This is a copy of the free text from BioMed Centrals Repository (PMC)

PMCID: PMC145795

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC145795/

published in

Kollmus, H., Flohé, L., McCarthy, J.E.G.

Analysis of eukaryotic mRNA structures directing
cotranslational incorporation of selenocysteine
Analysis of eukaryotic mRNA structures directing cotranslational incorporation of selenocysteine

Heike Kollmus, Leopold Flohé and John E.G. McCarthy*

Department of Gene Expression, National Biotechnology Research Centre (GBF), Mascheroder Weg 1, D-38124 Braunschweig, Germany

Received January 24, 1996; Revised and Accepted February 20, 1996

ABSTRACT

Translation of an mRNA encoding a selenoprotein requires that at least one UGA codon in the reading frame is recoded as a site for the insertion of selenocysteine. In eukaryotes, the termination codon recoding event is directed by a cis-acting signal element located in the 3' untranslated region of the gene. This 'selenocysteine insertion sequence' (SECIS) comprises conserved sequences in a region of extensive base-pairing. In order to study the structure–function relationships of the SECIS structure, we have applied a newly developed reporter gene system which allows analysis of stop codon suppression in animal cell lines. This system obviates the need for enzymatic or immunological estimation of selenoprotein synthesis, relying instead on the simple quantification of translational readthrough from the lacZ gene into the luciferase gene. The 3'-UTR of the phospholipid hydroperoxide glutathione peroxidase (PHGPx) gene was shown to contain a highly active SECIS element. Mutations in the base-paired sequences of other SECIS elements were used to analyse the significance of primary structure, secondary structure and pairing stability in the stem regions. The results demonstrate that the exact sequences of the paired nucleotides are comparatively unimportant, provided that a consensus combination of length and thermodynamic stability of the base-paired structures is maintained.

INTRODUCTION

Selenium is found covalently associated with certain prokaryotic and eukaryotic proteins called selenoproteins (1–3). The carrier of this group VI element is the twenty-first amino acid, selenocysteine, which can be incorporated cotranslationally into nascent polypeptide chains. The site of selenocysteine incorporation in both prokaryotes and eukaryotes was found to be the UGA codon (4–7), originally identified as one of the nonsense codons in the genetic code (8). This alternative interpretation of the UGA codon (‘recoding’; 9) requires the presence of additional signals. In Escherichia coli, UGA is only read as a selenocysteine incorporation site if it is immediately followed in the reading frame by a region that has been shown to form a stem–loop structure (10). An element with an apparently analogous function, referred to as the selenocysteine insertion sequence (SECIS; 11) or selenium translation element (STE; 12), exists in selenoprotein-encoding mRNAs of eukaryotes (7,13). However, unlike the prokaryotic signal, the eukaryotic element is found in the 3' untranslated region (3'-UTR; 7,11,13,14), and must therefore be effectively acting at a distance. The mechanism of action of the SECIS elements is as yet unknown. In E.coli, recoding of UGA involves a number of specific cellular components, including a special type of seryl-tRNA (tRNA^Sec) and a unique translation factor (SELB) which is related to EF-Tu (6). Eukaryotic tRNA^Sec species have also been described (15,16), but no eukaryotic counterpart of SELB has been characterized so far.

The SECIS element is an example of a structural element in eukaryotic mRNA that most likely fulfills its function via interactions with an RNA-binding protein (analogously to the prokaryotic SELB). Although many processes of posttranscriptional gene expression involve RNA–protein interactions, very little is known about the structural properties and functional mechanisms that underlie them (17,18). An important approach to furthering our knowledge of the mechanisms of such interactions is to obtain a deeper understanding of the structural characteristics of mRNA binding sites that are necessary for function. Typically for a number of known specific RNA binding sites for proteins (17), the SECIS type of element comprises a small number of conserved nucleotides located in a stem–loop-like structure. The conserved sequences of the SECIS element are three short stretches: AAA in or near the apical loop of the overall structure, UG at an internal loop on the 3' side of the stem, and AUGA at an internal bulge site on the 5' side of the stem. The overall structure is generally large, comprising 80–150 nucleotides. This feature is also conserved, at least as far as can be judged from a theoretical assessment of the secondary structures likely to be formed by the known SECIS elements (12–14).

In the present work we have applied a newly developed assay system to test the relative levels of stop codon suppression directed by a range of SECIS elements. This reporter gene system, which is based on the fusion of two reporter genes, allows assessment of the functional properties of SECIS elements independently of the selenoprotein genes with which these sequences are normally associated. We have compared the activity of the recently described SECIS element of the PHGPx gene (19) with the activities of those reported previously.

* To whom correspondence should be addressed
Moreover, we have obtained results that provide insight into the significance of the molecular shape and dimensions of SECIS element structure, showing that these are more important than simply the thermodynamic stability of its constituent parts.

**MATERIALS AND METHODS**

**Plasmid constructs**

All DNA manipulations were carried out according to standard procedures (20). Synthetic oligonucleotides used in the plasmid constructions are listed in Table 1. An overview of the constructed plasmids is given in Table 2. In the first step towards the construction of pBPLUGA, the plasmid pBgalac-1 (21) was cleaved with HindIII and partially with BamHI. The annealed oligonucleotide pair GLBAM1/2 was inserted resulting in the plasmid DNA leading to elimination of the BamHI site, whereby a silent mutation in the N-terminal amino-acid sequence of the β-galactosidase was introduced. The resulting plasmid was called pBGL. After digestion with SalI and BamHI, pBGL was relicated with the oligonucleotide pair GLUGA1/2, creating an in-frame stop codon in the intergenic region between β-galactosidase and luciferase. In parallel, an oligonucleotide pair (GLPLCYS1/2) containing a cysteine codon instead of the stop codon was introduced into the same site. The polylinker oligonucleotide pair PL3U1/2 was inserted into the KpnI site of the 3′-UTR. The resulting plasmids were called pBPLUGA and pBPLPCLys. Insertion of the polylinker sequence in the reverse orientation resulted in the plasmid pBPLUGA. Insertion of the oligonucleotides GLSTOP1/2 (SalI/BamHI) between the two reporter genes fused the luciferase gene in the −2 frame relative to the lacZ gene (pBSTOP). pBPLUGA (Fig. 1A) was the master plasmid used for the insertion of different SECIS elements, which were ligated into the polylinker in the 3′-UTR. The plasmids pBPHGpx3U and pBPHGpx3U3 contain the 3′-UTR of the pig heart PHGpx gene (19) in the natural and reversed orientation, respectively. This 3′-UTR was obtained by digesting the plasmid pMM7 using BglII and NsiI. The resulting fragment was ligated with the BglII/PstI-digested plasmids pBPLUGA and pBPLUGA. pM7 is derived from pHST7 (22) and contains the 3′-UTR of the PHGpx gene (nucleotides 509–734 in ref. 23), synthesized by PCR using the primers RPHG and LPHG and the vector pMM5 (23) digested with BglII and EcoRI. The other SECIS elements were cloned as synthetic oligonucleotides bearing BglII (5′) and BamHI (3′) ends. The plasmid with the minimal PHGpx SECIS element (pBPHGpxmin) comprises nucleotides 570–626 of the PHGpx 3′-UTR (Fig. 2B). The SECIS element of the rat 5′ deiodinase (pBDI, Fig. 2B) comprises nucleotides 1519–1596 of the gene sequence (11). The mutant forms of this structure (pBDIM1 and pBDIM2) are indicated in Figure 2B. The construct pBSEPL1min includes the nucleotides 1472–1519 (Fig. 2B) of the first SECIS element of the rat selenoprotein P gene. The sequence of the mutated form of this SECIS element (pBSEPL1M1) is indicated in Figure 2B. The plasmid pBSEC contains a semi-synthetic SECIS element, cloned by integration of a BglII/AvIII oligonucleotide (SECIS1/2) in pBPLUGA. The mutation of the SECIS element of the plasmid pBSECISM is marked in Figure 4B. Further extensions (E) of the SECIS element were created using oligonucleotides comprising naturally occurring sequences or other synthetic extension sequences. The 5′ extensions of the pBSEPL1min SECIS elements were constructed using BglII oligonucleotide pairs (NE151/2, NE251/2, ES1/2 or SE1/2) while the 3′ extensions were created using BamHI/PstI oligonucleotide.

![Table 1. Oligonucleotides](image)

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBHAV</td>
<td>ACTCAAGGACTGCTCGAGTTGAGTTGTTT</td>
</tr>
<tr>
<td>pHAV</td>
<td>AATGAGATCTAGATTGAGTTGTTT</td>
</tr>
<tr>
<td>GLBAM1</td>
<td>AAGTGCTTCAAGATATTGAGTTGTTT</td>
</tr>
<tr>
<td>GLBAM2</td>
<td>AAGTGCTTCAAGATATTGAGTTGTTT</td>
</tr>
<tr>
<td>GLSTOP1</td>
<td>GATGCTTCAAGATATTGAGTTGTTT</td>
</tr>
<tr>
<td>GLSTOP2</td>
<td>GATGCTTCAAGATATTGAGTTGTTT</td>
</tr>
<tr>
<td>GLMAM1</td>
<td>AAGTGCTTCAAGATATTGAGTTGTTT</td>
</tr>
<tr>
<td>GLMAM2</td>
<td>AAGTGCTTCAAGATATTGAGTTGTTT</td>
</tr>
<tr>
<td>GLPCLYS1</td>
<td>GATGCTTCAAGATATTGAGTTGTTT</td>
</tr>
<tr>
<td>GLPCLYS2</td>
<td>GATGCTTCAAGATATTGAGTTGTTT</td>
</tr>
</tbody>
</table>

Relevant restriction sites are indicated by underlining. All oligonucleotides are displayed 5′ to 3′. PCR oligonucleotides are marked with R and L, indicating right and left side, and complementary oligonucleotides are designated 1 and 2, respectively.

**Cell culture, gene transfer and enzyme assays**

BHK-21 cells (baby hamster kidney cells; ATCC CC110) were cultivated in DulBEcco’s modified Eagle’s medium supplemented with 10% fetal calf serum. Transient transfections with plasmids were carried out by means of calcium phosphate coprecipitation (23). Determination of activities of β-galactosidase and luciferase was performed in cells harvested 48 h after transfection. Cells were detached from culture plates by incubation with TEN (40 mM Tris–HCl, pH 7.5; 1 mM EDTA, 150 mM NaCl) after rinsing the cells twice with PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, 1.47 mM KH₂PO₄, pH 7.0). After centrifugation, the
Table 2. Summary of the plasmid constructions

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Short description</th>
<th>SECIS</th>
<th>50% (kb/kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBS-5TOP</td>
<td>Background control mixture &amp; 3′-UTR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBGLCys</td>
<td>Cysteine control mixture &amp; 3′-UTR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBGLCys</td>
<td>Cysteine control mixture &amp; 3′-UTR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBGLCysA</td>
<td>Cysteine control mixture &amp; 3′-UTR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBGLCysB</td>
<td>Cysteine control mixture &amp; 3′-UTR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBGLCysC</td>
<td>Cysteine control mixture &amp; 3′-UTR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBGLCysD</td>
<td>Cysteine control mixture &amp; 3′-UTR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBGLCysE</td>
<td>Cysteine control mixture &amp; 3′-UTR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The right hand columns indicate whether a SECIS element is present in the 3′-UTR and provide estimates of the stabilities of the base-paired structures predicted to be formed by the Zuker and Stiegler program (38). Where sequence numbers are given, the indicated reference provides the exact nucleotide sequence used in the plasmid construction.

cell pellet was resuspended and subjected to three cycles of freeze–thaw in 250 µl 250 mM Tris–HCl (pH 7.5). For measurement of the luciferase activity, 20–60 µl of the supernatant was mixed with 350 µl reaction buffer (25 mM glycylglycine, pH 7.8, 5 mM ATP, 15 mM MgSO4). The light emission (light units/10 s) of the reaction mixture was measured in a Berthold Biolumat (LB9501) after injection of 100 µl luciferin (0.2 mM; Sigma). For determination of β-galactosidase activity, 20–60 µl cell extract were incubated with 1 µl reaction buffer (60 mM NaPO4, 40 mM NaH2PO4, 10 mM KCl, 1 mM MgCl2, 50 mM β-mercaptoethanol) and 200 µl substrate (2 mg/ml o-nitrophenyl-β-D-galactopyranoside in 60 mM NaPO4). The reaction was stopped by addition of 500 µl 1 M NaCO3. The colorimetric change was measured at 405 nm.

Determination of stop codon suppression

The luciferase activity of cells transfected with the construct pBGLCysC (harbouring a cysteine codon between the coding regions of β-galactosidase and luciferase) was taken as the reference value (100%). Stop codon suppression efficiency was calculated by relating the enzymatic activities of cells containing the listed constructs (Table 2) to the values obtained with this control. The specific activity of the β-galactosidase enzyme segment was shown previously to be insensitive to the fusion of a further protein to its C-terminus (21). Thus the ratio of the

Figure 1. Analysis of stop codon suppression mediated by SECIS elements. (A) The lacZ and luc genes are fused via a synthetic linker sequence including an in-frame TGA codon flanked by the restriction sites SalI and BamHI, respectively. A synthetic polylinker region in the 3′-UTR can be used for the insertion of SECIS-related elements. Transcription from the SV40 promoter produces an mRNA that can be translated in either of two ways (B). In the absence of a SECIS element, translation terminates at the UGA in the linker region, yielding a C-terminally extended β-galactosidase. Suppression of the UGA mediated via a SECIS element leads to synthesis of a β-galactosidase–luciferase fusion protein (GAL–LUC) showing both catalytic activities.

The respective enzyme activities can be taken as a reliable indicator of the relative levels of SECIS-dependent readthrough.

RESULTS

A test system for the quantification of stop codon suppression efficiency in mammalian cell lines

We developed an assay system designed to facilitate analysis of elements in the 3′-UTR that influence stop codon suppression in mammalian cell lines. It is based on the reporter genes encoding β-galactosidase and luciferase, which are fused in frame via a TGA stop codon (pBPLUGA, Fig. 1A). After transfection of the respective plasmids into mammalian cell lines, the DNA is transcribed under the control of the SV40 promoter and translation leads to the synthesis of the reporter enzymes. Translation can either terminate at the C-terminus of lacZ or, upon suppression of the TGA codon, continue through the linker region to generate a fusion protein of β-galactosidase and luciferase (GAL–LUC) (Fig. 1B). The sequence context of the TGA codon can be altered by using the SalI and BamHI restriction sites either side of it. The polylinker region in the 3′-UTR serves as a site for insertion of known or putative SECIS elements, or alternatively derivatives of them, using the restriction sites shown (Fig. 1A). The encoded enzyme activities are readily measured and afford a high level of sensitivity. Calculation of suppression activity is based on the parallel transfection of the reference plasmid pBGLCysC. In this plasmid, the coding regions of β-galactosidase and luciferase are fused via an in-frame cysteine codon. Translation of the mRNA encoded by this plasmid leads exclusively to synthesis of the fusion protein GAL–LUC. The efficiency of any given SECIS element can be calculated by comparing the enzymatic activities of pBGLCysC-transfected cells, which are taken at 100% reference values, to those of cells transfected with the appropriate derivative of pBPLUGA. Further
experiments revealed that variation of the sequences in the 3'-UTR region does not have any influence on either the enzyme activities or the stability of the mRNA (data not shown).

**SEClS elements from different genes and organisms mediate stop codon suppression in BHK-21 cells**

We found the fusion gene reporter system to be suited for the analysis of wild-type and mutated forms of natural SEClS elements. Stop codon suppression activities were measured in transiently transfected BHK-21 cells. We inserted the complete 3'-UTR of the recently described pig heart phospholipid hydroperoxide glutathione peroxidase (PHGPx) gene (19). This supported a stop codon suppression activity of 2.8% (pBPHGPx3U, Fig. 2A). A minimal SEClS element only harbouring the upper part of the stem–loop structure, including the conserved nucleotides (pBPHGPxmin, Fig. 2B), supports a 10-fold reduced activity (0.3%). In contrast, a construct lacking a SEClS structure (pBPlUGA), or with the 3'-UTR inserted in the inverse orientation (pBPHGPxU3), mediates suppression with an efficiency of 0.1%. This value reflects the naturally occurring, non SEClS-dependent, stop codon suppression activity observed in mammalian cell lines. A further control construct, in which the luciferase gene is fused in the reading frame that is –2 with respect to lacZ (pBSTOP, referred to ‘control’ in Fig. 2A), yields the background value of the test system (0.05%).

We compared other known SEClS elements to that of the PHGPx gene. The SEClS element of the gene encoding rat 5' deiodinase (nucleotides 1519–1596 in ref. 11; pBDI) led to a relative stop codon suppression efficiency of 1.1% (Fig. 2A). We also examined a loop mutant, previously tested by Berry and colleagues (13), which converts one of the conserved As into Gs (pBDIM1; Fig. 2B). It was found to reduce drastically the efficiency, although a residual activity remained (0.16%). This contrasts with the complete elimination of activity caused by a combination of this mutation with a further change in the 3' bulge which converts the conserved UG to UA (pBDIM2; Fig. 2B). Finally, we also investigated whether a minimal SEClS element, comprising the upper part of the first stem–loop of the selenoprotein P 3'-UTR (pBSELPImin; Fig. 2B), is functional, finding that it supports stop codon suppression at a level of 0.2%. A deletion mutant form of this minimal element lacking the three conserved As in the loop region (pBSELPI1min; Fig. 2) showed no activity.

The above experiments demonstrate that a minimal SEClS element retaining only the part with the conserved regions is sufficient to mediate stop codon suppression above the level of the naturally occurring suppression activity of the cell. However, activities typical of the complete wild-type sequences are evidently not attainable with such minimal structures.

**The significance of the overall shape and structural stability of SEClS elements**

Inspection of the different stabilities of the respective SEClS elements analysed (Fig. 2B) leads to the suggestion that there...
might be a correlation between activity and structural stability. In order to examine this possibility we made use of the synthetic minimal SELP1 element as the starting point for a further series of constructs (Fig. 3B). The stem was extended by introducing additional nucleotides (pBSELPIE). The synthetic extension sequences were chosen arbitrarily and are not derived from known SECIS elements. A 2-fold increase in the stability of this element resulted in a doubling of the efficiency (0.2–0.4%). A control in which only the 3' arm was extended (pBSELPIE3) had the same activity as the minimal selenoprotein P SECIS element. This approach left open the question as to the significance of the specific sequences that had been inserted. We therefore added back wild-type sequences to the minimal element that are present in the bottom part of the natural selenoprotein P SECIS element. The first extension (pBSELPIE1; Fig. 3B), which was predicted to lead to only a slight increase in stability compared to the minimal element, did not stimulate the rate of selenocysteine incorporation. In contrast, a further stabilization using additional wild-type sequences (pBSELPIE2), yielding a predicted increase in stability ($\Delta$G = $-46.8$ kcal.mol$^{-1}$), led to an enhancement of stop codon suppression activity (0.85%). Overall, this set of results indicated that extension of the minimal SECIS element to allow formation of a more stable structure enhances activity.

We next addressed the question why the naturally occurring SECIS elements are generally such large structures. Equivalent stabilities could theoretically also be attained using shorter stretches of mRNA that are more G/C rich. We therefore tested different types of structure which shared the same stability. These were created by inserting either a short G/C rich extension

Figure 3. The relationship between the folding stability of the SECIS structure and stop codon suppression. The stop codon suppression efficiency mediated by different plasmids after transient transfection of BHK-21 cells was determined as described in Figure 2. The SECIS elements are drawn schematically, whereby lines show wild-type sequences of the selenoprotein P SECIS element. The conserved regions are indicated by letters. The large filled points symbolize added base pairs.

Figure 4. The role of specific sequences in the SECIS elements. The stop codon suppression efficiency was determined as described in Figure 2. The symbols are identical to those used in Figure 3.
Exchange of base-paired nucleotides in the upper stem of the SelP SECIS element

We examined more closely the sequence requirements of the SelP SECIS minimal element by exchanging the base-paired stem region between the apical loop and the bulges with base pairs equal in base composition but different in sequence. This manipulation had only a minimal effect on the predicted stability and did not change the overall shape of the SECIS element (compare pbSELPlmin, Fig. 3B with pbSECIS, Fig. 4B). Despite the complete substitution of all the nucleotides in each arm of the stem region, there was little change in efficiency. Moreover, stabilisation of this rearranged minimal element led also to an enhancement of stop codon suppression efficiency (pBSECISEF, Fig. 4B; compare pSELPl1 in Fig. 3B). In contrast, the exchange of only one conserved base located in the apical loop of the RNA structure eliminated activity (pBSECISM and pBSECISME, Fig. 4).

DISCUSSION

We have described a novel reporter gene system for the analysis of selenocysteine incorporation in eukaryotes. This system can be used for structure–function studies of the mRNA elements involved in selenocysteine incorporation as well as for investigations of the mechanism of this process. Clearly, selenocysteine incorporation can be directed into a heterologous protein that normally does not contain this amino acid provided the appropriate 3’-UTRs are present. This principle could be of use in generating selenocysteine-containing derivatives of a range of proteins. Assuming that these derivatives could be produced in sufficiently large amounts, this would provide an alternative approach to solving the well-known phase problem encountered in X-ray diffraction analysis of protein crystals. The use of a reporter gene fusion for the estimation of stop codon suppression obviates the necessity to determine selenocysteine incorporation efficiency mediated by different SECIS elements on the basis of measurements of the activities of specific selenoproteins, radioactive labelling with 75Se, or immunochemical methods (11,12,14). The system could also be utilized to analyse the significance of the codon context of the suppressed TGA codon.

The new system has allowed us to assess the activity of the recently described 3’-UTR of the PHGpx gene (19) relative to the activities of previously identified SECIS elements. Our data clearly show that the PHGpx 3’-UTR contains an active SECIS element. Indeed, this element functions highly effectively (pBPpHgpx3U, Fig. 2) compared with the SECIS elements of both selenoprotein P and 5’ deiodinase. This observation is relevant to the physiological regulation of selenocysteine incorporation into different proteins in mammalian tissues. Under conditions of limited selenium availability, there is apparently selective incorporation of selenium into PHGpx in preference over GPx (26). The data presented here provide a potential explanation for this selective incorporation phenomenon. The enhanced ability of the PHGpx SECIS element to direct recoding of the UGA codon may allow selenocysteine incorporation into this protein to function relatively efficiently at limiting selenium concentrations. The same principle may also play a role in the organ-specific expression hierarchies that have been described for selenoproteins (27,28).

It should be noted that our results indicate similar efficiencies for the 5’ deiodinase and SelP SECIS elements, whereas Berry et al. (13) reported that the SelP element is the more effective of the two. This might be at least partially attributable to the different cell types used in the respective studies. Moreover, Berry and colleagues (13) used large parts of the 3’-UTRs of the respective selenoprotein-encoding genes, whereas we have focused on the sequences defining the SECIS element itself.

An interesting mechanistic problem arises in the case of a natural gene, such as SelP, which has 10 TGA sites at which selenocysteine has to be incorporated. If incorporation were to occur independently at each site with an efficiency as low as reported here, the yield of complete SelP protein would be negligible, since it would be a function of the multiplicative sum of all 10 insertion efficiencies. This clearly cannot be true, and we assume that suppression of translational termination in selenoprotein mRNAs follows a processive mechanism. One explanation of this would be that the presence of a SECIS element in the 3’-UTR leads to a ‘reprogramming’ of a small percentage of the ribosomes translating that particular mRNA. The degree of reprogramming of the translational apparatus would be equivalent to the suppression efficiencies measured in an assay system of the type we have presented. The ‘reprogrammed’ ribosomes would then interpret all UGAs on the mRNA that they are translating as selenocysteine incorporation sites. It should also be noted in this context that the low efficiency of SECIS-dependent UGA suppression is an intrinsic property of the incorporation mechanism. It was shown previously that translation of the complete reading frame of the 5’ deiodinase is 20- to 400-fold more efficient if the internal UGA is substituted by an UGU cysteine codon (29).

We have studied the components of SECIS elements directly involved in the formation of the secondary structure of this type of element. In particular, we have investigated the nature of the boundaries of an efficiently functioning SECIS element and the minimal structural and sequence requirements for activity. We find that the stability of the base-paired regions in the SECIS structure is, in itself, an inadequate criterion for the structural and functional significance of these regions (compare pbSELPl1SE to pbSELPl1EE and pbSELPl1E EI; Fig. 3B). Equally stable base-paired stem regions can be substituted into the SECIS element in which a smaller number of G-C base pairs achieves the same stability as that of the equivalent region in a natural SECIS element. Yet these shorter stem regions support only greatly reduced activity. This indicates that the length, and thus overall...
topology of the SECIS element, is essential for optimal function. Moreover, the less than maximal density of highly stable base pairing may allow a degree of flexibility that could be required for the SECIS to be able to perform its function, for example in a ‘looping back mechanism’, as proposed previously by Berry et al. (13). The content of distorted helical sections and A-U base pairs presumably contributes to specific properties of SECIS elements that cannot be achieved in more compact structures. For example, one functional requirement of the SECIS structure may be that at least sections of it need to be partially unwind or restructured.

While the correct balance between length and stability of base-paired stretches in the SECIS element is essential for optimal function, the exact nucleotide sequence in the respective arms can be varied with little effect. This principle even applies to the base-paired arms immediately adjacent to the apical loop (compare pBSEL1min. Fig. 3 with pBSECIS, Fig. 4). This suggests that apart from the nucleotides that are conserved between SECIS elements, the major criterion for identification of a SECIS element is the ability of the component regions to assume a specific shape and size.

Examination of other known protein-binding sites in mRNA reveals that a number of them, such as the iron responsive element (IRE; 30), the Rev-responsive element (RRE; 31), the trans-activation response element (TAR; 32) and the U1A 3'-UTR protein-binding site (33), comprise partly base-paired structures. The specifically recognized and conserved nucleotides generally lie in single-stranded regions, or in parts of the structures unlikely to form stable, undistorted helices (17). An informative comparison of sequence and structure can be made with the results of structure–function studies of the IRE. It has been shown that a minimal IRE structure comprising the conserved apical loop and a 10 bp A-U-rich stem containing a single-base bulge is sufficient to mediate strong translational repression when present in the 5'-UTR of mRNAs in animal cells (34,35) and yeast (36). This minimal structure lacks the base-paired ‘flanking regions’ that are otherwise present in ferritin mRNAs (37), and which would normally be expected to stabilize the IRE structure considerably. Yet binding of the iron regulatory protein (IRP) to the minimal IRE is evidently still sufficiently tight to lead to regulatable translational inhibition. This contrasts with the drastic loss of SECIS function observed upon size reduction. The minimal IRE required for binding of IRP, and thus for tight translational regulation via the 5'-UTR, is evidently considerably smaller than the SECIS structure necessary in the 3'-UTR for mediation of selenocysteine incorporation. One possible explanation for the drastic loss of function associated with reductions in overall size of the SECIS element might be the consequent changes in its binding behaviour towards protein factor(s) essential for selenium incorporation.

ACKNOWLEDGEMENTS

We thank Dr Joop van den Heuvel and Dr Matilde Maiorino for providing pMM7, as well as Dr Hansjörg Hauser for pBgalG1-1.

REFERENCES