RNA (1998), *4*:65–73. Cambridge University Press. Printed in the USA.

Functionality of mutations at conserved nucleotides in eukaryotic SECIS elements is determined by the identity of a single nonconserved nucleotide

GLOVER W. MARTIN III,[1] JOHN W. HARNEY,[2] and MARLA J. BERRY[1,2]

[1]Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, Massachusetts 02115, USA
[2]Thyroid Division, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA

ABSTRACT

In eukaryotes, the specific cotranslational insertion of selenocysteine at UGA codons requires the presence of a secondary structural motif in the 3′ untranslated region of the selenoprotein mRNA. This selenocysteine insertion sequence (SECIS) element is predicted to form a hairpin and contains three regions of sequence invariance that are thought to interact with a specific protein or proteins. Specificity of RNA-binding protein recognition of cognate RNAs is usually characterized by the ability of the protein to recognize and distinguish between a consensus binding site and sequences containing mutations to highly conserved positions in the consensus sequence. Using a functional assay for the ability of wild-type and mutant SECIS elements to direct cotranslational selenocysteine incorporation, we have investigated the relative contributions of individual invariant nucleotides to SECIS element function. We report the novel finding that, for this consensus RNA motif, mutations at the invariant nucleotides are tolerated to different degrees in different elements, depending on the identity of a single nonconserved nucleotide. Further, we demonstrate that the sequences adjacent to the minimal element, although not required for function, can affect function through their propensity to base pair. These findings shed light on the specific structure these conserved sequences may form within the element. This information is crucial to the design of strategies for the identification of SECIS-binding proteins, and hence the elucidation of the mechanism of selenocysteine incorporation in eukaryotes.

Keywords: recoding; selenocysteine; selenoproteins; translation; 3′ untranslated region

INTRODUCTION

In the universal genetic code, UGA signals the termination of translation of mRNA into protein. Given the convergence of several additional circumstances, the UGA codon may also signal the cotranslational insertion of selenocysteine. In prokaryotes, this alternative reading of the genetic code has been found to require four *trans*-acting gene products (Leinfelder et al., 1988). Three of these factors consist of a selenocysteine-specific tRNA ($tRNA^{ser[sec]}$), and two enzymes required for the synthesis of the amino acid. The fourth prokaryotic *trans*-acting factor, the product of the *Escherichia coli* selB gene, is a selenocysteyl tRNA-specific homologue of the translation elongation factor EF-Tu, and has the dual functions of binding the bacterial selenocysteine insertion sequence (SECIS) and bringing the charged $tRNA^{ser[sec]}$ to the ribosome at the appropriate time (Forchhammer et al., 1989, 1990; Ringquist et al., 1994).

Another necessity for selenocysteine incorporation at UGA codons is the presence of an mRNA secondary structural motif. In eubacteria, this structural element is predicted to form a stem-loop or hairpin, and lies in the peptide-coding region of the mRNA, within one codon of the selenocysteine UGA (Zinoni et al., 1990; Heider et al., 1992). Although structurally similar to their *E. coli* counterparts in that they are predicted to form stem-loops, the critical features of eukaryotic SECIS elements that are essential to their function are distinct from those of the prokaryotic elements. Specifically, there are three regions of high conservation, one of which consists of three A residues on the 5′ side of either the terminal loop or an internal bulge. The other two regions, AUGA and GA, lie on 5′ and 3′ sides, respectively, of the SECIS element stem, 10–12 base pairs below the conserved A residues. Initial mu-

Reprint requests to: Marla J. Berry, Thyroid Division, Harvard Institutes of Medicine, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA; e-mail: berry@rascal.med.harvard.edu.

tational analysis identified these nucleotides as critical to SECIS function (Berry et al., 1993), however, predictions rendered by computer fold programs depicted these regions as predominantly single-stranded, and ascribed no particular structural significance to these nucleotides and their relative positions (Berry et al., 1993; Hill et al., 1993; Shen et al., 1993; Kollmus et al., 1996). By manipulation of the computer-generated eukaryotic SECIS element structure, we derived a model in which the two conserved regions in the stem base pair with each other to form a novel structural motif (Fig. 1, core region on left; Low & Berry, 1996). This model makes two important predictions. First, a non-Watson–Crick base pair will form between the invariant G residues. Secondly, the nucleotide 3′ of the conserved AUGA sequence, a U residue in the rat D1 element, will be bulged out of the helix. We refer to this region as the SECIS element core, specifically consisting of AUGAN on the 5′ and NGAN on the 3′ sides of the stem. This proposed structure seeks to integrate mutagenesis data obtained previously with secondary structure predictions to form a coherent model of the eukaryotic SECIS element.

Concurrent with the development of the aforementioned model, Walczak et al. (1996) independently proposed a different structure on the basis of extensive chemical and enzymatic probing of SECIS element RNAs (Fig. 1, core region model on right). This model aligns the core nucleotides such that a tetra-purine GA-AG structure will form. A critical difference between these two structures, therefore, is the relationship of AUGAN to the NGAN sequence on the opposite side of the SECIS element; the register of base pairs below the fifth nucleotide differs in the two models (see Fig. 1 for numbering). Both structures take into account precedents established through studies of RNA-binding proteins and their respective RNA ligands. Unlike DNA, RNA double helices are predominantly of the A-form, a structure with a particularly deep and narrow major groove which, in its native conformation, may not allow direct recognition of bases by amino acid side chains (Rould et al., 1989). Therefore, local distortions in the helix, such as internal loops, bulges, or noncanonical base pairs, are thought to facilitate access to the distinguishing chemical groups of the bases by amino acid side chains. This has been shown experimentally for the HIV tat/TAR and rev/RRE interactions (Calnan et al., 1991; Weeks & Crothers, 1991; Battiste et al., 1996).

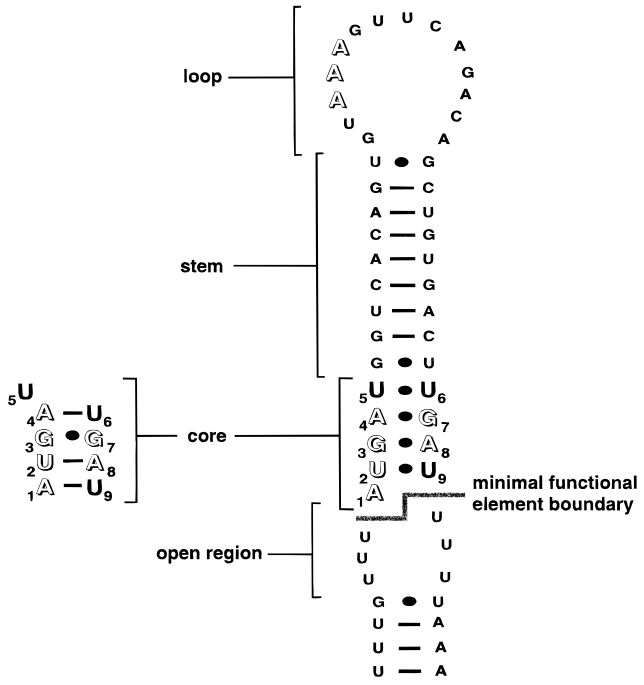

FIGURE 1. Anatomy of a typical eukaryotic SECIS element, that of the rat type 1 deiodinase, and two models of the core region. Invariant nucleotides are outlined and shadowed. Core nucleotides are numbered. Note the difference in register of base pairing within the core region. The base of the minimal functional element coincides with the base of the core region. Predicted non-Watson–Crick base pairs are denoted with a "ball" rather than a "bar."

Here, we used mutagenesis to determine the relative contributions of core nucleotides to the function of two SECIS elements. Our results showed that differential tolerances for particular mutations at certain positions varied from element to element. We then investigated the possibility that the identity of the non-conserved core position 5 residue was contributing to this effect. The data indicate that the residue at position 5 interacts directly with those at positions 4 and 6, as would be predicted by a model in which this nucleotide is an integral part of the helix and not "bulged" out. We also show that the nucleotides below the core (the "open region" shown in Fig. 1) are not required for function of the SECIS element, but can affect activity adversely if they have a high propensity to base pair. Finally, our earlier studies demonstrated that the minimal functional rat D1 SECIS element corresponds to the base of the core region (Martin et al., 1996; Fig. 1). We present data that indicate that this minimal SECIS element definition is generally applicable.

RESULTS

Individual nucleotides in the core region vary in their contributions to SECIS function

Shen et al. (1995a) reported the results of mutagenesis of the core nucleotides of the human cGPx SECIS element (Fig. 2B). We found several inconsistencies between the effects of certain mutations in the rat D1 SECIS element and those same mutations in the human cGPx element. Comparison of Figure 2A and B shows that changes of any kind to nucleotide A4 of the rat D1 element are tolerated very poorly. However, although changing A4 to C also abolishes activity of

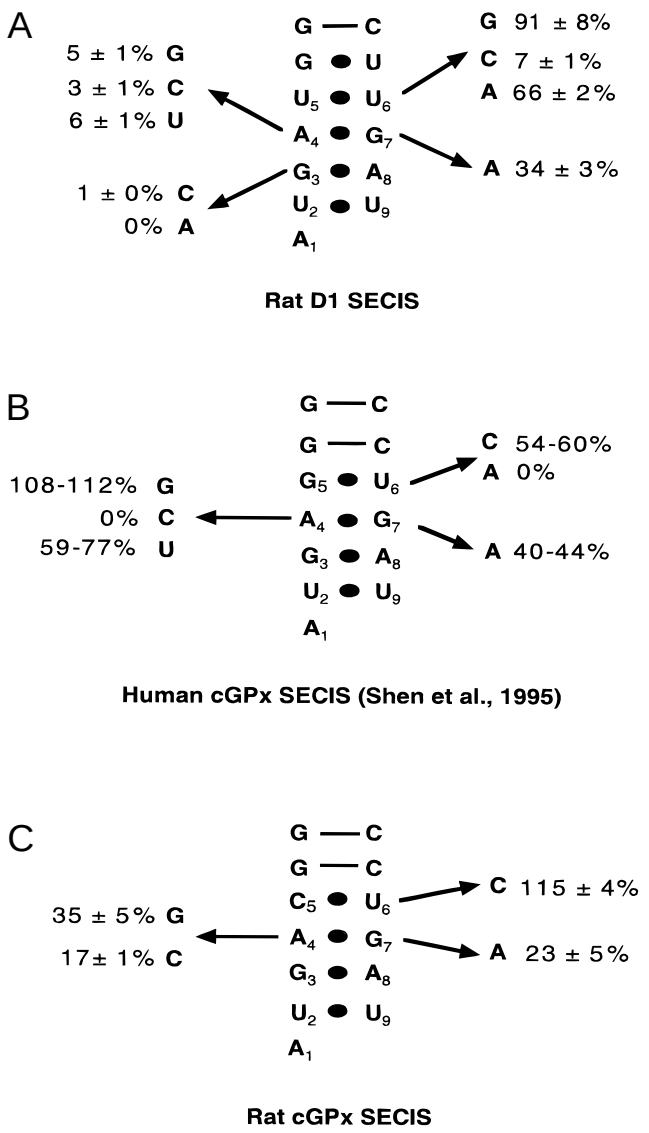

FIGURE 2. Mutagenesis of the highly conserved "core" region of D1 and cGPx SECIS elements. To facilitate interpretation, the data are depicted using the continuous helix model, which we conclude is most likely to be correct. **A:** Mutagenesis of the core of the rat D1 SECIS element. **B:** Mutants generated by Shen et al. (1995a) in the human cGPx SECIS element. Activities are indicated as a range of percentages of the wild-type human cGPx SECIS element. **C:** Mutagenesis of the core of the rat cGPx SECIS element. Mean activities, indicated as percentages of the activity of the corresponding wild-type SECIS element, are shown with standard errors. Subscripts indicate core nucleotide positions as referenced in the text.

the human cGPx element, substitution of U or G at this position has a modest effect, or no effect, respectively, on the activity of the element. Thus, mutations at this position, which is absolutely invariant in all elements identified to date, have widely differing effects in different elements. We next compared the effects of changes at position 6 in the two elements. Changing U6 to C results in a greater than 90% loss in activity of the rat D1 SECIS element, but the activity of the analogous human cGPx mutant was only reduced by approximately half. Conversely, mutation of U6 to A abolishes the activity of the human cGPx SECIS element, but reduces the activity of the rat D1 element by only one third. Interestingly, the conservative purine substitution of A for G7 had similar intermediate effects on activity in both elements.

The unexpected differential tolerances for specific mutations in the highly conserved core nucleotides of two SECIS elements prompted us to test the same mutations in a third SECIS element, that of rat cGPx (Fig. 2C). The A4 to G mutation in the rat cGPx SECIS element resulted in an activity intermediate to the analogous change in the previous two elements. Additionally, the A4 to C rat cGPx mutant exhibited significant (~17%) activity. This contrasts with the abolition of function resulting from the same mutation in the other two elements. We also found that changing the rat cGPx U6 to C produced a slight gain in activity compared to the wild-type element. The resultant activity is nearly twice that of the analogous human cGPx mutant and 20 times that of the same mutation of the rat D1 SECIS element, all in relation to their respective wild-type elements. The mutation of G7 to A, which affected the rat D1 and human cGPx SECIS elements to a similar degree, also decreased the activity of the rat cGPx SECIS to a similar extent. From the results of these experiments, it is clear that certain mutations within the highly conserved SECIS core, specifically at positions 4 and 6, have different effects in different elements. We next considered whether these effects might be due to the local context in which the nucleotides are positioned.

Identity of the position 5 nucleotide affects mutations at neighboring positions

Examination of the core regions and adjacent nucleotides of the three SECIS elements shown in Figure 2 reveals that the local environs of the core nucleotides are all highly similar. The nucleotide immediately 3′ of the invariant AUGA (position 5), however, is unique in each of these elements. Analysis of the available eukaryotic SECIS element sequences shows that there is no consensus identity for this nucleotide (Berry et al., 1993; Walczak et al., 1996), suggesting that it may not contribute significantly to the biological activity of the SECIS element. We tested this hypothesis by deleting U5 in the native rat D1 SECIS element or by substituting C or G, and comparing the activities of the resulting mutants to the wild-type element. Figure 3A shows that deletion of the nucleotide at position 5 abolished activity. The G substitution reduced activity by about half compared to wild-type. Strikingly, the C mutation led to an almost 50% increase in activity versus wild-type. These results reveal a significant contribution of this nonconserved nucleotide to function, possibly via interaction with other core nucleotides that might affect the structure of the SECIS element and/or recog-

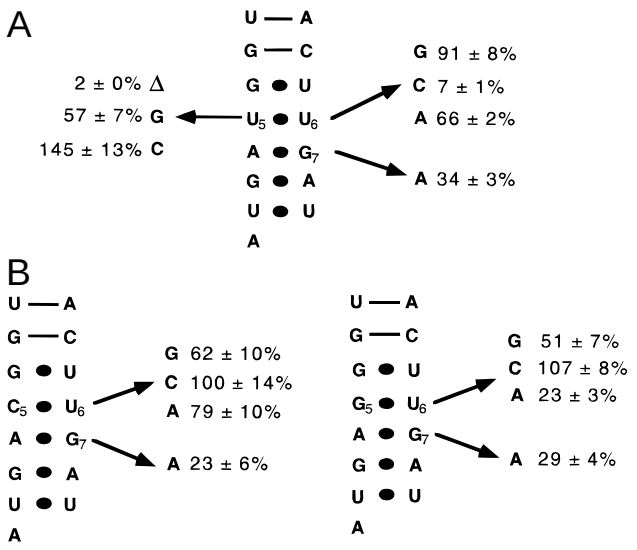

FIGURE 3. Contextual effects of mutations in the core region of the rat D1 SECIS element. **A:** Effects of deleting or altering the identity of N5, the putative bulged nucleotide in the core region of the SECIS element. **B:** Effects on the tolerance of mutations at other positions within the SECIS element core to N5 mutations. The Greek symbol Δ (Delta), signifies a deletion of the SECIS element core position 5 nucleotide. Mean activities, indicated as percentages of the activity of the corresponding wild-type SECIS element, are shown with standard errors. Subscripts indicate core nucleotide positions as referenced in the text.

nition of the element by an as yet unidentified protein factor.

The nucleotide at position 5 is the only one in the core region that varies among the three elements under investigation. This fact, taken together with the effects of this nucleotide on activity shown above, suggests that the differences at this position might account for some of the differential effects of mutations elsewhere in the core region. Our approach to examining this possibility was to make mutations at various positions in the core of the rat D1 element in the context of changes at position 5. We found that the identity of N5 has a significant impact on the effects of mutations at other positions. As shown in Figure 2, the rat D1 element with U at position 5 poorly tolerates substitution of C for U at position 6. In contrast, in the rat D1 elements containing G or C at position 5 (Fig. 3B), the U6 to C substitution had no effect on activity, approximating what was observed for the native rat and human cGPx SECIS elements. In the context of a G at position 5, the U6 to G mutant has an activity of ~50% compared with ~90% for the analogous change to the wild-type D1 element. The tolerance for A6 increases slightly with the U5 to C change and decreases by a factor of two with the U5 to G change. Although not tested directly, we may further speculate from the data in Figure 2 that, in addition to effects on position 6 mutants, position 5 also affects the activity of position 4 mutants. Changing A4 to any other nucleotide affects the activity of each SECIS element to a different degree, with the local context difference in the three elements being limited to the nucleotide at position 5. These data, together with those described above, suggest that the nucleotide at position 5 may interact directly with core nucleotides 4 and 6. One explanation for our observations regarding effects of N5 on its neighbors is that this nucleotide is not bulged out, but is base paired with the nucleotide at position 6 of the SECIS core, and thus stacked into the helix, as has been proposed by Walczak et al. (1996).

Definition of the minimal SECIS element applies to other eukaryotic elements

We reported previously that the 5′ boundary of the rat D1 SECIS element coincided with the first two nucleotides (AU) of the conserved core region (Martin et al., 1996). This contrasts with the findings of Shen et al. (1995a), who reported in a prior study that removal of the "basal stem," which consists of several nucleotides below and including the open region (see Fig. 1), results in a complete loss of activity of the human cGPx SECIS element. That is, sequences "below" the core appeared to be required for function in this element. This would suggest that the minimal constraints we defined for rat D1 do not hold true for the human cGPx element. We sought to determine whether the minimal sequence requirements we had defined for rat D1 might also apply to SECIS elements of other selenoproteins. Oligonucleotide primers were designed to amplify the SECIS elements of the rat and human selenoprotein W (selW; M. Beilstein & P. Whanger, pers. comm.), the rat phospholipid hydroperoxide glutathione peroxidase (phGPx), and rat cGPx mRNAs. The 5′ and 3′ boundaries of the amplified regions are analogous to those of the rat D1 SECIS element, with restriction sites for subcloning into the D1 expression vector introduced on either side of the core region. We find that the minimal rat cGPx is able to direct specific selenocysteine incorporation into D1 at a level comparable to that of the rat D1 element (Fig. 4), as do the minimal phGPx and selW elements (not shown). These data indicate that the native sequences below the core are not required for SECIS function, supporting the generality of the minimal element.

Of the eukaryotic SECIS elements that have been characterized to date, all are found in vertebrate genes. So far, one selenoprotein sequence from a nonvertebrate eukaryotic organism has been reported, but requirements for selenocysteine incorporation in this organism have not been determined. Examination of the 3′ untranslated region of the GPx sequence from *Schistosoma mansoni*, a parasitic trematode (Williams et al., 1991), reveals the presence of a putative SECIS element, containing all the conserved features of SECIS elements described previously. We amplified and tested

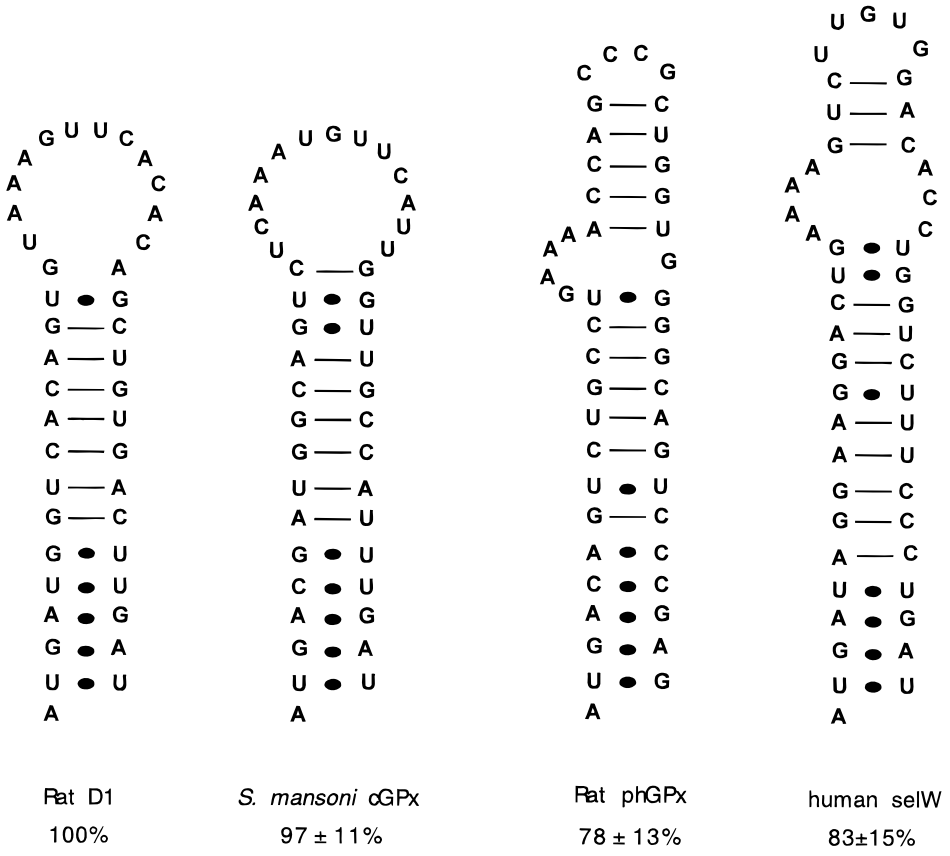

FIGURE 4. Predicted secondary structures and relative activities of several previously reported SECIS elements. Activities are given as percentages of the activity of the wild-type rat D1 minimal SECIS element. Numbers shown are the means of at least three independent experiments and the corresponding standard errors. Oligonucleotide primers used to generate each SECIS element are shown in Table 1.

this region for its ability to drive the incorporation of selenocysteine into an heterologous open reading frame and found its activity to be on par with that of the mammalian D1 and cGPx SECIS elements. This is the first demonstration that a nonvertebrate SECIS element will function in cells of vertebrate origin. Thus, the minimum sequence requirements we defined previously for a functional rat D1 SECIS element also hold true for SECIS elements of various eukaryotic seleno proteins in organisms as diverse as mammals and trematodes.

Sequences below the core are nonessential but constrained

Current models of eukaryotic SECIS elements predict that the region below the core conserved nucleotides is single-stranded (see Fig. 1). Chemical and enzymatic probing of the rat D1 and cGPx SECIS elements has shown that this is indeed the case, in vitro, for these two elements (Walczak et al., 1996). We investigated the effects of changing the nucleotides below the core region and their Watson–Crick base pairing potential on the activity of the rat D1 SECIS element in vivo. Using oligonucleotide-directed mutagenesis, the three unpaired bases on each side were changed. Replacement of the three U residues on the right with C's had no significant effect, as did the double substitution of AAA/CCC for UUU/UUU. This indicates that the specific nucleotide sequence below the core is not critical for SECIS function. These mutants are predicted to maintain the open structure of the region immediately below the core and thus would not address the issue of base pairing in this region. We therefore investigated the effects of substituting sequences with the potential to form base pairs of varying strengths. We observed that, as the predicted stability of base pairing interactions increases (i.e., G-U < A-U < G-C), the activity of the resulting mutant decreases (Fig. 5; mutants c, d, and e). Strikingly, disruption of the base pair immediately below the core region in mutant e (A-U to A-A or A-C, mutants f and g) resulted in a 10-fold increase in its activity compared with mutant e, underscoring the significance of this region being open. These results indicate that effects in this region are dependent on the strength of the interaction between

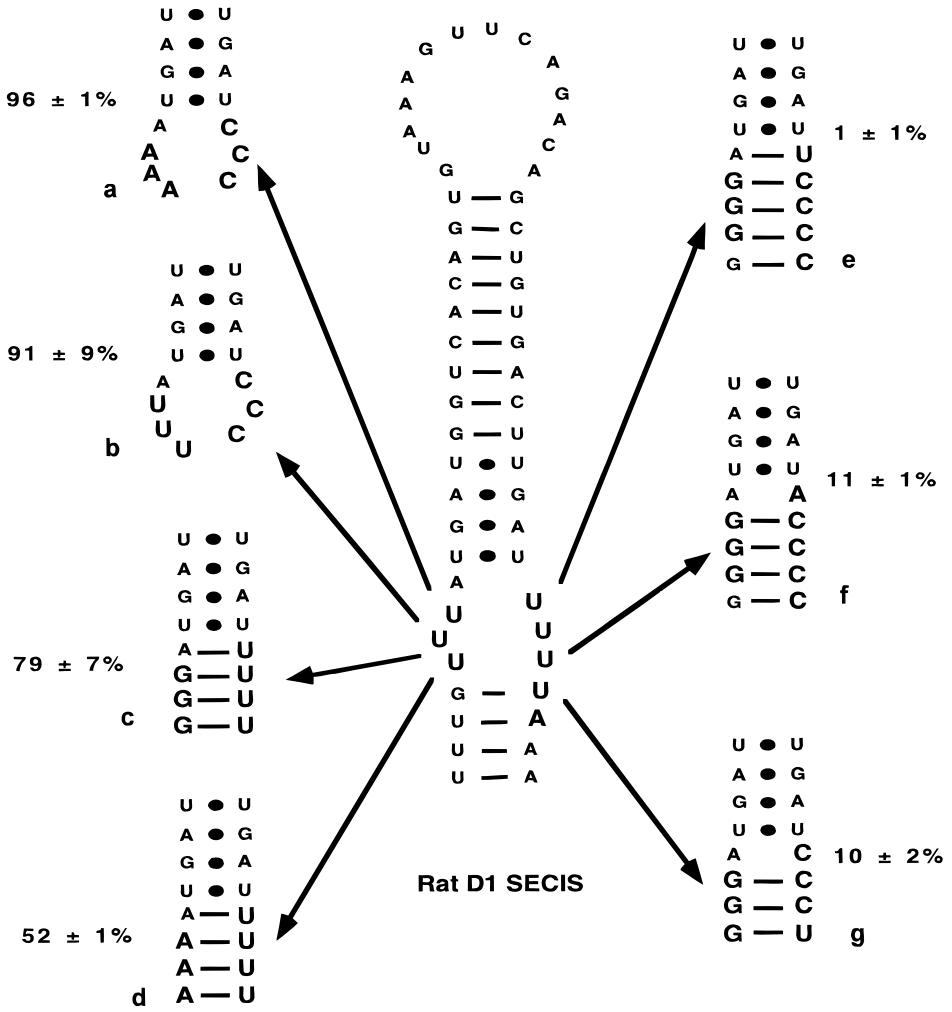

FIGURE 5. Effects of primary and secondary structural changes in the region below the core of the minimal rat D1 SECIS element. Oligonucleotide-directed mutagenesis of the nucleotides immediately below the highly conserved "core" region of the rat D1 SECIS element. For clarity, only the lower portion of the minimal SECIS element, containing the mutagenized region, is shown. Activities are given as percentages of the activity of the wild-type rat D1 minimal SECIS element. Numbers shown are the means of at least three independent experiments and the corresponding standard errors.

the nucleotides below the core region. Further, the data show that, although a specific sequence of nucleotides immediately below the minimal SECIS is not required for function, the identities of the bases at those positions and their potential to base pair with those on the opposite strand can significantly affect the function of the element.

DISCUSSION

The process of specific selenocysteine incorporation requires a translational recoding event, which we know to be dependent upon an mRNA secondary structural element, the SECIS. Presumably, proteins recognize a combination of characteristics of this element, which almost certainly includes the conserved core nucleotides. Here, we have taken steps toward understanding in greater detail the properties of a SECIS element that make it functional, and hence may be critical for its recognition by a eukaryotic SELB homologue or other putative SECIS-binding proteins.

The first goal of these studies was to ascertain the effects of particular mutations at highly conserved positions within the SECIS core. Our initial mutagenesis of the rat D1 SECIS element yielded results that were at variance with those reported by Shen et al. (1995b) for identical mutations made to the human cGPx SECIS element, and prompted us to investigate the rat cGPx element also. Considering the invariant nature of the core nucleotides among SECIS elements, we were not surprised by the effects of G3 and A4 mutations on the activity of the D1 SECIS element. However, the finding that G could be substituted for A4 in the human cGPx element without compromising its activity

was unexpected, and was the first hint that the N5 context may be the cause of this differential tolerance. These data were further supported by the intermediate phenotype of the identical A4 mutation in the rat cGPx element.

In searching for a potential cause for the differential tolerances we had observed, we first looked at nucleotide differences near the SECIS core that could explain our data. The G-C base pair immediately above the core region was common to both the cGPx elements, but was replaced by G-U in the rat D1 element. In the two elements that have the G-C base pair in common, the effect of N6 mutations are quite different. We concluded, therefore, that this change could not account for differences in the activities of C6 mutations between the human and rat cGPx elements. There was only one nucleotide in the SECIS core region that was different in each of the three elements discussed, the nonconserved position 5. We therefore tested the hypothesis that nucleotide 5 was the source of the differential tolerance of the three SECIS elements for particular nucleotide substitutions. We found that changing U5 of the rat D1 SECIS element to G or C resulted in effects at position 6 that were highly similar to those observed in the human and rat cGPx SECIS elements, which contain G or C, respectively, at position 5. However, although the trends were generally similar, in some cases the effects were not mirrored. For example, the human GPx G5-C6 combination exhibits intermediate activity, versus full activity for the rat D1 G5-C6 combination. This suggests that, although N5 is a major determinant for these effects, other features of the elements, specifically the identity of the nucleotide 5′ of N6, may also contribute to the differential tolerances.

The high degree of allowable variation in a consensus sequence, which is ostensibly recognized by an RNA-binding protein, was unanticipated. This feature of the SECIS element suggests that many nucleotides and nucleotide pairs have important structural roles and may serve to put bases that are directly recognized by putative SECIS-binding proteins in an accessible orientation. This would especially apply to the nucleotides that lie adjacent to the GA-AG homopurine motif in the SECIS core. We have found that the U5-C6 mutation exhibits very low activity, and the fact that this combination of nucleotides is excluded from the eukaryotic SECIS elements identified so far indicates that nature has selected against this nucleotide pair. In fact, when C6 is present, position 5 is always a C. Our data for the C6 mutant of the rat cGPx element, which naturally contains a C at position 5, strongly corroborates the conclusion that the C-C pair is not simply allowed, but optimal. Our data also suggest that U5-G6 and U5-A6 pairs are tolerated, but these combinations have yet to be found in nature. The data of Shen et al. (1995a) indicate that G4-G5 is as good a combination as the natural A4-G5 found in the human cGPx element, but, intriguingly, SECIS elements deviating from the AUGA consensus have not been identified. We found that, although certain mutations had similar effects in all three elements (e.g., G7 to A), other nucleotide substitutions had a different effect in each of the three elements (A4 to G). These data are suggestive of a common, critical role for A4 in SECIS element activity, however, more detailed studies of the effects of positions 4 and 6 on this mutation are needed to support this conclusion. A clear understanding of why certain nucleotides are allowed and others are not, although they function in our assay system, awaits both the identification of SECIS-binding proteins and high-resolution structure determination of the SECIS core.

A second objective of our experiments was to distinguish between the two prevailing SECIS element models. Both models incorporate RNA structural features such as non-Watson–Crick base pairing and unpaired regions, features known to be critical to the function of several other regulatory RNA molecules. As Figure 1 shows, the primary difference between these two models is the position of nucleotide 5. An unpaired, bulged base at this position could potentially be required simply in order to increase the accessibility of functional groups of the invariant core nucleotides. However, we have presented evidence that the identity of this nucleotide influences the effects of mutations at adjacent nucleotides. Identical mutations in three elements having different N5′s can have different effects. In addition, changing U5 of the D1 SECIS element to G or C affects the tolerance for mutations at position 6. Furthermore, the deletion of U5 renders the rat D1 SECIS element ineffective, strongly suggesting that it is an important component of the RNA structure in the region. Taken together, these results indicate an integral role for N5, implicating a direct interaction between nucleotides 5 and 6. The data suggest that the structural prediction of Walczak et al. (1996), in which N5 is stacked into the helix, presents a more likely scenario of the SECIS core.

Berry et al. (1993) previously delimited the rat D1 SECIS element to a 175-nt region of the 3′ UTR. Our recent deletion analysis of this region defined the 5′ and 3′ boundaries of the minimal SECIS element as the invariant nucleotides AUGA and GA, respectively (Martin et al., 1996). Here we have extended these observations and have shown that the specific nucleotides below the minimal rat D1 SECIS element are not essential to the biological activity of the element. Substitution of the three nucleotides 5′ and 3′ to the minimal D1 SECIS element did not diminish the function of the element, provided the substituted nucleotides did not introduce base pairing potential. However, although the nucleotide sequence does not affect function, it is critical that an open structure in this region be preserved. We found that, as the energy of the base–

base interactions was increased from near zero in the wild-type element, the activity decreased. This finding provides in vivo evidence, which complements in vitro findings regarding the susceptibility of nucleotides below the core to single strand-specific cleavage and modifying reagents (Walczak et al., 1996). We believe this unpaired structure may be needed to make the invariant nucleotides accessible or, as Walczak et al. (1996) have proposed, allow for flexibility of the SECIS in this region, effectively acting as a hinge and contributing to some higher-order structure of the SECIS element.

Finally, we present evidence that supports the universality of the minimal rat D1 SECIS element defined previously. The putative minimal SECIS elements from several selenoprotein mRNAs, representing a range of eukaryotic species, exhibit activities near that of the wild-type rat D1 element. There may be, however, other factors within the 3′UTRs of these mRNAs that may alter their activities, such as stability/instability elements. That other such elements may exist should have no bearing on the definition of the absolute minimal SECIS element required to achieve selenocysteine incorporation.

MATERIALS AND METHODS

Generation of parent constructs and mutagenesis

To generate D10ΔH3, the unique *Hin*d III restriction site in the plasmid pUHD 10-3 was deleted by blunting and religating (Gossen & Bujard, 1992). The 2.1-kilobase *Eco*R I–*Xba* I fragment of the full-length rat D1 cDNA containing its entire coding region and SECIS element was cloned into the corresponding sites in D10ΔH3 to generate G16Δ10DH3 (Berry et al., 1991a). Oligonucleotide primers complementary to cDNA positions 1533–1555 and 1585–1558 of D1 (Table 1) were designed with terminal *Hin*d III and *Not* I sites, respectively. Analogous primers complementary to positions 1014–1036 and 1060–1039 of the rat cGPx and 718–737 and 779–760 of the rat phGPx cDNAs were designed. These primers were then used to amplify the minimal SECIS element from the respective wild-type rat cDNAs. The human SelW SECIS element, encompassing nt 367–442, was generated by the amplification of the SelW6 template with oligonucleotides SelW5 and SelW7 (P. Whanger, pers. comm.; Table 1). Oligonucleotides complementary to positions 560–595 and positions 610–575 (Table 1) of the *S. mansoni* cGpx cDNA were annealed to each other and extended to produce a double-stranded DNA.

PCR products were digested with the appropriate enzymes and subcloned into the corresponding sites of G16-D10ΔH3 (Martin et al., 1996), *Hin*d III at the 5′ end and *Not* I or *Xba* I at the 3′ end. Mutagenesis was performed by amplification from wild-type templates with oligonucleotides that contain the desired nucleotide change(s). PCR products were then subcloned into the *Hin*d III and *Not* I sites of G16-D10ΔH3 as described above. Once subcloned, all amplified regions were sequenced in their entirety.

TABLE 1. Oligonucleotide sequences.

SelW5	5′-ccaagctttagccgcttcatgataggaagg-3′
SelW6	5′-acatcagggaaagaccaggtgtccacaagacttttcagtccttcct atcatgaagcg-3′
SelW7	5′-ccgcggccgcagaacatcagggaaagacca-3′
SMcGpx1	5′-ccaagctttagcctatatgacgatggcagtctcaaatgttcattggtt-3′
SMcGpx2	5′-cctctagaatttcatcaaatggcaaccaatgaacatttgagact-3′
rphGpx1	5′-ccaagctttagcggcactcatgacggtctgc-3′
rphGpx2	5′-ccgcggccgcctggtcctcgggactgc-3′
wtD1-1	5′-ccaagctttagtttatgatggtcacagtgtaaag-3′
wtD1-2	5′-ccgcggccgctttttaaaaatcaagtcacagctgtgtg-3′

Cell culture and transfections

A human embryonic kidney cell line, 293-HEK, was grown and maintained by standard tissue culture techniques in Dulbecco's modified Eagle's medium supplemented to 10% with fetal bovine serum. Transient transfections were performed by calcium phosphate precipitation as described previously (Berry et al., 1991b) using 10 or 20 μg of each deiodinase construct. Three micrograms of an expression vector containing the human growth hormone cDNA under control of the HSV thymidine kinase promoter, or 1 μg of an expression plasmid containing the β-galactosidase cDNA were cotransfected in order to monitor transfection efficiency. In addition, 4 μg of the pUHD-15 plasmid (Gossen & Bujard, 1992), which encodes the tetracycline-repressor DNA binding domain/VP16 activation domain fusion protein necessary for transcriptional activation of the pUHD10-3 promoter, were cotransfected.

Deiodinase activity assays

Cells were transfected as described above and harvested two days after transfection. Cell sonicates were assayed for the presence of 5′ deiodinase activity as described previously (Berry et al., 1991b). Briefly, cells were harvested, washed in PBS, and resuspended in 0.25 M sucrose in 0.1 M sodium phosphate/1 mM EDTA buffer. Cells were then sonicated briefly and assayed for 5′ deiodination of ^{125}I-reverse T3. Reactions were performed with 10–250 μg protein in 0.1 M potassium phosphate, pH 6.9, 10 mM DTT, and 1 mM EDTA, for 30 min at 37 °C. ^{125}I release was quantitated as described previously (Berry et al., 1993). Deiodinase activities were calculated per microliter of cell sonicate and normalized to either the amount of growth hormone secreted into the media or the amount of β-galactosidase activity in the cell sonicate. In a random sample of sonicates from various separate transfections, protein concentrations vary by less than 5% within a given week. All constructs were tested in at least three independent transfections and deiodinase assays were performed in duplicate from each transfection.

Protein expression was confirmed for all constructs by western blot analysis of cell sonicates using a polyclonal antiserum directed against an N-terminal peptide of rat D1 (Berry et al., 1994). This antibody detects both the functional full-length 28-kDa D1 product and the 14-kDa product resulting from termination, rather than selenocysteine incorporation, at the native in-frame UGA codon. For SECIS mutant

constructs exhibiting low deiodinase activity, that this result was due specifically to impairment of SECIS function and not to other effects of the mutations (e.g., RNA stability) was confirmed by the presence of the 14-kDa translation termination product and absence of the functional 28-kDa product on western blots. We interpreted the presence of roughly equivalent amounts of protein resulting from the usage of UGA as a stop codon, rather than a selenocysteine codon, among sonicates of cells transfected with different constructs as evidence that the steady-state levels of mRNA being generated were not significantly affected by SECIS mutations. Northern blot analysis of total RNA from representative transfections supports the validity of this conclusion. Relative transfection efficiency as measured by growth hormone RIA or β-galactosidase assay was also confirmed by comparison of the amounts of 28-kDa D1 protein.

ACKNOWLEDGMENTS

This work was supported by NIH grant DK47320 (M.J.B.) and an NSF Predoctoral Fellowship (G.W.M.). We thank Dr. Phil Whanger for the unpublished human selenoprotein W cDNA sequence and Shoshana Killian for excellent technical assistance.

Received August 5, 1997; returned for revision September 12, 1997; revised manuscript received October 7, 1997

REFERENCES

Battiste JL, Mao H, Rao NS, Tan R, Muhandiram DR, Kay LE, Frankel AD, Williamson JR. 1996. α-Helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science 273*:1547–1551.

Berry MJ, Banu L, Chen Y, Mandel SJ, Kieffer JD, Harney JW, Larsen PR. 1991a. Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3′ untranslated region. *Nature 353*:273–276.

Berry MJ, Banu L, Harney JW, Larsen PR. 1993. Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J 12*:3315–3322.

Berry MJ, Harney JW, Ohama T, Hatfield DL. 1994. Selenocysteine insertion or termination: Factors affecting UGA codon fate and complementary anticodon:codon mutations. *Nucleic Acids Res 22*:3753–3759.

Berry MJ, Kieffer JD, Harney JW, Larsen PR. 1991b. Selenocysteine confers the biochemical properties of the type I iodothyronine deiodinase. *J Biol Chem 266*:14155–14158.

Calnan BJ, Biancalana S, Hudson D, Frankel AD. 1991. Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA–protein recognition. *Genes & Dev 5*:201–210.

Forchhammer K, Leinfelder W, Böck A. 1989. Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature 342*:453–456.

Forchhammer K, Rucknagel KP, Bock A. 1990. Purification and biochemical characterization of SELB, a translation factor involved in selenoprotein synthesis. *J Biol Chem 265*:9346–9350.

Gossen M, Bujard H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci USA 89*:5547–5551.

Heider J, Baron C, Böck A. 1992. Coding from a distance: Dissection of the mRNA determinants required for the incorporation of selenocysteine into protein. *EMBO J 11*:3759–3766.

Hill KE, Lloyd RS, Burk RF. 1993. Conserved nucleotide sequences in the open reading frame and 3′ untranslated region of selenoprotein P mRNA. *Proc Natl Acad Sci USA 90*:537–541.

Kollmus H, Flohé L, McCarthy JEG. 1996. Analysis of eukaryotic mRNA structures directing cotranslational incorporation of selenocysteine. *Nucleic Acids Res 24*:1195–1201.

Leinfelder W, Forchhammer K, Zinoni F, Sawers G, Mandrand-Berthelot MA, Böck A. 1988. *Escherichia coli* genes whose products are involved in selenium metabolism. *J Bacteriol 170*: 540–546.

Low SC, Berry MJ. 1996. Knowing when not to stop: Selenocysteine incorporation in eukaryotes. *Trends Biochem Sci 21*:203–208.

Martin GW, Harney JW, Berry MJ. 1996. Selenocysteine incorporation in eukaryotes: Insights into mechanism and efficiency from sequence, structure, and spacing proximity studies of the type 1 deiodinase SECIS element. *RNA 2*:171–182.

Ringquist S, Schneider D, Gibson T, Baron C, Böck A, Gold L. 1994. Recognition of the mRNA selenocysteine insertion sequence by the specialized translation elongation factor SELB. *Genes & Dev 8*:376–385.

Rould MA, Perona JJ, Soll D, Steitz TA. 1989. Structure of *E. coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8 Å resolution. *Science 246*:1135–1142.

Shen Q, Chu FF, Newburger PE. 1993. Sequences in the 3′-untranslated region of the human cellular glutathione peroxidase gene are necessary and sufficient for selenocysteine incorporation at the UGA codon. *J Biol Chem 268*:11463–11469.

Shen Q, Leonard JL, Newburger PE. 1995a. Structure and function of the selenium translation element in the 3′-untranslated region of human cellular glutathione peroxidase mRNA. *RNA 1*:519–525.

Shen Q, McQuilkin PA, Newburger PE. 1995b. RNA-binding proteins that specifically recognize the selenocysteine insertion sequence of human cellular glutathione peroxidase mRNA. *J Biol Chem 270*:30448–30452.

Walczak R, Westhof E, Carbon P, Krol A. 1996. A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs. *RNA 2*:367–379.

Weeks KM, Crothers DM. 1991. RNA recognition by Tat-derived peptides: Interaction in the major groove? *Cell 66*:577–588.

Williams DL, Pierce RJ, Cookson E, Capron A. 1991. Molecular cloning and sequencing of glutathione peroxidase from *Schistosoma mansonii*. *Mol Biochem Parasitol 52*:127–130.

Zinoni F, Heider J, Böck A. 1990. Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. *Proc Natl Acad Sci USA 87*:4660–4664.