Supporting information

Redox active motifs in selenoproteins

Fei Li, Patricia B. Lutz, Yuliya Pepelyayeva, Elias S.J. Arnér, Craig A. Bayse and Sharon Rozovsky

MATERIALS AND METHODS

Materials. Elemental ^{77}selenium (99.20%) was purchased from Isoflex USA (San Francisco, CA). Enzymes used for molecular biology were acquired from New England Biolabs (Ipswich, MA). The pMHTDelta238 plasmid expressing Tobacco Etch Virus (TEV) protease fused to the cytoplasmic maltose binding protein (MBP) (1) was purchased from the Protein Structure Initiative: Biology Materials Repository (2). Chromatography media was supplied by GE Healthcare Bio-Sciences Corporation (Pittsburgh, PA) and Qiagen (Hilden, Germany). All other chemicals and reagents were supplied by Sigma-Aldrich (St. Louis, MO), Acros Organics (Geel, Belgium), and GoldBio (St. Louis, MO).

Genetic incorporation of Sec and ^{77}Se-Sec in proteins. Previously, the methods used for labeling of biological samples with ^{77}Se were solid state peptide synthesis, chemical reactions with small molecules (3), ^{77}Se enriched animal feed (4), or the substitution of sulfur to ^{77}Se in all sulfur-containing amino acids (5). Our studies required the specific incorporation of Sec at one single defined position relative to a partner Cys in a protein. Hence, we have here developed an isotopic enrichment protocol that is used to incorporate ^{77}Sec via heterologous expression in *E. coli* by manipulating its native selenium incorporation machinery (6). Sec is specifically incorporated during translation by decoding of a predefined UGA codon upstream of the actual termination codon, utilizing a dedicated RNA stem loop **se**leno**c**ysteine **i**nsertion **s**equence (SECIS) that is utilized by a Sec-specific translation machinery (7, 8). The key for expression of selenoproteins in *E. coli* is the coexpression of a desired target selenoprotein transcript carrying the UGA codon and a functional SECIS element, interacting with the native bacterial selenoprotein production proteins (6, 9). Using this approach, selenoproteins can be expressed in *E. coli* at suitable yields for NMR experiments (10, 11). To incorporate the ^{77}Se isotope we employed a procedure reported by the Stadtman research group for partial ^{77}Se enrichment of formate dehydrogenase H expressed in *E. coli* (12) in which ^{77}selenite was supplied directly in the growth media.

As a scaffold protein to carry the different Sec-containing redox motifs we selected Trx1 from *E. coli* with its active site Cys residues changed to Ser, which yielded a small, stable, redox inert globular protein devoid of additional Cys residues or other redox active motifs. For this, we first had an N-terminal hexahistidine-tagged *E. coli* Trx (trxA) gene synthesized by DNA2.0 (Menlo Park, CA), fused to a C-terminal tetrapeptide GGCUG sequence, followed by a non-encoding variant of the *E. coli* formate dehydrogenase (fdhH) SECIS element, which is commonly used for genetic insertion of Sec into proteins (13). A TEV cleavage site was introduced before Trx. The synthetic gene was cloned into the bacterial expression vector pET28a (Novagen, San Diego, CA). This vector was subsequently used as template for construction of the other redox motifs using PCR-based mutagenesis (see Figure S1 for sequences of all constructs). For protein expression, the appropriate plasmid was cotransformed into *E. coli* BL21(DE3), with the pSUABC plasmid expressing *E. coli* SelA, SelB, and SelC under the control of their endogenous promoters in order to increase Sec incorporation (6).

Cells were grown in a defined media adapted from Studier's MDAG media (14). The media contained: 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.2x metals, 0.5% glucose (27.8 mM), 0.25 mM aspartate (18.8 mM) and 200 μg/ml each of 17 canonical amino acids (not Cys, Met or Tyr). 1x concentration trace metals solutions contained 50 μM $FeCl_3$, 20 μM $CaCl_2$, 10 μM $MnCl_2$, 10 μM $ZnSO_4$, 2 μM $CoCl_2$, 2 μM $CuCl_2$, 2 μM $NiCl_2$, 2 μM Na_2MoO_4, and 2 μM H_3BO_3. Cells were grown at 37 °C, with good aeration and an antibiotics selection of 50 μg/mL kanamycin sulfate and 34 μg/mL chloramphenicol. When the Optical Density (OD) at 600 nm reached 2.0, the temperature was lowered to 18 °C and the cells were allowed to shake at the lower temperature for an additional hour. Protein expression with all variants was induced with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) when the OD at 600 nm reached 2.4. At induction, the cells were supplied with 5 μM ^{77}selenite and 100 μg/ml L-Cys. To generate ^{77}selenite, the appropriate amount of elemental ^{77}Se was oxidized to ^{77}selenite in a minimal volume of nitric acid (12). The identity of the expected selenite was confirmed by solution ^{77}Se NMR. The volume was kept to a minimum, in order to reduce the amount of acid added to the growth media. The cell paste, following an 18 h expression period, was resuspended in 20 mM Tris (pH 8.0) and 500 mM NaCl

(IMAC buffer) supplemented with 1 mM PMSF and 0.5 mM benzamidine and lysed using a high pressure homogenizer (EmulsiFlex-C5, Avestin, Ottawa, Canada). Cell debris was removed by centrifugation at 15000 g for 1 h at 4 °C. Clear supernatant was loaded onto a HisTrap FF column whereupon the column was washed with IMAC buffer, supplemented with 20 mM imidazole. The target protein was eluted using IMAC buffer with a linear gradient of 20-350 mM imidiazole. Following elution, the buffer was exchanged into 50 mM Tris (pH 8.0), 1 mM DTT, 0.5 mM EDTA by dialysis and the His_6 tag cleaved with His-tagged TEV protease, at 4 °C for 18 h (1, 15). After cleavage, the protein N-terminal retained a Gly at position 1. The protein was passed once more on a HisTrap column in the presence of 20 mM imidiazole concentrations to remove the TEV protease and uncleaved protein, both of which bind the column while the cleaved protein does not. The cleaved protein was then dialyzed against 20 mM Tris, 1 mM DTT, and 1 mM EDTA, pH 8.4 at 4 °C. The protein was loaded on a fast flow Q-sepharose and eluted with a salt gradient between 0-400 mM NaCl over 15 column volumes. The full-length Sec-containing protein eluted around 150 mM NaCl. The ratio of the full-length and truncated proteins in the individual fractions was assayed by LC-ESI-TOF mass spectrometry. Fractions containing at least a ratio of 0.2 full-length to truncated protein were combined and dialyzed again against 20 mM diethanolamine (pH 9.0), 1 mM DTT, and 1 mM EDTA at 4 °C. The protein was loaded on a mono-Q and eluted with a salt gradient between 0-400 mM NaCl over 10 column volumes. The full-length protein eluted around 150 mM NaCl. Fractions containing full-length Sec-containing protein were combined and dialyzed to 20 mM Tris (pH 8.0), 50 mM NaCl, and 1 mM EDTA at 4 °C. The protein was then concentrated to about 100 mg/ml, flash frozen in liquid N_2, and stored at -80 °C until further use. Protein purity, as determined by 18% SDS-PAGE Tris-glycine gels, was higher than 98%. Total yield for all protein species was about 200 mg/L growth media, and the yield of the full-length protein was ≈8 mg/L growth media. We tested extensively the yield of full-length Sec-containing protein as a factor of selenium concentration in media, growth temperature, induction level, auto-induction methods, cell line, expression length, and purification schemes. The maximum yield reached for this construct was about 8×10^{-4} mol full-length protein per liter growth media, following the here described protocol. The presence of selenium was evaluated using LC-ESI-QTOF mass spectrometry (Figure S2) and inductively coupled plasma (ICP) spectroscopy. The percentage of full-length selenium-containing protein varied from 50 to 75% in different batches of final samples.

Expression and purification of Cys variants. Expression constructs for the Cys variants were prepared by site-directed mutagenesis of the opal codon TGA into the cysteine codon TGC in the selenoproteins' constructs. (Figure S1). The proteins were expressed in BL21(DE3) in LB media (without the pSUABC vector). Protein expression was induced with IPTG when the OD at 600 nm reached 0.7. The purification protocol was as described above for the Sec variants.

Electrospray Ionization Mass Spectrometry. Mass spectra were obtained using a QTOF Ultima (Waters), operating under positive electrospray ionization (+ESI) mode and connected to an LC-20AD (Shimadzu, Kyoto, Japan). Protein samples were separated from small molecules by reverse phase chromatography on a C4 column (Waters XBridge BEH300), using an acetonitrile gradient from 30-71.4% and 0.1% formic acid as the mobile phase for 25 min, at a flow rate of 0.2 ml/min at room temperature. Data were acquired from m/z 350 to 2500, at a rate of 1 sec/scan.

Inductively Coupled Plasma (ICP) Spectrometry. ICP spectrometry (Thermo IRIS Intrepid II XSP Dual View) was used to determine the ratio of elemental sulfur to selenium in the protein samples. The instrument was calibrated using customized reference standard from AccuStandard (New Haven, CT) containing 2 µg/ml of elemental sulfur and selenium in 2% nitric acid.

NMR samples. The NMR tube was either Shigemi tube or air-tight tube manufactured by Wilmad-LabGlass. The samples were stable and could be maintained under aerobic conditions during acquisitions. While we originally carried out NMR experiments in air-tight NMR tubes under N_2 environment, it became apparent during our work that the samples were stable at 20 °C even when exposed to air. We found no evidence for selenium oxidation or elimination by either mass

spectrometry or NMR before or after the experiments. We found no differences between samples acquired with different protein concentrations and hence viscosity.

Measurements of spin lattice relaxation times. Measurements of the spin lattice relaxation times were carried out using inversion recovery experiments. These experiments were carried in full only once, due to their demands on spectrometer time. The measurement error was estimated from the signal to noise to be no higher than 5% (the upper bound in area measurement). We have also used the upper and lower bounds of the Chi-square fit to the measured values as well as their upper and lower error margins.

Measurements of Redox Potentials by ESI Mass Spectrometry. A 0.5 ml sample, which contained 0.2 μM fully oxidized protein, was dialyzed in buffer containing 50 mM sodium phosphate (pH 7.0), 1 mM EDTA, as well as various ratios of either reduced DTT (DTT_{red}) and oxidized DTT (DTT_{ox}), or glutathione (GSH) and oxidized glutathione (GSSG) at a total DTT_{red} + DTT_{ox} or GSH+GSSG concentration of 20 mM. The ratio of GSH/GSSG or DTT_{red}/DTT_{ox} was used to poise the redox potential of the buffer. To reduce oxidation, the protein solutions were degassed and subsequently flushed with nitrogen. The samples were incubated at 25 °C for 18 hours to ensure full equilibration. The equilibrated reaction mixtures were quenched by adding ice-cold 100% W/V trichloroacetic acid (TCA) to a final concentration of 20% and were mixed immediately. The quenched reaction mixture was then spun for 10 min at 16110 g. The supernatant was decanted and the pellet was washed with 0.25 ml ice-cold acetone twice and spun at 16110 g for 10 min at 4 °C after each wash. After the final acetone wash, the pellet was dried by exposing it to air for 10 min and was then re-suspended in 50 μl of 15 mM iodoacetamide solution. 10-20 μl of the re-suspended sample was used to acquire mass spectrum. Experiments were repeated three times, using two independent preparations, for each protein. Following deconvolution, the fractional ratio of reduced to oxidized protein was determined from the presence of mass increments of 116 Da (transfer of two acetamide groups). The percentage of reduced and alkylated protein to oxidized protein was determined by integrating the relevant peaks area in Excel (Microsoft, WA). Identical results were obtained when using N-ethylmaleimide (NEM) for labeling.

The reaction of the protein with the DTT_{ox}/DTT_{red} or GSSG/GSH redox pair (reactions 1 and 2) and the corresponding equilibrium constant (K_{eq}) is given in Equations 1 and 2:

$Protein_{red} + DTT_{ox} \leftrightarrow Protein_{ox} + DTT_{red}$ Reaction 1

$Protein_{red} + GSSG \leftrightarrow Protein_{ox} + 2GSH$ Reaction 2

$$K_{eq} = \frac{[Protein_{ox}][DTT_{red}]}{[Protein_{red}][DTT_{ox}]} \quad \text{(Eq. 1)}$$

$$K_{eq} = \frac{[Protein_{ox}][GSH]^2}{[Protein_{red}][GSSG]} \quad \text{(Eq. 2)}$$

The redox potential of the protein at pH 7.0 and 25 °C (E_0'(Protein)) was then calculated from the Nernst equation (Equations 3 and 4):

$$E_0'(Protein) = E_0'(DTT_{red}/DTT_{ox}) - [\frac{RT}{nF}] \times \ln(K_{eq}) \quad \text{(Eq. 3)}$$

$$E_0'(Protein) = E_0'(GSH/GSSG) - [\frac{RT}{nF}] \times \ln(K_{eq}) \quad \text{(Eq. 4)}$$

where $E_0'(DTT_{red}/DTT_{ox})$ = -323 mV (pH 7.0 and 25 °C),(16) $E_0'(GSH/GSSG)$ = -240 mV (pH 7.0 and 25 °C),(17) R is the gas constant (8.315 J K^{-1} mol^{-1}), T is the absolute temperature (298 K), n is the number of electrons transferred in the reaction (here n=2), and F is the Faraday constant (9.649×10^4 C mol^{-1}). Plots and curve fits were made with Origin (OriginLab Corp., Northampton, MA). The error bars in Figures S8 and S9 represent the range of measurements; that is, the highest and lowest values recorded in the series. The measurement error was estimated to be 1 mV by fitting extreme curves to the minimal and maximal fractions of reduced protein at each point.

Measurements of Redox Potential by Gel Shift Assay. The same sample preparation described above for ESI mass spectrometry-based assays was employed for gel shift assays and visualization by SDS-PAGE. Following an 18 h incubation in the respective redox buffer, the protein was precipitated with 100% TCA and dried by acetone. The pellet was dissolved in 50 µl 10 mM Methyl-PEG_{24}-Maleimide ($MM(PEG)_{24}$) solution in 2x non-reducing Tricine-SDS-PAGE loading buffer. The samples were boiled for 5 min prior to loading on 16% Tricine-SDS-PAGE gels. The gels were Coomassie-stained and the bands intensities were visualized and quantified using a FluorChem Q gel imager (ProteinSimple, Santa Clara, CA). Protein bands include: oxidized full-length protein, alkylated truncated protein and alkylated full-length protein, as well as dimers of the two last forms. The presence of the oxidized truncated protein that forms an intermolecular disulfide bond can be neglected due to its redox potential (the truncated protein is fully alkylated in all redox conditions tested by ESI mass spectrometry). The ratio of reduced protein to oxidized protein was used in conjunction with the Nernst equation to calculate the redox potential.

I. Trx C32S, C35S with SCUS at the C-terminal

DNA sequence:
ATGGGACATCATCATCATCATCACAGCAGCGGCGAAAACCTGTATTTTCAGGGCAGCGATAAA
ATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTC
GTCGATTTCTGGGCAGAGTGGTCCGGTCCGTCCAAAATGATCGCCCCGATTCTGGATGAAATC
GCTGACGAATATCAGGGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACT
GCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCG
GCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCG
GGTAGCTGCTGAAGCTAATAATCGGTTGCAGGTCTGCACCAATCGTTAACTTATGCGGCC

Protein sequence prior to TEV cleavage:
MGHHHHHHSSGENLYFQGSDKIIHLTDDSFDTDVLKADGAILVDFWAEWSGPSKMIAPILDEIADEY
QGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSCUS

Protein sequence following TEV cleavage:
GSDKIIHLTDDSFDTDVLKADGAILVDFWAEWSGPSKMIAPILDEIADEYQGKLTVAKLNIDQNPGTA
PKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSCUS

II. Trx C32S, C35S with GCUG at the C-terminal

DNA sequence:
ATGGGACATCATCATCATCATCACAGCAGCGGCGAAAACCTGTATTTTCAGGGCAGCGATAAA
ATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTC
GTCGATTTCTGGGCAGAGTGGTCCGGTCCGTCCAAAATGATCGCCCCGATTCTGGATGAAATC
GCTGACGAATATCAGGGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACT
GCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCG
GCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCG
GGTGGTTGCTGAGGTTAATAATCGGTTGCAGGTCTGCACCAATCGTTAGCCTATGCGGCC

Protein sequence prior to TEV cleavage:
MGHHHHHHSSGENLYFQGSDKIIHLTDDSFDTDVLKADGAILVDFWAEWSGPSKMIAPILDEIADEY
QGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGGCUG

Protein sequence following TEV cleavage:
GSDKIIHLTDDSFDTDVLKADGAILVDFWAEWSGPSKMIAPILDEIADEYQGKLTVAKLNIDQNPGTA
PKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGGCUG

III. Trx C32S, C35S with GCAUG at the C-terminal

DNA sequence:
ATGGGACATCATCATCATCATCACAGCAGCGGCGAAAACCTGTATTTTCAGGGCAGCGATAAA
ATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTC
GTCGATTTCTGGGCAGAGTGGTCCGGTCCGTCCAAAATGATCGCCCCGATTCTGGATGAAATC
GCTGACGAATATCAGGGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACT
GCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCG
GCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCG
GGTGGTTGCGCGTGAGGTTAATAATCGGTTGCAGGTCTGCACCAATCGTTAGCCTATGCGGCC

Protein sequence prior to TEV cleavage:
MGHHHHHHSSGENLYFQGSDKIIHLTDDSFDTDVLKADGAILVDFWAEWSGPSKMIAPILDEIADEY
QGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGGCAUG

Protein sequence following TEV cleavage:
GSDKIIHLTDDSFDTDVLKADGAILVDFWAEWSGPSKMIAPILDEIADEYQGKLTVAKLNIDQNPGTA
PKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGGCAUG

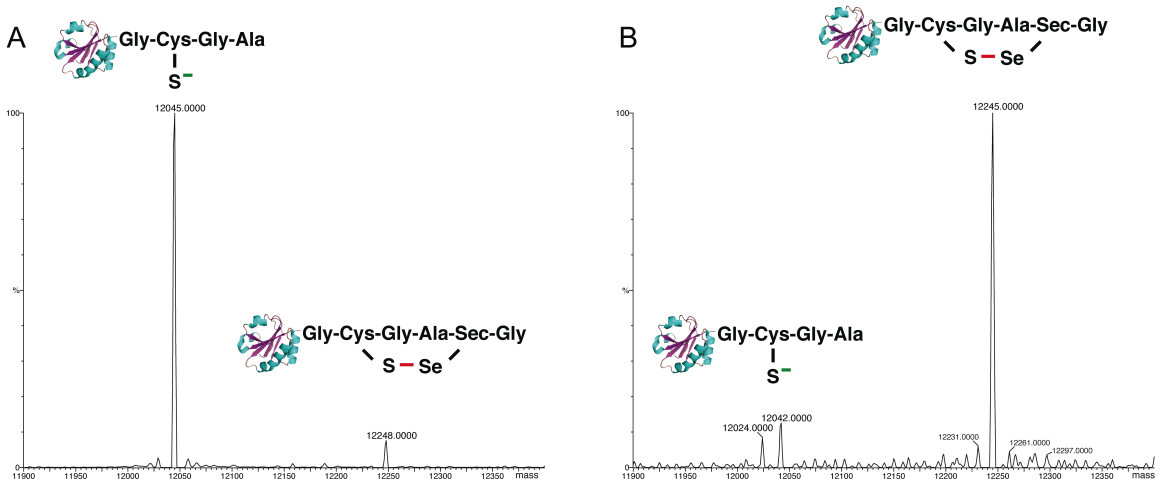
IV. Trx C32S, C35S with GCGAUG at the C-terminal

DNA sequence:
ATGGGACATCATCATCATCATCACAGCAGCGGCGAAAACCTGTATTTTCAGGGCAGCGATAAA
ATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTC
GTCGATTTCTGGGCAGAGTGGTCCGGTCCGTCCAAAATGATCGCCCCGATTCTGGATGAAATC
GCTGACGAATATCAGGGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACT
GCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCG
GCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGGCG
GGTGGGTTGCGGTGCGTGAGGTTAATAATCGGTTGCAGGTCTGCACCAATCGTTAGCCTATGC
GGCC

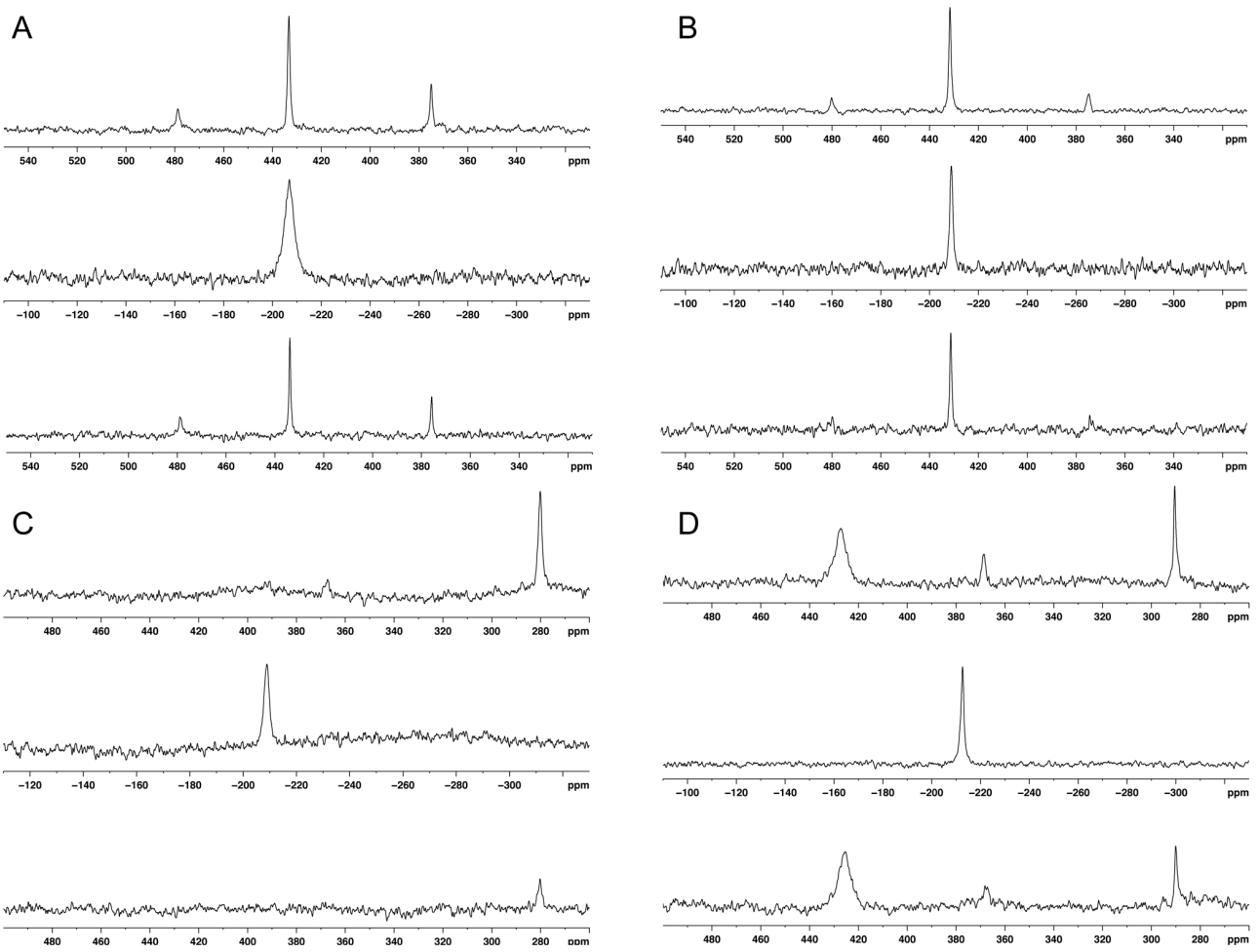
Protein sequence prior to TEV cleavage:
MGHHHHHHSSGENLYFQGSDKIIHLTDDSFDTDVLKADGAILVDFWAEWSGPSKMIAPILDEIADEY
QGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGGCGAUG

Protein sequence following TEV cleavage:
GSDKIIHLTDDSFDTDVLKADGAILVDFWAEWSGPSKMIAPILDEIADEYQGKLTVAKLNIDQNPGTA
PKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGGCGAUG

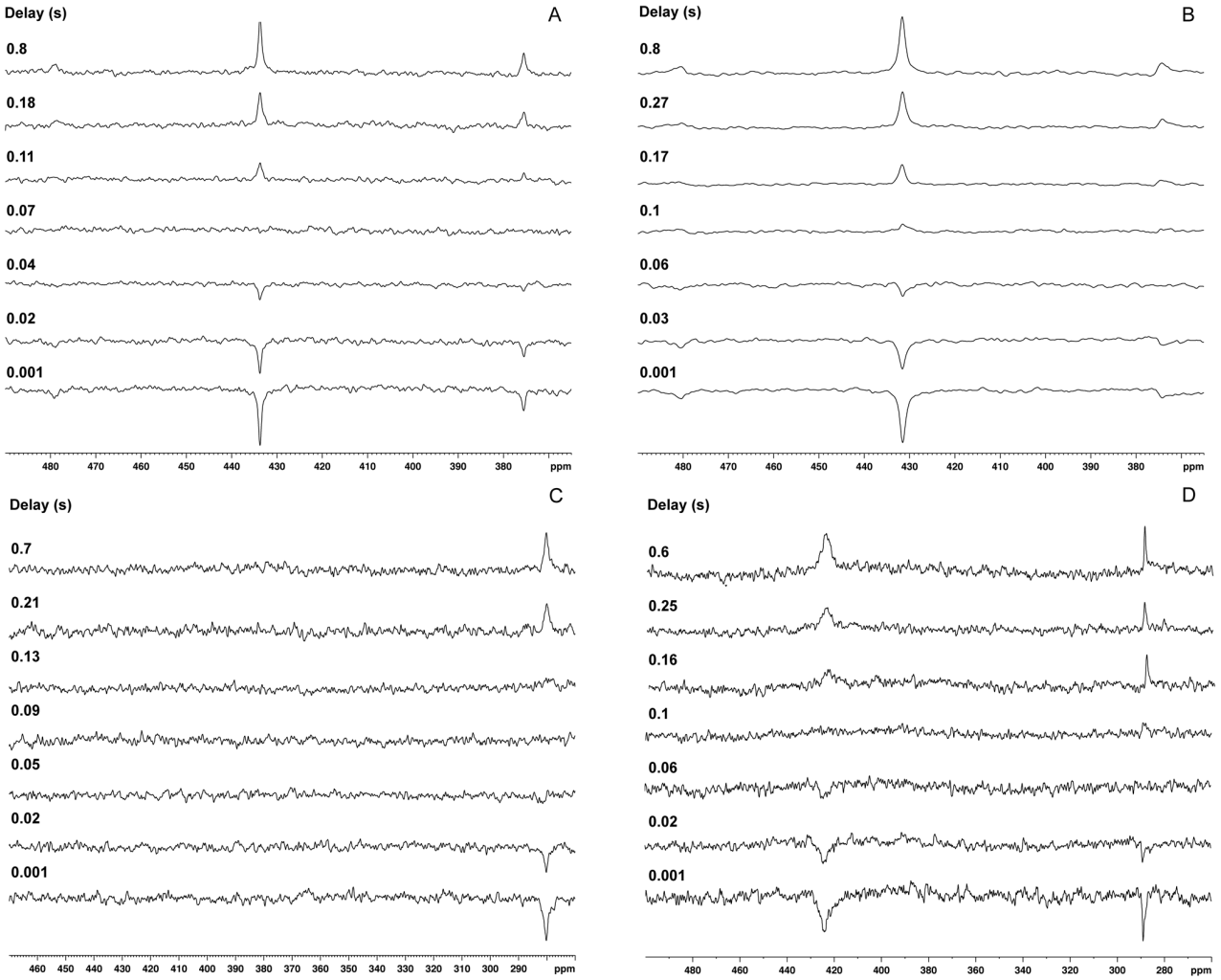
Supporting Figure S1. DNA and amino acids information for the constructs used in these studies. The information is color coded according to the following scheme:
Blue: His_6 tag
Green: TEV protease cleavage site
Purple: *E. coli* Trx C32S C35C (when fused with a His_6 tag these residues are numbered as C51S C54S)
Red: Inserted redox motif
Underline: SECIS element



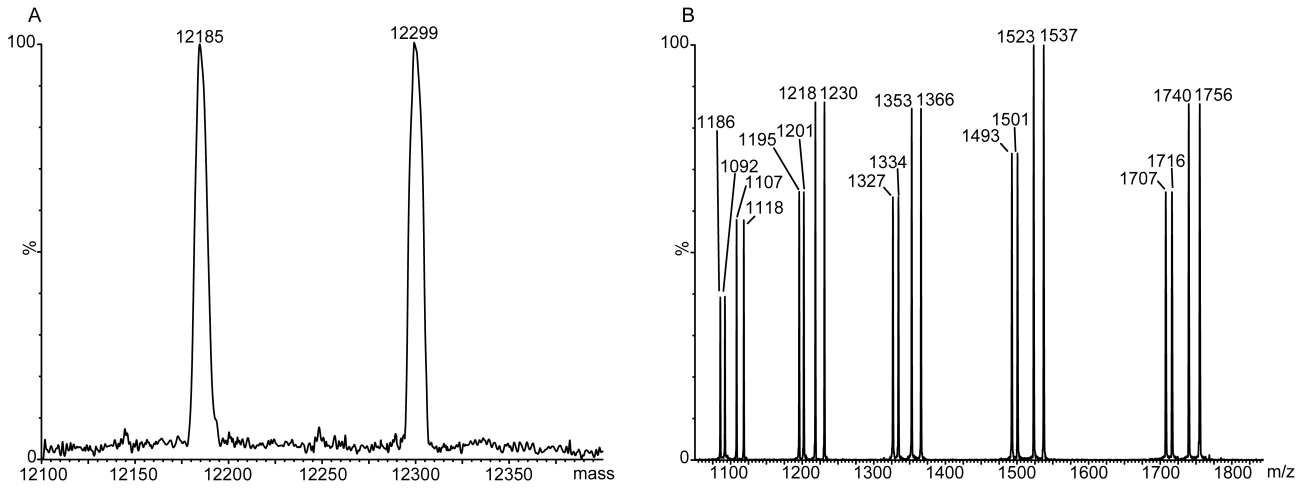
Supporting Figure S2. Mass spectrometry of a representative GCGAUG selenoprotein purification. a) Mass spectrometry of GCGAUG following purification with IMAC and subsequent cleavage of the N-terminal His_6 tag by TEV protease, b) Mass spectrometry of GCGAUG following purification by ion exchange. Additional peaks for the truncated form arise from binding of a sodium atom. The relative ratio of full length to truncated forms of the protein was confirmed by measurements of S/Se ratio by inductively coupled plasma atomic emission spectrometry, as described in the text.



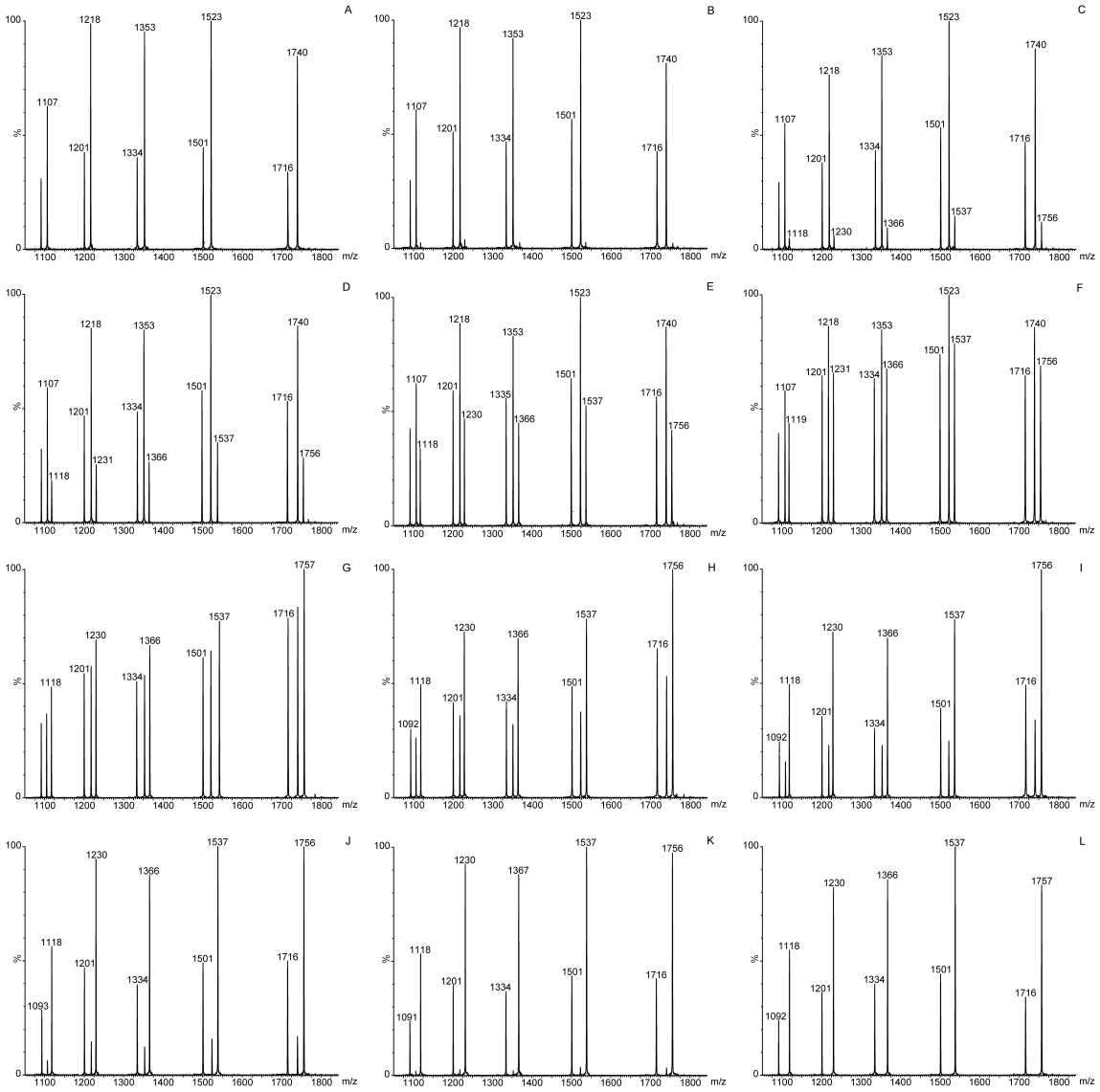
Supporting Figure S3. ^{77}Se spectra of Sec-containing motifs before reduction (top), in the reduced state (middle) and following reoxidation once the reducing agent, DTT, was removed. a) SCUS. b) GCUG. c) GCAUG. d) GCGAUG. The results demonstrate that when the reducing agent was removed by dialysis, the NMR spectrum was, within signal to noise, similar to that prior to reduction. In general, the resonance positions and heights were well reproducible across different sample preparations except the 370 ppm resonance of GCAUG and GCGAUG.



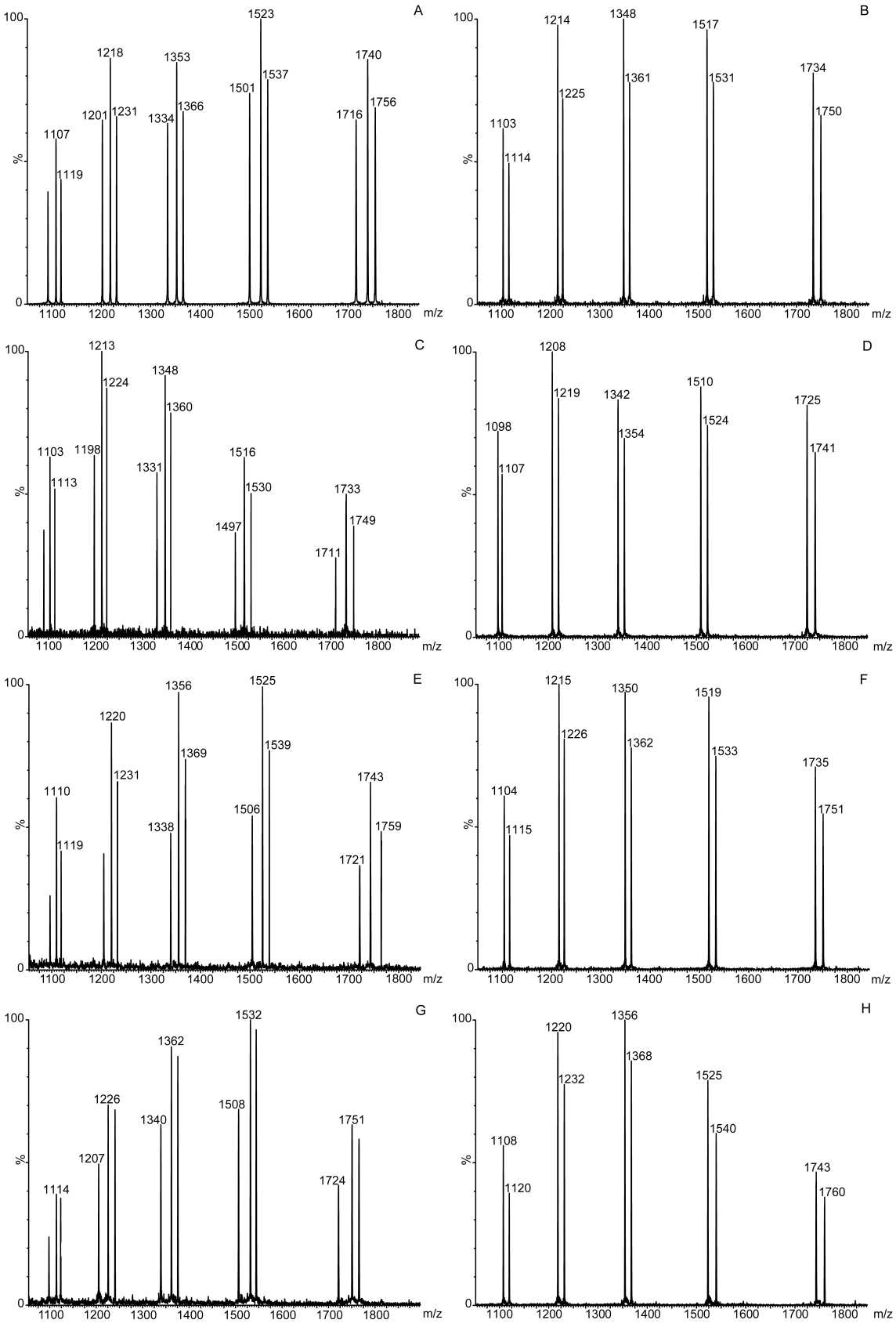
Supporting Figure S4. Inversion-recovery experiments performed on oxidized selenenylsulfide motifs at 20 °C and 14.1 T. a) SCUS. b) GCUG. c) GCAUG. d) GCGAUG.



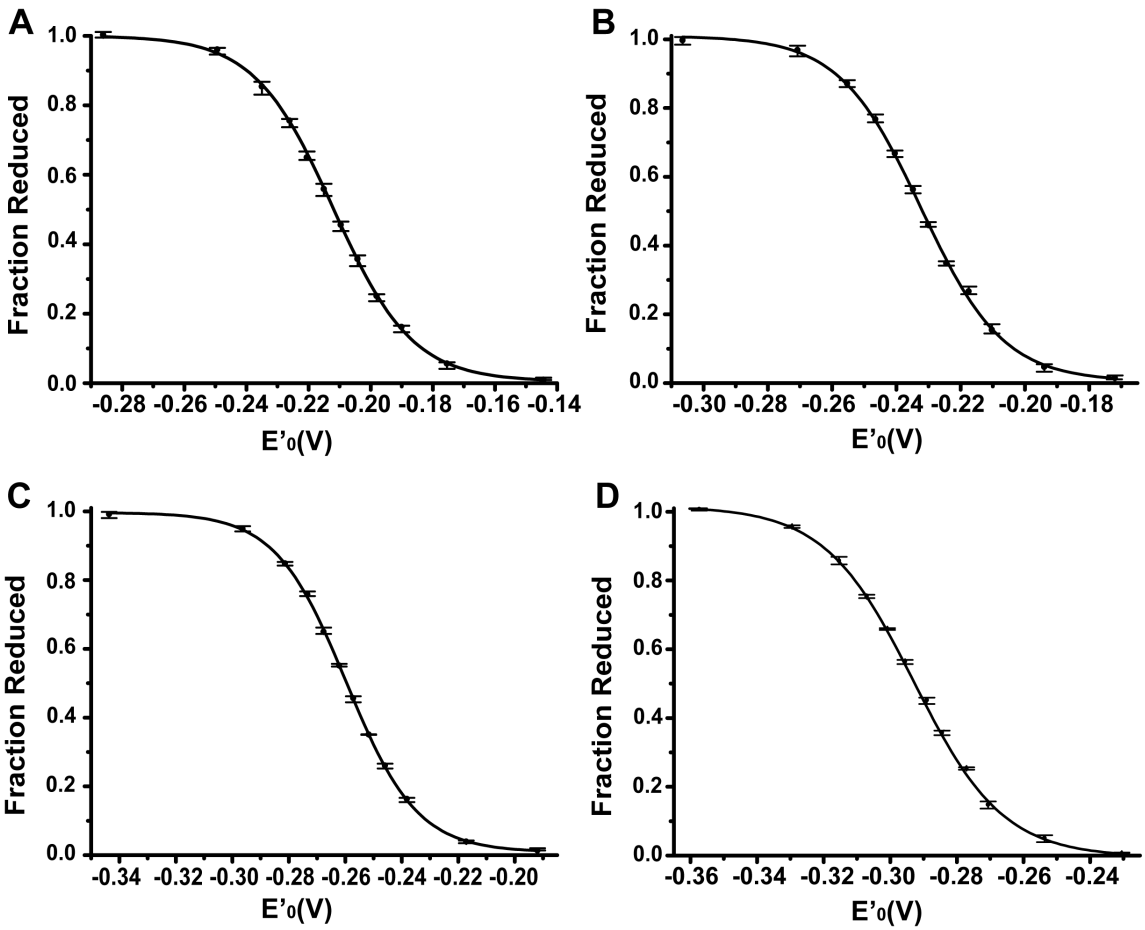
Supporting Figure S5. Ionization efficiencies of different protein forms by ESI mass spectrometry. A representative SCUS sample in which the oxidized and the alkylated forms of the protein were mixed at a 1:1 molar ratio. (A) m/z spectrum. (B) Deconvoluted spectrum. The integrated area of the truncated form (11948 Da), alkylated truncated form (12006 Da), oxidized SCUS (12185 Da) and alkylated SCUS (12299 Da) is in agreement to the ratio in which the forms were mixed. Hence, the mass spectrum reflects the relative ratios of the different forms of proteins in the sample.



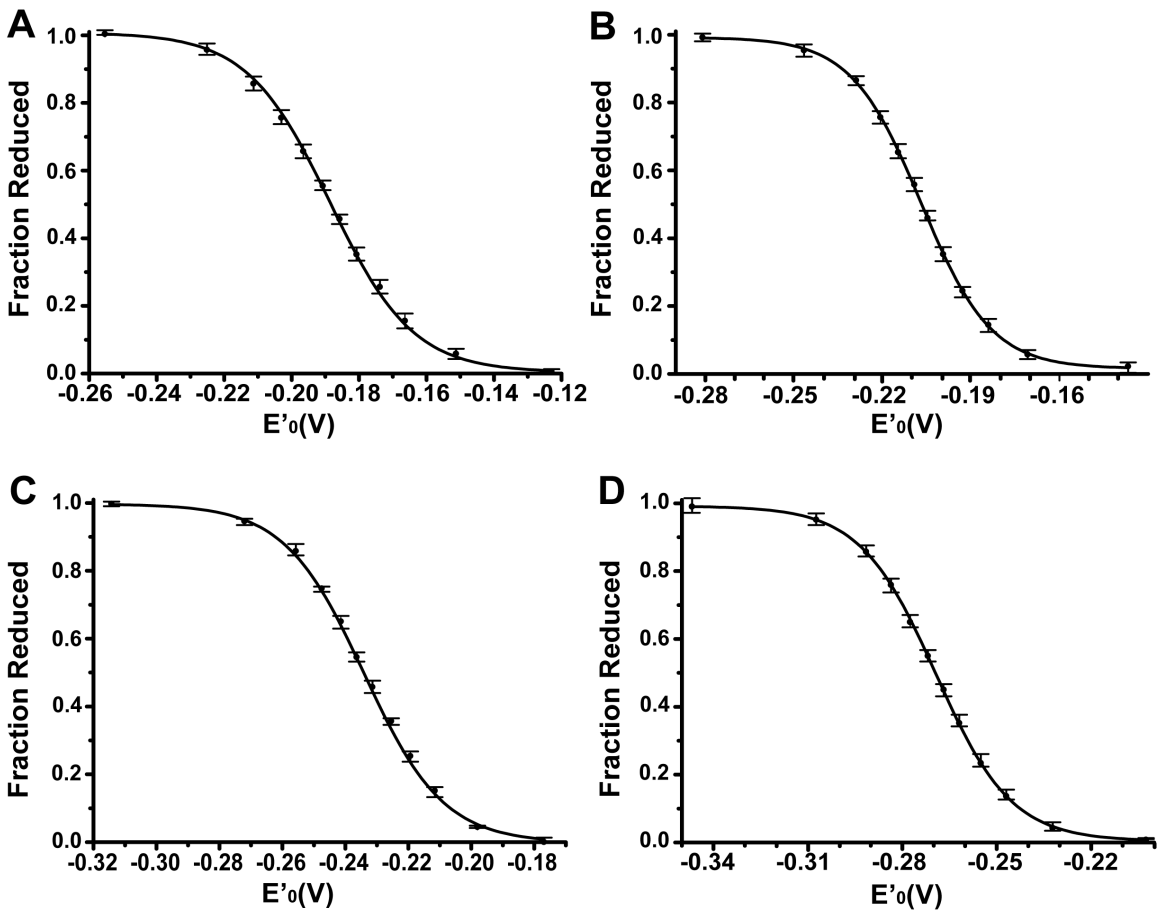
Supporting Figure S6. Representative titration curve used to measure the redox potential of the proteins (in here the example shows SCUS). ESI mass spectra of SCUS following incubation in various redox buffers and subsequent alkylation. The individual panels show the charge state distribution of multiply charged ions when the redox potential of the buffer was poised at (A) -147 mV, (B) -178 mV, (C) -193 mV, (D) -201 mV, (E) -207 mV, (F) -212 mV, (G) -218 mV, (H) -223 mV, (I) -229 mV, (J) -237 mV, (K) -252 mV, (L) -289 mV.



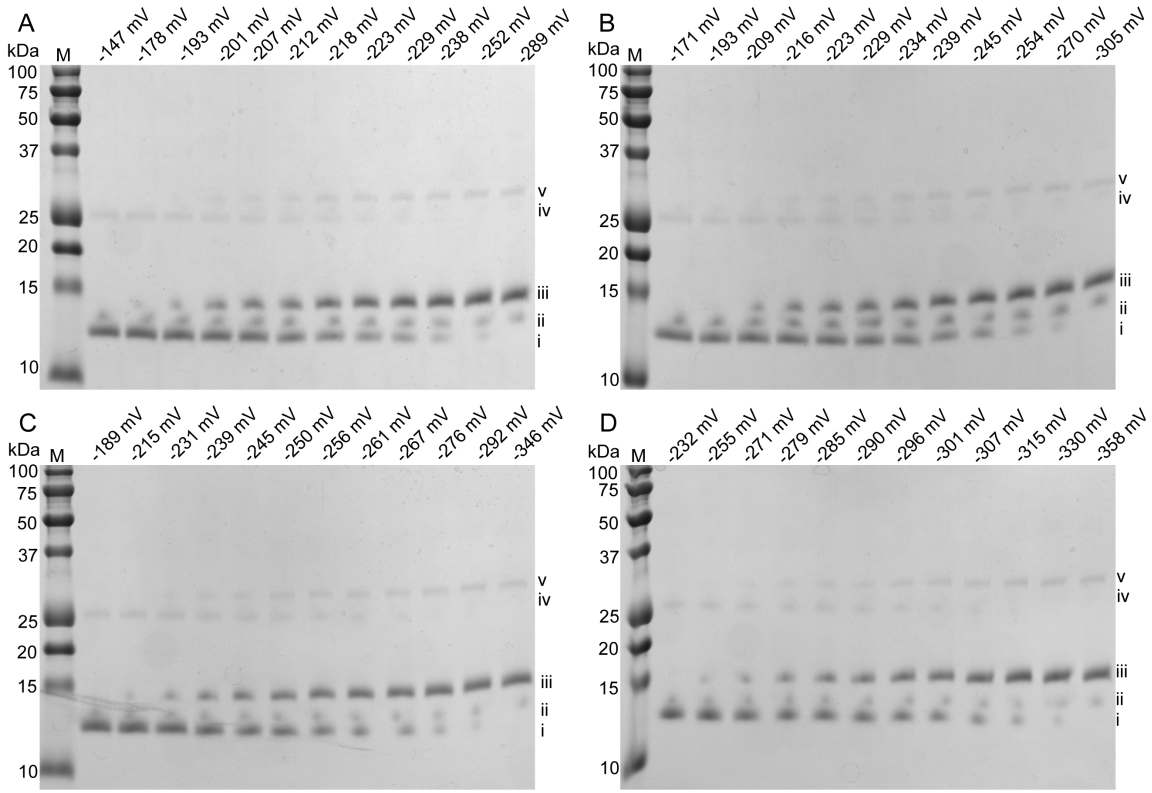
Supporting Figure S7. Representative ESI mass spectra of all proteins. Each panel shows the charge state distribution of multiply charged ions $[M+7H]^{7+}$ to $[M+11H]^{11+}$ from m/z 1050 to 1850. (A) SCUS poised at -212 mV, (B) SCCS poised at -187 mV, (C) GCUG poised at -229 mV, (D) GCCG poised at -205 mV, (E) GCAUG poised at -251 mV, (F) GCACG poised at -233 mV, (G) GCGAUG poised at -293 mV, (H) GCGACG poised at -266 mV.



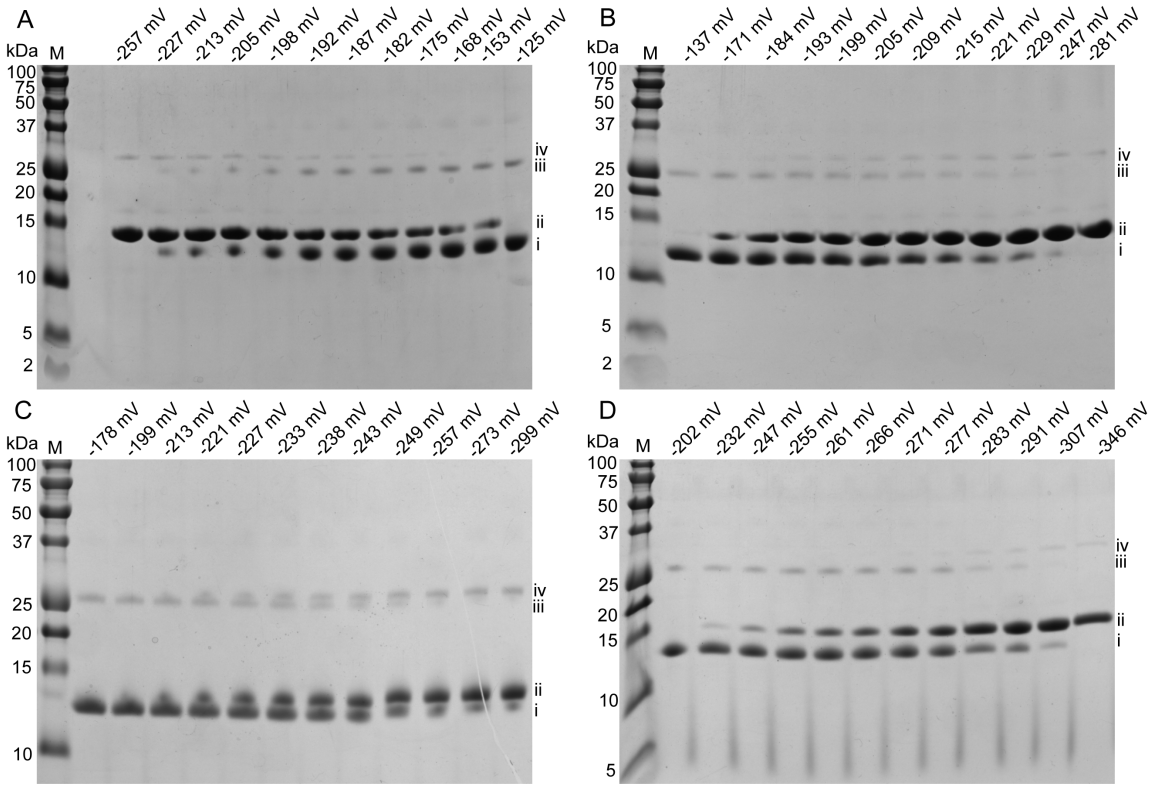
Supporting Figure S8. Determination of the redox midpoint potential of Sec-containing proteins. (A) SCUS, (B) GCUG, (C) GCAUG, (D) GCGAUG. The error bars represent the range of measurements (that is, highest and lowest measurements) among three repetitions, using two independent protein preparations.



Supporting Figure S9. Determination of the redox midpoint potential of Cys-containing proteins. (A) SCCS. (B) GCCG. (C) GCCUG. (D) GCGACG. The error bars represent the range of measurements (that is, highest and lowest measurements) among three repetitions, using two independent protein preparations.



Supporting Figure S10. Gel shift assays of Sec-containing proteins using alkylation with MM$(PEG)_{24}$. The proteins were separated by SDS-PAGE. The protein bands are labeled to the right of each image to simplify identification. M denotes molecular weight markers (biorad Precision Plus Protein™ Dual Xtra Standards). i denotes oxidized protein; ii alkylated truncated protein; iii alkylated full-length protein; iv dimer of oxidized protein; v dimer of alkylated full-length protein. (A) SCUS. (B) GCUG. (C) GCAUG. (D) GCGAUG.



Supporting Figure S11. Gel shift assays of Cys-containing proteins using alkylation with MM(PEG)$_{24}$. The proteins were separated by SDS-PAGE. The protein bands are labeled to the right of each image to simplify identification. M denotes molecular weight markers (biorad Precision Plus Protein™ Dual Xtra Standards). i denotes oxidized protein; ii alkylated protein; iii dimer of oxidized protein; iv dimer of alkylated protein. (A) SCCS. (B) GCCG. (C) GCACG. (D) GCGACG.

Table S1. DFT relative energies, GIAO ^{77}Se chemical shifts and dihedral angles for conformations of oxidized CU model 1.[a]

		ΔE+ZPE[b] (kcal/mol)	ΔG (kcal/mol)	^{77}Se ppm	Ω[c] ppm	κ[c]	ω_{pep}	χ_{SSe}	ψ_1	ϕ_2	χ_{11}	χ_{12}	χ_{21}	χ_{22}
1_{TTA}	trans twist	0.00	0.00	443	668	-0.46	-147.7	-97.0	96.0	97.1	-64.8	82.4	-59.2	78.8
1_{TTB}	trans twist	1.25	1.06	382	761	-0.51	149.1	100.5	-96.1	-89.3	59.8	-80.8	53.0	-77.4
1_{TCA}	trans chair	1.35	1.04	492	713	-0.61	149.8	-84.5	-64.1	-50.5	-44.4	94.1	-52.5	86.9
1_{TCB}	trans chair	2.68	2.37	435	877	-0.61	-151.3	86.6	59.0	59.3	45.3	-95.1	46.8	-85.1
1_{CA}	cis	0.12	-0.19	386	916	-0.79	7.9	93.0	-91.0	106.6	83.4	-79.0	-57.6	-45.6
1_{CB}	cis	0.35	-0.20	388	529	-0.56	-4.0	-92.7	-106.7	91.8	53.7	53.5	-79.6	70.3

[a] See Figure 2 of the main text for the schematics of the structures and the definition of angles.
[b] Energies were calculated relative to the 1_{TTA} structure.
[c] The span, Ω, and skew, κ, are provided in the Herzfeld-Berger convention.

Table S2. Redox potentials as measured by ESI mass spectrometry and by gel shift assays.

Protein	Redox potential determined by ESI mass spectrometry (mV)	Redox potential determined by gel shift assays (mV)
SCUS	-215	-215
SCCS	-190	-190
GCUG	-231	-231
GCCG	-207	-206
GCAUG	-255	-252
GCACG	-235	-235
GCGAUG	-293	-289
GCGACG	-269	-269

References

1. Blommel PG & Fox BG (2007) A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr Purif* 55(1):53-68.
2. Cormier CY*, et al.* (2010) Protein Structure Initiative Material Repository: an open shared public resource of structural genomics plasmids for the biological community. *Nucleic Acids Res* 38:D743-D749.
3. Gettins P & Wardlaw SA (1991) NMR relaxation properties of Se-77-labeled proteins. *J Biol Chem* 266(6):3422-3426.
4. Gettins P & Crews BC (1991) Se-77 NMR characterization of Se-77-labeled ovine erythrocyte glutathione-peroxidase. *J Biol Chem* 266(8):4804-4809.
5. Schaefer SA*, et al.* (2013) 77Se enrichment of proteins expands the biological NMR toolbox. *J Mol Biol* 425:222-231.
6. Arner ESJ, Sarioglu H, Lottspeich F, Holmgren A, & Bock A (1999) High-level expression in Escherichia coli of selenocysteine-containing rat thioredoxin reductase utilizing gene fusions with engineered bacterial-type SECIS elements and co-expression with the selA, selB and selC genes. *J Mol Biol* 292(5):1003-1016.
7. Böck A (2001) Selenium metabolism in bacteria. *Selenium: Its molecular biology and role in human health*, eds Hatfield DL, Berry MJ, & Gladyshev VN (Kluwer Academic Publishers, Boston), Vol 7-22, pp 7-22.
8. Yoshizawa S & Bock A (2009) The many levels of control on bacterial selenoprotein synthesis. *Biochim Biophys Acta* 1790(11):1404-1414.
9. Tormay P, Sawers A, & Bock A (1996) Role of stoichiometry between mRNA, translation factor SelB and selenocysteyl-tRNA in selenoprotein synthesis. *Mol Microbiol* 21(6):1253-1259.
10. Johansson L, Gafvelin G, & Arner ESJ (2005) Selenocysteine in proteins - properties and biotechnological use. *Biochim Biophys Acta* 1726(1):1-13.
11. Rengby O*, et al.* (2004) Assessment of production conditions for efficient use of Escherichia coli in high-yield heterologous recombinant selenoprotein synthesis. *Appl Environ Microbiol* 70(9):5159-5167.
12. Gladyshev VN, Khangulov SV, & Stadtman TC (1994) Nicotinic acid hydroxylase from Clostridium barkeri: Electron paramagnetic resonance studies show that selenium is coordinated with molybdenum in the catalytically active selenium-dependent enzyme. *Proc Natl Acad Sci USA* 91(1):232-236.
13. Cheng Q, Stone-Elander S, & Arner ESJ (2006) Tagging recombinant proteins with a Sel-tag for purification, labeling with electrophilic compounds or radiolabeling with C-11. *Nat Protoc* 1(2):604-613.
14. Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* 41(1):207-234.
15. Kapust RB*, et al.* (2001) Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng* 14(12):993-1000.
16. Szajewski RP & Whitesides GM (1980) Rate constants and equilibrium-constants for thiol-disulfide interchange reactions involving oxidized glutathione. *J Am Chem Soc* 102(6):2011-2026.
17. Aslund F, Berndt KD, & Holmgren A (1997) Redox potentials of glutaredoxins and other thiol-disulfide oxidoreductases of the thioredoxin superfamily determined by direct protein-protein redox equilibria. *J Biol Chem* 272(49):30780-30786.