April 2007

Selenoprotein H Is a Nucleolar Thioredoxin-like Protein with a Unique Expression Pattern

Sergey V. Novoselov
University of Nebraska-Lincoln

Gregory V. Kryukov
University of Nebraska-Lincoln

Xue-Ming Xu
NCI, National Institutes of Health, Bethesda, Maryland 20892

Bradley A. Carlson
NCI, National Institutes of Health, Bethesda, Maryland 20892

Dolph L. Hatfield
NCI, National Institutes of Health, Bethesda, Maryland 20892

See next page for additional authors

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 16, pp. 11960–11968, April 20, 2007
Printed in the U.S.A.

Selenoprotein H Is a Nucleolar Thioredoxin-like Protein with a Unique Expression Pattern*[S]

Received for publication, September 10, 2006, and in revised form, February 23, 2007 Published, JBC Papers in Press, March 1, 2007, DOI 10.1074/jbc.M701605200

Sergey V. Novoselov[‡1], Gregory V. Kryukov[‡2], Xue-Ming Xu[§], Bradley A. Carlson[§], Dolph L. Hatfield[§], and Vadim N. Gladyshev[‡3]

From the [‡]Department of Biochemistry, University of Nebraska, Lincoln, Nebraska 68588 and [§]Molecular Biology of Selenium Section, Laboratory of Cancer Prevention, Center for Cancer Research, NCI, National Institutes of Health, Bethesda, Maryland 20892

The human selenoproteome consists of 25 known selenoproteins, but functions of many of these proteins are not known. Selenoprotein H (SelH) is a recently discovered 14-kDa mammalian protein with no sequence homology to functionally characterized proteins. By sensitive sequence and structure analyses, we identified SelH as a thioredoxin fold-like protein in which a conserved *CXXU* motif (cysteine separated by two other residues from selenocysteine) corresponds to the *CXXC* motif in thioredoxins. These data suggest a redox function of SelH. Indeed, a recombinant SelH shows significant glutathione peroxidase activity. In addition, SelH has a conserved RKRK motif in the N-terminal sequence. We cloned wild-type and cysteine mutant forms of SelH either upstream or downstream of green fluorescent protein (GFP) and localized this fusion protein to the nucleus in transfected mammalian cells, whereas mutations in the RKRK motif resulted in the cytosolic protein. Interestingly, the full-length SelH-GFP fusion protein localized specifically to nucleoli, whereas the N-terminal sequence of SelH fused to GFP had a diffuse nucleoplasm location. Northern blot analyses revealed low expression levels of SelH mRNA in various mouse tissues, but it was elevated in the early stages of embryonic development. In addition, SelH mRNA was overexpressed in human prostate cancer LNCaP and mouse lung cancer LCC1 cells. Down-regulation of SelH by RNA interference made LCC1 cells more sensitive to hydrogen peroxide but not to other peroxides tested. Overall, these data establish SelH as a novel nucleolar oxidoreductase and suggest that some functions in this compartment are regulated by redox and dependent on the trace element selenium.

Selenium is an essential trace element with significant biomedical potential and roles in human health (1). As a dietary supplement, selenium has been shown to serve as a cancer chemopreventive agent and has roles in immune function, mammalian development, male reproduction, and in preventing heart disease and other cardiovascular and muscle disorders (for reviews, see Ref. 2). Its biological significance is attributed to the occurrence of this element in proteins in the form of selenocysteine (Sec),[4] the 21st amino acid in the genetic code (3). This rare, highly reactive amino acid (4) is encoded by the UGA codon and is incorporated into proteins co-translationally (5–7). Sec-specific incorporation requires the presence of *cis*- and *trans*-acting elements. In addition to UGA, a stem-loop structure known as the SECIS (selenocysteine insertion sequence) element operates as the *cis* element for Sec incorporation (8–10). In eukaryotes, SECIS is located in 3′-UTRs of selenoprotein mRNAs and interacts with *trans*-acting factors. Recently, an additional recoding element was identified in a subset of eukaryotic selenoprotein genes that is adjacent to Sec-encoding UGA codons (11). In eukaryotes, the *trans*-acting factors consist of Sec tRNA (12–15), selenocysteyl-tRNA-specific elongation factor (14, 15), and SBP2 (SECIS-binding protein 2) (16, 17). These three components function together to redefine the UGA codon to dictate Sec insertion (6, 18). Ribosomal protein L30 has also been shown to be a component of the eukaryotic Sec recoding machinery (19).

Selenoproteins are present in all kingdoms of life, and the full set of such proteins in a particular organism is known as the selenoproteome. The human selenoproteome consists of 25 proteins, whereas rodents have 24 such proteins (20). About half of these proteins (*e.g.* glutathione peroxidases (21) and thioredoxin reductases (22)) play roles in various redox reactions, but functions of the remaining selenoproteins are not known. Therefore, to understand the effects of selenium on human health and to explain how these effects are governed by dietary dependence on selenium, the functions of all selenoproteins need to be elucidated.

Selenoprotein H (SelH) is a recently discovered 14-kDa mammalian protein with no homology to functionally characterized proteins (20). However, certain similarities to other selenoproteins can be seen, particularly a conserved Cys-Xaa-Xaa-Sec (*CXXU*) motif wherein cysteine and Sec are separated

* This study was supported by National Institutes of Health Grant GM061603 (to V. N. G.) and, in part, by the Intramural Research Program of the NCI, Center for Cancer Research, National Institutes of Health. The use of ^{75}Se was supported by DoE DE-FG07-02ID14380. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S3, Table S1, and references.

[1] Present Address: Institute of Cell Biophysics, Russian Academy of Sciences, Pushchino, Mascow region 142292, Russia.

[2] Present address: Div. of Genetics, Dept. of Medicine, Brigham and Women's Hospital and Harvard Medical School, New Research Bldg., 77 Ave. Louis Pasteur, Boston, MA 02115.

[3] To whom correspondence should be addressed. E-mail: vgladyshev1@unl.edu.

[4] The abbreviations used are: Sec, selenocysteine; SelH, selenoprotein H; GFP, green fluorescent protein; GPx, glutathione peroxidase; *t*-BOOH, *tert*-butyl hydroperoxide; PBS, phosphate-buffered saline; NLS, nuclear localization signal; DAPI, 4′,6-diamidino-2-phenylindole.

by two other amino acids. This feature is also present in several other mammalian selenoproteins (*i.e.* SelW, SelT, SelM, and SelV), but the physiological roles of these selenoproteins are also not known. Structures of SelM (23) and its distant homolog Sep15 (24) have recently been determined (25). Both proteins possess a thioredoxin-like domain and reside in the endoplasmic reticulum, suggesting that these two proteins are thiol-disulfide isomerases with a role in disulfide bond formation.

Among other C*XX*U-containing selenoproteins, SelW is a better characterized protein, but its function also has not been defined (26). It is the smallest mammalian selenoprotein existing in the form of several isoforms that differ by post-translational modifications (27, 28). SelW is expressed in various tissues and is abundant in muscle. Reduced levels of SelW have been attributed to white muscle disease (29). SelV is a testis-specific distant homolog of SelW having an additional domain of unknown function (20). Here we used computational, molecular, and cell biology techniques to characterize SelH, which revealed that this protein is a nucleolar oxidoreductase with an unusual expression pattern.

EXPERIMENTAL PROCEDURES

Materials—Chemicals were from Sigma, restriction enzymes from Amersham Biosciences, DNA purification kits from Qiagen, and mammalian cell culture reagents from Invitrogen. Primers used in the study are shown in supplemental Table S1.

Multiple Sequence Alignment, Topology Prediction, and Structural Motif Search for SelH—All detected SelH sequences were extracted from GenBank™ nonredundant and EST databases and analyzed by BLAST (30) and ClustalW software (31). Localization of SelH was predicted using PSORT II (32). PSI-BLAST/PHI-BLAST (33) searches with the default setting were carried out to detect conserved protein domains. The HHpred web-based program (34) was used for fold prediction.

Recombinant Protein Expression in Escherichia coli and Antibody Production—Mouse SelH cDNA clone was purchased from Research Genetics. The Sec codon was substituted with that coding for Cys using QuikChange II kit (Stratagene) and the coding sequence was PCR-amplified and cloned into pET28a (Novagen). The resulting expression construct was further mutated to obtain C*XX*S and S*XX*C variants.

Expression in *E. coli* BL21 DE3 cells (Novagen) and purification of the recombinant His-tagged protein using TALON resin (Clontech) was carried out according to the manufacturer's protocols. Rabbit polyclonal antibodies against the recombinant protein were produced by Covance.

Activity Assays—Glutathione peroxidase (GPx) activity was measured by a coupled reaction with glutathione (35) using hydrogen peroxide or *tert*-butyl hydroperoxide (*t*-BOOH) as substrates and monitoring the decrease in NADPH absorbance at 340 nm according to the manufacturer's protocol (BOIXYTECH GPx-340 assay kit, Oxis Research). The enzyme unit was defined as the amount of protein needed to catalyze the oxidation of 1 μmol NADPH/min. We assayed both Cys and Ser mutant forms of recombinant SelH and used mouse liver lysate as control.

Identification of Target Proteins—Targets of SelH were identified by trapping mixed disulfides formed between affinity resins containing mutant SelH forms and proteins from a nuclear fraction of LCC1 cells. The affinity resins were prepared by cross-linking recombinant SelH proteins (C*XX*S and S*XX*C forms) to *N*-hydroxysuccinimide ester-activated Sepharose (GE Healthcare) according to the recommendations of the manufacturer. Nuclear fraction of LCC1 cells was prepared by differential centrifugation, homogenized in phosphate-buffered saline (PBS) containing protease inhibitors (Roche Applied Science), and incubated with SelH-immobilized NHS-Sepharose resins for 1 h at 25 °C. The resins were extensively washed with PBS to remove nonspecifically bound proteins. The resins were then resuspended in 50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 10 mM dithiothreitol, and incubated for 30 min at 25 °C. The eluted samples were fractionated by SDS-PAGE, and proteins were silver-stained. The stained bands were excised and subjected to liquid chromatography tandem mass spectrometry. Data base searches of the acquired tandem mass spectra were performed using Mascot (Matrix Science, version 1.9.0, London, UK).

Green Fluorescent Protein (GFP) Fusion Constructs—A set of chimera proteins included SelH fused C- or N-terminally with GFP. In brief, we amplified wild-type mouse SelH cDNA (construct number 5), prepared its Sec38Cys mutant, and cloned it into pEGFPN2 and pEGFP-C3 vectors (Clontech). We also prepared a construct by cloning a sequence coding for a 3-kDa N-terminal portion of SelH upstream of GFP (construct number 3). A putative nuclear localization signal (NLS) was disrupted by mutating arginines 6 and 8 to serines using the QuikChange II kit (Stratagene), and for each chimera, the corresponding mutant with disrupted NLS was obtained. Transfection of NIH 3T3 cells was performed using Lipofectamine (Invitrogen), and the images were captured using an Olympus FV500 inverted focal microscope at the University of Nebraska-Lincoln Microscopy Core Facility.

Immunohistochemistry—NIH 3T3 cells were seeded on coverslips and transfected with SelH-GFP fusion constructs. After 24 h, the cells were washed with PBS and fixed by flashing with ice-cold methanol containing 0.1% Triton X-100 followed by a 5-min treatment with 4% paraformaldehyde. The cells were then treated with blocking solution (Roche Applied Science) and incubated with rabbit anti-nucleolin antibodies (Novus Biologicals) and mouse monoclonal anti-UB2 antibodies (RDI Research Diagnostics) concurrently followed by washing with PBS and treatment with donkey anti-mouse Cy5-conjugated and monkey anti-rabbit Cy3-conjugated antibodies (Jackson ImmunoResearch Laboratories). The samples were then washed with PBS and distilled water and mounted with Gel/Mount (Biomeda). Dilutions of all commercial antibodies were as suggested by the manufacturers.

Gene Expression Analyses—An *in silico* expression analysis was performed using the SAGE Genie web resource (cgap.nci.nih.gov/SAGE) (36). Pre-made Northern blot membranes (mouse adult tissue blot and mouse conceptus embryonic tissue blot) were purchased from Seegene and probed with 0.3-kb ^{32}P-labeled (Redivue [^{32}P]dCTP, Amersham Biosciences) SelH cDNA. The probe was generated using the Rediprime II random prime labeling system (Amersham Biosciences) in accordance with the manufacturer's protocol. To analyze the expres-

sion in cell lines, total RNA was isolated using the RNAqueous kit (Ambion), separated on an agarose gel, transferred onto a Zeta-Probe blotting membrane (Bio-Rad, protocol from (37)), and probed as described above.

Western Blot Analyses—Tissues were extracted from euthanized mice, frozen in liquid nitrogen, and stored at −80 °C until the day of analysis. The tissues were homogenized in PBS containing complete protease inhibitors (Roche Applied Science), and the lysates were normalized with regard to protein concentration. 10% BisTris NOVEX gels (Invitrogen) were used, and each well was loaded with 25 μg of protein. Following SDS-PAGE, the proteins were transferred onto polyvinylidene difluoride membranes (Invitrogen) and probed with anti-SelH antibodies. Secondary horseradish peroxidase-conjugated antibodies were from Amersham Biosciences, and chemiluminescent peroxidase substrate was from Sigma. Lysates from various human cancer cell lines (a generous gift from Dr. Katerina Gurova, Cleveland BioLabs) were examined as described above.

Cell Culture Labeling, Subcellular Fractionation, and Nucleoli Preparation—Experiments involving metabolic labeling of selenoproteins were carried out by adding ^{75}Se ([^{75}Se]selenious acid; specific activity 1,000 Ci/mmol, Research Reactor Facility, University of Missouri, Columbia, MO) to the culture medium as described previously (38). Separation of cytoplasm and nuclear fractions was performed as described elsewhere (39). In brief, following labeling for 48 h, cells were trypsinized, washed with PBS, homogenized using a tight pestle in 0.25 M sucrose buffer containing complete protease inhibitor mixture (Roche Applied Science), and cytosolic and nuclear fractions were separated by differential centrifugation. The resulting cellular fractions were normalized in regard to protein concentration, subjected to SDS-PAGE, and transferred onto polyvinylidene difluoride membrane, and then ^{75}Se-labeled proteins were visualized with a Storm PhosphorImager system (Amersham Biosciences). Further subfractionation of nuclear compartments to obtain nucleoli and nucleoplasm fractions was conducted as described previously (40) with modifications (41).

RNA Interference in LCC1 Cell Line and Oxidative Stress—To knock down SelH gene expression, five separate 21-nucleotide sequences were selected from SelH cDNA as being unique using the online service "si*DESIGN" of* Dharmacon Research, Inc. These small interfering RNA constructs of pU6-m3 were prepared as described previously (42). The sequences of all five SelH small interfering RNA constructs were confirmed. The five constructs and the pU6-m3 negative control plasmids were separately transfected into LLC1 cells. After stabilizing the transfected cells with 0.8 mg/ml hygromycin, SelH mRNA levels were determined by Northern blot analysis. Based on mRNA levels, the best small interfering RNA target sequences of SelH were identified as 5′-GAA TTG AAG AAG TAC CTT TCA-3′. These small interfering RNA constructs were used in subsequent experiments.

Sensitivity of LCC1 cells to oxidative stress was analyzed by the CellTiter 96 AQ_{ueous} One Solution cell proliferation assay (Promega). 5×10^3 cells were seeded into each well in 96-well plates and grown overnight in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. The cells were washed twice with PBS and incubated in 100 μl of serum/phenol red-free medium containing the indicated (Fig. 6*B*) concentrations of hydrogen peroxide or 0–1 mM *t*-BOOH, cumyl hydroperoxide or paraquat for 1 h at 37 °C. To examine sensitivity to menadione, we used 0–30 μM treatment with this compound for 6 h. 20 μl of CellTiter 96 AQ_{ueous} One Solution reagent were added to each well and incubated for an additional 4 h. The A_{490} value was recorded using an ELx808 Ultra microplate reader (Bio-Tek Instruments). Cell viability was calculated as the percentage of untreated controls, and the wells without cells were used as a blank.

Evolutionary Analysis—Amino acid sequence alignment of SelH sequences was generated using the PileUP program. The distances between sequences were calculated from these alignments using the DISTANCE program with the Kimura distance-measuring method, and then the unrooted phylogenetic tree with unscaled distance branches was generated using the GrowTree program with the unweighted pair-group method. All programs were part of the Wisconsin Package, Genetic Computer Group, Madison, WI (43).

RESULTS AND DISCUSSION

SelH Is a Putative Nuclear Oxidoreductase—Available SelH sequences were collected and assembled into a multiple sequence alignment (Fig. 1*A*). SelH sequences contain a conserved C*XX*U motif, which is flanked upstream by a predicted β-strand and downstream by a predicted α-helix. This arrangement is highly similar to thiol/disulfide oxidoreductases of the thioredoxin fold; however, the latter proteins have a C*XX*C motif instead of C*XX*U (44, 45). The use of advanced BLAST programs (PSI-BLAST/PHI-BLAST) (33) with several iterations revealed a distant homology to thioredoxin fold proteins. Indeed, further analysis with HHPred identified a thioredoxin-like fold in SelH. Structural alignment places the C*XX*U motif in SelH in the same place as the C*XX*C motif in thioredoxin. These data suggest that SelH is an oxidoreductase that uses its C*XX*U motif as the reversible thiolselenol/selenenylsulfide redox group.

Sequence analysis also revealed nuclear localization signals (46) in SelH sequences. Remarkably, SelH proteins have a high content of basic residues (18.3% in mouse and 22.2% in zebrafish) due to the high occurrence of arginines and lysines. By using PSORT II, we found six putative NLSs in the SelH sequences of chickens (all monopartite), three in zebrafish (all monopartite), and eight in mice and humans (six monopartite and two bipartite). Comparison of NLS location and conservations using the multiple sequence alignment showed that only two NLS were conserved, of which one was in the N-terminal and one in C-terminal sequences. Within the N-terminal NLS, the two strictly conserved residues were Arg-6 and Arg-8 (Fig. 1*A*) (numbering based on the mouse SelH sequence), suggesting that these residues have a critical role in nuclear import as a part of NLS.

Peroxidase Activity of Recombinant SelH—To further elucidate whether SelH is an oxidoreductase, we performed several biochemical tests with recombinant mutant forms of the protein in which the Sec residue was substituted with Cys or Ser. Specifically, these mutants were assayed for thioredoxin, per-

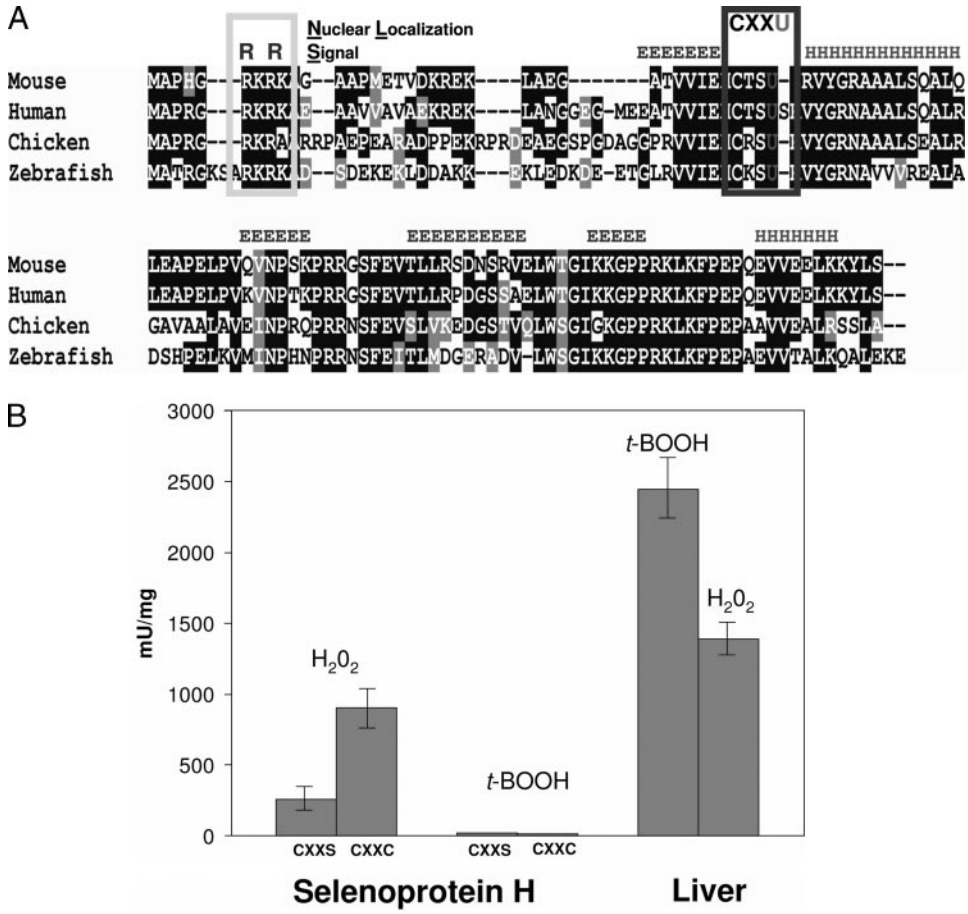


FIGURE 1. **Redox properties of SelH.** *A*, multiple sequence alignment of SelH sequences. The locations of C*XX*U (which corresponds to the C*XX*C motif in thioredoxins) and NLS motifs are indicated by *boxes*, respectively. The predicted secondary structure is shown *above* the sequence. *B*, glutathione peroxidase activity of SelH. The indicated recombinant forms of SelH (C*XX*S and C*XX*C forms) or mouse liver lysate were assayed for GPx activity as described under "Experimental Procedures." Reactions were initiated by adding 300 μM H_2O_2 or 233 μM *t*-BOOH as the substrate. *mU*, milliunits.

oxiredoxin, and GPx activities. In the thioredoxin assay, mammalian thioredoxin reductase 1 was used as the reductant and 5,5′-dithiobis-(2-nitrobenzoic acid) as the oxidant (47, 48), and a separate assay involved insulin reduction, as in the corresponding assay of thioredoxin function (49). We could detect no thioredoxin activity of SelH (data not shown). The peroxiredoxin activity was examined in an assay of protection of glutamine synthetase activity (47). SelH also showed no activity in this assay. Interestingly, SelH had a significant GPx activity (Fig. 1*B*). This activity was detected only with hydrogen peroxide, whereas *t*-BOOH was not active as a substrate. The Cys mutant had four times higher activity than the Ser mutant of SelH. We further compared the GPx activity of SelH with that of mammalian GPx1 using mouse liver lysate and quantitative Western blot analysis of GPx1 amounts in the sample (not shown). This comparison revealed that the GPx activity of the Cys mutant was 2,000–4,000-fold lower than that of natural mouse liver GPx1 under the conditions used in the assay. It should be pointed out that, although SelH had a much lower activity than GPx1, we assayed the Cys form of the natural Sec-containing SelH. In addition, the peroxidase activity detected was similar to the activities of peroxiredoxins in the reduction of hydroperoxides. Thus, these data provided further evidence for the oxidoreductase activity of SelH. The natural oxidoreductase activity of SelH, however, remains unknown, and further studies will be needed to identify specific substrates/targets of this protein.

Identification of Target Proteins—The finding that SelH possesses a thioredoxin fold and has a putative C*XX*U redox motif suggests analogy with thioredoxin-like proteins that form transient disulfides with target proteins. Mutation of a resolving Cys in the C*XX*C motif of thioredoxin often results in stabilization of such complexes (51–53). We utilized this approach to trap the target proteins of SelH. Affinity resins were prepared by linking C*XX*S and S*XX*C forms of recombinant SelH and incubating with a nuclear fraction prepared from LCC1 cells. A silver-stained gel showing proteins eluted from the resins is shown in supplemental Fig. S1. Proteins enriched on the SelH resins were then identified by tandem mass spectrometry sequencing. Among those detected in the samples were several thiol oxidoreductases (*e.g.* peroxiredoxins 1 and 2, thioredoxin, and glutaredoxin), which is consistent with the redox nature of interactions between SelH and target proteins. Interestingly, most other identified proteins have previously been described as nucleolar proteins (*e.g.* pre-mRNA processing protein 8, putative pre-mRNA splicing factor RNA helicase (DEAH box protein 15), nucleolin, 60 S ribosomal protein L18, ribosomal protein S28). It should be noted that the identified proteins should only be viewed as preliminary targets. Further studies will be required to functionally characterize these interactions and determine their specificity.

Localization of SelH-GFP Fusion Constructs—To test whether SelH is a nuclear protein and whether Arg-6 and Arg-8 are responsible for its location, we prepared and transiently expressed various SelH-GFP fusion proteins in NIH 3T3 cells (Fig. 2*A*). In these experiments, we used a Sec38Cys mutant form of SelH, which was placed upstream (construct 1) and downstream (construct 5) of GFP. In addition, we fused a 3-kDa N-terminal portion of the protein that contains the putative NLS with GFP (construct 3). We also placed a full-length SelH cDNA sequence including a 3′-UTR containing a SECIS element downstream of GFP as described previously (20). For each of these four constructs, we developed mutants in which Arg-6 and Arg-8 were mutated to serines. Western blot analysis (Fig. 2*B*) of cells transfected with these constructs revealed correspondence of the actual sizes of the

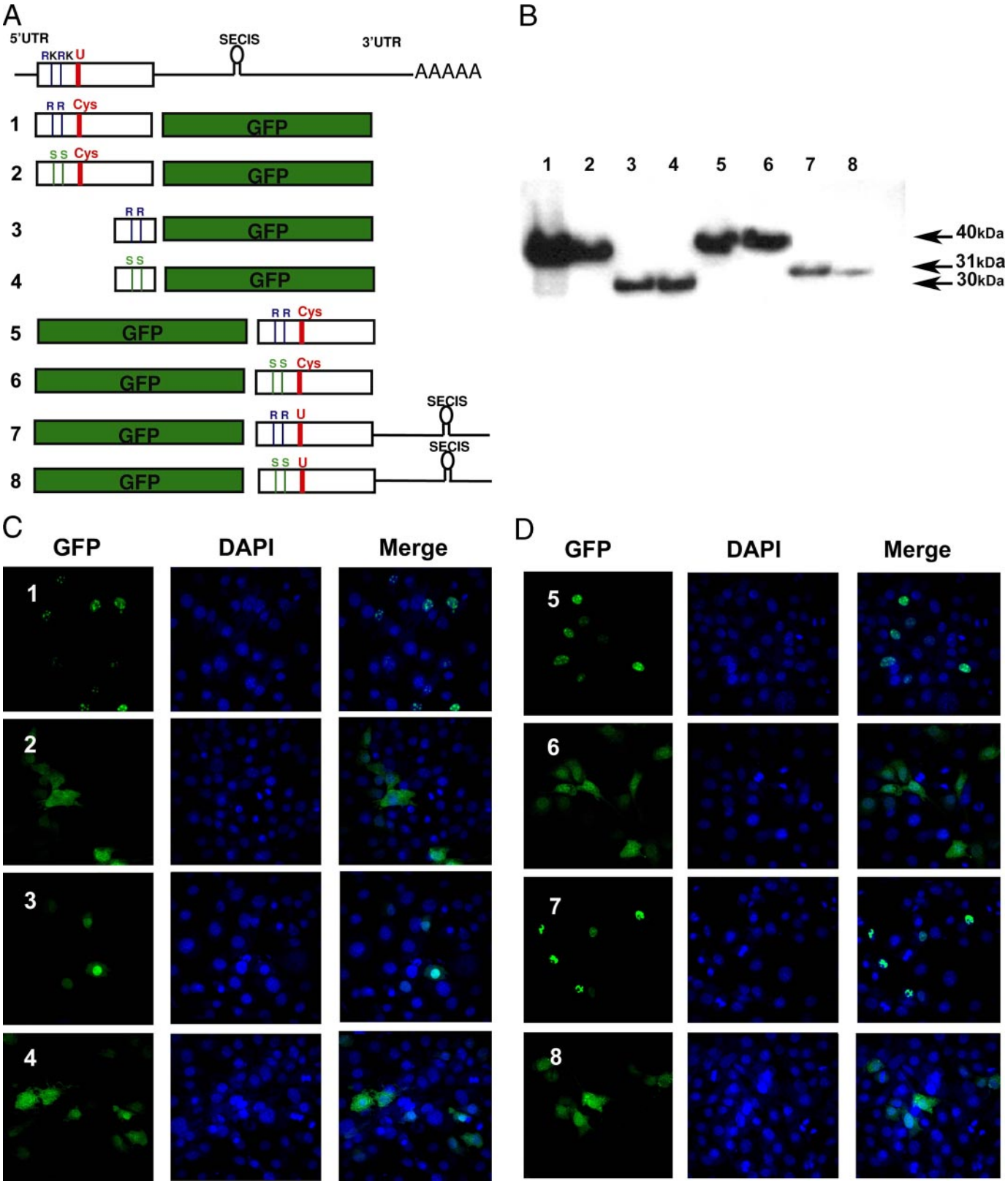


FIGURE 2. **Localization of SelH.** *A*, organization of SelH-GFP fusion constructs utilized in the localization experiments. SelH was fused either upstream or downstream of GFP, and for each construct, the two conserved arginines in the NLS motif were mutated to serines. In addition, to increase expression levels, Sec was mutated to Cys in several constructs as shown in the figure. In the constructs expressing a selenoprotein form of SelH, a 3′-UTR was included that contained a SECIS element. *B*, detection of SelH-GFP fusion proteins by Western blots using anti-SelH antibodies. Migration of expressed proteins corresponded to their predicted masses and is indicated by *arrows* on the *right*. *Lanes 1–8* contain proteins generated from the corresponding constructs shown in *A* above. *C*, confocal microscopy of GFP-SelH fusion protein constructs containing an N-terminal SelH. Three images are shown for each construct. *Left panels* show *green* (GFP) fluorescence, *middle panels* DAPI staining (nuclear marker), and *right panels* images obtained by merging the *left* and *middle panels*. *Numbering* from *1* to *4* corresponds to *A* above. *D*, Same as in *C*, except that SelH was fused downstream of GFP in the SelH-GFP protein. *Numbering* from *5* to *8* corresponds to *A* above.

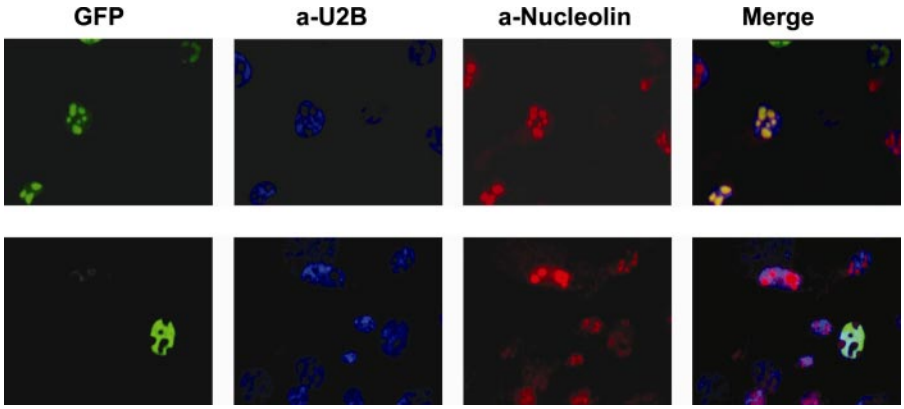


FIGURE 3. **SelH is a nucleolar protein.** Cells transfected with SelH-GFP (*upper panel*, construct 1 in Fig. 2) or GFP-SelH (*lower panel*, construct 7) were double-labeled with α-U2B antibodies (*blue*) and α-nucleolin antibodies (*red*). The *right* image is a merge of the three other images.

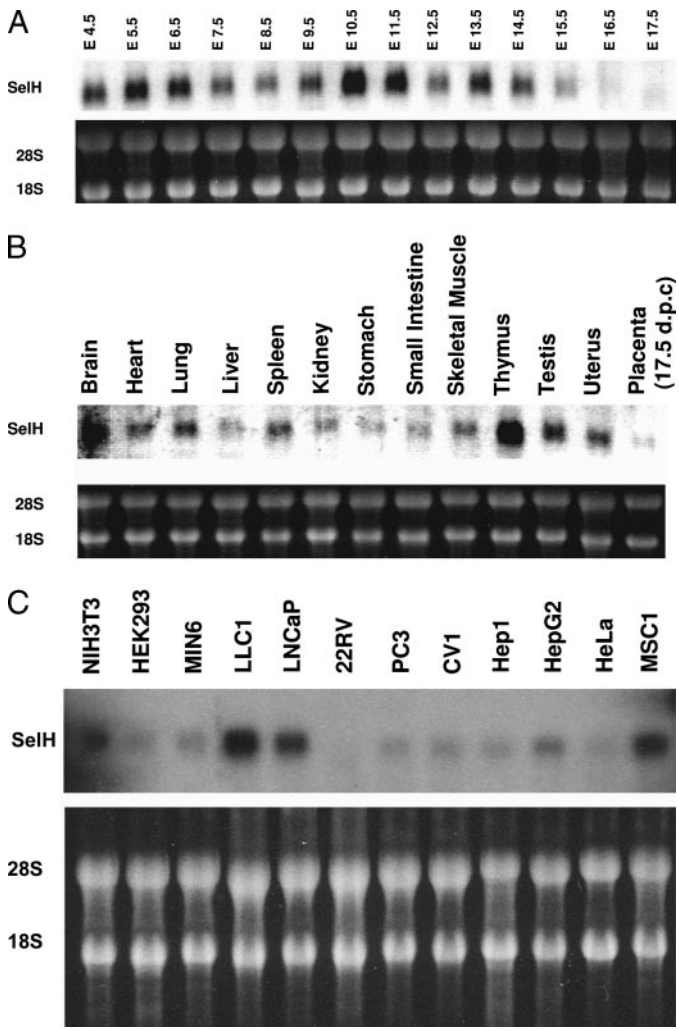


FIGURE 4. **Expression of SelH during development and in cancer cells.** Northern blot analysis of SelH mRNA expression is shown. *Upper panels* in *A–C* show SelH expression and lower panels (controls), expression of 28 and 18 S ribosomal RNAs. Expression during embryonic development (*A*), in mouse tissues (*B*), and in the indicated cells (*C*) is shown.

fusion proteins to the calculated masses, except for the wild-type SelH-GFP fusion proteins (*constructs 7* and *8* on Fig. 2*A* and *lanes 7* and *8* on Fig. 2*B*, respectively), because of significant premature termination at the Sec position. Full-length Sec-containing SelH-GFP fusion proteins were not detectable using available anti-SelH antibodies.

GFP fluorescence in transfected cells resulting from SelH-GFP fusion proteins was determined by dual wavelength confocal microscopy in parallel with a nuclear marker (DAPI). We found that, when NLS was present in proteins, fluorescence was localized to the nuclear fraction (Fig. 2, *C* and *D*; *panels 1*, *3*, *5*, and *7*), and when the arginines were mutated, the fluorescence signal was confined to the cytosol (Fig. 2, *C* and *D*; *panels 2*, *4*, *6*, and *8*). Thus, SelH is a nuclear protein possessing a RKRK monopartite NLS, and the arginines in this motif are essential for nuclear targeting of the protein.

SelH Is a Nucleolar Protein—As shown in Fig. 2, *C* and *D*, fluorescence inside nuclei that reflected localizations of proteins derived from constructs 1, 3, 5, and 7 was heterogeneous. In fact, both the SelH-GFP and GFP-SelH forms appeared to be enriched in nucleoli, whereas when only the NLS-containing N-terminal peptide was fused to GFP (construct 3), fluorescence was diffused throughout the nuclear compartment. Interestingly, in the Sec-containing SelH-GFP fusion protein (construct 7), which migrated as a 310-kDa protein in Western blot assays (Fig. 2*B*) (indicating significant premature termination at the UGA codon), the fluorescent signal was not diffused but still differed from the typical nucleolar staining. We subjected the cells transfected with constructs 1 and 7 to immunohistochemical analysis and used markers for nucleolar and nucleoplasm staining in parallel. As shown in Fig. 3, *upper panel*, the full-length SelH co-localizes with nucleolin, a nucleolar marker, indicating that SelH is indeed a nucleolar protein. However, the location of the 3-kDa portion of SelH fused to GFP (construct 7) was surprising, as it appeared to be present in the nucleoplasm. Taken together, these data indicate that SelH is a nucleolar protein and that its correct localization required the full-length folded protein. The nucleolar location of SelH is also consistent with the abundance of nucleolar proteins that were enriched on the SelH resins.

The nucleolus has long been known as a dense subnuclear structure where ribosome biogenesis occurs (for reviews, see Refs. 54 and 55). Unlike the nucleus and other membrane-restricted organelles, there is no evidence for the occurrence of a frontier separating the nucleolus from the surrounding nucleoplasm (56, 57). It is thought that any molecule could migrate in and out of the nucleolus and that targeting of particular cellular components to the nucleolus depends on interactions with the co-called nucleolar binding blocks (58).

Expression of SelH mRNA—To analyze SelH gene expression, we initially carried out an *in silico* expression analysis using mouse and human expression data that is available in the form of supplemental Figs. S2 and S3. Both mouse and human SelH

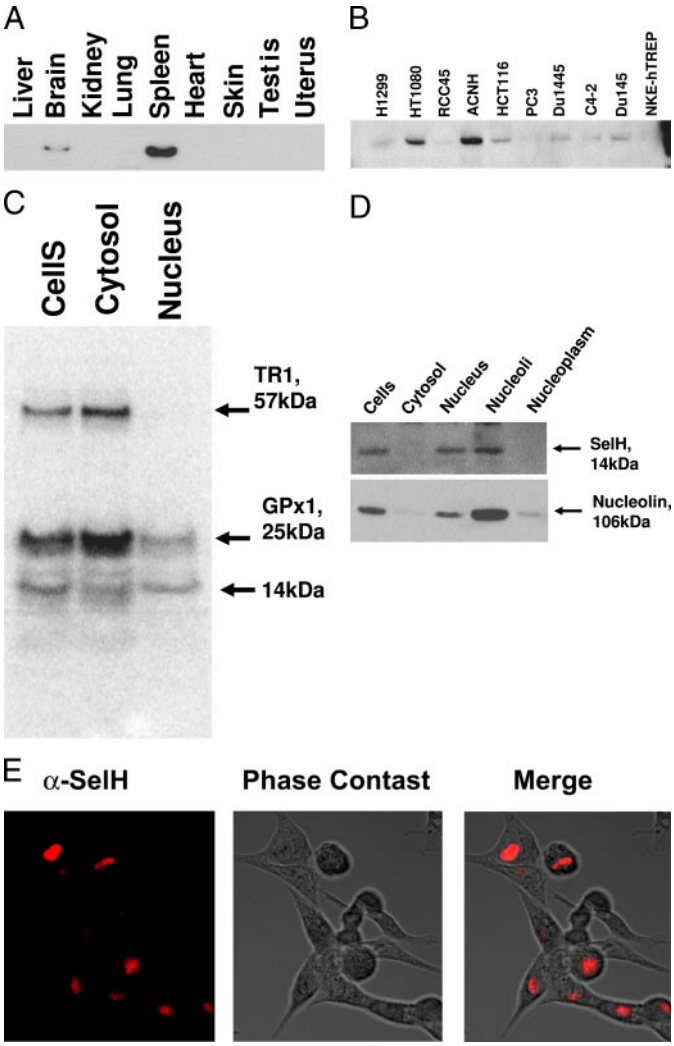


FIGURE 5. **Expression of SelH in mouse tissues.** Western blot analysis of SelH expression in mouse tissues (*A*) and indicated cell lines (*B*) is shown. *C*, subcellular fractionation of ^{75}Se-labeled LCC1 cells. Note enrichment of thioredoxin reductase 1 and GPx1 in the cytosolic fraction, whereas the nuclear fraction had 14- and 25-kDa selenoproteins. *D*, SelH localizes specifically to the nucleolar compartment. Western blot analysis using SelH (*upper panel*) and nucleolin (*lower panel* (control)) is shown. *E*, direct immunohistochemistry of SelH in LCC1 cells.

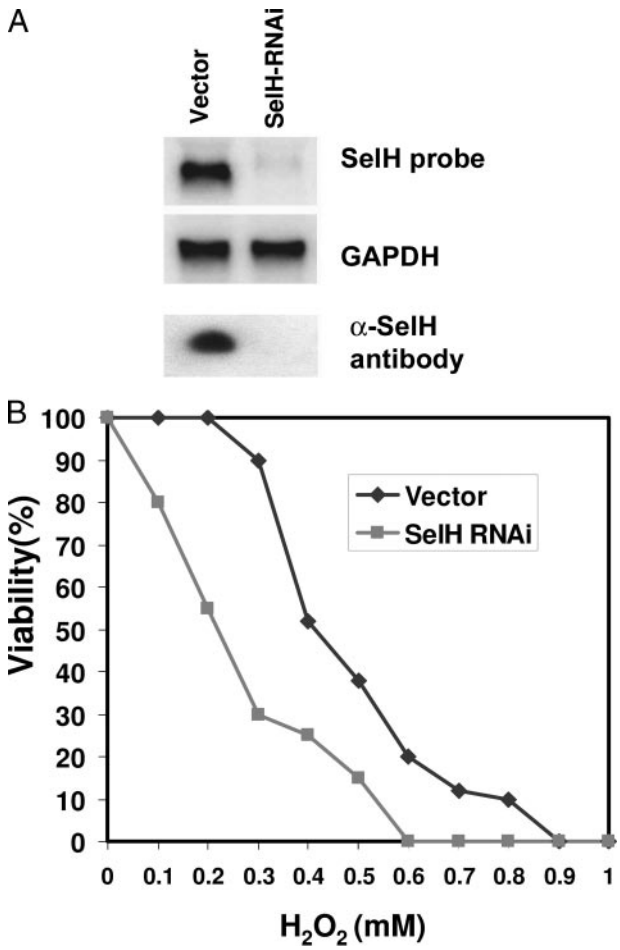


FIGURE 6. **SelH-deficient cells are sensitive to hydrogen peroxide treatment.** *A*, verification of knockdown cells. The *upper panel* shows Northern blot analysis using the SelH probe, *middle panel* (control) shows Northern blot analysis using the GAPDH probe, and the *lower panel* shows Western blot analysis using anti-SelH antibodies. The *left lane* shows cells transfected with a control vector and the *right lane*, SelH knockdown cells. *B*, viability of control and SelH knockdown cells treated with the indicated concentrations of hydrogen peroxide is shown.

mRNA appeared to be moderately expressed in many, if not all, tissues and organs. However, the majority of expressed sequence tags were derived from either embryonic or tumor cells, particularly from carcinomas. Moreover, analysis of SAGE datasets suggested that expression of human SelH is elevated in some tumors, especially thyroid, lung, stomach, and liver cancers. In cell lines, SelH appeared high in embryonic samples, carcinomas, and cell lines derived from bone marrow. In mouse, it appeared to be elevated in developing brain. Interestingly, we previously found that, during zebrafish development, SelH localized within highly proliferative tissues such as the brain ventricular zone, part of the retina and tectum at 24 hpf, and later in branchial arches and pectoral fin buds and the proliferative zone of the retina (59). In addition, zebrafish SelH was identified as essential for embryonic and early larval development (60).

To directly examine SelH mRNA expression, we carried out Northern blot analyses using mouse embryonic and adult tissues and a number of available cell lines (Fig. 4). As expected, we detected SelH mRNA in various tissues. The signal was particularly strong in samples from early stages of embryogenesis and in the adult tissues of thymus, brain, testis, and uterus (Fig. 4, *A* and *B*). In cell lines, SelH mRNA expression was elevated in LCC1 (lung cancer), LNCaP (human prostate carcinoma), MCS1 (mouse Sertoli cells), and NIH 3T3 (mouse fibroblasts) cell lines.

Expression of SelH in Mouse Tissues and Human Cancer Cell Lines—To further elucidate the expression of SelH, we carried out Western blot analysis (Fig. 5). SelH was abundant in mouse spleen but was expressed in lower levels in the brain and was not detected in the other organs examined (Fig. 5*A*). In addition, we analyzed cancer cell lines and found that SelH expression was dissimilar in various cells, but as expected, it was high in LNCaP, LCC1, and several other cell lines (Fig. 5*B*).

We metabolically labeled LNCaP cells with ^{75}Se and then prepared nuclear and cytoplasmic fractions (Fig. 5*C*). A 14-kDa band corresponding in size to SelH was enriched in the nuclear fraction. The nuclear fraction was further fractionated into

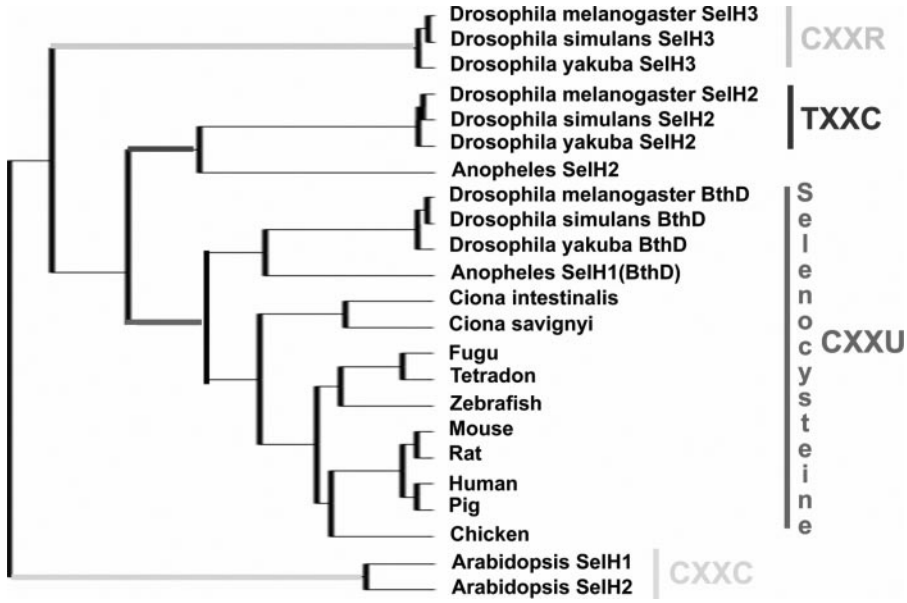


FIGURE 7. **Phylogenetic analysis of SelH sequences.** Animal SelH sequences are split into three branches, the major of which is the SelH branch, and the two other groups include fruit fly sequences representing non-selenoprotein SelH homologs. More distantly related plant sequences have a C*XX*C motif. The accession numbers (GenBank™) are as follows: *Gallus gallus* SelH (46017578), *Danio rerio* SelH (31342376), *Tetradon* SelH (56305185), *Anopheles gambiae* SelH (27645579), *Drosophila melanogaster* BthD (14718671), *Arabidopsis thaliana* SelH1 (30689011), *Arabidopsis thaliana* SelH2 (50058858), *Drosophila melanogaster* SelH3 (24584864), *Drosophila melanogaster* SelH2 (113194556), Mouse SelH (82546880), Rat SelH (109470184), Pig SelH (87230686), Human SelH (3516960), *Ciona savignyi* SelH (51695500), *Ciona intestinalis* SelH (56083593), *Drosophila yakuba* SelH3 (84681766), *Drosophila simulans* SelH3 (61736437), *Drosophila yakuba* SelH2 (84681627), *Drosophila simulans* SelH2 (111276688), *Drosophila yakuba* BthD (84679567), and *Drosophila simulans* BthD (61714200).

nucleolar and nucleoplasm fractions. By Western blot analysis, SelH was detected in the nucleolar fraction, in similar manner as the control protein nucleolin (Fig. 5*D*). Subsequent immunohistochemistry analysis further verified the presence of SelH in nucleoli (Fig. 5*E*).

Nucleoli are best known as a site for rRNA biosynthesis and modification. However, nucleoli have additional functions, such as virus infection control, maturation of non-nucleolar RNAs and RNPs, regulation of telomerase function and cell cycle, tumor suppressor and oncogene activities, and cell stress sensing and signaling (61–63). Moreover, by proteomic approaches, many uncharacterized proteins have been identified in nucleoli, suggesting additional unknown functions for this compartment (41, 64).

RNA Interference of SelH in LCC1 Cells—Because we found that SelH is a thioredoxin-like protein with significant GPx activity, this protein was further examined for antioxidant functions. We employed RNAi to suppress SelH expression and tested the sensitivity of knockdown cells to oxidative stress. Decreased expression of SelH in LCC1 knockdown cells was verified by both Northern and Western blot analyses (Fig. 6*A*). We subjected these cells to treatment with hydrogen peroxide and found that the knockdown cells were more sensitive to this form of oxidative stress then control LCC1 cells (Fig. 6*B*). However, no difference in sensitivity of knockdown and control cells was observed when they were treated with *tert*-butyl hydroperoxide (*t-BOOH*), cumyl hydroperoxide, paraquat, and menadione (data not shown). Higher sensitivity of SelH knockdown cells to hydrogen peroxide was consistent with a role of this protein in redox reactions and observed specificity of SelH for hydrogen peroxide compared with *t*-BOOH.

Nucleoli have been shown to be involved in oxygen-dependent regulatory mechanisms via the hypoxia-inducible factor (HIF-1). This protein is degraded under conditions of normal oxygen tension but is stabilized in hypoxia by VHL (von Hippel-Lindau, a tumor suppressor protein) activating a set of genes implicated in oxygen homeostasis, tumor vascularization, and ischemic preconditioning (65, 66). It has been shown that hypoxia can neutralize the function of VHL by triggering its nucleoli sequestration (67–69). This mechanism underlines the importance of redox status in nucleoli, and we speculate that SelH, being a nucleolar oxidoreductase, may play a related role. Moreover, being a selenoprotein, SelH may mediate certain effects of dietary selenium on nucleolar function.

Evolutionary Analyses—A SelH phylogenetic tree was constructed based on the available SelH sequences. As shown in Fig. 7, various *Drosophila* species have evolved two additional homologs of the protein, which form separate branches from the selenoprotein forms of the protein. Interestingly, additional fruit fly homologs have neither C*XX*C nor C*XX*U motifs. Instead, SelH3 possesses a C*XX*R and SelH2 a C*XX*T motif, the latter more typical of peroxiredoxins and glutathione peroxidases, which are large families of thiol peroxidases involved in various biological processes (50, 70). These data suggest functional differences between SelH and the two *Drosophila* homologs. It is also possible that various functions of mammalian SelH are carried out by separate proteins in fruit flies.

Acknowledgments—We thank Dr. Katerina Gurova for providing human cancer cell lysates and Anton Turanov for assistance with thioredoxin activity assays.

REFERENCES

1. Burk, R. F. (2002) *Nutr. Clin. Care* **5,** 75–79
2. Hatfield, D. L., Berry, M. J., and Gladyshev, V. N. (2006) *Selenium: Its Molecular Biology and Role in Human Health*, Springer, New York
3. Stadtman, T. C. (1996) *Annu. Rev. Biochem.* **65,** 83–100
4. Johansson, L., Gafvelin, G., and Arner, E. S. (2005) *Biochim. Biophys. Acta* **1726,** 1–13
5. Hatfield, D. L., and Gladyshev, V. N. (2002) *Mol. Cell. Biol.* **22,** 3565–3576
6. 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) *EMBO Rep.* **1,** 158–163
7. Copeland, P. R., and Driscoll, D. M. (2001) *Biofactors* **14,** 11–16
8. Martin, G. W. III, Harney, J. W., and Berry, M. J. (1998) *RNA* **4,** 65–73

9. Berry, M. J., Banu, L., Harney, J. W., and Larsen, P. R. (1993) *EMBO J.* **12,** 3315–3322
10. Hubert, N., Walczak, R., Sturchler, C., Myslinski, E., Schuster, C., Westhof, E., Carbon, P., and Krol, A. (1996) *Biochimie* (*Paris*) **78,** 590–596
11. Howard, M. T., Aggarwal, G., Anderson, C. B., Khatri, S., Flanigan, K. M., and Atkins, J. F. (2005) *EMBO J.* **24,** 1596–1607
12. Diamond, A., Dudock, B., and Hatfield, D. (1981) *Cell* **25,** 497–506
13. Hatfield, D., Diamond, A., and Dudock, B. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 6215–6219
14. Amberg, R., Urban, C., Reuner, B., Scharff, P., Pomerantz, S. C., McCloskey, J. A., and Gross, H. J. (1993) *Nucleic Acids Res.* **21,** 5583–5588
15. Park, S. I., Park, J. M., Chittum, H. S., Yang, E. S., Carlson, B. A., Lee, B. J., and Hatfield, D. L. (1997) *Biomed. Environ. Sci.* **10,** 116–124
16. Low, S. C., Grundner-Culemann, E., Harney, J. W., and Berry, M. J. (2000) *EMBO J.* **19,** 6882–6890
17. Copeland, P. R., Fletcher, J. E., Carlson, B. A., Hatfield, D. L., and Driscoll, D. M. (2000) *EMBO J.* **19,** 306–314
18. Copeland, P. R. (2003) *Gene* **312,** 17–25
19. Chavatte, L., Brown, B. A., and Driscoll, D. M. (2005) *Nat. Struct. Mol. Biol.* **12,** 408–416
20. Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R., and Gladyshev, V. N. (2003) *Science* **300,** 1439–1443
21. Aumann, K. D., Bedorf, N., Brigelius-Flohe, R., Schomburg, D., and Flohe, L. (1997) *Biomed. Environ. Sci.* **10,** 136–155
22. Arner, E. S., and Holmgren, A. (2000) *Eur. J. Biochem.* **267,** 6102–6109
23. Korotkov, K. V., Novoselov, S. V., Hatfield, D. L., and Gladyshev, V. N. (2002) *Mol. Cell. Biol.* **22,** 1402–1411
24. Korotkov, K. V., Kumaraswamy, E., Zhou, Y., Hatfield, D. L., and Gladyshev, V. N. (2001) *J. Biol. Chem.* **276,** 15330–15336
25. Ferguson, A. D., Labunskyy, V. M., Fomenko, D. E., Arac, D., Chelliah, Y., Amezcua, C. A., Rizo, J., Gladyshev, V. N., and Deisenhofer, J. (2006) *J. Biol. Chem.* **281,** 3536–3543
26. Whanger, P. D. (2000) *CMLS Cell. Mol. Life Sci.* **57,** 1846–1852
27. Beilstein, M. A., Vendeland, S. C., Barofsky, E., Jensen, O. N., and Whanger, P. D. (1996) *J. Inorg. Biochem.* **61,** 117–124
28. Gu, Q. P., Beilstein, M. A., Barofsky, E., Ream, W., and Whanger, P. D. (1999) *Arch. Biochem. Biophys.* **361,** 25–33
29. Gu, Q. P., Sun, Y., Ream, L. W., and Whanger, P. D. (2000) *Mol. Cell. Biochem.* **204,** 49–56
30. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
31. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
32. Nakai, K., and Horton, P. (1999) *Trends Biochem. Sci.* **24,** 34–36
33. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
34. Soding, J., Biegert, A., and Lupas, A. N. (2005) *Nucleic Acids Res.* **33,** W244–W248
35. Paglia, D. E., and Valentine, W. N. (1967) *J. Lab. Clin. Med.* **70,** 158–169
36. Boon, K., Osorio, E. C., Greenhut, S. F., Schaefer, C. F., Shoemaker, J., Polyak, K., Morin, P. J., Buetow, K. H., Strausberg, R. L., De Souza, S. J., and Riggins, G. J. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 11287–11292
37. Sambrook, J., Fritsch, E. F., and Maniatis, T. (eds) (1989) *Molecular Cloning, A Laboratory Manual.* 2nd Ed., pp. 16.33–16.37, Cold Spring Harbor Laboratory Press, New York
38. Kryukov, G. V., Kryukov, V. M., and Gladyshev, V. N. (1999) *J. Biol. Chem.* **274,** 33888–33897
39. Berrios, M. (1997) *Nuclear Structure and Function (Methods in Cell Biology)*, pp. 6–15, Academic Press, New York
40. Busch, H., Muramatsu, M., Adams, H., Steele, W. J., Liau, M. C., and Smetana, K. (1963) *Exp. Cell Res.* **24,** Suppl. 9, 150–163
41. Andersen, J. S., Lam, Y. W., Leung, A. K., Ong, S. E., Lyon, C. E., Lamond, A. I., and Mann, M. (2005) *Nature* **433,** 77–83
42. Xu, X. M., Mix, H., Carlson, B. A., Grabowski, P. J., Gladyshev, V. N., Berry, M. J., and Hatfield, D. L. (2005) *J. Biol. Chem.* **280,** 41568–41575
43. Senger, M., Flores, T., Glatting, K., Ernst, P., Hotz-Wagenblatt, A., and Suhai, S. (1998) *Bioinformatics* (*Oxf.*) **14,** 452–457
44. Martin, J. L. (1995) *Structure* (*Lond.*) **3,** 245–250
45. Kinch, L. N., Baker, D., and Grishin, N. V. (2003) *Proteins* **52,** 323–331
46. Hicks, G. R., and Raikhel, N. V. (1995) *Annu. Rev. Cell Dev. Biol.* **11,** 155–188
47. Kim, K., Kim, I. H., Lee, K. Y., Rhee, S. G., and Stadtman, E. R. (1988) *J. Biol. Chem.* **263,** 4704–4711
48. Krause, G., and Holmgren, A. (1991) *J. Biol. Chem.* **266,** 4056–4066
49. Bjornstedt, M., Hamberg, M., Kumar, S., Xue, J., and Holmgren, A. (1995) *J. Biol. Chem.* **270,** 11761–11764
50. Rhee, S. G., Kang, S. W., Jeong, W., Chang, T. S., Yang, K. S., and Woo, H. A. (2005) *Curr. Opin. Cell Biol.* **17,** 183–189
51. Holmgren, A., Johansson, C., Berndt, C., Lonn, M. E., Hudemann, C., and Lillig, C. H. (2005) *Biochem. Soc. Trans.* **33,** 1375–1377
52. Fernandes, A. P., and Holmgren, A. (2004) *Antioxid. Redox Signal.* **6,** 63–74
53. Motohashi, K., Kondoh, A., Stumpp, M. T., and Hisabori, T. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 11224–11229
54. Granneman, S., and Baserga, S. J. (2005) *Curr. Opin. Cell Biol.* **17,** 281–286
55. Raska, I., Shaw, P. J., and Cmarko, D. (2006) *Curr. Opin. Cell Biol.* **18,** 325–334
56. Jacobson, M. R., Cao, L. G., Wang, Y. L., and Pederson, T. (1995) *J. Cell Biol.* **131,** 1649–1658
57. Shaw, P. J., and Jordan, E. G. (1995) *Annu. Rev. Cell Dev. Biol.* **11,** 93–121
58. Carmo-Fonseca, M., Mendes-Soares, L., and Campos, I. (2000) *Nat. Cell Biol.* **2,** E107–E112
59. Thisse, C., Degrave, A., Kryukov, G. V., Gladyshev, V. N., Obrecht-Pflumio, S., Krol, A., Thisse, B., and Lescure, A. (2003) *Gene Expr. Patterns* **3,** 525–532
60. Amsterdam, A., Nissen, R. M., Sun, Z., Swindell, E. C., Farrington, S., and Hopkins, N. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 12792–12797
61. Olson, M. O., Hingorani, K., and Szebeni, A. (2002) *Int. Rev. Cytol.* **219,** 199–266
62. Pederson, T. (1998) *Nucleic Acids Res.* **26,** 3871–3876
63. Handwerger, K. E., and Gall, J. G. (2006) *Trends Cell Biol.* **16,** 19–26
64. Scherl, A., Coute, Y., Deon, C., Calle, A., Kindbeiter, K., Sanchez, J. C., Greco, A., Hochstrasser, D., and Diaz, J. J. (2002) *Mol. Biol. Cell* **13,** 4100–4109
65. Maxwell, P. H., Wiesener, M. S., Chang, G. W., Clifford, S. C., Vaux, E. C., Cockman, M. E., Wykoff, C. C., Pugh, C. W., Maher, E. R., and Ratcliffe, P. J. (1999) *Nature* **399,** 271–275
66. Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., Kriegsheim, A., Hebestreit, H. F., Mukherji, M., Schofield, C. J., Maxwell, P. H., Pugh, C. W., and Ratcliffe, P. J. (2001) *Science* **292,** 468–472
67. Mekhail, K., Khacho, M., Carrigan, A., Hache, R. R., Gunaratnam, L., and Lee, S. (2005) *J. Cell Biol.* **170,** 733–744
68. Mekhail, K., Khacho, M., Gunaratnam, L., and Lee, S. (2004) *Cell Cycle* **3,** 1027–1029
69. Mekhail, K., Gunaratnam, L., Bonicalzi, M. E., and Lee, S. (2004) *Nat. Cell Biol.* **6,** 642–647
70. Hofmann, B., Hecht, H. J., and Flohe, L. (2002) *Biol. Chem.* **383,** 347–364

Supplementary data

Table S1. Primers used in the study.

1. Primers used for mutating CxxU motif
Forward_CxxC -5'-GCATTGTACGAGCTGTCGCGTGTACGGCCGCCATGC-3'
Reverse_CxxC -5'-GCATGGCGGCCGTACACGCGACAGCTCGTACAATGC-3'
Forward_CxxS -5'-GCATTGTACGAGCTCACGCGTGTACGGCCGCCATGC-3'
Reverse_CxxS -5'-GCATGGCGGCCGTTGACGCGACAGCTCGTACAATGC-3'
Forward_SxxC -5'-CCGTGGTCATTGAGCATAGTACGAGCTGTCGCGTGTAC3'
Reverse_SxxC -5'-CCGTACACGCGACAGCTCGTACTATGCTCAATGACCAC3'
2. Primers used to mutate arginines 6 and 8 to serines:
ATGGCCCCCCACGGAAGTAAGAGTAAGGCGGGGGCCGCG
CGCGGCCCCCGCCTTACTCTTACTTCCGTGGGGGGCCAT
3. Primers used for cloning into pET:
Forward -5'-TAAACGA CAT*ATGGCCCCCCACGGAAG AAAG*-3'
Reverse -5'-CTACC'TCGAG*TTATGAAAGGTACTTCTTCAATTC*-3'
4. Primers used for cloning into pEGFP-N2:
Forward -5'-CTAAGAATTC *ATGGCCCCCCACGGAAGAAAG*-3'
Reverse 1 -5'-CGTAGGATCCA*TGAAAGGTACTTCTTCAATTCTTCAACC*-3'
Reverse 2 -5'-CGTAGGATCC*TCGCGCCCTCCGCCAGTTTCTC*-3'
5. Primers used for cloning into pEGFP-C3:
Forward -5'-CTACCTCGAG*TTATGAAAGGTACTTCTTCAATTC*-3'
Reverse 1 -5'-GCAAGGATCCGAAACACTTTTACAATGTCTTGC-3'
Reverse 2 -5'-CTACC'TCGAG*TTATGAAAGGTACTTCTTCAATTC*-3'

Fig. S1

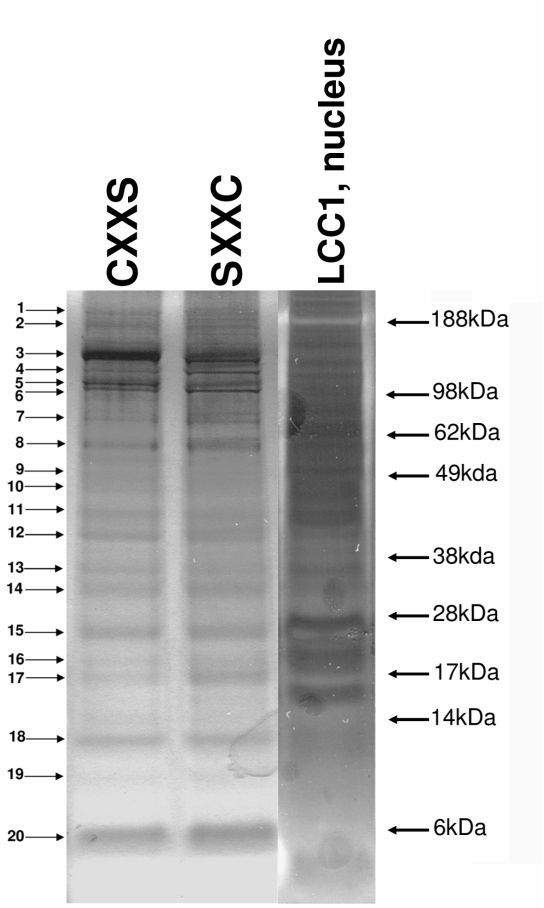
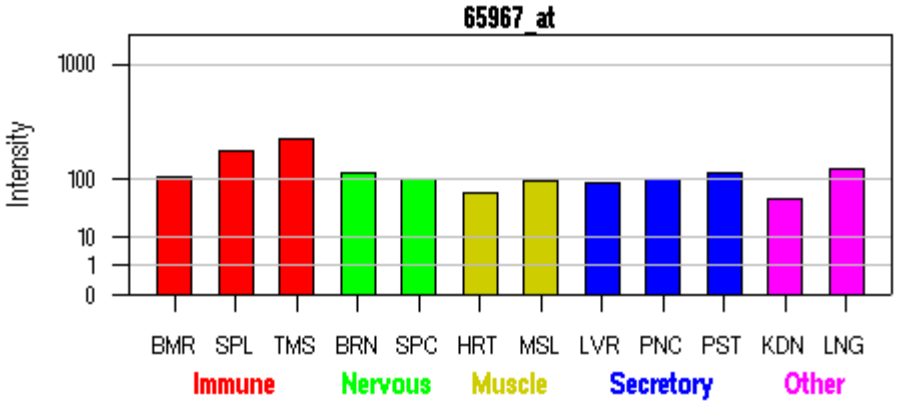


Fig. S1. Identification of SelH target proteins in LCC1 cells. This silver stained SDS PAGE gel shows nuclear proteins from LCC1 cells enriched on SxxC-SelH and SxxC-SelH resins. Due to close similarity of targets identified in each of the variants, the targets are labeled only for one mutant. Migration of molecular weight standards is shown on the right. The following proteins were identified by MS/MS sequencing of protein bands extracted from the gel: 1 - pre-mRNA processing protein 8, 2 - putative pre-mRNA splicing factor RNA helicase (DEAH box protein 15), 3 - nucleolin, 4 - nuclear cap binding protein subunit 1, 5 - PTB-associated splicing factor, 6 - heterogeneous nuclear ribonucleoprotein R, 7 - IGF-II mRNA-binding protein 2, 8 - nucleolar protein, 9 - glyoxylate reductase/hydroxypyruvate reductase, 10 - poly(rC) binding protein 1 and heterogeneous nuclear ribonucleoprotein X, 11 - Rbbp5 protein, 12 - spliceosome-associated protein SAP 62, 13 - nucleophosmin 1 and nucleolar protein NO38, 14 - heterogeneous nuclear ribonucleoprotein A1 (helix-destabilizing protein), 15 - peroxiredoxin 1, natural killer-enhancing factor A and proliferation-associated gene, 16 - peroxiredoxin 2, 17 -60S ribosomal protein L18, 18 - glutaredoxin, 19 - thioredoxin 1, and 20 - ribosomal protein S28.

Fig. S2

A



B

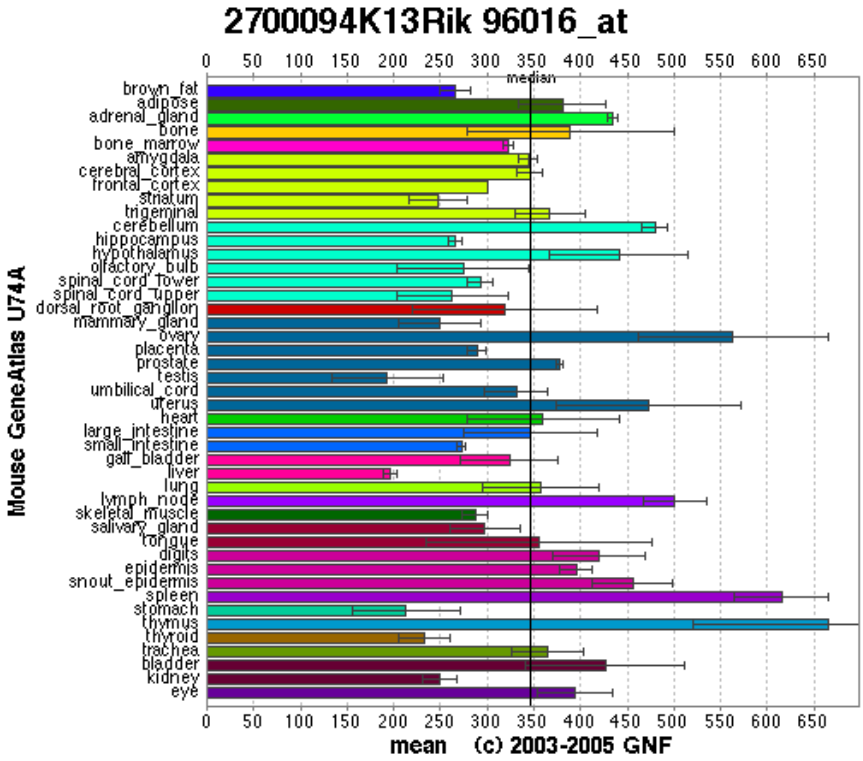


Fig. S2. Computational analysis of SelH mRNA expression profiles using publicly available expression data. A. GeneNote database contains data on whole genome mRNA expression profiles in normal human tissues. Two organs that showed the highest SelH mRNA levels in humans are thymus and spleen. **B**. Gene Atlas database contains data on whole genome mRNA expression profiles in normal human and mouse tissues. Similar to human, two organs with the highest SelH mRNA levels in mouse are thymus and spleen.

Fig. S3

Library	Total Tags in Library	SelH Tags per 200,000
Embryonic stem cell HES3 normal p16 CL SHE10	205353	105
Bone marrow normal AP CD34+/CD38-/lin-	86391	90
Embryonic stem cell H9 normal p38 CL SHES1	151735	90
Embryonic stem cell H1 normal p54 CL SHE16	218214	82
Liver cholangiocarcinoma B K2D	46853	76
Breast carcinoma CL ZR75 1 untreated	32303	74
Bone marrow normal AP CD34+/CD38+/lin+	81595	73
Breast carcinoma CL MCF7 control 0h	59877	73
Emryonic stem cell H14 normal p22 CL SHE14	212170	73
Lung adenocarcinoma B 1	36256	71
Colon carcinoma CL Hct116 wildtype p53 Normal Oxygen	80398	69
Prostate carcinoma CL LNCaP-T	43542	68
Colon carcinoma CL Hct116 p53 knockout Anoxia	79053	65
Pancreas carcinoma CL ASPC	31224	64

Fig. S3. SAGE Genie *in silico* expression analysis of SelH expression.

Supplementary References

Shmueli, O., Horn-Saban, S., Chalifa-Caspi, V., Shmoish, M., Ophir, R., Benjamin-Rodrig, H., Safran, M., Domany, E. and Lancet, D. (2003) *C. R. Biol*. **326**, 1067-72.

Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M.P., Walker, J.R. and Hogenesch, J.B. (2004) *Proc. Natl. Acad. Sci. U S A*. **101**, 6062-7