RNA-binding Proteins That Specifically Recognize the Selenocysteine Insertion Sequence of Human Cellular Glutathione Peroxidase mRNA*

Qichang Shen, Patricia A. McQuilkin, and Peter E. Newburger†

From the Departments of Pediatrics and Molecular Genetics/Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts 01655

Translational incorporation of the unusual amino acid selenocysteine in eukaryotes requires a coding region UGA codon (which otherwise serves as a termination signal), a selenocysteine insertion sequence (SECIS) in the 3′-untranslated region of the mRNA, and selenocysteyl-tRNA. The mechanisms involved in SECIS recognition by the eukaryotic translational machinery remain unknown. We report the detection of RNA-binding proteins that specifically recognize the SECIS from human cellular glutathione peroxidase (GPX1) transcripts. RNA gel shift assays showed three retarded bands after incubation with COS-1 whole cell lysate or S-100 cytosol fraction or with extracts from hepatoma cell lines HepG2 and Hep3B. The specificity of the binding was demonstrated by competition by cold unlabeled SECIS RNA and by lack of competition by other RNA species with similar stem-loop secondary structures, such as the human immunodeficiency virus (HIV) transactivation-response region of HIV mRNA element, and mutated SECIS constructs. UV cross-linking and SDS-polyacrylamide gel electrophoresis revealed at least two proteins, with estimated molecular masses of 55,000 and 65,000 Da, that bind to the SECIS. Examination of a series of insertion and deletion SECIS mutants indicated recognition of the SECIS primarily through the basal stem region, although the upper stem, loop, and two of three short conserved sequences also appear to contribute to the affinity of the binding.

A wide variety of sequence elements have recently been identified in the 5′- and 3′-untranslated regions (UTR)1 of eukaryotic mRNAs. The resultant RNA secondary structures and the RNA-binding proteins that recognize them play important roles in the regulation of mRNA stability and translation (1-6).

One such sequence is the selenocysteine insertion sequence (SECIS; also termed the selenium translation element) in the 3′-UTR of eukaryotic gene transcripts encoding selenoproteins (7-9). Members of this unique group of proteins contain one or more selenocysteine residues, often at their active sites, and several catalyze important oxidation/reduction reactions (10). The gene transcripts for these selenoproteins encode the atypical amino acid by a UGA codon, which normally functions as a termination signal (11, 12). In eukaryotes, incorporation of selenocysteine into the polypeptide chain at this UGA codon requires the presence of at least one SECIS, located in the 3′-UTR as much as 1200 nucleotides downstream (8, 11).

Comparison of the sequences of SECIS in rat and human iodothyronine deiodinase, cellular glutathione peroxidase, and selenoprotein P mRNAs (7, 8, 13) has revealed only three very short conserved sequences, but all share a common computer-predicted secondary structure featuring a long stem, several bulges, and an apical loop with three short conserved sequences. These characteristic features are illustrated in the diagram of the SECIS from the human cellular glutathione peroxidase (GPX1) gene transcript in Fig. 1A.

We have shown that this SECIS is necessary for translation of GPX1 mRNA as well as sufficient to direct the translation of an opal (UGA) mutation as selenocysteine in heterologous proteins (9).‡ We also demonstrated that the integrity of the overall stem-loop secondary structure and the short conserved sequences are essential for the element’s function (14).

The mechanisms involved in SECIS recognition by the eukaryotic translational machinery remain unknown. In prokaryotes, selenoenzyme gene transcripts, such as that encoding Escherichia coli formate dehydrogenase, contain a critical sequence with a computer-predicted stem-loop secondary structure immediately downstream from the UGA codon (15). A translation elongation factor, SELB, has been demonstrated to form, in vitro, a quaternary complex with this stem-loop structure and selenocystyl-tRNA (16, 17). The RNA-protein complex is postulated to interact in vivo with the ribosome, resulting in the decoding of the coding region UGA as selenocysteine.

On the basis of these studies in prokaryotes, much effort has been focused on the search for a putative eukaryotic SELB-like elongation factor and other potential proteins involved in the combined recognition of the UGA in the coding region and the SECIS in the 3′-UTR to allow selenocysteine insertion rather than translational termination at the “dual purpose” codon.

One candidate SELB equivalent is a 48-kDa polypeptide recognized by autoantibodies present in a subgroup of patients with a severe form of autoimmune chronic active hepatitis (18). This protein is co-immunoprecipitated with the selenocystyl-tRNA and does not exhibit tRNA synthetase activity. A second candidate is a 50-kDa protein from bovine liver extract, which protects selenocystyl-tRNA against alkaline hydrolysis (19).

*This work was supported by a child health research grant from the Charles H. Hood Foundation, a resident research grant from the American Academy of Pediatrics, and by United States Public Health Service Grant DK41625. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡To whom correspondence should be addressed: Dept. of Pediatrics, University of Massachusetts Medical School, 55 Lake Ave. N., Worcester, MA 01655. Tel.: 508-856-4225; Fax: 508-856-5500; E-mail: PNewburger@BANGATE1.UMMED.EDU.

1 The abbreviations used are: UTR, untranslated region of mRNA; HIV, human immunodeficiency virus; SECIS, selenocysteine insertion sequence(s); TAR, trans-activation-response region of HIV mRNA; PAGE, polyacrylamide gel electrophoresis; nt, nucleotide(s).

2 Leonard, J. L., Leonard, D. M., Shen, Q., Farwell, A. P., and Newburger, P. E., (1995) J. Clin. Biochem., in press.
However, neither of these protein factors is reported to be able to bind to the SECIS. The present studies report the detection, in COS-1 cell extract, of RNA-binding proteins that specifically recognize the SECIS of GPX1 mRNA.

**EXPERIMENTAL PROCEDURES**

_Plasmid Construction—_Wild type and deletion mutant GPX1 SECIS sequences, described elsewhere (14), were inserted into pHBluescript KS (Stratagene), lacking the polylinker sequence between BamHI and Sad, at the Aval site, with proper orientation for in vitro synthesis of GPX1 SECIS RNAs from the T7 promoter. These constructs were linearized by EcoO1098 digestion or XhoI digestion for the deletion mutant RNA probe.

Construction of another mutant GPX1 SECIS with an extra unpaired U was obtained by inserting into pCMV4 a CiaI and XbaI digestion of a polymerase chain reaction product, which was amplified from a GPX1 cDNA subclone in pHBluescript KS using the following primers: 5'-ATATATACGATATATAACCCCTCCTCAAAGGGGATCTCGGGGGGTGTCATC-3' and 5'-AAATTATCTAGATTTTCGTTTGGTGGATATTCTTCTGTT-3'. The plasmid was linearized by XbaI digestion, and the synthesis of mutant GPX1 SECIS RNA was performed from the T3 promoter built into the first oligonucleotide.

Construction of a truncated GPX1 SECIS was performed by annealing the following four oligonucleotides and ligating the product into the CiaI restriction sites in pCMV4: 5'-MGATAATTAAACCCCTCCTAAAGGGGATCTCGGGGGGTGTCATC-3', 5'-GGGGGGTTGGTTCAAAATACCCCTCCTAGGAAATATTTTCTCGAGGTGGTATTTTGAAAA-3', and 5'-CCCCCCCCGAGATCTCCTGTTGTTGATTGTGTTGTTTAAA-3', and 5'-CCCCCGGAGATCTCCTGTTGTTGATTGTGTTGTTTAAA-3'. The construct was linearized by XbaI and the mutant GPX1 SECIS RNA synthesized from the T3 promoter in the first oligonucleotide.

Plasmids pSP64 containing the human immunodeficiency virus (HIV) trans-activation-response (TAR) element and exhp*CAT containing a TAR element and exhp*CAT were obtained as a kind gift from Dr. Joel Richter. Each RNA probe was synthesized from the SP6 or T7 promoter. These constructs were linearized by HindIII or BamHI, respectively.

In Vitro RNA Synthesis—32P-Labeled RNA probes were synthesized in 50 ml of [γ-32P]CTP (400 Ci/ml) and 1 ml of cold CTP using the Stratagene RNA transcription kit or the Boehringer Mannheim SP6/T7 transcription kit; unlabeled RNA competitors were synthesized using the Ambion Megascript kit. All RNA probes and competitors were purified by electrophoresis in 8 M urea-polyacrylamide gels, overnight elution at 4°C in 10 mM Tris-HCl, pH 7.4, 0.3 M NaCl, and 1 ml EDTA, phenol extraction, and ethanol precipitation.

_COS-1 Cell Extract Preparation—_The COS-1 cell extract on the migration, in non-denaturing PAGE, of a GGPX1 SECIS probe, with predicted secondary structures generated by the Genetics Computer Group program FOLDRNA. A, wild type GPX1 SECIS, with locations of deletion and insertion mutations indicated. Nucleotide numbering starts at the first base of the SECIS. CS, conserved sequence. B, GPX1 SECIS truncation mutant containing the basal stem sequences. Nucleotide numbers correspond to those in A, C, synthetic hairpin RNA, D, HIV TAR element.

**RESULTS**

In order to identify RNA-binding protein(s) that recognize the SECIS, we examined the effect of incubation with COS-1 cell extract on the migration, in non-denaturing PAGE, of a 32P-labeled RNA probe containing the 87-nt GPX1 SECIS (Fig. 1A). The probe contained minor changes caused by the presence of the polylinker sequence of the template vector: two extra base pairs U-G and C-G plus a C-G to U-A base pairing change at the bottom of the basal stem. However, the Genetics Computer Group program FOLDRNA predicts that this probe forms a correct stem-loop secondary structure analogous to the native GPX1 SECIS.

The RNA gel shift assay, illustrated in Fig. 2, showed three retarded bands on the native polyacrylamide gel after incubation with either a COS-1 whole cell lysate (lanes 2 and 2A) or the more defined S-100 cytosol fraction (not shown). The upper, minor bands of unshifted transcript, visible in lanes 1 and 3, probably represent aggregates of the probe that persisted or reformed after gel purification. They contained only a small proportion of probe radioactivity, were not consistently detected, and their presence or absence did not affect the gel shift pattern. The finding of the gel shift indicates the formation of at least three complexes between the SECIS probe and cellular factor(s) present in the COS-1 extract. Similar experiments with extracts from hepatoma cell lines HepG2 and Hep3B revealed identical patterns (data not shown), indicating that the binding activity is not unique to COS-1 cells.

Several lines of evidence indicated that formation of these complexes was specific. First, two irrelevant 32P-labeled RNA probes showed no similar binding in the gel shift assay. A probe containing an RNA sequence predicted to form a hairpin of 26 base pairs and a 4-nt loop (Fig. 1C) produced only a single band.
with much weaker intensity (Fig. 2, lanes 3 and 4). This band appeared at a location close to the middle retarded band formed with the wild type probe, which could represent either coincidental migration of a different RNA-protein complex or nonspecific binding to one of the SECIS recognition factors.

The HIV TAR element (6) contains a stem-loop secondary structure roughly similar to that of GPX1 SECIS (Fig. 1D). Gel shift analysis of this probe revealed no retarded bands (data not shown). Cold RNA competition (Fig. 3) showed that the unlabeled GPX1 SECIS competitor (A), but not the unlabeled HIV TAR (B), was capable of efficiently competing with the \(^{32}\)P-labeled probe for binding with the cellular factor(s).

Indirect evidence against nonspecific RNA binding activity in the gel shift was derived from variations in the conditions of the assay. There were no significant changes in the positions or intensity of the retarded bands despite changes in the amount of E. coli tRNA from 5 to 10 \(\mu\)g, the concentration of heparin from 2.5 to 15 \(\mu\)g/\(\mu\)l, or of MgCl\(_2\) from 2.5 to 5 mM (data not shown). tRNA and heparin are used in the assay to suppress nonspecific binding between the RNAs and RNA-binding proteins, and high magnesium ion concentrations promote nonspecific binding. Altogether, these data indicate that the gel shift represents specific recognition of the GPX1 SECIS by COS-1 cellular factor(s).

Further evidence for the specificity of recognition as well as delineation of binding target region(s) within the GPX1 SECIS came from the results of gel shift assays using mutant GPX1 SECIS probes. In one mutant probe, the last six nucleotides of the GPX1 SECIS were removed, as shown in Fig. 1A (“6-nt deletion”). This mutation was predicted by the FOLDRNA program to form the same secondary structure as the wild type, except for the remaining six unpaired nucleotides at the base of the stem. When this probe was incubated with COS-1 cellular extract, extremely faint retarded bands could be seen compared with the very strong bands formed by the wild type GPX1 SECIS probe under the same assay conditions (Fig. 4). This result suggested that the basal stem is a critical region for the binding of COS-1 cellular factor(s) to the SECIS. To further test this hypothesis, another mutant GPX1 SECIS probe was made, which was also predicted to form the same secondary structure as the wild type probe but with an extra, unpaired U after the 4th base pair of the basal stem region (Fig. 1A, “U insertion”). This probe failed to produce any retarded bands in the gel shift assay (data not shown). These findings indicate that the basal stem of the SECIS and its perfect base-paired region in particular are major determinants for specific binding.

The negative result of gel shift assay with the hairpin RNA probe suggests that the COS-1 cellular factor(s) does not simply recognize double-stranded RNA. To further examine the specificity of binding, gel shift analysis was applied to a truncated GPX1 SECIS probe containing most of the basal stem, including all the perfect base pairing region (Fig. 1B). Incubation of this probe with COS-1 extract did not produce any shifted bands (data not shown), indicating that other region(s) of the SECIS are also required for the recognition of the element by the binding factors.

To investigate more subtle differences in binding affinity,
additional cold GPX1 SECIS constructs with deletion mutations were used to compete with the radiolabeled wild type probe in the gel shift assay. The deletions included each of the three conserved short nucleotide sequences, the non-conserved apical loop sequence, and the upper stem of the GPX1 SECIS (indicated in Fig. 1A). FOLDRNA analysis predicted local secondary structure perturbations surrounding each deletion area but not affecting the basal stem or other non-contiguous secondary structure elements. These mutant probes were able to form the same retarded bands on the gel shift assay as the wild type SECIS but with somewhat less intensity (data not shown). In the competition assay (Fig. 3) cold wild type SECIS RNA at 100-fold molar excess dramatically decreased the intensity of the retarded bands of labeled wild type probe; at 500-fold excess, the bands almost completely disappeared. As noted above, cold HIV TAR RNA failed to compete even at 5000-fold molar excess.

Of the GPX1 SECIS mutant competitors, deletions of the conserved sequences AUGA and UG (CS 2 and 3 on Fig. 1A), the non-conserved apical loop sequence, and the upper stem sequence showed less efficient competition than the cold wild type probe; at molar ratios of 100- or even 500-fold excess, there was little effect on the intensity of the retarded bands (Fig. 3, D-G). The SECIS mutant lacking the conserved sequence AAA (CS 1 on Fig. 1A) competed with the labeled wild type probe as well as the wild type competitor. The data indicate that all of the tested deletions except that of the AAA conserved sequence caused a decrease in the binding of the SECIS to the cellular factor(s). The loss of binding affinity suggests that these other elements, distant from the basal stem in primary and secondary structure analyses, also contribute, although not as decisively, to the recognition of the GPX1 SECIS by the COS-1 cellular factor(s).

The chemical nature of the binding factors was investigated using a combination of UV cross-linking and SDS-PAGE. Labeled wild type and six-nucleotide deletion SECIS probes were incubated with COS-1 cell extract, as for the gel shift assay, and then subjected to UV irradiation in order to cross-link the cellular factors with the probes. After RNase T1 treatment, the mixture was analyzed by SDS-PAGE to visualize the cross-linked product(s). The deletion mutant served as a negative control for any nonspecific UV cross-linking. As shown in Fig. 5, two separate 32P autoradiographic bands, with estimated molecular masses of 55,000 and 65,000 Da, appeared bound to the wild type SECIS (lane 2). No UV cross-linking product was found when the COS-1 extract was pretreated with protease K (lane 3), when a 2000-fold molar excess of cold SECIS competitor was added to the incubation mixture prior to UV cross-linking (lane 4), or when the deletion mutant SECIS probe was used (lane 5). These findings suggest that the SECIS binding factors are cellular proteins and further confirm the specificity of SECIS recognition. Due to the broad appearance of the more slowly migrating band on the SDS-PAGE, it is possible that the band actually represents more than one protein species.

**DISCUSSION**

In eukaryotes, the synthesis of selenoproteins depends upon the alternate reading of the codon UGA as selenocysteine rather than as a termination signal (12). This process requires the recognition of a selenium translation element (SECIS) in the 3′-UTR of the selenoprotein gene transcript (7–9). We have used RNA gel shift assays to detect cytoplasmic factors that specifically bind the SECIS of mRNA from the human cellular glutathione peroxidase gene, GPX1. UV cross-linking and SDS-PAGE indicate that the three complexes observed on non-denaturing gels represent the binding of at least two proteins with estimated upper limits of molecular mass of 65,000 and 55,000 Da.

A rough estimate of the abundance of these SECIS-binding proteins in the COS-1 cell extract, based on the binding of 1 fmol of SECIS by 5 μg of extract protein, suggests that they account for approximately 10⁻⁵–10⁻⁶ of total protein in the COS-1 extract. The detection of similar gel shift patterns in hepatoma cell lines Hep B and Hep G2, which are rich in the major intracellular selenoenzyme glutathione peroxidase, implies that the proteins are widely distributed components of the selenoprotein translation mechanism.

The specificity of the binding activity for the SECIS is based on the following observations. First, unlabeled SECIS RNA is able to compete efficiently with the radiolabeled SECIS probe for recognition by the binding proteins. In addition, we have found that other RNA competitors with a similar secondary structure, such as HIV TAR and hairpin RNA, as well as a truncated SECIS incorporating most of the basal stem fail to form detectable complexes in the gel shift assay.

Examination of a series of insertion and deletion SECIS mutants for direct gel shift activity or for competition with wild type SECIS in the assay indicates that the binding proteins recognize the SECIS stem-loop primarily through the basal stem region, in particular its perfect base pairing of the loop and two of three short conserved sequences (AUGA and UG) (Fig. 1). The loss of binding affinity suggests that these other elements, distant from the basal stem in primary and secondary structure analyses, also contribute, although not as decisively, to the recognition of the GPX1 SECIS by the COS-1 cellular factor(s).

The importance of the base-paired basal stem suggests that the binding proteins might possess one or more double-stranded RNA binding motifs (5) such as those described in mammalian interferon-induced protein kinase (22), staufen protein of Drosophila oocytes (23), p25 of Vaccinia virus (24), and TAR RNA-binding protein (6, 25). All of these proteins are involved in post-transcriptional gene regulation, although with diverse cellular functions. They recognize double-stranded RNA binding targets within more complex, specific secondary structures.

At the present time, there is no direct evidence to demonstrate that SECIS-binding proteins are involved in the in vivo translation of GPX1 or other selenoprotein gene transcripts. However, results of previous experiments support the inference that the recognition of the GPX1 SECIS by these binding proteins is functionally important (9, 14). In structure-function studies of the effects of SECIS mutations on glutathione peroxidase translation, the mutations that interfered with GPX1 expression in COS-1 cells were those that disrupted the stem-loop secondary structure, including several that interfered with binding protein affinity in the current study (i.e. changes in the basal stem, loop, and conserved sequences 2 and 3). The loss of...
SECIS function in those experiments probably reflects, at least in part, the consequences of diminished affinity of the binding proteins for the GPX1 SECIS.

However, the structure-function studies also showed major functional effects of changes in the conserved sequences, including substitutions in the loop AAA sequence, that have no effect on secondary structure. Thus, not surprisingly, events other than the binding activity reported in the present study must also play a role in SECIS function. One possible model would be for one set of proteins to bind to the SECIS by recognition of its basal stem region, with the formation of this initial complex then facilitating the addition of a eukaryotic SELB-like elongation factor and other factor(s) that recognize both selenocysteyl-tRNA and the other conserved determinants of the SECIS, such as the loop AAA sequence. The resultant mature complex would then interact with the ribosome for the final decoding. Alternatively, the SECIS-binding proteins could also recognize selenocysteyl-tRNA, and the other determinants on the SECIS could play a role in the process only after conformational changes in the binding proteins or in recognition by other components of the translational apparatus.

Both models predict the participation of factor(s) other than the SECIS-binding proteins in the formation of a mature, functioning multiprotein complex. The elucidation of the functional role of the SECIS-binding proteins and the identification of their specific interactions with RNA and other proteins will depend upon their eventual molecular cloning and expression.

Acknowledgments—We thank Dr. Joel Richter for advice and for provision of hairpin and HIV TAR element probes and Carolyn Padden for technical assistance.

REFERENCES

1. Shaw, G., and Kamen, R. (1986) Cell 46, 659–667
2. Kühn, L. (1991) Br. J. Haematol. 79, 1–5
3. Jaffrey, S. R., Haile, D. J., Klausner, R. D., and Harford, J. B. (1993) Nucleic Acids Res. 21, 4627–4631
4. Mattaj, I. W. (1993) Cell 73, 837–840
5. Burd, C. G., and Dreyfuss, G. (1994) Science 265, 615–621
6. Gait, M. J., and Karp, J. (1993) Trends Biochem. Sci. 18, 255–259
7. Berry, M. J., Banu, L., Chen, Y. Y., Mandel, S. J., Kieffer, J. D., Harney, J. W., and Larsen, P. R. (1991) Nature 353, 273–276
8. Berry, M. J., Banu, L., Harney, J. W., and Larsen, P. R. (1993) EMBO J. 12, 3315–3322
9. Shen, Q., Chu, F.-F., and Newburger, P. E. (1993) J. Biol. Chem. 268, 11463–11469
10. Burk, R. F., and Hill, K. E. (1993) Annu. Rev. Nutr. 13, 65–81
11. Berry, M. J., Banu, L., and Larsen, P. R. (1991) Nature 349, 438–440
12. Böck, A., Forchhammer, K., Heider, J., and Baron, C. (1991) Trends Biochem. Sci. 16, 463–467
13. Berry, M. J., and Larsen, P. R. (1993) Biochem. Soc. Trans. 21, 827–832
14. Shen, Q., Leonard, J. L., and Newburger, P. E. (1995) RNA 1, 519–525
15. Zinoni, F., Heider, J., and Böck, A. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 4660–4664
16. Baron, C., Heider, J., and Böck, A. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 4181–4185
17. Ringquist, S., Schneider, D., Gibson, T., Baron, C., Böck, A., and Gold, L. (1994) Genes & Dev. 8, 376–385
18. Gelpli, S., Sontheimer, E. J., and Rodriguez-Sanchez, J. L. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 9739–9743
19. Yamada, K., Mizutani, T., Ejiri, S., and Totsuka, T. (1994) FEBS Lett. 347, 137–142
20. Dignam, J. D., Lebowitz, R. M., and Roeder, R. G. (1983) Nucleic Acids Res. 11, 1475–1489
21. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) J. Biol. Chem. 193, 265–275
22. Feng, G. S., Chong, K., Kumar, A., and Williams, B. R. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 5447–5451
23. Ferrandon, D., Elphick, L., Nusslein-Volhard, C., and St. Johnston, D. (1994) Cell 79, 1221–1232
24. Watson, J. C., Chang, H. W., and Jacobs, B. L. (1991) Virology 185, 206–216
25. Gatignol, A., Budker-White, A., Berkhout, B., and Jeng, K. T. (1991) Science 251, 1597–1600
RNA-binding Proteins That Specifically Recognize the Selenocysteine Insertion Sequence of Human Cellular Glutathione Peroxidase mRNA
Qichang Shen, Patricia A. McQuilkin and Peter E. Newburger

J. Biol. Chem. 1995, 270:30448-30452.
doi: 10.1074/jbc.270.51.30448

Access the most updated version of this article at http://www.jbc.org/content/270/51/30448

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 25 references, 11 of which can be accessed free at http://www.jbc.org/content/270/51/30448.full.html#ref-list-1