synthesized to the amino acid sequences of the amino-terminal region and the ^{75}Se-labeled peptide from an Endo-Glu-C protease digestion of purified selenoprotein A: A, amino-terminus-predicted noncoding strand sequence, degeneracy of 4,608; B, selenopeptide-predicted gene sequence, degeneracy of 32,768. The oligonucleotide bases are represented by the following designations: A, deoxyadenosine; C, deoxycytidine; G, deoxyguanosine; T, deoxythymidine; N is A, C, G, or T; P is either A or G; Q is either C or T; 3 is A, C, or T. *Eco*RI and *Sac*I indicate restriction enzyme sites added to provide vector directional cloning sites, if necessary.

A 10-s extension was added to each subsequent polymerization step. The reaction was allowed to proceed for 30 cycles. The reaction was linked to a final polymerization step of 15 min at 72°C. The amplified DNA was analyzed on a Phastsystem, using a homogeneous 20% polyacrylamide gel (Pharmacia). The DNA was visualized by silver staining. A 150-bp major product was detected. The 150-bp product was purified by electrophoresis on polyacrylamide gels. The purified product was labeled with [α-^{32}P]dCTP by nick translation (Bethesda Research Laboratories kit instructions).

Nucleotide sequence accession number. The sequence reported in this paper was deposited in the EMBL/GenBank data base and given accession no. M55254.

RESULTS

Amino-terminal and peptide amino acid sequence. Selenoprotein A from *C. purinolyticum* is not blocked at the amino terminus, and therefore the amino-terminal sequence could be determined directly (Fig. 1, peptide A). ^{75}Se was a convenient marker for the isolation of the selenocysteine-containing peptide following digestion of purified protein with endoproteinase Endo-Glu-C. To aid in the identification of the gene, several other peptides also were isolated and sequenced. The amino acid sequences for the peptides isolated are shown in Fig. 1. Comparison of the selenocysteine peptide from *C. purinolyticum* (peptide H) with that of *C. sticklandii* (12) showed that the sequences are identical. In total, about 55% of the *C. purinolyticum* protein was sequenced.

Cloning and nucleotide sequence of the gene (*grdA*) for the selenoprotein component of the glycine reductase complex. The finding that the *C. purinolyticum* selenoprotein A has a free N terminus provided the opportunity to produce a probe by the polymerase chain reaction method. The probe termini having the gene sequence corresponding to the amino-terminal region (peptide A) on one end and a sequence corresponding to the internal selenocysteine peptide (peptide H) on the other were synthesized. The sequences of the primers used are shown in Fig. 2. The primers were used to amplify a 150-bp product from *C. purinolyticum* genomic DNA. Southern blot analysis revealed that the labeled probe hybridized to a 3,600-bp fragment from genomic DNA digested to completion with restriction enzyme *Hin*dIII. This fragment was inserted into the *Hin*dIII site of plasmid pUC-13 to generate plasmid pGG7. To generate a smaller DNA fragment for sequencing, pGG7 was digested with restriction enzyme *Sau*3AI. A 300-bp fragment was found to hybridize to the probe on a Southern blot. This fragment was subcloned into the *Bam*HI site of pUC-19 to produce plasmid pGG8.

The insert was found to contain about 150 bp of sequence upstream from the apparent site of translation initiation and 150 bp of the selenoprotein A gene. Appropriate oligonucleotide primers were synthesized and used to sequence the pertinent portion of the insert in pGG7. The overall sequencing strategy is shown in Fig. 3. Both strands of the gene were sequenced. The nucleotide sequence of the selenoprotein A gene is shown in Fig. 4. The amino acid sequence predicted from the gene sequence was in complete agreement with the known *C. purinolyticum* peptide sequences (Fig. 1), except for amino acid 14. The amino acid analysis indicated that Asp was at this position (Fig. 1), whereas the nucleotide sequence predicts an Arg (Fig. 4). Since it would require a double mutation in the two "nonwobble" positions to convert an Asp codon to an Arg codon, it seems likely that Arg is the true amino acid for this location. This finding is supported by amino acid composition analysis (13) in which three Arg residues were found in the protein. The gene sequence predicts a total of three Arg residues.

The gene encodes a protein of 150 amino acids with a calculated molecular weight of 15,757. This is in good agreement with the apparent molecular weight of 18,000 determined by SDS-PAGE analysis (7a). A putative ribosome-binding site (GGAGG) was present 8 bases upstream from the initiator ATG. An in-frame termination codon "TGA" was found at position 127. The codon for termination of translation was TAA. Even though the gene sequence was determined up to 80 bases downstream from the translation terminator, there was no apparent transcription termination site (23) detected. In fact, a second possible ribosome-binding site (GGAGG) was found within 16 bases of the translation termination site for selenoprotein A. The second putative binding site is located 11 bases upstream from a potential open reading frame (ORF 2) at base 484. Table 1 shows the codon usage in the selenoprotein A gene.

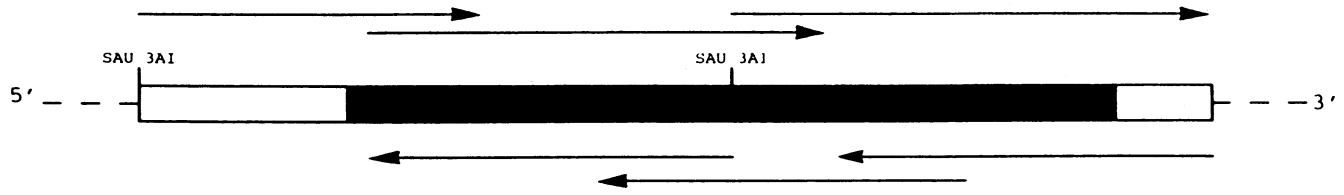

FIG. 3. Sequencing strategy for the selenoprotein A gene (*grdA*) and surrounding DNA. The gene location is indicated by the filled bar.

```
                                                    GATCAATAGAAGCAGAAATCCAAGTTAT

      AACTGGAGCTACTTGCGAAGCGAAGTTGGATTTGGATACTTAACAGCTAAGACATATTA

      ATTTAATTAAATAAAATCGAATTATATATAATATTATAAATTTTTAGGAGGTACTTGTT

  1   ATG ATA TTA CAA GGA AAA AAA GTT ATA GCA ATA GGA GAT AGA GAT
  1   Met Ile Leu Gln Gly Lys Lys Val Ile Ala Ile Gly Asp(Arg)Asp

 46   GGT ATT CCA GGA CCA GCT ATC GAA GAA TGT GTA AAA TCA GCA GGA
 16   Gly Ile Pro Gly Pro Ala Ile Glu Glu Cys Val Lys Ser Ala Gly

 91   GCA GAA ATT GCT TTC TCA TCA ACT GAA TGT TTT GTC TGA ACT GCA
 31   Ala Glu Ile Ala Phe Ser Ser Thr Glu Cys Phe Val Sec Thr Ala

136   GCA GGA GCT ATG GAT CTG GAG ATC CAA CAA AAA GTT AAA GAT GCT
 46   Ala Gly Ala Met Asp Leu Glu Ile Gln Gln Lys Val Lys Asp Ala

181   GCT GAG TCA ATC GGA GCA GAC AAT TTA GTA GTT GTT CTT GGA GGA
 61   Ala Glu Ser Ile Gly Ala Asp Asn Leu Val Val Val Leu Gly Gly

226   GCT GAA GCT GAA TCA TCA GGA CTA TCA GCA GAA ACA GTA ACA ACA
 76   Ala Glu Ala Glu Ser Ser Gly Leu Ser Ala Glu Thr Val Thr Thr

271   GGA GAC CCA ACA TAT GCA GGA CCT TTA GCT GGA GTA GAG CTA GGA
 91   Gly Asp Pro Thr Tyr Ala Gly Pro Leu Ala Gly Val Glu Leu Gly

316   CTA AAA GTT TAC CAC GTT GTA GAA GAT GAA CTA AAG GCT GAA TTT
106   Leu Lys Val Tyr His Val Val Glu Asp Glu Leu Lys Ala Glu Phe

361   GAT GAA GCT ATA TAT GAA GAT CAA TGC GGA ATG ATG GAA ATG GTT
121   Asp Glu Ala Ile Tyr Glu Asp Gln Cys Gly Met Met Glu Met Val

406   CTT GAC GTT GAC GGA ATA AAG GAA GAA ATG AAT AGG GTT AGA GGA
136   Leu Asp Val Asp Gly Ile Lys Glu Glu Met Asn Arg Val Arg Gly

                                                    →
451   TAA TCCACCCAAAATAAAGGAGGGGCTAATAAGATGATTCGTGTAGTACACTATATAA
      ***

509   AGGAATTCTTCGCTAATATTGGTGGAGAAGAAATGGCTCATGTTGCACCAGAAGTTAGA

567   GAAGGTGTAGTAGGACCAGGTAT
```

FIG. 4. Nucleotide sequence and predicted amino acid sequence of the *grdA* gene from *C. purinolyticum*. The site of translation initiation is designated base 1. The codon corresponding to selenocysteine is found at base 127. The potential second open reading frame site for initiation of translation is indicated by the overline. Underlining indicates homology to peptide amino acid sequence (Fig. 1), and parentheses indicate ambiguity.

DISCUSSION

The codon usage found for the selenoprotein A gene (Table 1) indicates that the G+C content is 39.7%. This is due to the high (23%) alanine-plus-glycine content of the protein. The "wobble" position bases for all amino acids are overwhelmingly A or U. Only aspartate and histidine codons had >50% G or C (4 of 7 C for aspartate and 1 of 1 C for histidine) at the wobble position. Arginine, cysteine, and tyrosine codons had between 30 and 50% G or C at the wobble position. In the case of leucine, with which there are two possibilities for the first position (U or C), 7 of 10 codons were C, but the wobble positions for those codons beginning with C were 6 of 7 A or U.

Interestingly, a TGA codon was found in-frame at position 127 (Fig. 4) in the selenoprotein A gene, and this position corresponds to the location of the selenocysteine residue in the protein. This is the first report of this coding feature in clostridia. Only two other selenoprotein genes have been reported to have TGA (UGA) codons that direct the specific incorporation of selenocysteine: a formate dehydrogenase of *E. coli* (25) and mammalian glutathione peroxidase (3). Normally, this codon is a translation termination signal that is recognized by release factor 2.

The best-understood system for the specific incorporation of selenocysteine into protein has been described in *E. coli* in which a cotranslational mechanism has been suggested (24). Why selenocysteine is incorporated at this site as opposed to the binding of release factor 2 to terminate translation is unknown. Böck and coworkers (26) have proposed an mRNA secondary-structure regulation hypothesis for selenocysteine insertion at the UGA site. In the case of *E. coli* formate dehydrogenase (*fdhF*), a computer-assisted examination of the nucleotide sequence around the site showed the possibility that stem-loop structure transitions could be required for the insertion of selenocysteine. The model predicts that initially a stem-loop structure exists 20 bases upstream from the UGA codon, with a calculated free energy value of −8.6 kcal (ca. −36 kJ)/mol, and this could play a role in slowing ribosomes as they approach the UGA region. A second stem-loop structure follows almost immediately, with the UGA codon present in a loop of 5 bases. The calculated free energy value of the second stem structure is −10 kcal (ca. −42 kJ)/mol. It is proposed that a

TABLE 1. Codon usage in selenoprotein A gene of *C. purinolyticum*

Amino acid	Codon	Usage	Amino acid	Codon	Usage	Amino acid	Codon	Usage	Amino acid	Codon	Usage
Phe	UUU	2	Ser	UCU	0	Tyr	UAU	2	Cys	UGU	2
	UUC	1		UCC	0		UAC	1		UGC	1
				UCA	7						
Leu	UUA	3		UCG	0	Ter	UAA	1	Trp	UGG	0
	UUG	0		AGU	0		UAG	0			
	CUU	2		AGC	0				Arg	CGU	0
	CUC	0								CGC	0
	CUA	4	Pro	CCU	1	His	CAU	0		CGA	0
	CUG	1		CCC	0		CAC	1		CGG	0
				CCA	3					AGA	2
				CCG	0	Gln	CAA	4		AGG	1
Ile	AUU	2					CAG	0			
	AUC	3	Thr	ACU	2				Gly	GGU	1
	AUA	5		ACC	0	Asn	AAU	2		GGC	0
				ACA	4		AAC	0		GGA	16
Met	AUG	6		ACG	0					GGG	0
						Lys	AAA	6			
Val	GUU	9	Ala	GCU	10		AAG	2	Glu	GAA	15
	GUC	1		GCC	0					GAG	3
	GUA	5		GCA	8	Asp	GAU	7			
	GUG	0		GCG	0		GAC	4	Sec[a]	UGA	1

[a] Sec, Selenocysteine.

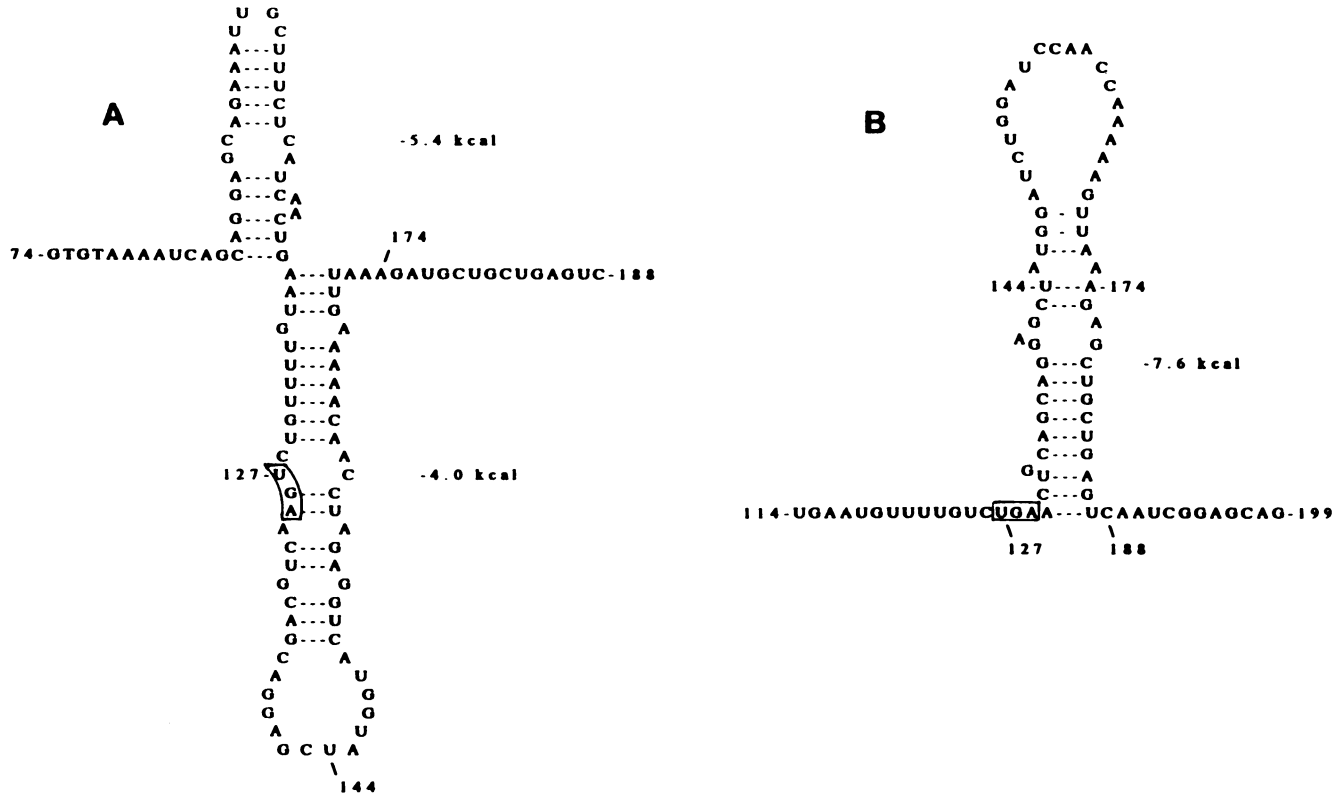

FIG. 5. Proposed secondary structure of *grdA* mRNA around the UGA_{127} codon for selenocysteine. The structures were predicted by the Microgenie (Beckman) program. The calculated free energy values for the stem-loop structures are indicated and were determined by the method of Tinoco et al. (21).

ribosome would melt the structure during translation to allow the downstream half of the stem to participate in the formation of a new stem-loop structure further downstream. In the new stem-loop structure, the A residue of the UGA codon is the first base in the stem. The calculated free energy value of the stem structure is −17.6 kcal (ca. −74 kJ)/mol. The proposed function of the new structure is to allow the recognition of this site by the aminoacylated tRNA-elongation factor complex (7) or to prevent the recognition of the site by release factor 2.

To see whether the mRNA secondary-structure hypothesis for the regulation of selenocysteine insertion could apply to the selenoprotein A mRNA, a computer-assisted stem-loop search was used to examine the potential secondary structures around the UGA codon. The results of the search are shown in Fig. 5. A potential stem-loop structure exists upstream from the UGA site (Fig. 5A) which is similar to that found in the *fdhF* mRNA. The calculated free energy value for this structure is −5.4 kcal (ca. −23 kJ)/mol. Although a structure in which the UGA codon would exist in a loop was not found, a potential stem-loop structure was found in which the GA portion of the codon could participate in a stem (Fig. 5A). The calculated free energy value of the overall stem is −4 kcal (ca. −17 kJ)/mol. If this structure is melted as a ribosome traverses the site, a portion of the loop could base pair with the nucleotide sequence further downstream. A new stem-loop structure could form in which the stem begins 1 base past the UGA codon (Fig. 5B). The calculated free energy value of the new stem is −7.6 kcal (ca. −32 kJ)/mol. The second stem-loop structure is similar to that described in *E. coli* except the UGA codon may not be involved in the final stem structure. Although these predicted structures for the clostridial mRNA are weaker energetically, the structures do resemble those predicted for *E. coli* and suggest that, in this case also, mRNA secondary structure may govern the insertion of selenocysteine at the UGA codon.

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. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Chambers, I., J. Frampton, P. Goldfarb, N. Affara, W. McBain, and P. R. Harrison.** 1986. The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the "termination" codon, TGA. EMBO J. **5**:1221–1227.
4. **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.
5. **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.
6. **Durre, P., and J. R. Andreesen.** 1982. Selenium-dependent growth and glycine fermentation by *Clostridium purinolyticum*. J. Gen. Microbiol. **128**:1457–1466.
7. **Forchhammer, K., K.-P. Rucknagel, and A. Böck.** 1990. Purification and biochemical characterization of SelB, a translation

factor involved in selenoprotein synthesis. J. Biol. Chem. **265**:9346–9350.

7a. **Garcia, G.** Unpublished data.

8. **Gould, S. J., S. Subramani, and I. E. Scheffler.** 1989. Use of the DNA polymerase chain reaction for homology probing: isolation of partial cDNA or genomic clones encoding the iron-sulfur protein of succinate dehydrogenase from several species. Proc. Natl. Acad. Sci. USA **86**:1934–1938.

9. **Maniatis, T. E., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

10. **Saito, H., and K.-I. Miura.** 1963. Preparation of transforming deoxyribonucleic acid by phenol treatment. Biochim. Biophys. Acta **72**:619–629.

11. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

12. **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.

13. **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.

14. **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.

15. **Stadtman, T. C.** 1987. Specific occurrence of selenium in enzymes and amino acid tRNAs. FASEB J. **1**:375–379.

16. **Stadtman, T. C.** 1990. Selenium biochemistry. Annu. Rev. Biochem. **59**:111–127.

17. **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.

18. **Sunde, R. A.** 1990. Molecular biology of selenoproteins. Annu. Rev. Nutr. **10**:451–474.

19. **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.

20. **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.

21. **Tinoco, I., P. N. Borer, B. Dengler, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246**:40–41.

22. **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.

23. **Whitehead, T. R., and J. C. Rabinowitz.** 1988. Nucleotide sequence of the *Clostridium acidiurici* ("*Clostridium acidiurici*") gene for 10-formyltetrahydrofolate synthetase shows extensive amino acid homology with the trifunctional enzyme C_1-tetrahydrofolate synthase from *Saccharomyces cerevisiae*. J. Bacteriol. **170**:3255–3261.

24. **Zinoni, F., A. Birkmann, W. Leinfelder, and A. Böck.** 1987. Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon. Proc. Natl. Acad. Sci. USA **84**:3156–3160.

25. **Zinoni, F., A. Birkmann, T. C. Stadtman, and A. Böck.** 1986. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogenase-lyase-linked) from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83**:4650–4654.

26. **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.