*5868–5880* *Nucleic Acids Research, 2009, Vol. 37, No. 17* *Published online 3 August 2009*
*doi:10.1093/nar/gkp635*

# Novel structural determinants in human SECIS elements modulate the translational recoding of UGA as selenocysteine

**Lynda Latrèche[1,2], Olivier Jean-Jean[2], Donna M. Driscoll[3,4] and Laurent Chavatte[1,3,*]**

[1]Centre de recherche de Gif-sur-Yvette, FRC 3115. Centre de Génétique Moléculaire, CNRS, FRE 3144, Gif-sur-Yvette, [2]UPMC Univ Paris 06, FRE 3207, CNRS, F-75005 Paris, France, [3]Department of Cell Biology, Lerner Research Institute, Cleveland Clinic Foundation and [4]Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, OH 44195, USA

Received June 8, 2009; Revised July 15, 2009; Accepted July 16, 2009

## ABSTRACT

**The selenocysteine insertion sequence (SECIS) element directs the translational recoding of UGA as selenocysteine. In eukaryotes, the SECIS is located downstream of the UGA codon in the 3′-UTR of the selenoprotein mRNA. Despite poor sequence conservation, all SECIS elements form a similar stem-loop structure containing a putative kink-turn motif. We functionally characterized the 26 SECIS elements encoded in the human genome. Surprisingly, the SECIS elements displayed a wide range of UGA recoding activities, spanning several 1000-fold *in vivo* and several 100-fold *in vitro*. The difference in activity between a representative strong and weak SECIS element was not explained by differential binding affinity of SECIS binding Protein 2, a limiting factor for selenocysteine incorporation. Using chimeric SECIS molecules, we identified the internal loop and helix 2, which flank the kink-turn motif, as critical determinants of UGA recoding activity. The simultaneous presence of a GC base pair in helix 2 and a U in the 5′-side of the internal loop was a statistically significant predictor of weak recoding activity. Thus, the SECIS contains intrinsic information that modulates selenocysteine incorporation efficiency.**

## INTRODUCTION

Selenium is an essential trace element for human health. Epidemiological studies report that selenium protects against cancer, delays the development of neurodegenerative disorders, reduces the progression of HIV in infected patients and diminishes mortality in the aging population (1–4). Most of the beneficial effects of selenium are likely due to its presence as selenocysteine in a specific group of proteins, called selenoproteins. Extensive computer analyses of the human genome identified what is now thought to be the complete set of selenoprotein genes (5). The human selenoproteome is composed of at least 25 selenoproteins, although some genes may lead to the production of more than one isoform (6). Many human selenoproteins are enzymes involved in oxidoreduction reactions. Enzymatic activities have been characterized for approximately one half of the selenoproteome, including the glutathione peroxidases (Gpx1–Gpx6), thioredoxin reductases (Trxr1–Trxr2), iodothyronine deiodinases (Dio1–Dio3), methionine sulfoxide reductase (MsrB1 or SelX) and selenophosphate synthetase (Sps2). In many cases, the selenocysteine residue plays a vital role for enzymatic activity, given its known or predicted location in the catalytic site (7–9). However, the remaining human selenoproteins have not been assigned a precise function due to their recent bioinformatic identification in the genome.

Selenocysteine is considered as the 21st amino acid, since its incorporation into the protein sequence depends on the translational recoding of an in-frame UGA codon as reviewed in (10–14). Eukaryotes have evolved an intricate strategy to use the UGA as a sense codon for selenocysteine in selenoprotein messenger RNAs (mRNAs), while maintaining its use as a stop codon in other cellular mRNAs. The 3′-untranslated region (3′-UTR) of mammalian selenoprotein mRNAs always harbor a selenocysteine insertion sequence (SECIS) downstream of the UGA codon, which is essential for selenocysteine incorporation. *Trans*-acting recoding factors identified during the past decade include a selenocysteine $tRNA^{Sec}$ ($Sec\text{-}tRNA^{Sec}$) (15,16), a dedicated elongation factor (EFsec) (17,18), and two proteins that bind to the SECIS element: SECIS binding protein 2 (SBP2) (19,20) and ribosomal protein L30 (21). It has been proposed that SBP2, which is limiting in cells, functions as a discriminatory factor and

*To whom correspondence should be addressed. Tel: (33) 1 69 82 32 13; Fax: (33) 1 69 82 31 40; Email: laurent.chavatte@cgm.cnrs-gif.fr

dictates the expression of the selenoproteome (20–23). The mechanism of selenocysteine incorporation, particularly the sequence of events and the nature of the complexes leading to UGA recoding, is not fully understood yet and is discussed in (10–14).

Eukaryotic SECIS elements are not highly conserved at the nucleotide level, but their secondary structures can be represented as a consensus stem-loop–stem-loop folding, as illustrated in Figure 1A. Two conserved motifs in the SECIS are essential for faithful recoding of UGA as selenocysteine: the SECIS core that contains a quartet of non-Watson–Crick base pairs and an AAA/G motif (24–28). The precise distance between these two regions is maintained by a base paired structure (Figure 1A). The SECIS core encompasses two sheared tandem G–A pairs (hereafter referred to as tandem GA pairs), which are conserved in all eukaryotic SECIS elements (29) and are required for the binding of SBP2 and L30 (19–21). In other RNAs with a known 3D structure, the tandem GA pairs are characteristic of a kink-turn motif, which can adopt distinct conformations in solution (30,31). We proposed that the putative kink-turn motif in the SECIS allows the element to act as a molecular switch by undergoing dynamic conformational changes in response to protein binding (21). In contrast to the SECIS core, the essential AAA/G motif has no known function or interacting partner. This motif is replaced by CCX in some SECIS elements (5).

Although it is well-established that the SECIS core and AAA/G motif are required for selenoprotein synthesis, little is known about the determinants in SECIS elements from endogenous selenoprotein mRNAs that might regulate the efficiency of selenocysteine incorporation *in vivo*. The region surrounding the tandem GA pairs, which is highly variable among SECIS elements, could affect the kink-turn folding of the RNA and/or the binding activities of SBP2 and L30. In addition to the SECIS, the efficiency of UGA recoding is affected by other *cis*-acting elements in the selenoprotein mRNA, including (i) the selenocysteine codon context (32), (ii) the distance between UGA codon and the SECIS (26,33) and (iii) the presence of a stem-loop structure adjacent to the UGA codon in a subset of selenoprotein mRNAs (34,35). However, in all these cases, the modulation of UGA recoding efficiency did not exceed a 10- to 20-fold difference.

The purpose of this work is to investigate a potential regulatory effect of the SECIS on selenoprotein expression on the scale of the human selenoproteome. We analyzed the UGA recoding efficiencies of the 26 human SECIS elements in transfected cells and in an *in vitro* translation system. Unexpectedly, we found that SECIS elements vary dramatically in their ability to direct UGA recoding, with the differences ranging 4900- and 200-fold in cells and *in vitro*, respectively. Using chimeric SECIS elements, we showed that the internal loop and helix 2 contain major determinants that control the strength of UGA recoding. Sequence conservation analysis of the regions surrounding the kink-turn motif in human SECIS elements identified statistical predictors of weak UGA recoding activity.

## MATERIALS AND METHODS

### DNA constructs

Coding sequence of firefly luciferase (*Photinus pyralis*) was previously cloned in pcDNA3.1 vector (Invitrogen). PacI and NotI restriction sites are present immediately downstream of the UAA stop codon. Additionally, the luciferase sequence was modified at position 258 to contain an in frame UGA codon. This construct has been designed to quantitatively analyze the influence of any SECIS element on UGA recoding efficiency *in vivo* and *in vitro* as described in (21,33,36). The 26 SECIS element sequences were found in the human genome database and used to design primers for RT-PCR amplification and cloning. To have SECIS elements of similar size that would encompass the minimum active domain for all selenoprotein mRNAs, we selected 103-nt long sequences starting ~30 bp before the highly conserved ATGA motif (as listed in Figure 1B). PCR primers were designed to contain either a PacI or NotI restriction site (listed in Supplementary Figure S1), in the forward and reverse primers, respectively, for subcloning into pcDNA3.1 luciferase $UGA^{258}$ vector. Total RNAs were extracted from Hek293 and HepG2 cells using TRIzol reagent (Invitrogen) following the manufacturer's instructions. Total RNAs (5 µg) were annealed with 50 pmol of oligodT and used for reverse transcription using M-MLV retrotranscriptase (Promega) in 40 µl reaction. Two microliter of this reaction was used for the PCR amplification using Platinum *Pfx* DNA polymerase (Invitrogen) and 4 µM of adequate primers. PCR products and pcDNA3.1 luciferase $UGA^{258}$ vector were digested with PacI and NotI enzymes, purified and ligated with T4 DNA Ligase after dephosphorylation of the vector. Quick change site-directed mutagenesis (adapted from Stratagene) were performed to generate SECIS (ΔAUGA, ΔAAAC and ΔAAAG) and luciferase ($UAA^{258}$ and $UGU^{258}$) mutants using overlapping oligonucleotides containing the designed mutations (see Supplementary Figure S2) and the *PfuTurbo* DNA polymerase (Stratagene). To generate the SECIS chimeric constructs of SelX and Gpx3 listed in Supplementary Figure S3, we used pairs of overlapping oligonucleotides with a 3′-overhang that were completed with Pwo DNA polymerase. PCR amplifications were performed using 4 µM of either the set of Gpx3 or SelX SECIS primers, designed to add PacI and NotI restriction sites. For RNA electromobility shift assay (REMSA) experiments, luciferase $UGA^{258}$-Gpx3 and luciferase $UGA^{258}$-SelX constructs were digested with PacI and NotI enzyme to clone the SECIS element downstream of a T7 RNA polymerase promoter in a pUC19 vector. All the construct sequences were verified by automated DNA sequencing.

### Cell culture and transfection

Hek293 (Invitrogen) and HepG2 (ATTC) cells were grown and maintained in 100 mm plates in Dulbecco's Modified Eagle Medium (D-MEM) and in minimum essential medium (MEM) supplemented with non-essential amino

acids, respectively. Media were supplemented with 10% fetal bovine serum, 100 μg/ml streptomycin, 100 UI/ml penicillin, 1 mM sodium pyruvate and 2 mM L-glutamine. Media and supplements were purchased from Invitrogen. Cells were cultivated in 5% $CO_2$ at 37°C and humidified atmosphere. Calcium precipitation methods were performed to transiently transfect the cells with 6 μg of plasmid (5 μg of luciferase and 1 μg of β-galactosidase vectors). Cells were split the day prior to transfection. Media was changed 24-h post-transfection and cells collected at 48 h with 300 μl of lysis buffer (25 mM Tris–phosphate pH 7.8, 2 mM DTT, 2 mM EDTA, 1% Triton X-100, 10% glycerol). Cell extracts were assayed for luciferase and β-galactosidase activities by chemiluminescence (Promega Luciferase and Beta-Glo assay systems, respectively), in triplicate using a microplate reader FLUOstar OPTIMA (BMG Labtech).

### *In vitro* transcription and translation

Luciferase-SECIS pcDNA3.1 plasmids were linearized by NotI enzyme and used as templates for *in vitro* transcription using T7 RNA polymerase (Ribomax kit, Promega) according to the manufacturer's instructions. Cap analog was added to the reaction at a 4 : 1 ratio (versus GTP) to generate 5′-capped mRNAs. *In vitro* translation reactions were assembled in a total volume of 12.5 μl that includes 100 ng of luciferase mRNA (0.19 pmol), 50 ng of recombinant SBP2 (amino acids 399–846), 8.25 μl of nuclease-treated rabbit reticulocyte lysate (RRL) (Promega), complete amino acid mixture and RNA guard as described in (37). After 30 min incubation at 37°C, translation reactions were assayed in triplicate for luciferase activity as described earlier.

### RNA probes synthesis, electrophoretic mobility shift assays and competitions

The SECIS probes were synthesized *in vitro* from linearized templates with T7 RNA polymerase using 1 mM GTP, 1 mM ATP, 1 mM CTP and 1 μM UTP and 10 μCi of $^{32}$P-labeled UTP for 3 h at 37°C. The transcription reactions were treated with DNase RQ1 for 15 min at 37°C and then phenol : chloroform extracted. The aqueous phase was passed through a Micro Bio-Spin column according to the manufacturer's instructions (Bio-Rad). The REMSA binding buffer includes 30 mM Tris, pH 7.8, 8% glycerol, 75 mM KCl, 1 mM EDTA, 0.25 μg/μl BSA and 0.20 μg/μl tRNA. The $^{32}$P-labeled SECIS RNAs (10 fmol) were incubated with the indicated amount of recombinant SBP2-RBD (amino acids 517–777), which was expressed in bacteria and purified as described in (36). After incubation at 37°C for 10 min, the complexes were resolved on 8% nondenaturing polyacrylamide gels as described in (21). The dried gels were exposed to autoradiographic screens, which were analyzed by a Typhoon 9400 scanner (GE Healthcare). Quantitation of radioactive species was performed with ImageQuant software (Molecular Dynamics).

## RESULTS

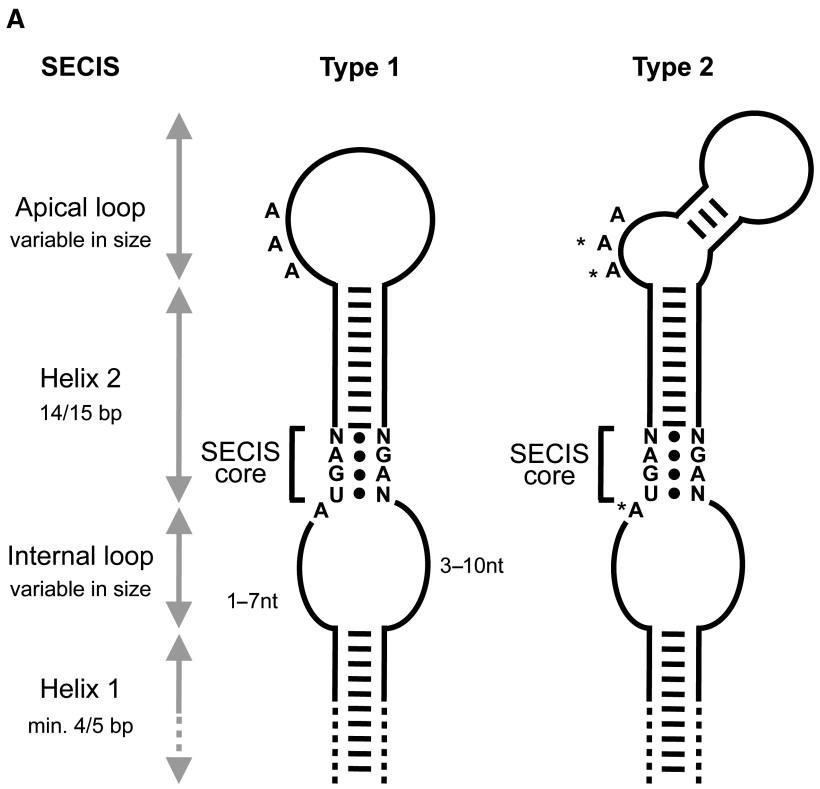
### The 26 SECIS elements found in the human genome were selected and cloned in a luciferase-based reporter construct

Previous studies that analyzed the efficiency of UGA recoding tested only SECIS elements from a subset of selenoprotein mRNAs or used SECIS elements of dissimilar lengths (23,38,39). The human selenoproteome is composed of 25 selenoproteins that contain a single selenocysteine residue, with the exception of selenoprotein P (SelP), which has 10 selenocysteines. For each selenoprotein, a SECIS element has been identified in the 3′-UTR. In the case of SelP, two functional SECIS have been characterized (thereafter referred as SelP1 and SelP2). It has been proposed that SelP2 recodes the initial UGA of SelP mRNA, while SelP1 is dedicated to the processive recoding of the downstream UGA triplets (40). To gain comprehensive insight into the intrinsic influence of the SECIS element on UGA recoding efficiency, we cloned the 26 human SECIS elements listed in Figure 1B. All the sequences were 103 nt in length, starting ~30 nt before the highly conserved AUGA motif, which should encompass the minimum active domain (24–28).

As illustrated in Figure 1A, the human SECIS elements all contain a four-motif secondary structure (stem-loop–stem-loop) despite the poor primary sequence conservation. To direct faithful UGA recoding, the SECIS requires the highly conserved nucleotides shown in the schematic Figure 1A and represented in bold in Figure 1B. These include the 5′-AUGA and 3′-GA at the base of helix 2 and the AAA/G motif (or CCX) in the apical region. The AUGA and GA motifs close the internal loop by forming the tandem GA pairs that resemble an RNA kink-turn motif. As shown in Figure 1A, the AAA/G (or CCX) motif is found either in a loop or in a bulge in the apical region of the SECIS element. This difference has been used to classify the SECIS elements as a Type 1 (loop) or Type 2 (bulge) SECIS, respectively (25,28,41). Another conserved feature is the length of helix 2 (14–15 base pairs) that ensures a precise distance between the AUGA and AAA/G (or CCX) motifs. The internal loop is closed by helix 1, which is variable in size and composition but important for UGA recoding (27).

### Human SECIS elements display a wide range of UGA recoding activities *in vivo* and *in vitro*

To analyze UGA recoding efficiency, the 26 human SECIS elements were cloned downstream of a luciferase coding sequence, which has been modified to contain an in frame UGA codon at position 258 (luciferase UGA$^{258}$-SECIS constructs). With this construct previously used in (21,33,36), an active luciferase is made only when the UGA is recoded as a selenocysteine codon. This reporter construct has been validated for selenocysteine incorporation in transfected cells and in an *in vitro* translation system. Plasmids containing the luciferase UGA$^{258}$-SECIS reporter constructs were transiently transfected in human kidney (Hek293) and liver (HepG2) cell lines. To analyze data from three independent transfections,




B

```
Gpx1   AGTGCTGCTGTCTCGGGGGGGTTTTCATCTATGAGGGTGTTTCCTCTAAACCTACGAGGG      AGGAACACCTTGATCTTACAGAAAATACCACCTCGAGATGGGT
Gpx2   AGACTTGGGTAAGCTCTGGGCCTTCACAGAATGATGGCACCTTCCTAAACCCTCAT          GGGTGGTGTCTGAGAGGCGTGAAGGGCCTGGAGCCACTCTGCTAGAA
Gpx3   CACGGACCCCATGGCAGGGGTGGCGTCTTCATGAGGGAGGGGCCCAAAGCCCTTGTGGGC      GGACCTCCCCTGAGCCTGTCTGAGGGGCCAGCCCTTAGTGCAT
Gpx4   CCACGCCCTCGGAGCCTTCCACCGGCACTCATGACGGCCTGCCTGCAAACCTG            CTGGTGGGGCAGACCCGAAAATCCAGCGTGCACCCCCGCCGGAGGAAGGTC
Gpx6   CCCACCTCACATGAAGGGAAGGGCATCTCCATGATGGTGGATCCCAAAACCCCTCTGGGTC      GCACCCTGCCAGAGCCTTCCTTGGTGCCTGTCCCTTAGTGCA
Trxr1  CATTTGGCAGGGCATCGAAGGGATGCATCCATGAAGTCACCAGTCTCAAGCCCATGTGGT      AGGCGGTGATGGAACAACTGTCAAATCAGTTTTAGCATGACCT
Trxr2  CACCCCCCCCCAGGCTCCTGGTGCCAGATGATGACGACCTGGGTGGAAACCTACCCTGTGGG    CACCCATGTCCGAGCCCCCTGGCATTTCTGCAATGCAAATA
Trxr3  CTGACAGCGAGAAGCAGTGGGACTGCTTCCTTGACGCCTTAGCTTGGAGCCCCGTTATGAGGT    GAGCCAAGGCTGACTCTCGCAAGCCAGGACTGAGCTTCCC
Dio1   TTTTAACTCTGTGTCTTTACATATTTGTTTATGATGGCCACAGCCTAAAGTACACAC         GGCTGTGACTTGATTCAAAAGAAAATGTTATAAGATGCAGTAAACTA
Dio2   CAGAGATGTGCAGAGTTGACCAGTGTGCGGATGATAACTACTGACGAAAGAGTCATCGACTCA    GTTAGTGGTTGGATGTAGTCACATTAGTTTGCCTCTC
Dio3   TGTTTTGGGTGCACAGGAGCCCCACTGCTGATGACGAACTATCTCTAACTGGTCTTGACAC     GAGCTAGTTCTGAATTGCAGGGGCCTCAAAGCAGCACCTAAA
SelH   CTTTGTGTCCCTGGTGATGTTGGAACATTAATGATGGAACATGGCCAAACTTCA            GTCATGATCCTGAAGCCATGGTTTCTTCCCTGCCAGAAATGAAGGTTCA
SelI   TATTTTTCACTGAATGAAGTTTGTGCTTGAATGAAGAGTGTATCTTAAACCCCCTTTTTTTGGACAGGCTGCACTTGGATAAAATAGGCACCACTGTGTTGAT
SelK   ACAAGGACTGCTCTGTGTCCTCACAGATGAATGAGGTCATGCTGGGAATTCCCTCTGCAGGGAA    CTGGCCTGACTGACATGCAGTTCCATAAATGCAGATGTT
SelN   AGGGGCAGTGGCTTCCCCGGCAGCAGCCCCATGATGGCTGAATCCGAAATCCTCGAT         GGGTCCAGCTTGATGTCTTTGCAGCTGCACCTATGGGAAGAAGTAG
SelO    TGCCCTGGCCCATGCACACCCGTCTTTCCATGATGGCAGAGACATCCAGTCAGGACCTGACC     CGTCTCTGTCTGAGGCCGGCTCAGCAGTGCAGCCTGGTCCC
SelP1  TTTTCTTTTTCCAGTGTTCTATTTGCTTTAATGAGAATAGAAACGTAAACTATGACCTAGG     GGTTTCTGTTGGATAATTAGCAGTTTAGAATGGAGGAAGAAC
SelP2  TATATTGCTTAGTAAGTATTTCCATAGTCAATGATGGTTTAATAGGTAAACCAAAC          CCTATAAACCTGACCTCCTTTATGGTTAATACTATTAAGCAAGAATG
SelM     GACCTACCTGCCTGAGTCCTGGAGACAGAATGAAGCGCTCAGCATCCCGGGAATACTTCTCT    TGCTGAGAGCCGATGCCCGTCCCCGGGCCAGCAGGGATGGG
SelS   TAGGACAGTCTCTGTGACAGGTTGCGTTGAATGATGTCTTCCTTATCAATGGTGAGCCCACCAG   TGAGGATTACTGATGTGGACAGTTGATGGGGTTTGTTTC
SelT   GATCATTGCAAGAGCAGCGTGACTGACATTATGAAGGCCTGTACTGAAGACAGCAAGCTGTT     AGTACAGACCAGATGCTTTCTTGGCAGGCTCGTTGTACCTC
SelV GCAAGGGGTGGAGCTGGAGGAGTCTCAGCTGGATGATGAGAAGGGCTGAAATGTTGCCAAGTCA    GGTCCTTTTCTGATGGTGGCTGGGGCTGGGGTGAGCT
SelW   CTTGACCCAGCCCCTCTCAGCAGACGCTTCATGATAGGAAGGACTGAAAAGTCTTGTGGACACCT  GGTCTTTCCCTGATGTTCTCGTGGCTGCTGTTGGGGGC
SelX   CCTGCCAGCCGCCCTGGCCCTGGTCACTGCATGATCCGCTCTGGTCAAACCCTTCCAGGCCA     GCCAGAGTGGGGATGGTCTGTGACCTGCTGGGAAGGCAGGC
Sel15  AGAGTGAAACATTCACAAAGATTTGCGTTAATGAAGACTACACAGAAAACCTTTCTAGGGATT   TGTGTGGATCAGATACATACTTGGCAAATTTTTGAGTTTT
Sps2   ACCTGCAACCATCTGACTTGGTCTCTGTTAATGACGTCTCTCCCTCTAAACCCCCATT         AAGGACTGGGAGAGGCAGAGCAAGCCTCAGAGCCCAGGCCTCAGTG
```

**Figure 1.** SECIS elements from human selenoproteome. (**A**) Schematic of eukaryotic SECIS elements. Human type 1 elements include Gpx1, Gpx2, SelN, Dio1 and SelV, whereas the other elements are type 2, containing an additional bulge in the apical loop. The four main structural motifs are identified by arrows. A bracket indicates the position of the SECIS core that contains a quartet of non-Watson–Crick base pairs. Highly conserved nucleotides are represented in letters. The dots symbolize the two sheared tandem GA base pairs that are essential for kink-turn structures. The asterisks indicate the variations from the consensus sequence (AA) that are present as CC in SelM, SelO and Trxr3 SECIS elements. (**B**) Sequence alignment of the 26 human SECIS elements that were cloned downstream of luciferase UGA$^{258}$ coding sequence and used in the reporter constructs. They were obtained from SelenoDB website (51). Highly conserved nucleotides are represented in bold. Nucleotides involved in helix 2 are underlined. Gpx, glutathione peroxidase; Trxr, thioredoxin reductase; Dio, iodothyronine deiodinase; Sel, selenoprotein; Sel15, 15 kDa selenoprotein, Sps2, selenophosphate synthetase.

we arbitrarily expressed the luciferase activities relative to the activity measured for SelX, which was expressed as 100%. The results obtained in Hek293 and HepG2 cell lines showed that, in contrast to previous studies, the nature of the SECIS element had a dramatic influence on the efficiency of UGA recoding *in vivo* (Figure 2A and B). We observed an ~2250-fold difference between the strongest (SelN) and the weakest (Gpx6) constructs in Hek293 cells (Figure 2A). The amplitude is even larger in HepG2 cells with an ~4900-fold difference between the reporters containing the Dio3 and SelS SECIS elements (Figure 2B). Moreover, in both cell lines, we

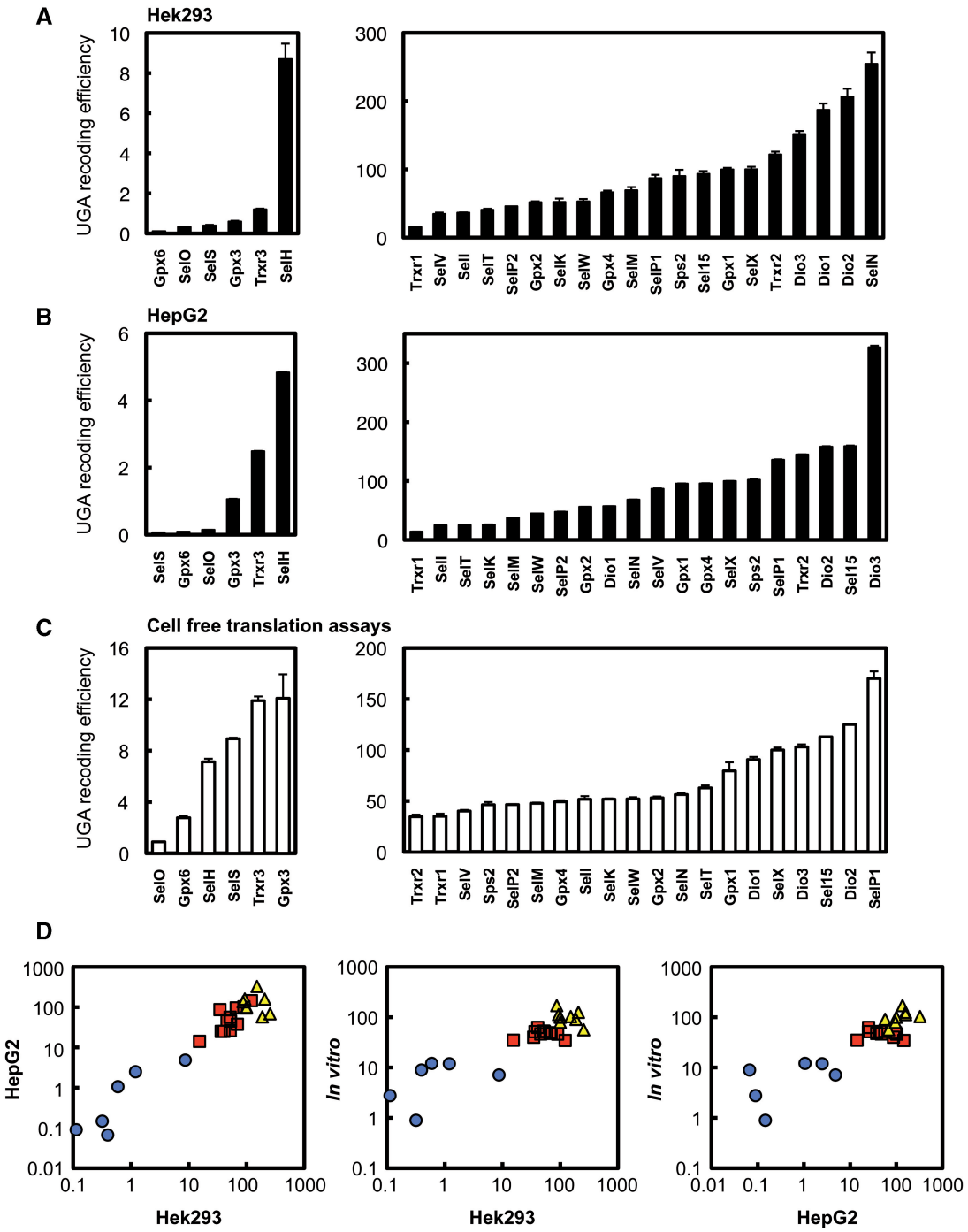



**Figure 2.** The UGA-selenocysteine recoding efficiency is strongly influenced by the nature of the SECIS element *in vivo* and *in vitro*. Luciferase UGA$^{258}$-SECIS reporter gene constructs were generated with all human SECIS elements and tested in transiently transfected Hek293 (**A**) or HepG2 (**B**) cells (black bars), or in the cell free translation system (**C**, white bars). For *in vivo* experiments (A and B), each luciferase plasmid was transiently co-transfected with β-galactosidase plasmid. Enzymatic activities were measured 48-h post-transfection on cell protein extracts. Transfection efficiencies were normalized by calculating the ratio between luciferase relative to β-galactosidase activities. For *in vitro* experiments (C), the luciferase UGA$^{258}$-SECIS synthetic mRNAs were translated using RRL as described in 'Materials and Methods' section. To analyze data from three independent experiments *in vivo* and *in vitro*, the UGA recoding efficiencies were arbitrarily expressed relative to the activity from the luciferase UGA$^{258}$-SelX construct, which was set as 100%. The group of constructs that show weak efficiencies (SelH, SelO, SelS, Gpx3, Gpx6 and Trxr3) are represented at a magnified scale in left panel. (**D**) Comparison of UGA–selenocysteine recoding efficiencies of our luciferase UGA$^{258}$-SECIS reporter gene constructs between the different experimental conditions: HepG2 versus Hek293 transfection (left panel), *in vitro* translation versus Hek293 tranfection (middle panel) and *in vitro* translation versus HepG2 tranfection (left panel). The SECIS elements are represented as a function of recoding efficiency group according to Figure 6: blue circles (weak), red square (moderate) and yellow triangles (strong).

identified a group of six weak SECIS elements that includes SelS, SelO, SelH, Gpx3, Gpx6 and Trxr3 (Figure 2A and B, left panels). The remaining 20 luciferase UGA$^{258}$-SECIS constructs were in the moderate and strong categories of UGA recoding efficiency, with a 16- and 30-fold range observed in Hek293 and HepG2 cells, respectively (Figure 2A and B, right panels). As shown in Figure 2, the difference between the highest of the weak elements (SelH) and the lowest of the moderate elements (Txr1) was 2- to 3-fold in both Hek293 and HepG2 cells. Importantly, the ranking between the constructs is very similar between the two cell lines used in this study as illustrated in Figure 2D (left panel) on a logarithmic scale. The only movements that occurred were within the moderate and strong categories, as observed for SelN and Dio1.

To verify that the different SECIS elements drive UGA recoding with varying strengths, capped luciferase reporter construct RNAs were translated in a cell-free *in vitro* system using RRL. Aliquots of the translation products were analyzed for luciferase activity, which was expressed relative to the activity from the reporter containing the SelX SECIS. Similar to the results in transfected cells, a wide variation in UGA recoding efficiencies was observed in the *in vitro* assay, although with a smaller range (Figure 2C). Indeed the strongest SECIS (SelP1) was ~200-fold more efficient in recoding UGA than the weakest (SelO). The lower amplitude of UGA recoding efficiencies observed in RRL was likely due to the unnatural non-competitive environment, the translation reactions being primed with a single mRNA species. In transfected cells, the luciferase reporter mRNA containing the SECIS element would compete with endogenous selenoprotein mRNAs for limiting components of selenocysteine insertion machinery, which is likely to magnify the differences observed *in vitro*. Strikingly, despite the lower amplitude, the same group of six SECIS elements that displayed weak UGA recoding activity in cells also gave low levels of luciferase activity in the *in vitro* translation assay. As for the transfected cells, an almost 3-fold difference was observed between Gpx3, the highest of the weak elements and Trxr2, the lowest of the moderate elements (Figure 2C). The moderate and strong elements were spread out over a 5-fold range (right panel of Figure 2C). The classification of SECIS elements into three groups (weak, moderate and strong) is best illustrated when the data are plotted on a logarithmic scale (Figure 2D). Our results suggest that the differences in UGA recoding activity resulted from intrinsic properties of the SECIS. Taken together, our data indicate that distinct properties of the different SECIS elements may modulate the efficiency of UGA recoding both *in vivo* and *in vitro*.

### Gpx3 and SelX as representative weak and strong SECIS elements, respectively

For the subsequent studies, we chose to focus on Gpx3 and SelX as representative weak and strong SECIS elements, respectively, since they share common features despite their differences in their ability to direct UGA recoding. In humans, Gpx3 and SelX mRNAs produce detectable amounts of selenoproteins, with Gpx3 being the most abundant selenoprotein in the blood, while SelX is found in many cell lines (6). Both elements are predicted to adopt a type 2 structure (Figure 3A), with major differences between the Gpx3 and SelX SECIS elements occurring in the size and composition of the internal loop. The other structural features of the two elements differ only in their nucleotide composition but are similar in size.

We first confirmed that the luciferase activity of the Gpx3 and SelX SECIS contructs was due to UGA recoding and not due to random stop codon readthrough. We generated two sets of mutants, the first one disabling essential motifs from the SECIS, either the tandem GA pairs (ΔAUGA mutant of luciferase UGA$^{258}$-SECIS constructs) or the AAA/G motif (ΔAAAC and ΔAAAG for SelX and Gpx3, respectively), the other, replacing the UGA at position 258 of the luciferase coding sequence by the UAA stop codon, which does not support selenocysteine incorporation. These mutant constructs were analyzed in transient transfection experiments and *in vitro* translation assays (Figure 3B). As previously shown in Figure 2, the SelX SECIS is 350- and 8-fold more efficient in UGA recoding than Gpx3 in transfected Hek293 cells and in the *in vitro* translation assays, respectively. In every case, deletion of the AUGA or the AAA/G motif led to a dramatic decrease in luciferase activity (Figure 3B). Furthermore, the luciferase UAA$^{258}$-SECIS mutants showed a complete loss of luciferase activity, thus confirming that our *in vitro* and *in vivo* assays were specific for selenocysteine insertion even in the case of the Gpx3 construct that contains a very weak SECIS element. We have tested mutants in which UGA$^{258}$ was changed to a UGU cysteine codon. As shown in Figure 3B, the luciferase UGU$^{258}$-SECIS constructs gave similar activities for both the Gpx3 or SelX elements in transfected Hek293 cells and the *in vitro* translation assay (Figure 3B). This result excluded the possibility that the weak Gpx3 SECIS interfered with normal protein synthesis and also indicated that the Gpx3 SECIS construct did not contain any mRNA instability elements.

### SBP2 binds the Gpx3 and SelX SECIS elements with similar affinity

The simplest explanation for the wide range of UGA recoding efficiencies would be that various SECIS elements are bound with different affinities by SBP2, which is a limiting factor for selenocysteine incorporation (20–23). To test this hypothesis, we performed REMSA analyses using radiolabeled Gpx3 and SelX SECIS elements (Figure 4A and B, respectively). As a source of protein, we used increasing amounts of the purified RNA-binding domain of SBP2 (SBP2-RBD, amino acids 517–777), which has been shown to exhibit selective SECIS binding activity (36). The SBP2-SECIS complexes were separated from free probes on native acrylamide gels and quantified to determine apparent $K_d$ values. As shown in Figure 4, the SBP2-RBD bound to the SelX and Gpx3 SECIS elements with similar affinities, with $K_d$ values of

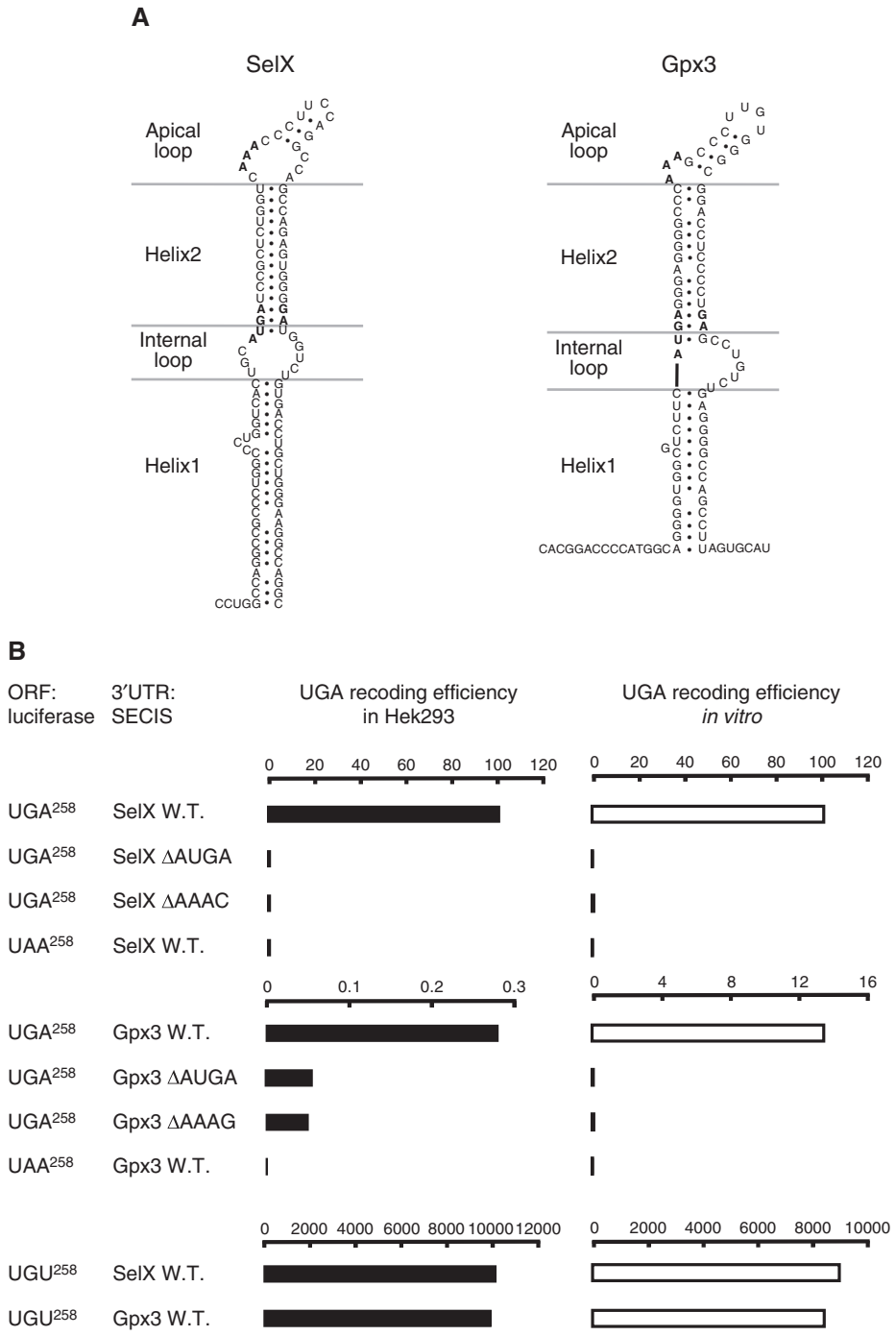



**Figure 3.** The UGA recoding efficiencies of the SelX and Gpx3 SECIS elements are specific for selenocysteine insertion *in vivo* and *in vitro*. (**A**) The secondary structures of SelX and Gpx3 SECIS elements are represented in the frame of type 2 category of elements. Consensus motifs are shown in bold. The A–U, G–C, G–U and noncanonical G–A base pairs are represented by dots. The main four structural motifs are separated by gray bars. (**B**) Mutant luciferase-SelX and luciferase-Gpx3 constructs were assayed for UGA recoding activity in transiently transfected Hek293 cells (black bars) or in the rabbit reticulocyte cell-free translation system (white bars). Mutations targeted either essential motifs the SECIS elements (ΔAUGA, ΔAAAC and ΔAAAG) or codon 258 in the luciferase coding sequence ($UAA^{258}$ and $UGU^{258}$). UGA recoding activities were expressed relative to the activity of the wild-type luciferase $UGA^{258}$-SelX construct (set as 100%).

8 nM and 14 nM, respectively. These very similar affinities for SBP2 cannot explain the difference in UGA recoding between Gpx3 and SelX SECIS. Thus, we hypothesized that other determinants in the SECIS may modulate the efficiency of selenocysteine incorporation.

## The internal loop and helix 2 of the SECIS determine UGA recoding efficiency

To define the motifs in the SECIS that modulate the efficiency of UGA recoding, we performed structure–function

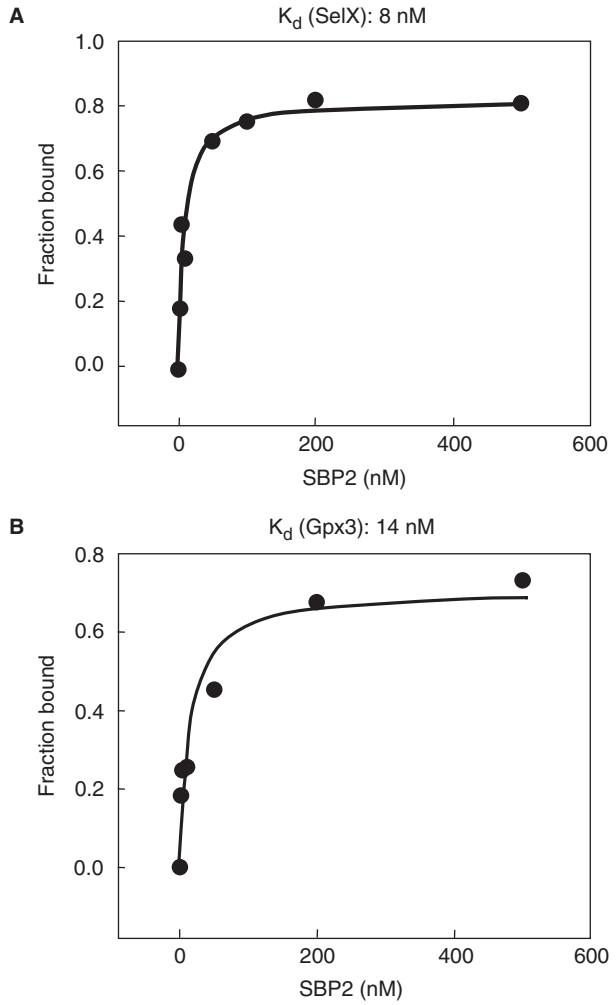



**Figure 4.** SBP2 binds with similar affinity to the SelX (**A**) and Gpx3 (**B**) SECIS elements. The $^{32}$P-labeled SECIS elements were incubated with increasing concentration of the RNA binding domain of SBP2 (SBP2-RBD). The RNA–protein complexes were analyzed on a native 8% polyacrylamide gel. The apparent $K_d$ values were determined by plotting the fraction of protein–SECIS complex as a function of protein concentration.

analyses of the SelX and Gpx3 SECIS elements using a chimera approach (see Supplementary Figure S3 for sequences, structures and nomenclature). Using this strategy, we were able to swap individual domains in the SelX SECIS with the corresponding regions from Gpx3 (Figure 5A and B); convert the weak Gpx3 SECIS into a stronger element (Figure 5C and D); and systematically transform the strong SelX SECIS into a weak one (Figure 5E and F). The chimeric SECIS elements were cloned downstream of luciferase UGA$^{258}$ coding region. The constructs were tested in transiently transfected Hek293 cells (Figure 5, left panels) and in the *in vitro* translation assay (Figure 5, right panels).

The first set of experiments (Figure 5A and B) consisted of replacing domains in the SelX SECIS (shown in black) with their Gpx3 counterparts (shown in gray). A change in luciferase activity should be detected when the domains contained important features for UGA recoding efficiency. Replacement of helix 1 in SelX with the Gpx3 counterpart (chimera 1) had very little effect both *in vivo* and *in vitro* (Figure 5A and B). Thus, helix 1 may only be important for closing the bottom of the internal loop (see the secondary structure of the SECIS in Figure 3A) as previously suggested. The most dramatic effects were observed with the individual swapping of the internal loop (chimera 7), helix 2 (chimera 12) or both domains (chimera 10), which decreased UGA recoding efficiency >300-fold, similar to the level observed for the Gpx3 parental construct. Our data also clearly indicate that helix 2, which was previously expected to only physically separate the AUGA from the AAA/G, harbors determinant(s) for selenocysteine insertion efficiency. Although the length of helix 2 is similar in the two SECIS elements, this region is completely base paired in the SelX SECIS while it contains two A–C mismatches in Gpx3 that can form non-WC base pairs. Finally, we analyzed the importance of the essential AAA/G motif, which is located in a bulge in the apical loop region of the SelX and Gpx3 SECIS elements. In contrast to what might be expected for a domain containing an essential motif, the insertion of a weak apical loop into the strong SECIS context (chimera 4) led to only a modest 2- to 5-fold decrease in luciferase activity. Thus, this region contains minor determinants that regulate UGA recoding efficiency.

In a second set of experiments, when the motifs from the strong element were inserted in the weak SECIS context, very few effects were observed (Figure 5C and D). Even the swapping of the internal loop induced little, if any increase in recoding efficiency (chimera 8). These results suggested that the determinants for the weak efficiency of Gpx3 are dominant over the strong determinants in the SelX SECIS. However, the simultaneous insertion of the SelX internal loop and helix 2 into the Gpx3 SECIS (chimera 9) resulted in almost full recovery of UGA recoding activity in cells (Figure 5C) and *in vitro* (Figure 5D). Taken together, our data indicate that the internal loop and helix 2, which both flank the tandem GA pairs, are necessary and sufficient to switch a weak SECIS element into a strong one, and vice versa. To verify this finding, we gradually transformed the strong SelX SECIS into a weak element from bottom to top and top to bottom. As shown in Figure 5E and F, the largest effects were observed when changes occurred in the internal loop and the helix 2 (chimeras 2 and 3). It is also clear from these results that negative determinants won over the positive determinants, as previously observed with the swapping of individual domains.

### Identification of nucleotide sequences that predict weak UGA recoding activity

To identify nucleotide sequences that may modulate UGA recoding efficiency, we first analyzed the sequence conservation of the internal loop and helix 2 in the 26 human SECIS elements using Weblogo (42) (Figure 6). In addition to the previously identified highly conserved motifs (i.e. ATGA, AAA/G and GA), we noticed that other positions with lower conservation were present near the tandem GA pairs in the SECIS core. More precisely, a

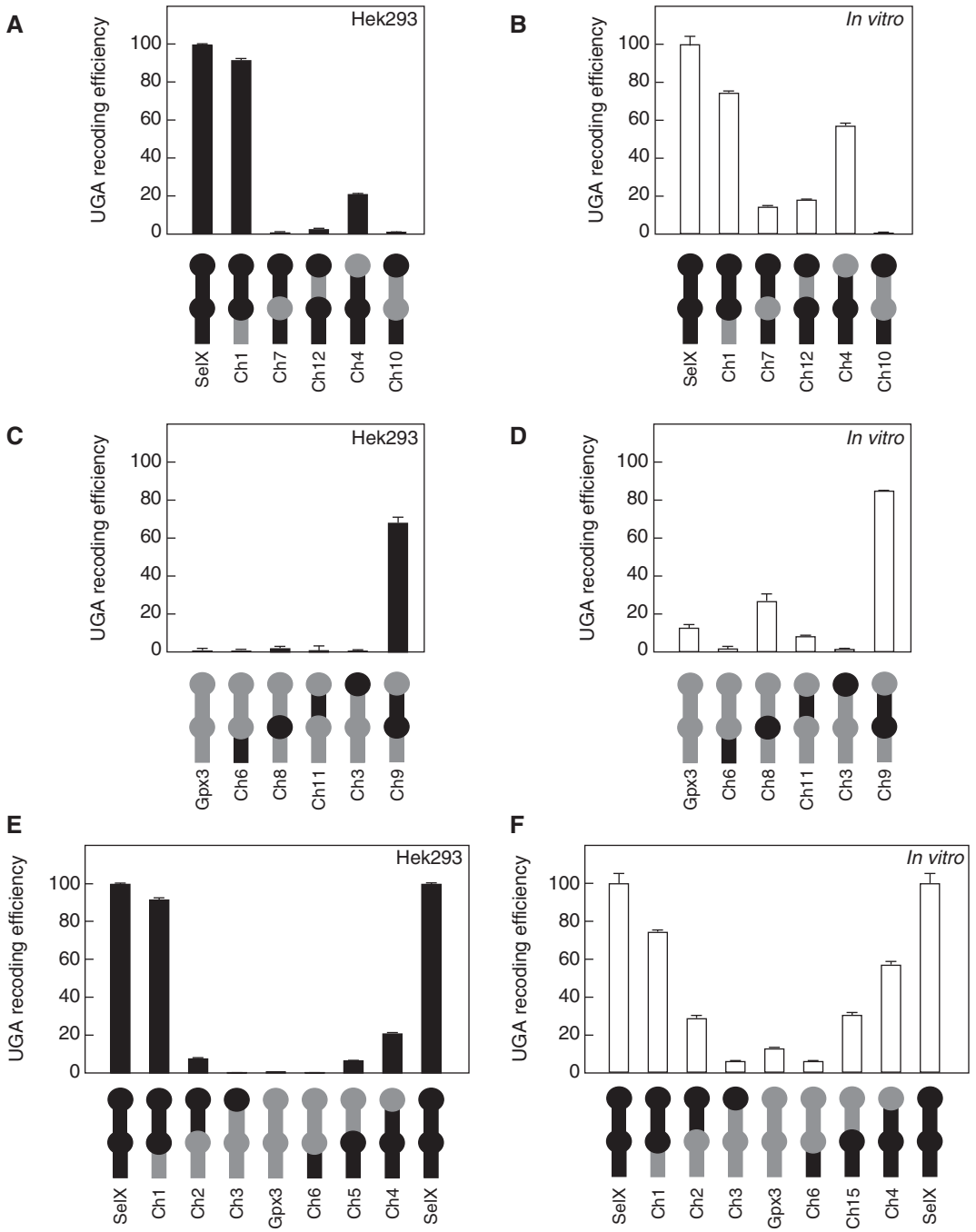



**Figure 5.** Structure-function analyses of SelX and Gpx3 SECIS elements. SelX and Gpx3 were used as a representative weak and strong element, respectively. Chimeric constructs of these elements were generated based on the four domains shown in Figure 3A and cloned downstream of luciferase UGA$^{258}$ coding sequence. The composition and nomenclature of the chimeric SECIS elements are represented beneath histograms, using gray and black to indicate domains originating from Gpx3 and SelX SECIS elements, respectively. The histograms represent data from transiently transfected Hek293 cells (black bars) or from cell free translation assays (white bars), which were performed as described in the Figure 2 caption. The UGA recoding efficiencies are expressed relative to the luciferase UGA$^{258}$-SelX construct (set as 100%). (**A** and **B**) Domain swapping of Gpx3 SECIS domains into the strong SelX element context. (**C** and **D**) Domain swapping of SelX SECIS domains into the weak Gpx3 context. (**E** and **F**) Gradual transformation of a strong SECIS element into a weak one, and vice versa.

potential GC base pair in the lower half of helix 2 is present in 62% of the human SECIS elements (illustrated by a sharp in Figure 6B and indicated as position +2 in Figure 6C). Interestingly, this GC base pair was found at the same distance from the tandem GA pairs in many of the synthetic RNAs selected for *in vitro* binding to SBP2 using SELEX methodology (43). Thus, the GC is an important but not essential RNA feature for SBP2 recognition. Furthermore, the genomic sequences of 65% of human SECIS elements contain a T, 3 nt before the ATGA on the 5′-side of the internal loop (illustrated by an asterisk in Figure 6B and indicated as position –7

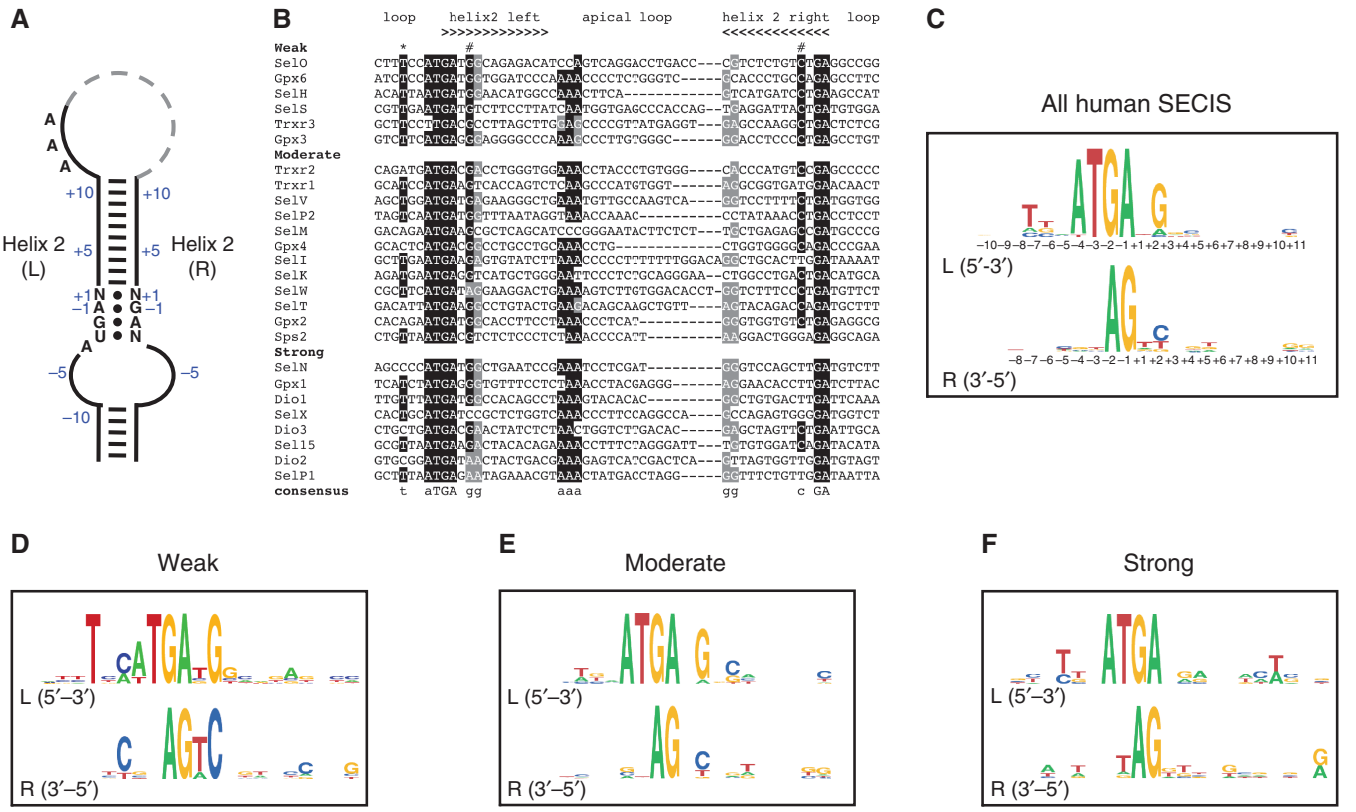



**Figure 6.** Sequence alignment of the regions surrounding the kink-turn motif in human SECIS elements as a function of UGA recoding efficiency. (**A**) Schematic representation of the consensus SECIS element structure. The left (L) and right (R) sides of helix 2 are numbered in blue relative to the AUGA and GA nucleotides, respectively. (**B**) DNA sequences that were used for sequence analyses started 6 nt before the ATGA (position-10) and ended 6 nt after the GA motif (position-8). The classification SECIS elements were classified having as weak, moderate and strong UGA recoding activity based on *in vitro* experiments (Figure 2C). Highly and moderately conserved nucleotides are boxed in black or gray, respectively. Consensus nucleotides are represented underneath the sequence alignment. Additional conserved nucleotides are indicated by a sharp and an asterisk, for the GC base pair and the T, respectively. Sequence logos were generated using Web logo software (42) with all human SECIS (**C**) or the group of weak (**D**), moderate (**E**) or strong (**F**) elements. The left part of helix 2 is represented in 5′- to 3′-orientation above the right counterpart represented in 3′- to 5′-orientation to highlight possible consensus base pairing. The overall height of the stack indicates the sequence conservation at each position, while the height of symbols within the stack indicates the relative frequency of each nucleic acid at that position.

of the left helix in Figure 6C). Both the GC and the T nucleotides have also been pointed out as being conserved features in a recent report, which aligned 286 SECIS elements from various eukaryote species (29).

We then investigated whether these conserved motifs are functionally relevant by generating the sequence logos separately with the weak, moderate or strong categories of elements (Figure 6D–F). Interestingly, we observed that the GC and T motifs are over-represented in weak SECIS elements versus the strong ones. As shown in Figure 6B, the simultaneous presence of the T and the GC was found in 100% of weak SECIS elements, compared to only 25 and 12.5% for the moderate and strong classes, respectively. The association of both motifs was a statistically significant predictor of the weak versus the strong/moderate class of SECIS element (Fisher test, $P < 0.002$).

## DISCUSSION

The SECIS element, which is essential for selenocysteine insertion, is found in all selenoprotein mRNAs that comprise the mammalian selenoproteome. However, the primary sequence conservation is poor and the four-motif secondary structure of the SECIS is subject to variability, notably in the internal and apical loops as illustrated in Figures 1 and 6. While previous studies reported significant differences among SECIS elements in their ability to drive UGA recoding, these analyses tested only specific subsets of selenoprotein mRNAs, i.e. Gpx1, Gpx4, Trxr1, Dio1 and SelP1 in (23) or Gpx1, Gpx2 and Gpx4 in (38,39) and did not identify the determinants in the SECIS that were responsible for this variability. In this work, we report the first comprehensive analysis of the 26 SECIS elements encoded in the human genome. We found that human SECIS elements displayed a wide range of UGA recoding activities *in vivo*, spanning several 1000-fold in transfected Hek293 and HepG2 cells and several 100-fold *in vitro* (Figure 2), which is much broader than that has been previously observed. Surprisingly, this variability was likely not due to variation in SBP2 binding activity, an essential and limiting factor for selenocysteine incorporation, since we found that the weak Gpx3 and strong SelX SECIS elements were bound with similar

affinity by SBP2. The consistency of the results is striking as the SECIS elements were classified almost identically as weak, moderate or strong by all three experimental approaches (Figure 2D). Taken together, our results strongly support the hypothesis that the SECIS element contains intrinsic information that modulates selenocysteine incorporation efficiency in an SBP2-independent manner.

Previous mutagenesis studies have focused on identifying the nucleotides or structures in an individual SECIS element that are essential for selenocysteine incorporation. To pinpoint the RNA structural domain(s) that define the UGA recoding efficiency of a SECIS, we generated chimeric molecules using SelX and Gpx3 as a representative strong and weak element, respectively. Our data indicate that helix 1 and the apical loop are the minor determinants for UGA recoding efficiency (Figure 5) although those regions are essential for the insertion of selenocysteine (24–28,44). Instead we found that the internal loop, which varies in composition and size among SECIS elements, was one important determinant in Gpx3 and SelX, respectively. Although both SECIS elements contain internal loops of comparable size (8 and 9 nt), they differ in their predicted structures, with an almost symmetric (4 + 5) loop in SelX and an asymmetric (1 + 7) loop in Gpx3. In our constructs, this domain was necessary and sufficient to transform a strong SECIS into a weak one. The converse experiment of inserting the SelX internal loop into the Gpx3 context did not rescue the weak element, indicating that more than one region is involved. Indeed, we found that helix 2, which was initially thought to act as a spacer to ensure a precise distance between the tandem GA pairs and the AAA/G (or CCX) motif, also modulated UGA recoding activity. Interestingly, the helices are of similar length (14 nt) in these two SECIS elements. However, the SelX helix 2 is fully base paired, whereas Gpx3 is predicted to contain two C–A mismatches in this region (Figure 3A). Finally, we showed that these two RNA structural domains, i.e. the internal loop and helix 2, together are necessary and sufficient to regulate the efficiency of UGA recoding *in vivo* and *in vitro* (Figure 5, chimeras 9 and 10).

We also found that the simultaneous presence of a GC base pair in helix 2 and a T 3′-nucleotides before the ATGA is a statistically significant predictor of weak UGA recoding activity. Although the two motifs may not be sufficient to transform a weak into a strong element or vice versa, they may contribute to selenocysteine incorporation efficiency by influencing the function of the adjacent tandem GA pairs, which are essential for binding SBP2 and L30 (19–21,45). We previously proposed that the tandem GA pairs in the SECIS core may represent a non-canonical kink-turn motif. The conserved GC and T nucleotides could also influence the angle or dynamics of the kink-turn folding. Precise structural analyses will be needed to understand the implications of the context of the tandem GA pairs on the functional activity of the SECIS.

Although SBP2 has been proposed to dictate the expression of the selenoproteome, the affinity of SBP2 for the Gpx3 and SelX SECIS elements differed by only 2-fold despite their 350-fold difference in directing UGA recoding in transfected cells. Thus, a wide range of UGA recoding efficiencies is not solely explained by SBP2 binding affinity. This observation is rather surprising as the SBP2 binding site encompasses helix 2 and the internal loop (45), the regions we identified as the major determinant of the selenocysteine insertion efficiency. Several studies have found that SBP2 exhibits selective SECIS binding activity. *In vitro* analyses using REMSA have found that SBP2 exhibits selective high affinity binding activity to Gpx4 and Trxr1 SECIS (36) while *in vivo* binding of selenoprotein mRNAs by SBP2 via immunoprecipitation of the proteins and quantitation of bound mRNAs have shown that SBP2 exhibits strong preferential binding to SelW, Sel15, Gpx4, SelH, Trxr1, Dio2 and SelK mRNAs over others (22). Indeed the difference in SBP2 binding affinity does explain the preferential synthesis of Gpx4 over Gpx1 and Dio2 in a competitive *in vitro* translation assay (36). However, the regulation of the selenoproteome appears to be more complex as there is no obvious correlation between the reported binding properties of SBP2 and the UGA recoding efficiencies that we observed, suggesting the involvement of additional *cis*- or *trans*-acting factors.

The most surprising finding from our study was the identification of a small group of SECIS elements, including Gpx3, Gpx6, SelH, SelO, SelS and Trxr3, that directed UGA recoding with extremely low efficiency both *in vivo* and *in vitro*. For Gpx3, we used mutations in either the apical loop or the tandem GA pairs of the SECIS elements to demonstrate that the low level of luciferase activity that we detected was due to UGA recoding as selenocysteine and not to a low level of translational readthrough as observed for a genuine stop codon (Figure 3). Metabolic labeling studies demonstrated that three other weak SECIS elements (SelH, SelO and SelS) directed selenocysteine incorporation in a reporter construct in transfected cells although no quantitative analysis was reported in this study (5). These results raise the question as to how selenoproteins with an inefficient SECIS element are made in significant amounts *in vivo*. One possibility is that the abundance of a selenoprotein mRNA may compensate for the weak efficiency of its SECIS element. Interestingly, Gpx3 is mainly produced in the kidney and the Gpx3 mRNA is ranked as the most abundant selenoprotein mRNA in this tissue in mice (46). However, the other selenoprotein mRNAs with weak SECIS elements are expressed at low or moderate levels in mouse tissues (46).

The more likely explanation for our results is that the selenocysteine incorporation efficiency of a weak SECIS element may be enhanced by other *cis*-acting elements within the transcript. Several selenoprotein mRNAs contain RNA structures in the coding region that modulate UGA recoding efficiency (34,35). However, none of these structures are present in the selenoproteins mRNAs that we identified as containing weak SECIS elements. A largely unexplored area in the field is whether the activity of a SECIS element is influenced by other sequences in the 3′-UTR of the transcript. Previous studies have analyzed minimal SECIS elements of 50–100 nt, which are sufficient

to support UGA recoding *in vitro* and in cells (19,41,47). We designed our SECIS elements to encompass ~100 nt, which is the minimal sequence required for binding of SBP2 and L30 to the Gpx4 SECIS (21,45). For some selenoprotein mRNAs, additional sequences flanking the SECIS may be required for efficient selenocysteine incorporation. Indeed, a very recent study analyzing the expression of human Gpx3 in transfected cells reported that selenocysteine incorporation was supported by a 3′-UTR of 500 nt, but not by a 100 nt 3′-UTR that encompassed the SECIS (48). While the underlying basis for this difference was not investigated, we speculate the additional sequences may be required to stabilize the stem-loop structure of the SECIS or to recruit other factors that may enhance the efficiency of UGA recoding. Finally, one must consider the possibility that the activity of a weak SECIS element may be enhanced *in vivo* by cell-type specific or selenoprotein-specific *trans*-acting factors that are not present in our experimental systems. These may be proteins, co-factors, subcellular localization of proteins and mRNAs, or differing oxidation states of UGA recoding components [as observed for SBP2 in (49)]. Indeed, SBP2 interact with proteins involved in the assembly of ribonucleoprotein complexes, including the R2TP complex and Nufip, which could influence selenocysteine incorporation *in vivo* (50). Such differences may also account for the wider range of UGA recoding activities that we observed *in vivo*, compared with *in vitro*. Regardless of the explanation, our results clearly show that the 26 human SECIS elements contain intrinsic information that modulates their ability to direct selenocysteine incorporation. Our findings have implications not only for designing experiments to analyze SECIS function but also for understanding how the expression of the selenoproteome is regulated *in vivo*.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We are grateful to Claude Thermes and Yves d'Aubenton-Carafa for helpful discussions on weblogo sequence interpretation and statistical analyses and to Jodi Bubenik for providing the SBP2-RBD plasmid and for critical reading of the manuscript.

## FUNDING

CNRS (ATIP program to L.C.); the Fondation pour la Recherche Médicale (to L.C.); the Ligue Contre le Cancer (Comité de l'Essonne, to L.C.); the Association pour la Recherche sur le Cancer (grants 4849 to L.C. and 4891 to O.J.-J.); and the National Institutes of Health (grant HL29582 to D.M.D.); fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche (to L.L.). Funding for open access charge: CNRS.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Rayman,M.P. (2005) Selenium in cancer prevention: a review of the evidence and mechanism of action. *Proc. Nutr. Soc.*, **64**, 527–542.
2. Patrick,L. (2004) Selenium biochemistry and cancer: a review of the literature. *Altern. Med. Rev.*, **9**, 239–258.
3. Birringer,M., Pilawa,S. and Flohe,L. (2002) Trends in selenium biochemistry. *Nat. Prod. Rep.*, **19**, 693–718.
4. Rayman,M.P. (2000) The importance of selenium to human health. *Lancet*, **356**, 233–241.
5. Kryukov,G.V., Castellano,S., Novoselov,S.V., Lobanov,A.V., Zehtab,O., Guigo,R. and Gladyshev,V.N. (2003) Characterization of mammalian selenoproteomes. *Science*, **300**, 1439–1443.
6. Fomenko,D.E., Novoselov,S.V., Natarajan,S.K., Lee,B.C., Koc,A., Carlson,B.A., Lee,T.H., Kim,H.Y., Hatfield,D.L. and Gladyshev,V.N. (2009) MsrB1 (methionine-R-sulfoxide reductase 1) knock-out mice: roles of MsrB1 in redox regulation and identification of a novel selenoprotein form. *J. Biol. Chem.*, **284**, 5986–5993.
7. Kim,H.Y., Fomenko,D.E., Yoon,Y.E. and Gladyshev,V.N. (2006) Catalytic advantages provided by selenocysteine in methionine-S-sulfoxide reductases. *Biochemistry*, **45**, 13697–13704.
8. Rocher,C., Lalanne,J.L. and Chaudiere,J. (1992) Purification and properties of a recombinant sulfur analog of murine selenium–glutathione peroxidase. *Eur. J. Biochem.*, **205**, 955–960.
9. Berry,M.J., Maia,A.L., Kieffer,J.D., Harney,J.W. and Larsen,P.R. (1992) Substitution of cysteine for selenocysteine in type I iodothyronine deiodinase reduces the catalytic efficiency of the protein but enhances its translation. *Endocrinology*, **131**, 1848–1852.
10. Berry,M.J., Tujebajeva,R.M., Copeland,P.R., Xu,X.M., Carlson,B.A., Martin,G.W. III, Low,S.C., Mansell,J.B., Grundner-Culemann,E., Harney,J.W. *et al.* (2001) Selenocysteine incorporation directed from the 3′UTR: characterization of eukaryotic EFsec and mechanistic implications. *Biofactors*, **14**, 17–24.
11. Driscoll,D.M. and Copeland,P.R. (2003) Mechanism and regulation of selenoprotein synthesis. *Annu. Rev. Nutr.*, **23**, 17–40.
12. Hatfield,D.L., Carlson,B.A., Xu,X.M., Mix,H. and Gladyshev,V.N. (2006) Selenocysteine incorporation machinery and the role of selenoproteins in development and health. *Prog. Nucleic Acid Res. Mol. Biol.*, **81**, 97–142.
13. Papp,L.V., Lu,J., Holmgren,A. and Khanna,K.K. (2007) From selenium to selenoproteins: synthesis, identity, and their role in human health. *Antioxid. Redox Signal.*, **9**, 775–806.
14. Allmang,C., Wurth,L. and Krol,A. (2009) The selenium to selenoprotein pathway in eukaryotes: more molecular partners than anticipated. *Biochim. Biophys. Acta.* (in press).
15. Hatfield,D., Lee,B.J., Hampton,L. and Diamond,A.M. (1991) Selenium induces changes in the selenocysteine tRNA[Ser]Sec population in mammalian cells. *Nucleic Acids Res.*, **19**, 939–943.
16. Lee,B.J., Worland,P.J., Davis,J.N., Stadtman,T.C. and Hatfield,D.L. (1989) Identification of a selenocysteyl-tRNA(Ser) in mammalian cells that recognizes the nonsense codon, UGA. *J. Biol. Chem.*, **264**, 9724–9727.
17. Fagegaltier,D., Hubert,N., Yamada,K., Mizutani,T., Carbon,P. and Krol,A. (2000) Characterization of mSelB, a novel mammalian elongation factor for selenoprotein translation. *EMBO J.*, **19**, 4796–4805.
18. Tujebajeva,R.M., Copeland,P.R., Xu,X.M., Carlson,B.A., Harney,J.W., Driscoll,D.M., Hatfield,D.L. and Berry,M.J. (2000) Decoding apparatus for eukaryotic selenocysteine insertion. *EMBO Rep.*, **1**, 158–163.
19. Copeland,P.R. and Driscoll,D.M. (1999) Purification, redox sensitivity, and RNA binding properties of SECIS-binding protein 2, a protein involved in selenoprotein biosynthesis. *J. Biol. Chem.*, **274**, 25447–25454.
20. Copeland,P.R., Fletcher,J.E., Carlson,B.A., Hatfield,D.L. and Driscoll,D.M. (2000) A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J.*, **19**, 306–314.
21. Chavatte,L., Brown,B.A. and Driscoll,D.M. (2005) Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes. *Nat. Struct. Mol. Biol.*, **12**, 408–416.

22. Squires,J.E., Stoytchev,I., Forry,E.P. and Berry,M.J. (2007) SBP2 binding affinity is a major determinant in differential selenoprotein mRNA translation and sensitivity to nonsense-mediated decay. *Mol. Cell Biol.*, **27**, 7848–7855.
23. Low,S.C., Grundner-Culemann,E., Harney,J.W. and Berry,M.J. (2000) SECIS-SBP2 interactions dictate selenocysteine incorporation efficiency and selenoprotein hierarchy. *EMBO J.*, **19**, 6882–6890.
24. Berry,M.J., Banu,L., Chen,Y.Y., Mandel,S.J., Kieffer,J.D., Harney,J.W. and Larsen,P.R. (1991) Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3′ untranslated region. *Nature*, **353**, 273–276.
25. Fagegaltier,D., Lescure,A., Walczak,R., Carbon,P. and Krol,A. (2000) Structural analysis of new local features in SECIS RNA hairpins. *Nucleic Acids Res.*, **28**, 2679–2689.
26. Martin,G.W. III, Harney,J.W. and Berry,M.J. (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.
27. Martin,G.W. III, Harney,J.W. and Berry,M.J. (1998) Functionality of mutations at conserved nucleotides in eukaryotic SECIS elements is determined by the identity of a single nonconserved nucleotide. *RNA*, **4**, 65–73.
28. Walczak,R., Westhof,E., Carbon,P. and Krol,A. (1996) A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs. *RNA*, **2**, 367–379.
29. Chapple,C.E., Guigo,R. and Krol,A. (2009) SECISaln, a web-based tool for the creation of structure-based alignments of eukaryotic SECIS elements. *Bioinformatics*, **25**, 674–675.
30. Goody,T.A., Melcher,S.E., Norman,D.G. and Lilley,D.M. (2004) The kink-turn motif in RNA is dimorphic, and metal ion-dependent. *RNA*, **10**, 254–264.
31. Matsumura,S., Ikawa,Y. and Inoue,T. (2003) Biochemical characterization of the kink-turn RNA motif. *Nucleic Acids Res.*, **31**, 5544–5551.
32. Gupta,M. and Copeland,P.R. (2007) Functional analysis of the interplay between translation termination, selenocysteine codon context, and selenocysteine insertion sequence-binding protein 2. *J. Biol. Chem.*, **282**, 36797–36807.
33. Mehta,A., Rebsch,C.M., Kinzy,S.A., Fletcher,J.E. and Copeland,P.R. (2004) Efficiency of mammalian selenocysteine incorporation. *J. Biol. Chem.*, **279**, 37852–37859.
34. Howard,M.T., Moyle,M.W., Aggarwal,G., Carlson,B.A. and Anderson,C.B. (2007) A recoding element that stimulates decoding of UGA codons by Sec tRNA[Ser]Sec. *RNA*, **13**, 912–920.
35. Howard,M.T., Aggarwal,G., Anderson,C.B., Khatri,S., Flanigan,K.M. and Atkins,J.F. (2005) Recoding elements located adjacent to a subset of eukaryal selenocysteine-specifying UGA codons. *EMBO J.*, **24**, 1596–1607.
36. Bubenik,J.L. and Driscoll,D.M. (2007) Altered RNA binding activity underlies abnormal thyroid hormone metabolism linked to a mutation in selenocysteine insertion sequence-binding protein 2. *J. Biol. Chem.*, **282**, 34653–34662.
37. Copeland,P.R. and Driscoll,D.M. (2002) Purification and analysis of selenocysteine insertion sequence-binding protein 2. *Methods Enzymol.*, **347**, 40–49.
38. Muller,C., Wingler,K. and Brigelius-Flohe,R. (2003) 3′UTRs of glutathione peroxidases differentially affect selenium-dependent mRNA stability and selenocysteine incorporation efficiency. *Biol. Chem.*, **384**, 11–18.
39. Wingler,K., Bocher,M., Flohe,L., Kollmus,H. and Brigelius-Flohe,R. (1999) mRNA stability and selenocysteine insertion sequence efficiency rank gastrointestinal glutathione peroxidase high in the hierarchy of selenoproteins. *Eur. J. Biochem.*, **259**, 149–157.
40. Stoytcheva,Z., Tujebajeva,R.M., Harney,J.W. and Berry,M.J. (2006) Efficient incorporation of multiple selenocysteines involves an inefficient decoding step serving as a potential translational checkpoint and ribosome bottleneck. *Mol. Cell Biol.*, **26**, 9177–9184.
41. Grundner-Culemann,E., Martin,G.W. III, Harney,J.W. and Berry,M.J. (1999) Two distinct SECIS structures capable of directing selenocysteine incorporation in eukaryotes. *RNA*, **5**, 625–635.
42. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
43. Clery,A., Bourguignon-Igel,V., Allmang,C., Krol,A. and Branlant,C. (2007) An improved definition of the RNA-binding specificity of SECIS-binding protein 2, an essential component of the selenocysteine incorporation machinery. *Nucleic Acids Res.*, **35**, 1868–1884.
44. Lesoon,A., Mehta,A., Singh,R., Chisolm,G.M. and Driscoll,D.M. (1997) An RNA-binding protein recognizes a mammalian selenocysteine insertion sequence element required for cotranslational incorporation of selenocysteine. *Mol. Cell Biol.*, **17**, 1977–1985.
45. Fletcher,J.E., Copeland,P.R., Driscoll,D.M. and Krol,A. (2001) The selenocysteine incorporation machinery: interactions between the SECIS RNA and the SECIS-binding protein SBP2. *RNA*, **7**, 1442–1453.
46. Hoffmann,P.R., Hoge,S.C., Li,P.A., Hoffmann,F.W., Hashimoto,A.C. and Berry,M.J. (2007) The selenoproteome exhibits widely varying, tissue-specific dependence on selenoprotein P for selenium supply. *Nucleic Acids Res.*, **35**, 3963–3973.
47. Shen,Q., McQuilkin,P.A. and Newburger,P.E. (1995) RNA-binding proteins that specifically recognize the selenocysteine insertion sequence of human cellular glutathione peroxidase mRNA. *J. Biol. Chem.*, **270**, 30448–30452.
48. Ottaviano,F.G., Tang,S.S., Handy,D.E. and Loscalzo,J. (2009) Regulation of the extracellular antioxidant selenoprotein plasma glutathione peroxidase (GPx-3) in mammalian cells. *Mol. Cell Biochem.*, **327**, 111–126.
49. Papp,L.V., Lu,J., Striebel,F., Kennedy,D., Holmgren,A. and Khanna,K.K. (2006) The redox state of SECIS binding protein 2 controls its localization and selenocysteine incorporation function. *Mol. Cell Biol.*, **26**, 4895–4910.
50. Boulon,S., Marmier-Gourrier,N., Pradet-Balade,B., Wurth,L., Verheggen,C., Jady,B.E., Rothe,B., Pescia,C., Robert,M.C., Kiss,T. *et al.* (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J. Cell Biol.*, **180**, 579–595.
51. Castellano,S., Gladyshev,V.N., Guigo,R. and Berry,M.J. (2008) SelenoDB 1.0: a database of selenoprotein genes, proteins and SECIS elements. *Nucleic Acids Res.*, **36**, D332–D338.